Proxemics of heterogeneous communicative affordances in social virtual environments

Yoann Bourse
Tutored by Nina Robin

November 25, 2024

Thèse professionnelle
Mastère Spécialisé designer d'expériences immersives, interactives et ludiques
Gobelins Imagery School - CNAM-ENJMIN
Written during a 4 months internship at Albyon Paris

TABLE OF CONTENTS

0.1 English Abstract . . . 3
0.2 Résumé Français . . . 4

1 Introduction **5**

2 Glossary **6**
2.1 Acronyms . . . 6
2.2 Key vocabulary . . . 7

3 Analysis of prior literature **8**
3.1 Main axes of Virtual Reality research . . . 8
3.1.1 Immersion and identity . . . 8
3.1.2 Therapeutic VR . . . 9
3.1.3 Empathic VR for arts and civics . . . 10
3.1.4 VR as teaching and training tool . . . 11
3.1.5 Coworking in VR . . . 11
3.2 Heterogeneous devices for VR . . . 13
3.2.1 Comparative studies . . . 13
3.2.2 Synchronous co-existence . . . 15
3.3 Social virtual worlds . . . 22
3.3.1 Heterogeneous devices in social virtual reality (SVR) . . . 22
3.3.2 Non-verbal communication (NVC) in social virtual reality (SVR) . . . 22
3.3.3 Social dynamics in SVR . . . 25
3.3.4 Life of desktop virtual worlds . . . 27
3.4 Proxemics and psychology of space . . . 29
3.4.1 A brief history of proxemics and the social space of real life . . . 29
3.4.2 Proxemics of virtual space . . . 30
3.5 Conclusions and synthesis . . . 34

4 VRChat as an example of successful cohabitation **35**
4.1 VRChat's population analysis . . . 35
4.2 Ethnographic study . . . 36
4.3 A resulting conceptual framework . . . 39

5 Prototyping work **43**
5.1 Rocket Launch Room . . . 43
5.1.1 Description of developed game world . . . 43
5.1.2 User testing and experiment protocol . . . 44
5.1.3 Observations and conclusions . . . 45
5.2 Influence of objects on personal space perception . . . 51
5.2.1 Experimental setup . . . 51
5.2.2 Observations and results . . . 52

6 Conclusions and future perspectives **55**

Appendices **57**

A Ethnographic Study Participants **57**

B Ethnographic Study Questions **58**

C Prototype 1 - Rocket Launch Room **59**

D Prototype 1 - User questionnaire **60**

E Prototype 2 - Medieval fair **63**

F Prototype 2 - User questionnaire **65**

Bibliography **66**

0.1 English Abstract

The early 2020s have seen an explosion of asymmetric experiences mixing virtual reality headsets and other devices such as smartphones and desktop computers. This mix of heterogeneous devices allows interesting gameplay experiences, efficient collaborative work, and access to social virtual worlds at lower cost than a full-VR setup. However, these devices have very different interactive affordances and communicative capabilities, making their fruitful coexistence challenging.

In this masters-level thesis, I analyze of the state of the art of research drawing from various scientific fields to highlight the many challenges posed by communication through heterogeneous devices. I then proceed to an ethnographic study on the platform VRChat to understand how these interactions happen in real-life settings. I draw a few hypotheses to explain the lack of communication problems observed in the study, which I further qualitatively evaluated with two prototypes during my internship. I conclude with perspectives for further research, notably that exposure to non-colliding objects might be a driving factor in making the spatial cognition of virtual reality users more fluid and open to heterogeneous devices.

Keywords: Social Virtual Reality, Virtual Environments, Computer-Mediated Communication, Non-Verbal Communication, Asymmetric Virtual Reality, Digital Proxemics

0.2 Résumé Français

Le début des années 2020 a vu une explosion d'expériences asymétriques mêlant casques de réalité virtuelle et autres dispositifs tels que les smartphones et les ordinateurs de bureau. Ce mélange d'appareils hétérogènes permet des expériences de jeu intéressantes, un travail collaboratif efficace, et l'accès à des mondes virtuels sociaux à moindre coût qu'un entièrement dédié à la réalité virtuelle. Cependant, ces dispositifs ont des affordances interactives et des capacités communicatives très différentes, rendant leur coexistence fructueuse difficile.

Dans cette thèse professionnelle, j'analyse l'état de l'art en m'appuyant sur diverses disciplines scientifiques afin de mettre en lumière les nombreux défis posés par la communication à travers des appareils hétérogènes. Je procède ensuite à une étude ethnographique sur la plateforme VRChat pour comprendre comment ces interactions se déroulent dans des contextes réels. J'élabore quelques hypothèses pour expliquer l'absence de problèmes de communication observée au cours de l'étude, que j'ai ensuite évaluées qualitativement avec deux prototypes développés durant mon stage. Je conclus en ouvrant des perspectives de recherche future, notamment l'hypothèse que l'exposition à des objets non-collisionnants pourrait être un facteur déterminant pour rendre la cognition spatiale des utilisateurs de réalité virtuelle plus fluide et plus réceptive aux dispositifs hétérogènes.

Mots-clés: Réalité virtuelle sociale, Environnements virtuels, Communication médiée par ordinateur, Communication non-verbale, Réalité virtuelle asymétrique, Proxémie numérique

Chapter 1

Introduction

Science-fiction has long been fascinated by the idea of living together in virtual worlds, from Stanley Weinbaum's early "*Pygmalion's Spectacles*" [138] to recent blockbusters like "*Ready Player One*" [107]. The prospect is enticing: in a world simulated by computers, the only limit is one's own imagination.

However, despite significant investment from Meta since 2021 [28], Virtual Reality remains a relatively niche market. The reasons are complex, but the cost and cumbersomeness of hardware as well as the rarity of dedicated content are chief among them [91].

Yet, prior examples like *World of Warcraft* [8] or *Second Life* [62] are great examples showing that virtual reality equipment is not necessary to have a vibrant virtual world. In fact, the word metaverse is often associated to platforms like Fortnite which does not even support virtual reality at time of writing [76].

This has lead to a push for more interoperability, and in particular virtual worlds that could combine different types of hardware, as exemplified by Meta's recent announcement to open its platform to desktop users [73].

As we will see in chapter 3, orchestrating the cohabitation of users with diverse hardware capabilities is a major challenge, and the academic literature could even make such a task seem impossible.

However, VRChat [135], the leading platform of social virtual reality (SVR), is famous for its very active community where computer users and virtual reality users spend a lot of time socializing successfully together.

In this thesis, after an analysis of academic literature on the topic, I will present an ethnographic study of VRChat population to try and understand the success of this cohabitation in chapter 4. In chapter 5, I will further explore these conclusions with experimental settings that I have been working on during my 5 months long internship at Albyon [2].

Chapter 2

Glossary

In this chapter, I will present well established terms and acronyms from the academic literature that I will be using in this document. For the sake of clarity, I will keep this list as short as possible, but I will briefly mention here related concepts for readers who might be interested in diving deeper in the literature.

2.1 Acronyms

VR: Virtual Reality
A computer-generated simulation of a 3D environment.

SVR: Social Virtual Reality
A VR experience centered around socialization/chit chat. This is related to the notion of Shared Virtual Environment (SVE) that is also widely discussed in the literature. The main difference is that SVR qualifies an experience centered around socialization and chitchat, whereas SVE qualifies an environment in which discussions as well as other activities can take place. For the sake of readability, we will only use the SVR acronym.

HMD: Head-Mounted Display
A piece of hardware, also known as headset or head-worn display (HWD), used to take part in VR by replacing the user's visual field with a virtual environment. Common examples include the Meta Quest line or the PlayStation VR.

NVC: Non-Verbal Communication
The part of human communication that is not explicitly linguistic, such as gestures or facial expressions. Note that NVC can also sometimes stand in the literature for non-verbal cues, putting the emphasis on the cues used in non-verbal communication.

2.2 Key vocabulary

Desktop: This term is ambiguous with respect to VR. In this document, I will only and exclusively use it to refer to users without HMD that access the virtual environment with a computer screen, mouse and keyboard. This condition is sometimes referred to as **non-immersive VR**.
In other contexts, desktop can also refer to people using computers as engines to simulate virtual environments paired with simple HMD, as opposed to some HMD like Meta Quest 2 which have their own internal processing and do not need a computer to operate. This is sometimes referred to as **PC-VR**.

Immersion: A core concept in VR describing the psychological state of existence within an environment. It is closely related to the concepts of **embodiment** and **presence** but this document will not need this level of granularity. You can refer to the works of Tham et al. [121] or Evans et al. [27] for more.

Affordance: The inherent capabilities that a piece of technology grants its user. For instance, HMD and Desktop have widely different affordances when it comes to immersion: HMD afford the user with a full visual-field immersion.

Proxemics: A branch of anthropology dedicated to the study of the use of space in social interactions.

Heterogeneous: The term refers of course to a diversity in character or content and can be applied to many factors. In this document, it will always refer to the use of desktop and HMD at the same time in the same virtual environment. This condition is sometimes referred to as **asymmetrical**.

Coexistence: I use this term to designate the simultaneous presence of users in the same environment. Similar terms are found in the literature like **co-presence**, which focuses on the feeling of presence, or **collaboration**, which usually implies a defined goal or task. Special care should be taken with the word **co-location** as it is mostly used to refer to users that are in the same room in real life.

Chapter 3

Analysis of prior literature

3.1 Main axes of Virtual Reality research

Although VR as a consumer product might seem relatively recent, we can find related research in the fields of Human Computer Interaction (HCI) and Computer Supported Cooperative Work (CSCW) going back well into the 20th century. We can trace the beginning of this field to Sutherland's Sketchpad [115].

However, the rarity of hardware has long limited research to laboratory settings and abstract considerations. A great example is the work from Sutcliffe et al. [114] which pioneered in 2000 a framework for the evaluation of usability of VR interfaces.

The brilliant meta-analysis from Monteiro et al. [75] reveals a clear skyrocketing of related research in the second half of the 2010s decade, but they do not study the main themes of VR-related research. Let us first start with a brief informal review to help situate the present thesis in this vast panorama, while taking care to articulate how related works can inform and benefit from this work.

3.1.1 Immersion and identity

Most research around VR is dedicated in one way or another to understanding the implication of the heightened sense of immersion and how it compares to the real world. They usually center around the inter-related concepts of immersion, presence and embodiment [27, 121]. One can find a good summary of this exploration in Cummings et al. meta-analysis [19], which extracts practical considerations for hardware designers to optimize immersive experiences.

Although diving into the details of this topic can quickly fill several books, it is nonetheless important to briefly explore what lies beyond what I'll simply qualify as "various degrees of immersion" in the rest of this document. Indeed, the psychological effects of presence and immersion will greatly vary between heterogeneous devices, and this imbalance in user experience is the basic foundational context underlying all asymmetrical interactions considered in the present work.

Most papers treat immersion as a success metric but do not critically examine why it even matters in the first place. An interesting perspective comes from Slater et al. [105], who suggest that the "place illusion" can be strong enough to change the user's notion of reality itself. They comment that this is ideal, for instance, to study realistic behavior for research purposes.

But perhaps the strongest impact of such a total illusion is on self-perception, and ultimately identity. Indeed, the meta-analyses of Beyea et al. [6, 96] confirm an important effect of the degree of immersion offered by VR on the Proteus Effect, a much discussed psychological effect where a virtual avatar representation ends up influencing the person controlling the avatar.

It is interesting to note that this effect is further reinforced in the context of social virtual reality (SVR) where the avatar is not only a self-representation but takes a whole new social dimension. This explains the importance of mirrors in SVR which serve as a constant real-time feedback enhancing the Proteus Effect even further [3, 122].

The takeaway here is that the gap between heterogeneous devices can be massive. Not only will hardware condition the user's experience and sensations, it will also influence their behavior and their very identity. It's therefore no surprise to see the prevalence of research on this topic, since this effect necessarily underlies any use of VR.

That is why, although immersion might often be treated as an end in itself, it seems more accurate to think of it as the instrumental means through which VR operates. It is worth keeping this perspective in mind as I review in the rest of this chapter the main use cases of VR with explicit goals. Any benefit of VR can be understood in relation to this psychological effect of heightened immersion.

3.1.2 Therapeutic VR

VR has been studied for its therapeutic potential. The immersion derived from this technology makes it a safe environment to explore phobias, anxiety or depression. It can be a cheap and controlled medium for all sorts of exposure therapies. Fodor et al [31] offer a meta-analysis of such approaches, concluding that even though positive effects can be exhibited, the actual outcome is not as flagrant as one might expect compared to other forms of therapy. In particular, they warn against the complexity of the topic which results in many biases in studies, not the least of which is publication bias.

Another major use case of VR in the medical field is rehabilitation: VR provides an excellent support for immersive exercises which can be customized to the user and offer real-time feedback. A meta-analysis from Rutkowski et al. [99] shows that the majority of VR therapy research is focused on neurological disorders like strokes. They echo the conclusions of earlier meta-analyses [102] that significant evidence towards the superiority of VR therapies is rather hard to exhibit. The paper concludes that upper limb recovery after a stroke is the only context with sufficient evidence in favor of VR-based approaches.

Note that this kind of meta-analysis focused on the domain of physical therapy, but VR is also widely praised for being helpful in teaching social skills, not only in general, but also specifically for patient suffering from autism spectrum disorder. As early as 2002, Cobb et al. [17] were picking up on the promising potential of roleplay within controlled realistic settings. Although hardware has evolved a lot since, contemporary research like Tsai et al. [124] show that it remains one of the most effective use cases of therapeutic VR.

The picture that emerges from these reviews is that VR assisted therapy is a complex and wide field. It appears very effective in some specific cases, but their boundaries are still poorly understood. Many confounding factors make conclusions relatively muddy.

One of these factors is the social dynamic between the patient and the therapist. While the majority of past VR assisted therapies use single-user VR applications, some like [57] involved real-time intervention from the therapist. It is precisely this kind of specifically tailored experiences that seem the most promising to best leverage the benefits of therapeutic VR [31].

Therefore, as therapeutic VR grows as a domain, I would expect customization to become a core focus, and to see more and more tailored systems relying on real-time collaboration between a patient in VR and a therapist, who would potentially use a different hardware.

The current thesis is a step towards understanding the implicit impact of such systems, which can both inform their development and exhibit biases and confounding factors to help future research draw clearer conclusions.

3.1.3 Empathic VR for arts and civics

The impact of VR on psychology is also leveraged and studied outside of purely therapeutic goals. VR appears particularly effective at triggering empathy and changing social attitudes. This has wide ranging implications ranging from political to artistic.

Ventura et al. [131] offer a meta-analysis which sheds a more precise light on the capabilities of VR in that regard. They did not find significant impact of VR on empathy as a whole as defined and measured by the Interpersonal Reactivity Index (IRI), a framework developed by Mark Davis [20] around four dimensions:

- Fantasy, the tendency to transpose oneself onto fictional settings.
- Empathic Concern, feelings of sympathy for unfortunate others.
- Personal Distress, feelings of unease in tense social settings.
- Perspective Taking, the tendency to spontaneously adopt the psychological point of view of others.

However, they did find a significant impact of VR on this last dimension of perspective-taking. It seems that the extreme sense of immersion and embodiment allows people to fully inhabit different perspectives, resulting in more pro-social behavior and better attitude towards specific target groups.

These findings are confirmed by the meta-analysis of Nikolaou et al. [86] which focuses on attitudinal change and observes a significantly stronger positive effect of VR on attitudes compared to any other form of media intervention.

VR can therefore be a strong force for social good. Although most VR experiences considered in this section are single-user, multi-user experiences are a natural extension. One such example is The Machine To Be Another, from the BeAnotherLab [24], which allows a body-transfer illusion between two participants who exchange their perspective through VR. I hope the asymmetrical perspectives developed in this thesis will serve as a base to inspire future artworks. We could imagine, for instance, a HMD user inhabiting avatars transformed in real time by others on a desktop computer.

Before moving on, it is interesting to note that Ventura et al. [131] highlight as one of the major advantages of VR the possibility to study empathy with controlled conditions in a simulated environment that poses a lot less ethical limitations than the real world (the experiment can rely on virtual environments and virtual humans). This is a fascinating perspective that paves the way for more studies in the field of moral philosophy, which may involve several people coexisting in virtual spaces.

3.1.4 VR as teaching and training tool

The pedagogical virtues of VR can be considered on a broader scope than social and civic education. VR has been widely praised and studied for uses in education and training. Meta-analyses have confirmed positive impact of VR technology both on general education [71] and in specialized education such as health professional training [60] or language learning [14].

At the same time, they also highlight the difficulty of finding moderator variables or statistically significant insights to go beyond this broad positive effect, leading to the conclusion that this effect is still poorly understood.

Howard et al. [45] exhibits task-technology fit as an important moderating factor, highlighting the need for custom training programs adapted to the specificity of the topics. In parallel, Van der Meer et al. [129] present a meta-analysis which suggests that VR provides a sense of immediacy, customization, safety and motivation. Interestingly, they include a lot of papers using a non-HMD desktop setup and conclude that more research is needed on the specific point of HMD vs desktop.

This is precisely where this thesis is relevant. The asymmetry between teacher and student fits very well the model of heterogeneous devices. A common setup in this literature is a teacher on desktop guiding immersed students. It is imperative for a safe, pleasant and fruitful learning experience that the underlying social dynamics of coexistence in virtual space are well understood so as to avoid potential issues. However, this very setup is rarely studied.

3.1.5 Coworking in VR

In a similar context, but without any pedagogical goal, VR is used and studied for allowing remote work collaboration. Ens et al. [26] offer a very helpful taxonomy of existing solutions for collaboration in the broader setting of Mixed Reality, which includes VR as well as other systems. Their historical analysis confirms the boom of VR research in the 2010s, as well as the increasing importance of asymmetric systems relying on heterogeneous devices. This kind of complex collaborative structure is also identified in their conclusion as a foreseeable direction for future research.

However, despite the wide variety of papers focusing on solving technical challenges, it is relatively hard to find a summary of the benefits of VR for coworking. Most research about remote work collaboration is focused on desktop users.

Helder do Espírito Santo Vicente's master thesis [133] presents a thorough meta-analysis (as well as an additional expert user study) that results in a great summary of the benefits of VR coworking: through immersion and presence, VR allows an increase of concentration and communication that results in increased productivity.

Interestingly, this improvement of communication solves one of the main causes for stress identified by Jean-Francois Stich in his meta-analysis of virtual workspaces [110]. However, both sources identify social interactions as one of the main challenges of working in VR. In particular, [133] explicitly mentions the difficulty of communication linked to the inability of hardware affordances to capture properly all non-verbal communicative cues.

This is accentuated by numerous asymmetries that play out in a corporate context where employees occupy various positions in the hierarchy but also possess a wide range of skill set. These considerations are outside the scope of the current thesis, but since it is unlikely that all employees would use the same hardware to connect to a remote workplace, this work should help alleviate some of the communication difficulties, and perhaps even inform us more generally on how to manage asymmetries in virtual environments in general.

3.2 Heterogeneous devices for VR

It appears from our above review that the vast majority of research is focused on simple setups with either a single HMD headset or identical devices for all participants. Yet, there is a branch of the literature dedicated to asymmetrical setups using heterogeneous devices.

3.2.1 Comparative studies

There is a branch of research dedicated to comparing various HMDs with other setups along different dimensions.

Impact of differing spatial capabilities

Hepperle et al. [41] proposes one of the most exhaustive meta-analyses on the topic of comparing HMD and desktop. Interestingly, they also include real life as a point of comparison. This allows them to give more support to VR being very similar to the real world, validating the approaches described in the previous section (3.1). They take the angle of psychology, concluding that VR is especially promising for behavioral studies.

They found HMD to outperform desktop on efficiency at timed tasks, satisfaction, presence, immersion, and data visualisation. The impact on learning, workload and overall quality of experience is a lot more mixed and poorly understood.

This mixed impact on learning is echoed in the study of Yoshimura et al. [143] comparing desktop and HMD for remote class attendance. They reach the conclusion that interpersonal variability in the students (notably motion sickness) might be too high to allow a categorical conclusion.

Yet, HMD seems helpful for memorization: Ventura et al. [132] exhibit a significant improvement of memory of HMD over desktop in a task where participants had to recognize objects from a virtual environment they had been exposed to under the two modalities. Immersive HMDs allow to leverage the powerful memorization capabilities of spatial cognition [92].

The spatial capabilities of HMDs are also what makes them attractive for data visualisation. The project CodeCity, a representation of software code as buildings, is a pretty telling example. Moreno-Lumbreras et al. [78, 79] measured a much shorter completion time for various analytics and information retrieval tasks. They demonstrate that as complexity of the dataset grows so does the advantage of HMD over desktop, since HMD allows more fluent manipulation of interconnected spatial visualizations.

This effect has been observed even in other fields like the Journal of Construction Engineering and Management in which Han et al. [37] measure a significant performance improvement at error detection in design reviews from HMD users.

However, the advantage in spatial cognition of HMD might be coming at a cost. There are cases where the two-dimensional overview of the desktop screen becomes the most efficient. The work of Dong et al. [23] on map use is illuminating in that respect. They distinguish information searching and information processing based on eyes saccades and fixation. They managed to exhibit that although HMD users process information more efficiently (spend less time looking at the map), they did pretty poorly on information retrieval leading to desktop users having significantly shorter response time. Although spatial awareness allowed efficient data processing with HMD, the ease of use of desktop (not to mention the fixed small view port) seemed to have favored it for fast browsing and searching operations.

We find a similar outcome in the work of Feng et al. [30]. Studying pathfinding strategies in a vast virtual building, they measured significantly better performance coming from desktop at finding an exit route. They relate this to slight differences in pathfinding strategies: HMD have a tendency to move horizontally first, to spend more time observing, and to be slowed down crossroads. In a nutshell, it would seem that their spatial awareness negatively biased their performance.

Similarly, Berkman et al. [5] compared the HMD and desktop conditions in the puzzle game "Keep talking and nobody explodes". Although they measured a heightened sense of presence and spatial awareness in the HMD condition, it was not enough to make it a clear better device in their eyes. Not only did they not find any significant impact on game satisfaction, they also exhibited some downsides of HMD such as physical exhaustion.

Cao et al. [13] explore this further by analyzing biological signals of fatigue as well as motion sickness over long sessions of play of a driving simulation. Although they note that play sessions with HMD are shorter than desktop and traditional screen, the A/B test they ran comparing the two conditions for fixed session length only exhibited significant results with respect to motion sickness. The authors join me in wishing for more detailed investigation into the physical strain of HMD in comparison to desktop.

In conclusion, if HMD seem to present many advantages over desktop stemming from a heightened spatial awareness and immersion, there are a lot of cases where desktop can be more efficient, especially when fast interaction or overall point of view is required. Desktop has also advantages when it comes to ease of use, comfort, exhaustion or motion sickness. This is the kind of asymmetrical capabilities that will underlie interaction between heterogeneous devices.

Posture as confounding variable

Heterogeneous devices might have very different usage context. Typically, a desktop user is seated at a desk, whereas a HMD user could adopt more diverse postures. In order to understand and mitigate the role of body posture as a confounding factor, we can take a brief look at studies that focus on posture in VR.

Zielasko et al. [145] provide a thorough examination of the different factors impacted by seating or standing while wearing a HMD. Apart from the obvious gain of comfort and safety for seated users, they noted that the standing position affords more flexibility and opportunities for immersion: the user can more easily match their avatar, reducing motion sickness and increasing locomotion precision.

It is worth keeping in mind that a mismatch between the avatar and the body of the user could cause extra bias in their non-verbal communication (eye level, limb positions...). This is notably discussed by Coomer et al. [18] with respect to distance perception. They nevertheless did not find a significant difference between seated and standing for their exploration task. Tehreem et al. [120] also did not find a significant impact on posture for their task of training inside a virtual chemical plant. This suggests that the impact of body posture might not be very strong.

However, only the standing position allows the user to walk around with their whole body. A study by Srivastava et al. [108] compares HMD and desktop on a space exploration task where participants were asked to draw a map of a virtual environment or to answer a few memory questions. In this paper, HMD participants were non-ambulatory (i.e. fixed in place) and their performance were no better than using desktop computers. This suggests that ambulatory HMD is the best condition to fully leverage the gains in spatial cognition afforded by HMD.

Posture definitely impacts the spatial experience of the HMD user to varying degrees. For the purposes of the present document, we should simply keep in mind that the posture of the HMD user can have some impact, which can be summarized as bringing the experience of the sitting HMD user a bit closer to the experience of a desktop user. In the rest of this document, we will assume that HMD users are standing.

Heterogeneous devices for a single user

A few studies reflect on heterogeneous devices used sequentially by the same person. An excellent example comes from Hubenschmid et al. [47] who focus on the challenges posed by the transitions between devices. They give the example of a data analyst who might explore the data spatially with a HMD after having gathered and plotted it using a desktop computer.

This point was especially relevant during my internship (see chapter 5) where my work with heterogeneous devices was qualified of "cross-screen". This term is widely used in the industry but very ambiguous.

The term "cross-screen" is generally used in academia to refer to single-user many-devices experiences. It is often mentioned in the context of smart multimedia centers, combining more traditionally a TV viewing experience with metadata on a second screen [82]. In VR, the equivalent would be giving access to an immersed HMD user to their desktop computer or phone [144].

This is outside the scope of the present work, where each user is limited to a single device. To distinguish my work clearly from these cases, I have decided to stay away from the term "cross-screen", in order to make it clear that we are dealing with several simultaneous users.

3.2.2 Synchronous co-existence

From the point of view of this thesis, the most relevant papers are the ones that study experiences in which HMD and desktop users coexist in real time in asymmetric setups.

Research prototypes and technical challenges

Some papers center around the presentation of a prototype and focus on technical details of its implementation instead of the user psychology. They often push the boundaries of the possible, giving us a fairly broad panorama of what we can expect heterogeneous devices to be used for.

Some systems use mobile phones or tablets to allow non-HMD users to better interact with a person wearing a HMD. For instance, Marks et al. [65] propose a system where a HMD and a tablet display views of the same virtual environment, allowing an expert to guide an immersed user through bioengineering or geoscience lessons. A thesis by Tahsin Tausif [119] further explores other similar experiences.

VRInvite from Freiwald et al. [32] also allows smartphone viewports into the virtual world of a HMD user, this time in the context of rehabilitation in retirement homes. The setting of real-world collocation made the measurements of social presence rather complex, but they were able to exhibit that convenience and comfort was preferred over the increased interactivity offered by the system. They hypothesize that this is due to the fact that the gain in agency is negligible as the smartphone user is mostly a spectating bystander.

Fortunately, a viewport into a virtual world is only the most basic interaction between asymmetric devices. Dollhouse VR [49] from Ibayashi et al. allow participant on a tablet to modify the layout of a dollhouse where HMD users are immersed. Party animals by Hsu et al. [46] pushes the complexity further by developing asymmetrical party-type minigames. Interestingly, they use Microsoft Kinect camera for motion capture of non-HMD users, offering gesture controls to both conditions.

This culminates in systems like the one developed by Kim et al. [56] to allow all kinds of devices to coexist in the same virtual environment in real time. This opens the door to rich and complex multi-device asymmetrical experiences like The Last Play by Polydorou et al. [95], which presents a transmedia storytelling experience where a HMD user collaborates with a desktop user and a book reader in order to solve a game akin to an escape room.

What stands out from the brief overview of asymmetrical experiences is that they are becoming more and more common and easier to implement. This opens the door to a whole world of new experiences relying on the co-operation of heterogeneous devices, allowing each hardware to fully exploit the specificities of their affordances.

Asymmetric lab experiments

Further research has been made focusing on specific key aspects of these asymmetric experiences in order to study more precisely their intricacies. These experiments often focus on interaction between two participants collaborating in real time. In this thesis, I will put a special emphasis on cross-analyzing these studies with a psychological and sociological lense.

Benefits of asymmetry Tong et al. [123] compared symmetrical and asymmetrical experiences for data visualisation. They asked pairs of participants to collaborate on data analysis tasks, either using both the same device (HMD-HMD or Desktop-Desktop) or heterogeneous devices (HMD-Desktop). Although they found similar performances for all conditions (possibly due to the fact that the tasks were too easy), they were able to exhibit some specificities of asymmetrical settings. Asymmetrical experiences may be valuable ways to exploit the different strengths of different devices, but they require specific optimisations and accommodations which they had to put in place between two different studies. If well designed, asymmetric experiences can be very effective, illustrated by the fact that the heterogeneous conditions had the least mental load of all conditions.

ShareVR from Gugenheimer et al. [35] is another example of asymmetric experience that combines floor projection and tablet screen to give users outside the HMD more visibility into the virtual environment. They compare this peculiar setting to a more traditional controller and screen symmetrical baseline, over a competitive, collaborative and exploratory minigames. They measure significantly more enjoyment and social interaction in the case of ShareVR, while replicating the well documented boost of presence of HMD systems. It's interesting to note that they observed non-HMD users forming a certain bond with the experimenter since they were both physically outside of VR. This can be put in perspective with our hypotheses of section 3.5 about the potential tendency to segregate based on device.

Karaosmanoglu et al. [52] designed and implemented an asymmetric VR game where a desktop player guides a HMD player through a mine field. Their work also explores the potential of biometric feedback as gameplay mechanic, but the most important point for this document is that they measured a significant quantitative impact of asymmetric roles on player experience.

Interviews of participants highlighted that asymmetric settings do require time for adaptation, but that it can be a key factor of their appeal. Indeed, this adaptation is accompanied by an evolution of the social dynamics between the two players. The mutual interdependence strongly affects players' perceptions of agency, dominance and control in various ways that enhanced the game experience and increased teamwork.

McCready et al. [66] designed an escape room for three people with differing devices (desktop, HMD and smartphone) and measured no statistically significant difference in player experience between the devices. They concluded that for cooperative asymmetric gameplay, the immersion coming from devices such as HMD is not the main driver of enjoyment. Instead, they found player experience to be greatly influenced by the team dynamics and communication, the narrative of the experience, and the ludo-narrative fit between the device and the experience.

These examples demonstrate that a well designed asymmetric experience can be more valued than a fully immersive experience, both for games and more serious tasks, by leveraging different capabilities and fostering teamwork. They also highlight design challenges and social implications of such experiences that I will explore in the next paragraphs.

Design challenges The data visualisation tasks of Tong et al. [123] brought them to the conclusion that verbal communication and explicit visual cues were key to help users collaborate along different interfaces. For instance, their first study underlined the need to highlight the location of the desktop cursor to help immersed HMD user understand the current point of view of a desktop user, which they implemented in their followup studies.

This is further illustrated by the work of Saffo et al. [100] who investigate the impact of visual metaphors on group awareness, the ability to understand globally the activity of a team. They pioneer the "eyes-and-shoes" principle inspired by real world metaphors. They recommend leveraging the environment ("landscape") to provide a common reference frame, using visual cues abundantly to represent each other's attention, making sure that actions of each user have visible effects for the others, and encouraging the explicit sharing of point of views through screen sharing. They use this framework as a grid to analyze team dynamics in data analyses task to try and estimate each factor's contribution to group awareness.

Olin et al. [87] propose a very interesting study of the collaboration across heterogeneous devices. They recruited pairs of participants novice to VR and monitored them in a session articulated around two phases: simple chit-chat followed by a construction task.

Similar to Tong et al. [123], their pilot evaluations revealed a lack of shared spatial perception between the HMD user and the smartphone user. In fact, they observed a tendency for the smartphone users to invade the personal space of their HMD collaborator in the virtual environment.

In order to mitigate this, they stress out the importance of shared awareness cues. They convey deictic information (e.g. relative position) that is vital to properly act independently. They noticed it to be especially important for fast-moving or off-screen people, in order to foster a common understanding of a shared context. This lead them to research the ideal field-of-view setting to make the smartphone viewport more similar to the HMD, for instance.

Avatar's appearance is another point Olin et al. [87] focus on, since it can convey affordances and limitations of the used device, as well as enable non-verbal communication through viewing direction, hand gestures, body posture, etc.. This was especially important for users in the social interaction setting. In addition to this, they developed a compass pointing in the direction of the collaborator, and a trailing effect to indicate the direction and speed of movements.

All these studies point to the direction that proper support from the developers is required for an optimal asymmetrical experience. It seems crucial to make immediately clear the capabilities and attentional focus of each user.

Group dynamics Interestingly, Tong et al. [123] also include a user study measuring preference between symmetrical and asymmetrical setups, which indicated that user preference was directly related to the implicit social structure that is implied by the devices, as heterogeneous devices necessarily prevent a truly egalitarian structure.

It does not have to be a bad thing, asymmetries in group dynamics is not necessarily a reflection of worth or power. For instance, in the game of Karaosmanoglu et al. [52], HMD players tended to relinquish control and agency to desktop players to focus on the physical aspects of the game. This effect is of course conditioned by the game design, but the tendency for desktop users to be in a position of leadership has been reproduced and studied by Olin et al. [87].

Starting with the prior result that HMD increases social engagement, they try to analyse the details of social interactions during an unsupervised conversational phase. It turns out that users in their asymmetrical setup of HMD + smartphone adopt in virtual environment similar spatial formations than in real life. By adding varying obstacles to the virtual environment, they concluded that users position themselves face to face unless it's blocked, regardless of the device they use. This lead them to the postulate that this experience allowed even the smartphone user to feel more embodied in the virtual space, which they confirmed through user questionnaire.

In the other phase centered around a building task, they noticed two types of interaction, either preemptive or corrective, depending on whether instructions from the emerging (implicit) leader came before or after the task. As mentioned previously, they noted that the emerging leader was always the smartphone user, which seems to contradict prior work citing presence as an important factor for leadership, like the study from 2000 by Slater et al. [106].

In so doing, they exhibited factors that dwarf the importance of presence when it comes to leadership: environment mastery, information asymmetry and vertical positioning. Indeed, the smartphone users quickly adopted a top-down view while HMD users remained more stationary. Olin et al. [87] incidentally catalogued their movement into two broad categories: asteroid, meaning darting from point to point, and satellite, hovering more steadily around the task at hand. Work remains to be done to understand the interplay between those different factors.

In his master's thesis, Daniel Enriquez [25] studied both data visualisation and manipulation in asymmetric settings. The first study, centered around a data analysis task of choosing a hotel, centered around a HMD + desktop setting varying along the dimension of data representation (2D or 3D).

It seems that the 2D representation is more efficient for everyone, but the dimensionality adequation made the 3D representation a better fit for HMD than desktop. Daniel Enriquez did not find variations in group awareness or collaborative effects. However, he noticed that sharing the same data representation tended to foster a hierarchical group dynamic classified by Saffo et al. [100] as "follow the leader", whereas different representations promoted hypothesis discussions and lead to more collaborative effect.

This last example is a good cautionary tale about the effect of asymmetrical interactions on group dynamics. While it appears that non-HMD users, from their more "global" and "observatory" vantage point, seem to be favored for leadership, this last study shows that this dynamic could be reversed, for instance if no user has a clear advantage in information processing. Group dynamics is therefore a product of the hardware capabilities and the software handling of them.

Conceptual frameworks for asymmetric experiences

In order to reason about asymmetric experiences and heterogeneous devices, some researchers have lead meta-analyses and developed conceptual frameworks in which they situate the kind of projects I've mentioned above.

The pioneer of this approach is undoubtedly the Reality-Virtuality Continuum developed by Milgram and Kishino in 1994 [74]. They coined the term "mixed reality" and define an axis between the real world and fully virtual to situate all experiences. Using this terminology, we might frame the topic of this current document as "What happens when several people at distinct point of this continuum coexist in the same virtual space?".

The Reality-Virtuality Continuum is focused on single-user experiences, but Brudy et al. [10] proposed more recently one of the most exhaustive taxonomy of cross-device experiences. They take a very global point of view, placing interaction along 6 dimensions:

- Temporal: whether the different devices are used synchronously or not
- Configuration: the network relationship between the different devices
- Relationship: the number of users and their social structure
- Scale: the context of use, from personal accessories to public spaces
- Dynamics: the potential movement of the system
- Space: whether the participants are co-located or remote

They position 510 research papers from over 30 years of research in this framework in hope to bring some visibility into promising potential research agenda. This very helpful catalogue allows us to position the present thesis precisely: we're focused on multi-users synchronous systems where the users have equal permissions and are not co-located.

One of their conclusions is that most studies focus on controlled lab setups, often leaving unclear how the findings translate to real world conditions. The present thesis aim to contribute to bridging that gap.

Rogers et al. [97] established another meta-analysis, this time focused more specifically on games and gameplay. After refining an initial corpus of 481 papers to 25, they project them onto 4 famous analysis frameworks to try and identify opportunities for more complete coverage of the design space.

What stands out is an over-representation of co-located experiences, possibly echoing the importance of lab experiments. They also highlight an under-representation of player behavior analysis. While the scope of this document is not concerned with gameplay per se, it is interesting to situate ourselves in this blind spot of an already relatively small research field (25 papers).

Incidentally, Gugenheimer et al. in their ShareVR paper [35] attempt to map the space of asymmetrical experiences using four dimensions along which the interactions between players can vary: asymmetry in visualisation and interaction, coordination and dependency, power distribution and physical proximity.

This is similar to the Composite framework for Asymmetric VR (CAVR) developed by Ouverson et al. [89]. Drawing from 257 research papers, they extracted the underlying conceptual frameworks and synthesized them along 5 key axes:

- Spatial Co-presence: the feeling of being in a mixed-reality space together, extending the strict concept of physical co-location
- Transportation: the means of navigation of the shared space
- Informational Richness: the difference in information between the participants, often perceptual information offered by the device
- Team Interdependence: whether participants' goals align or not
- Balance of Power: potential differences in participants powers and abilities

They also note that asymmetry is more of a spectrum than a binary. Daniel Enriquez's work [25] is a perfect example of their conclusion: "Whereas low and high asymmetry foster dependence, medium asymmetry is uniquely suited for positive interdependence, just as is typical for collaborative teams".

While this framework is still being worked on at the time of writing, especially when it comes to evaluation strategies, it is nonetheless a tremendous resource to make sense of the current state of the art. In their conclusion, the authors note that asymmetric experiences are becoming more and more popular for a variety of reasons, not the least of which are the drawbacks of HMD such as motion sickness or adoption cost. They hope that the CAVR framework can act as a common language to describe asymmetric experiences. Interestingly, they also note the lack of understanding when it comes to the psychological and sociological implications of asymmetries. The current document hopes to shed some light on this, when information, interdependence and balance of power are equal.

In a followup paper, Ouverson et al. [90] proceed to a massive crawl of Reddit to validate their framework by comparing it to the naive understanding of asymmetric VR games. They extracted 12450 posts from Reddit and tagged them along the CAVR framework dimensions, thereby exhibiting 8 profiles of games including real world examples. Although the rest of this document will focus mostly on social hangout scenarii without any gameplay (and therefore fits more into the "cooperative" pattern of the paper), this rich panorama paints an inspiring picture of the possible levers a designer of asymmetrical experiences has at their disposal. This is especially true since the many activities offered by real world social VR platforms [72, 135] blur the line between games and non-game.

Locating this work in the broader scope of the conceptual frameworks we mentioned not only helps us not only organize our thoughts in a principled way, it also validates that we are working in a research gap. Although this thesis is a very modest contribution from a short internship, I hope that this work may help shed some light on the understudied field of synchronous remote collaboration between heterogeneous devices in real life virtual environments.

3.3 Social virtual worlds

Laboratory experiments are great for controlling the conditions at play, and we've reviewed in the previous section how informative they can be, but real life can be pretty different. In particular, the social aspects emerging from the coexistence of more than two humans in a shared space, real or virtual, cannot be understated. To understand them, we need to turn to the fields of psychology and ethnography.

3.3.1 Heterogeneous devices in social virtual reality (SVR)

There is sadly relatively little research done on existing social VR platforms like VRChat [135], and even fewer pertaining to the question of heterogeneous devices. Yet, it is a core tennant of Meta's new VR strategy [73] and a phenomenon that VRChat in particular is famous for.

Indeed, VRChat is home to a wide diversity of use cases. Far from an invisible minority, a lot of people connect to the VR platform using desktop instead of HMD. They are sometimes affectionately or pejoratively referred to as "deskies" and are a frequent topic of discussion [125–128].

I was unable to find a principled examination of this population, which is why I'll present my own in chapter 4. The closest I could find was the "Social VR Lifestyle Survey" run by independent researchers Nem and Mila (Ludmila Bredikhina, Geneva University) [83, 84]. With thousands of answers and dozens of questions, it is one of the most in-depth survey of social VR population. Although there are some sampling biases to keep in mind (the 2021 edition was almost fully Japanese), they are slowly mitigated as the sample size keeps growing.

They do not touch on the question of desktop usage directly, but some results are worth pointing out in the context of heterogeneous devices. The respondents are quite regular users: around 70% use social VR more than once a week, and almost all sessions last more than 1 hour. This gives us a glimpse into the life of established social VR.

What stands out is that it is quite different from real-life socialization: around 25% of respondents do not use their natural voice, and 67% recognize using smaller social distance in VR than real life. This last aspect is especially important since avatar contact appears to be something pretty intimate. Over 50% of respondents engage in skin-to-skin contact with their friends, a phenomenon that correlates with playtime. 43% declare being uncomfortable with unsolicited avatar touching.

After having seen in previous sections how HMD and desktop differ with respect to spatial cognition and non-verbal cues, it seems that heterogeneous devices could be the source of all sorts of problems. We need to pay special attention to these questions in order to prevent negative interactions or even new forms of harassment.

3.3.2 Non-verbal communication (NVC) in social virtual reality (SVR)

Non-verbal cues

Socialization in VR differs from real world because it is necessarily mediated, not only through hardware but also through software. Some speak of "avatar-mediated" communication. These layers of mediation are lossy and struggle to capture all the nuances of non-verbal communication (NVC): when HMD often simply capture voice and head direction, the human body naturally radiates many non-verbal cues that play a non negligible role in communication. In fact, Olin et al. [87] estimates the fraction that NVC play in communication to more than 60%, a figure they get from Hogan et al. [44].

It is hard to quantify NVC's importance in a single number. Instead, Wei et al. [137], after a meta-analysis of 32 papers over the last decade, singled out Avatar and NVC as the two main factors affecting user's communication behavior. They offer a deep review of the methodology used when studying communication in SVR which is an excellent starting point for anyone interested in this domain. They conclude that research should explore larger user groups and pay closer attention to marginalized users. The majority of papers included focused on a well defined task, though, whereas we are interested in communication without a specific goal.

An interesting approach to evaluate NVC in SVR is the one adopted by Aburumman et al. [1]. They observed participants interacting with non-human characters in SVR through several phases of discussions. They programmed those robot characters with non-verbal cues like eye movement or nodding. They were thus able to measure that the NVC cues they programmed had a significantly positive impact on the interactivity and engagement perceptions of the user. This specific paper focuses on the presence or absence of nodding behavior, but one could imagine further quantitative study trying to estimate this way the importance of different NVC cues.

Indeed, NVC is complex and relies on many different modalities. A major challenge of this field of inquiry is that most of this communication is unconscious, so technological sensors as well as researchers can only measure what they expect to find. Fang et al. [29] observed 8 participants in VR social events to draw a taxonomy of non verbal cues that they later used to survey 51 people. They focus on starting and ending social interactions, and observe a lot of similarity between non verbal cues in real life and in SVR, which makes the differences stand out. It appears than in SVR, people are more direct about starting interaction, and do it through spatialisation (simply moving closer to a desired interlocutor). This can be understood through the hypothesis that spatial position is perceived as more purposeful in a virtual environment.

However, the perceived similarity between SVR and real life might be influenced by the setting of public events that Fang et al. [29] studied. Tanenbaum et al. [118] sent groups of 3/4 researchers on the all major SVR platforms to draw a thorough inventory of NVC cues, that they divide in 4 categories.

- **Movement and Proxemic Spacing**: position and direction of the body overall
- **Facial Control**: most platforms stick to a preset selection of facial expression, but some offer lip sync or eye tracking
- **Gesture and Posture**: most platforms are limited to hand and head position, but some support torso, legs or even mood posture
- **Virtual Environment Specific NVC** is a miscellaneous category containing collisions, emotes or abstract meta-data like consent.

Noting that NVC in SVR is entirely conditioned by design decisions of platforms developer, they identify lacks ripe for future developments: facial expression control and unconscious body posture.

Maloney et al. [63] expand on this taxonomy, notably by digging deeper into gestural behavior. They proceeded to unobtrusive ethnographic observation in the SVR platform

Altspace for four weeks, followed by qualitative interviews of 30 participants. They observed a lot of similarity to offline face-to-face interaction in term of spatial behavior, hand gestures and facial expressions. However, and most interestingly, they exhibited a lot of NVC cues that do not correspond to a direct import from offline natural behavior. Instead, they seem to be extrapolated, evolved from their real life counterpart. They present their finding in a taxonomy keyed by intent:

- **Attention**: nodding, gaze direction, hand gestures
- **Approval**: emoji, applause
- **Direction attention**: pointing, patting chest
- **Social grooming**: waving, dancing, kissing
- **Provocation**: poking, pushing, moving too close
- **Disruption**: flying, extensive movement
- **Entertainment**: dancing, emojis, playing with objects

This list indeed contains a lot of behaviors that do not exist in real life context. Some are negative, and can even be vectors of harassment, especially when repeated. For this thesis, it is interesting to note that social distance, which we've seen to differ greatly between heterogeneous devices, is a way to provoke or bully people.

If NVC is a real challenge for SVR, it is perceived as positive for immersion and embodiment when done right, and valued as a rich informational bandwidth able to convey subtle and intimate information. Maloney et al. [63] found it to be a natural noninvasive way to initiate conversation, corroborating the observations of Fang et al. [29].

This paper shows brilliantly that although social norms and communication in SVR derive from the ones in the offline real world, they can take a life of their own, and they conclude that contemporary SVR now present social norms and expectations that are distinct from real life. One telling example that they dedicate some time to study is the phenomenon of muteness. It is not rare to observe on SVR platforms people who have opted out of voice chat [21]. Nem and Mila's survey [84] numbers them at around 10% of respondents.

Maloney et al. [63] analyze the benefits of this phenomenon: it affords privacy and protects marginalized users. Of course, it relies heavily on NVC, but this comes with the added effect of allowing communication through language barriers. The authors conclude that better support of this kind of use case through finer tracking and alternative accessibility options would be beneficial: sign language and text communication are still pretty challenging in SVR. Interestingly, Wei et al. [137] reach the same conclusions.

Muteness is just one part of the rich evolution of communication in SVR. In a very recent paper, Kukshinov et al. [59] study how users reinterpret affordances of SVR to develop new forms of communication. Applying Reflexive Thematic Analysis to a survey of 120 participants, they exhibit emerging behaviors that result from an appropriation of the medium by its users beyond intended capabilities: more than a third of the participants engaged in solo use of SVR, for instance. They report that while the promise of embodiment in SVR is mitigated by the limiting affordances of HMD, users developed their own workarounds, like expressing emotion in a simplified and exaggerated way. More research is needed to catalogue new emerging behaviors.

In summary, if NVC in SVR is heavily compromised by the limited affordances of HMD, we can see hints pointing at the evolution of coping strategies and the development of innovative new forms of communication. Most of it evolved naturally, but considering how communication is shaped by the affordances of the hardware, we can also wonder what purposeful intelligent design of hardware could mean for resulting interactions.

Augmented socialization

Mediated communication in SVR can be a challenge from the point of view of non-verbal communication, but the mediation is also an opportunity to augment the user experience beyond real-life human capabilities.

It is interesting in this context to consider the prevalence of mirrors in contemporary SVR. They have become a staple of VRChat. Their actual effects are complex and require more studies like Fu et al. [33] or Thibalut et al. [122]. They contribute to the Proteus effect and its modification of self perception (positively). For our purposes, however, they are at the very least crucial to give visual feedback to HMD users about their facial expression and other non-visual cues. Fu et al. [33] also show that they are only a very rudimentary palliative to missing non-verbal cues, as their 19 participants showed significantly lower conversation performance with a mirror. They seem pretty distracting and shift attention away from the interlocutor.

Researchers are opening up new perspectives when it comes to augmented socialization. We already mentioned a basic example from Olin et al. [87] which give users some visual cues to understand their partner's position and awareness all the time.

Roth et al. [98] go one step further with a quantitative study mobilizing 125 participants, visiting a virtual museum by groups of 5. Half of the participants got augmented signals such as particles marking people's attention, coloring denoting social groups, and bubbles representing eye contact. The augmented condition resulted in significantly more social presence and thought provoking experiences, and showed promises of more pro-social behavior.

This is only scratching the surface, however, since software is extremely flexible. For instance, McVeigh et al. [69] present a proof of concept presenting users with visual feedback of conversational balance between the participants. Techniques like the detection of intent to speak [15] could also be integrated in the same way.

However, although theses technologies seem promising and will be key to smooth communication through heterogeneous devices, they are still far from being widely adopted in contemporary SVR platforms.

3.3.3 Social dynamics in SVR

We can nonetheless look at real SVR platforms to extract insights about the social dynamics resulting from this mediated communication. In the cases when they include heterogeneous devices, we could even validate in the wild the behaviors reported in laboratory in section 3.2.

Jonas et al. [51] draw a taxonomy of 29 existing applications of SVR, mapping their design choices along three axes: the self, the environment, and interaction with others. Although they note the prevalence of NVC, their contribution is mostly a conceptual framework that helps structure further research on the topic. The work of McVeigh et al. [68] is another interesting take in this direction, putting special emphasis on the relationship between the design intentions of SVR environments and the resulting behaviors, often demonstrating forms of re-appropriations by the users [59].

Moustafa et al. [80] observed 17 participant groups in a longitudinal study over a 4 week period to study how real life social dynamics would transfer to the space of SVR. They conclude that the social behavior observed is very similar between real world and SVR, for which they credit the level of presence and immersion of HMD. Interestingly, they reported issues related to NVC and personal space invasion. This lead them to the conclusion that contemporary SVR is more suited to group activities than intimate conversations. However, it is important to note that these findings relate to groups that are new to SVR.

To get insights on regular users of SVR, we can turn to the work of Sykownik et al. [116] who conducted a survey over 273 participants with a special focus on the motivations driving participation in SVR. They note that participation in SVR is often intrinsically motivated, split between entertainment and socializing. Over 75% of the respondent indicated that they can experience in VR social closeness that they cannot experience in real life. They find that the main motivational driver in SVR is meeting people and staying in contact, with an emphasis on the diversity of people you can meet on the platform.

The content of their interaction might be best represented in the master's thesis of Christoffer Stockselius [111]. After semi-structured interviews with 9 users, he proceeded to one of the rare ethnographic field studies on VRChat for 10 hours of observation. Findings highlight that social interaction varies a lot between the users, making the choice of the interlocutor paramount. Socialization on SVR lean towards more anonymity and casualness, resulting in more comfortable and nonjudgmental conversations. Interestingly, he explicitly studies the question of hardware limitations and imperfect NVC. His study "showed that VRChat users rarely get affected by different kinds of limitations and difficulties. Most participants did not express any experiences of limitations or difficulties when it comes to the act of socializing in the game.". In the scope of this thesis, I hope to find out if this result can be reproduced in the specific context of heterogeneous devices.

However, it is certain that more research in digital ethnography is necessary to fully explore social dynamics of SVR, since they appear to pose significant differences from real life. In a qualitative meta-analysis, Han et al. [38] gives us an overview of the transformative effect that SVR can have on social behavior. Avatars are of course a big component of this effect: various degree of behavioral realism have been shown to affect people's NVC. Studies like Maloney et al. [63] hint at a different relation to tactile interactions inside SVR. Galace et al. [34] explore this phenomenon through the prism of available haptic technology, but the phenomenon known as "phantom touch", a synesthetic illusion where the HMD user feels a tactile sensation from a visual stimulus, hints that dedicated hardware might not even be necessary. More than a third of the respondents of the Social VR Survey [84] report at least some experience of phantom touch. Kato et al. [53] ran a survey on 341 respondents to find that pseudophysical contact under these circumstances seems to have a facilitating effect in interpersonal communication, making communication significantly easier after touch in SVR (self-reported).

Sadly, the perspective under which this has most been studied is that of moderation, governance and abuse. Virtual touch and phantom touch may well be social enhancers, but they are also vectors of bullying and harassment. Blackwell et al. [7] studied harassment in SVR through 25 participant interviews and exhibited, on top of verbal abuse and environmental abuse (images, objects...) a category of harassment in SVR that they qualify of "physical": unwanted touching, standing too close, obstructing movement, obscene gestures... I should point out that, as we've seen in previous sections, most of these can be accidental products of the coexistence of heterogeneous devices, which makes this problem potentially tricky. Fortunately, Blackwell et al. [7] reports that users are pretty good at figuring out intent and forgive mistakes of naive newcomers.

The study also notes that the SVR is especially prone to harassment as identity cues are salient and interactions very ephemeral. They noticed that communities in SVR tend to be pretty fractured, which makes harder the establishment of social norms, but also the ethnographic study of the population. They go on to discuss moderation and governance techniques that are outside the scope of this work and best addressed by the work of my predecessor schoolmate Stella Jacob [50].

Before moving on, I want to briefly address the latest work of Schulenberg et al. [103] who proceeded to a large scale survey of 223 respondents to investigate the relationship between identity and harassment in SVR. It is especially relevant for this thesis as they include the device used by participants in the different dimensions of identity they study. They were able to exhibit significant impact of this factor on how much women and non-white were harassed. They conclude that more research is needed to fully understand the extent with which device used is a moderating factor for harassment, as this is quite a novel perspective. We should keep in mind that in addition to being the potential cause of communication problems, heterogeneous devices as such could be a cause of ostracism.

3.3.4 Life of desktop virtual worlds

Fortunately, we can complement the lack of ethnographic field studies in SVR with the literature about social virtual worlds that predate VR. Games like World of Warcraft [8] or platforms like Second Life [62] were already home to persistent social communities. Even at time of writing, the word metaverse is often associated to desktop-only platforms like Fortnite or Roblox [76]. Lessons from a long history of desktop virtual worlds could be very informative for the context of this thesis.

Nardi et al. [81] is one of the most notable ethnographic study on the game World of Warcraft. Combining an ethnographic observation for weeks with 26 semi-structured interview, they observed the types of interaction and organization in the virtual world. Although most of socialization happened around activities set by the game, they noticed that downtime such as travel time gave rise to emerging discussions. They also noted that many of these social links persisted outside the game. But contrary to SVR, all social relationships within the game are influenced by the prism of the game rules and competitiveness. Most social links are created through coalescing of players around an in-game objective, which highlights how the structure of the software is crucial to create pro-social behaviors.

This bias is not present in Second Life, a social platform which does not have explicit objectives. Koutra et al. [58] proceeded to an ethnographic field study of more than a year, including players interviews. They summarize the various activities of the platform as commercial, entertainment, gaming or social, which corresponds to modern SVR platforms. They found usability of Second Life to be a problem and a significant hindrance to socialization, but the anonymity and diversity of backgrounds and geographical origins of the players to enhance sociability. Poor moderation and social rules was also detrimental to in-game socialization. In fact, the situation is very reminiscing of contemporary VRChat.

Diehl et al. [22] invite us to see Second Life as an Activity System from Cultural Historical Activity Theory. It is an analytical framework which considers human actions as complex, dynamic systems comprising interrelated components such as the subject, tools, object, community, rules, and division of labor, all interacting to achieve a shared goal. By putting the accent on the mediation structure (rules, software capabilities, participation time), it reveals intended and unintended outcomes. Within this framework, they surveyed 29 users, putting a particular emphasis on their cultural identity. They discovered that the game provided players with an ideal environment to explore intercultural communication, nurturing a cultural flow that results in the development of new forms of identity. The multicultural cross-pollination resulted in a new culture, extrapolated from the real world but which cannot be reduced to it. This dynamic seems to still be at play in contemporary SVR.

The main takeaway from former desktop worlds is that the design assumptions built into platforms shape the kind of social interaction one can see in virtual worlds. Moore et al. [77] combined field observation of 7 online games over several weeks, focusing on places they perceived to be public spaces. They studied how environment design impacts social activity in virtual space, identifying accessibility, social density, activities, and potential hosts as key factors and creating successful virtual spaces. They already highlighted as early as 2009 that avatar-mediated communication could potentially be better suited for sociability or intimacy because of the suppression of social and physical characteristics into an abstract form of interaction. This level of abstraction combined with anonymity and multiculturalism seem to be key element of the richness of social life that we observe nowadays in SVR.

In fact, from this paper to their more recent work on Fortnite [76], Moore et al. invite us to consider these virtual worlds from the point of view of "worldness", arguing that they constitute places in their own right. From semi-structured interviews with 24 children playing Fortnite, they extracted several characteristics that are key to making it feel like a world: the sense of temporal evolution, a persistent geography and a social capital accrued through play time. This perspective helps contextualize new players onboarding experience and explain the high fraction of regular users of these platform, since the resulting sense of belonging creates a long lasting attachment. Platforms become places where "participatory cultures" can develop, accompanied by their own languages and customs. But this point of view also leads us to consider virtual worlds through the lense of space, be it real or virtual, and the impact it has on human psychology and socialization.

3.4 Proxemics and psychology of space

Finally, we can turn to the real world for more context. If studies like Ventura et al. [131] or Hepperle et al. [41] draw a strong parallel between virtual reality and real life from a behavioral point of view, we might be able to find some insights in the vast literature from psychology or architecture that have studied for centuries how the environment we're in influences the way we socialize. This is especially relevant since the environment is the common frame shared by our heterogeneous devices, and the place where developers have most control over.

3.4.1 A brief history of proxemics and the social space of real life

Anthropologist Edward T. Hall introduced in his book The Hidden Dimension [36] the concept of proxemics to qualify the study of how humans perceive and use personal space in their interactions. Hall's groundbreaking work highlighted the cultural and psychological dimensions of spatial behavior, identifying distinct zones of personal space: intimate, personal, social, and public. These concentric zones govern how individuals relate to one another in different contexts. By investigating the invisible boundaries that shape human interactions, Hall revealed how cultural norms influence comfort levels and communication across various distances.

Following up on this work, Bill Hillier and Julienne Hanson introduce in their 1989 book The Social Logic of Space [43] the concept of "space syntax", a theory and method for analyzing the spatial configuration of built environments and how these configurations influence human behavior and social interactions. One notable example is their analysis of urban street networks, where they conclude that more "integrated" spaces—areas with higher visibility and more direct routes—tend to foster higher levels of social activity, while "segregated" spaces, which are less connected or visible, experience lower social engagement. This is exemplified in their study of traditional towns, where central squares and main streets, designed with greater accessibility, naturally become social hubs. They demonstrated that space is never neutral, but rather an active agent in the formation of social networks and cultural practices.

It is interesting to put this in the perspective of the pioneer art and political movement of Situationism mostly known through the works of Guy Debord. He developed the practice of derive, a random walk in a city designed to bring to light the implicit affordances present in city planning. By paying particular attention to the political implications of space, they exhibited its influence on the human psyche and socialization.

In recent years, architecture has drawn lessons from this, implicitly or explicitly, in order to design spaces more purposefully. One interesting and very telling example is the architectural design of reality television spaces, which stand as a middle-ground between environmental design in real life and in video games ([42]).

Sylaiou et al. [117] explore how this kind of practice, applied to the internet through artworks like those of Je Bek, reveals the psychogeography at work in abstract virtual spaces. Works like the theses of Allison Ong [88] or Witold van Ratingen [130] make it clear that the design of most of these spaces reflect the capitalist extractivist context in which they were built. However, contemporary social virtual reality platforms are often compared to the early internet [112] insofar as they feature heavily user generated content. Interestingly, world hopping and loitering are practices prevalent on these platforms that can be compared to situationist derive.

Combined with the obvious prevalence of spatialisation in VR, this makes the analytical lens of psychogeography especially relevant when it comes to VR. Yet, very few efforts have been made in that direction. All I could find was an essay by Rossella Salerno in Springer Tracts in Civil Engineering [101] who only briefly considers VR. The work of Andrea Shinyoung Kim [54] on the sociopolitical aspects of cyberspace also brings very interesting perspectives towards cybershamanic worldmaking without mentioning explicitly situationism. Although it is outside the topic of this thesis, I personally wanted to point out this open area of research which I believe to be especially promising.

3.4.2 Proxemics of virtual space

This is not to say that VR was completely neglected from the point of view of proxemics. The meta-analysis from Yassien et al. [142] is a good starting point to see how academic research on VR has tackled proxemics. They map 77 papers of the last decade onto a framework they propose to describe social presence.

- The first level is **self-embodiment**, referring to the style and completeness of self-representation in the virtual space.
- Proxemics is the second level, where the self encounters another being. They note that in a VR situation, it is split between the proxemics of the body of the user (co-located VR) and the proxemics of the virtual avatar. This is also the level where they situate symmetrical versus asymmetrical experiences (homogeneous or heterogeneous devices).
- They situate non-verbal communication in a third level, "interactive presence".
- Finally, the shared social presence is established on top of this base with high-level semantic considerations like the type of activity or the interdependency between the participants. A higher social presence means an increased engagement.

They confirm the research gap that this thesis is geared towards: only 9% of their corpus investigated asymmetric interactions. They attribute this to the novelty of the field, and notice a growing popularity that they predict to continue. In parallel, they deplore that gaze and facial expressions, among other non-verbal cues, are relatively understudied. Their interaction appears as a promising future research opportunity.

They note that the topics of proxemics and personal space have only recently come into focus. After it became evident that personal space played an important role in griefing and harassment [80], they observed a renewed interest in exploring the effect of SVR interactive capabilities on the participant's personal space conception. Yet, in their words, "the effect of action parameters (e.g. gestures, gaze) on personal space preferences needs further explorations, especially in asymmetric interactions".

Proxemics of VR is indeed a pretty recent field, which spun off from real life proxemics. One early example is Wilcox et al.'s study from 2006 [139] which observed the viewing distance of 16 participants watching virtual people and objects in stereoscopic vision. This archaic apparatus was enough for them to show that intrusion in a virtual personal space can cause discomfort, and that this effect is stronger if the intrusion is from a person compared to an object. They reported indications that seemed to point towards virtual personal space being similar to its real life counterpart.

Future research would confirm this trend. Hecht et al. [40] compare how proxemics from the real world translate to virtual environments. They observed the distance intuitively adopted by 20 participants during conversations in various orientations. They were thus able to conclude that personal space is circular, both in reality and in VR. They studied the influence of gender and personality of their participants but were unable to reach a clear conclusion on that front. However, they were able to show that social distance was the same when the user was moving their avatar towards a target (first person) and when they were moving the target towards them (third person). This hints at a possible conservation of personal space between heterogeneous devices.

Interestingly, they found an interpersonal distance of about 1.1m in both VR and reality. It is important to note that these findings were obtained by asking participants to position themselves for "conversation with a stranger". Kim et al. [55] extended this study to the 4 concentric zones identified in Hall's model of proxemics [36]: intimate (46cm radius), personal (120cm), social (370cm) and public (760cm). They recruited 69 subjects, some of which being lovers or friends, and compared their social distance in VR and real life.

Contrary to Hecht et al. [40], they observed that the social distance were increased in VR compared to real life, for all types of relationships. Intimate distance went from 92cm to 133cm, personal from 122cm to 207cm, and social from 171cm to 276cm. This constitutes a relatively uniform increase of 160% keeping the general shape of the personal space intact. It is crucial to keep in mind the fact that their participants were all new and unfamiliar to VR, meaning this scaling effect might be affected by apprehension towards an unknown medium. It contrasts with their observation of an increased amount of direct contact in VR (260% compared to physical space) which could suggest closer proximity in VR. In any case, expertise therefore appears as an factor significantly influencing personal space in VR, although more research on expert user would be required to complement this. We could venture the hypothesis that personal space requirement could diminish as users gain familiarity with VR, and might even result in personal space smaller in VR than real life for experts.

Other factors will influence the user's perception of their personal space. Petrizzo et al. [93] exhibited an effect of activity on personal space perception: a VR task to pull marbles towards the user resulted in a significant enlargement of personal space (more accurately, peripersonal space) by 20%, an increase they did not observe in other tasks like pushing objects away. Their results suggest that personal space does evolve, but it is a complex phenomenon conditioned by multiple factors still poorly understood.

Environment is obviously a major factor at play. The meta-analysis of Han et al. [38] reminded us that virtual environments have been shown to have transformative effects on walking behavior, group formation, physiological arousal, and more. This can be witnessed concretely, for instance, in the work of Maozhu Mao [64] who recorded people's position on VRChat and observed a significant effect of the presence of mirrors on interpersonal distance (users gather closer around the mirror).

In another paper, Han et al. [39] exhibited an effect of the presence of bystanders on VR proxemics: following student groups meeting either in isolation or in public with other groups, they found people to be moving slower and standing more closely together in the public condition. This suggests that other groups (public bystanders) acted as repellent and social inhibitors.

One of the most studied variable when it comes to personal space is gender. Iachini et al. [48] studied the comfort distance (somewhere between intimate and personal) of

40 participants to virtual characters of varying age and gender. Briefly summarized, they found that comfort distance grows with the age of the target, and is significantly bigger for men than women. They were able to reproduce these results in real world settings. Future studies like [11] show that the situation is much more complicated when considering all possible gender combinations between participants to an interaction (not to mention the distinction between player gender and avatar gender, or the impact of romantic dispositions).

Choudhary et al. [16] studied the impact of avatar scale and voice over social distance: they did not find an impact of verbal volume, but they did find that a scaled down avatar is perceived as being further away. Bonsch et al. [9] demonstrated that their participants were less keen to approach groups than individuals, and kept a larger distance to people displaying anger instead of happiness. Mello et al. [70] exhibit a significant link between non verbal cues and social distance: participants approached significantly closer to computer agents that simulated an inviting gaze instead of averting or fixed.

Out of all the studies investigating various factors of digital proxemics, the most relevant for this thesis is the work of Williamson et al. [140, 141]. In a first study [140], they analyzed the logs of 26 users during an academic workshop in Mozilla Hubs. This quantitative approach was complemented by qualitative interviews of 9 users. They observed participants standing closer to each other in breakout sessions than during the keynote, reminiscing of the results of Han et al. [39]. Heatmaps of the position data during the workshop revealed a clear and unexpected influence on the decoration on spatial positioning, as people tended to avoid standing on certain floor decals.

Interestingly for us, this early study contained participants using HMD and desktop. Williamson et al. noticed major behavioral differences between the devices, notably through the use of non-verbal cues, which prompted them to study this parameter in a followup study [141]. The paper begins with a proposed framework to understand and study digital proxemics along four main factors:

- **Activity**: interactions focused on a specific target will result in users coalescing around it, for instance.
- **Social signals**: we've already discussed the importance of non-verbal cues such as body position, head and limbs articulation, facial expression, physical appearance...
- **Audio design**: background noise and volume have been shown to impact distance judgements, and with them proxemics.
- **Environment**: they mostly discuss the results of Moore et al. [77] highlighting that more research is needed in that novel area. Indeed, if proxemics seems at first glance to translate from the real world to VR, digital environments can differ significantly from the real world (some are non-euclidean for instance).

They then leverage this framework to structure a quantitative experiment on 24 participants talking in groups of 6 or 3 for 45 minutes. The two parameters they varied between groups are the device used (HMD vs desktop) and the audio model of the world (i.e. the function that conditions how audio volume falls off as people get further away). The results are enlightening, exhibiting strong differences between HMD and desktops.

Desktop users tended to maintain a static field of view, moving less and positioning themselves in the middle of the group. In contrast, HMD users were more mobile, kept themselves further away watching from the periphery, claiming more personal space, and keeping other participants (especially speakers) in their field of view. Desktop users interacted mostly in the personal zone and even collided frequently in the intimate zone, whereas HMD users were much more likely to interact in the social zone. HMD users demonstrated more awareness of their personal space, orienting their body and actively avoiding collisions. In a nutshell, HMD users behaved much more closely to the real world than the desktop users. The paper concludes by recommending to bridge the gap between devices using software such as interface cues to make up for the imbalance in communication capabilities.

Proxemics in VR are only beginning to be understood. In first approximation, they seem imported from the proxemics of reality, confirming the strength of the "place illusion" reported by Slater et al. [105]. Yet, they are influenced by many factors, and have been showed to evolve over time. If nothing else, it seems that newcomers to VR will tend to require larger personal space. However, the context of heterogeneous devices makes this challenging. The recent work of Williamson et al. demonstrated that devices have a major impact on proxemics in VR, and desktop users will tend to invade the personal space of HMD users. New HMD users might be especially vulnerable to this kind of intrusions. In this thesis, I propose to build up on these results by comparing it with qualitative ethnographic observations in real platforms, with a focus on expert users that is sorely missing in the literature on digital proxemics.

3.5 Conclusions and synthesis

This review of the literature might be broader than it needs to be, but cross-referencing research from different domains already provides us with valuable insights.

We have reviewed how important heterogeneous device collaborations are already for the cutting edge of VR research, and asymmetric setups are bound to become more and more important as the technology progresses. Yet, they are still a research gap, especially when it comes to human behavior and real-world socialization. Fortunately, current platforms offer us successful examples to study, and our history is filled with research that informs us about the interplay between socialization and environment, be it a real or a virtual world.

In the midst of all of this, it appears pretty clear that HMD and desktop offer widely different non-verbal cues and spatial cognition. It might be most clearly seen in Williamson et al. [141], which shows us the different proxemics of HMD and desktop users. We could extrapolate from their work that desktop users will stand closer to HMD users than is acceptable, invade their personal space, or even bump into them frequently, which we've seen to be a frequent factor of discomfort or even a means of bullying.

This, combined with the importance of non-verbal cues in human communication, leads us to conclude by the hypothesis:

H0*: The difference of affordances between desktop and HMD users will cause communication problems.*

In fact, some papers like Olin et al. [87] already make some observations in this direction. Pushing this reasoning a step further, we can hypothesize the following potential consequence:

H0.1*: HMD users and desktop users will tend to congregate into segregated communities defined by hardware affordances.*

In the rest of this thesis, I will attempt to provide a modest contribution to the study of these hypotheses through the observation of real life platforms and the building of prototypes accompanied with small qualitative experiments.

Chapter 4

VRChat as an example of successful cohabitation

Although this literature review paints a picture where coexistence of heterogeneous devices in SVR seems challenging at best, there is a platform where desktop users and HMD users have been coexisting successfully for years: VRChat [135]. I started by a modest ethnographic study to understand the intricacies of this situation.

4.1 VRChat's population analysis

First of all, I wanted to make sure that VRChat had indeed the vibrant population of both HMD and desktop users that I had heard it had. VRChat is developed by the company VRChat Inc who does not have a habit of communicating their user count. Journalist Wagner James Au confirmed in early 2024 with an inside source a figure of about 10 million monthly active users [4], an order of magnitude corroborated by industry experts [12].

One of the reasons for the lack of official communication from VRChat Inc about its user count is that it is not really necessary: the community of users is very dedicated and has close ties with the company, allowing them to monitor user counts in real time directly through API access statistics [134]. At time of writing, the platform boasted an average of 80k concurrent users with frequent peaks above 100k.

API statistics also provide us with an important breakdown between two different software clients used to connect to VRChat: around 30% of users use the VRChat "Steam client", and the other 70% use the "Oculus client". These numbers are NOT the split between HMD and Desktop: the "Oculus client" is the client that runs on HMD but it also includes a marginal fraction of smartphone users (a feature in alpha test since end of 2023 [136]). The "Steam client", on the other hand, is the client that runs on desktop computers, but it includes a sizeable fraction of people using a desktop computer to run the VRChat program with a HMD attached.

There is absolutely no way to get a real desktop/HMD breakdown of the "Steam client" usage, so I proceeded to a very rough "back of the envelope" computation estimate. From the overall charts for VR games in steam [109], we can see that VRChat in VR represents around 10% of all VR usage on Steam. Playtracker [94] estimates a daily user count of VR on Steam at around 5 millions, which would put VRChat usage in VR on steam at around 0.5 million daily users. Furthermore, Playtracker estimates in the same way a daily user count of around 1 million users for the VRChat "Steam client", which suggest a roughly equal proportion of HMD users and desktop players using the "Steam client".

Put together, this suggests a 85% HMD usage on VRChat against 15% desktop. This can be contrasted with an interview of Graham Gaylor, CEO of VRChat Inc [61], who revealed in 2020 that the split at the time was 70/30 in favor of desktop users. The interview already highlighted back at the time the difference in growth speed between desktop and HMD, the latter growing more than twice as fast as desktop. A 2023 breakdown by independent researchers Nem and Mila [83, 85] confirms this trend. The actual proportion of desktop users is therefore somewhere between 10 and 70 percent.

Although this is a very wide estimate that is in the process of changing, the takeaway here is that they remain on the same order of magnitude. We are not dealing here with an overwhelming majority and a handful of marginals, but at two sizeable populations of different hardware types coexisting. And if desktop users are now a minority, they were the majority of the population during most of the platform's existence.

4.2 Ethnographic study

Ethnographic study setup

After asserting the existence of my object of study, I set out to proceed to an ethnographic study of VRChat population to understand the modalities of coexistence of HMD and desktop users. I have taken inspiration from other ethnographic studies on VRChat [3, 63, 104], although being a single contributor during a short 4 month internship I could obviously not be as thorough and deep as those researchers.

In order to mitigate these logistical constraints, I decided to stick to an exploratory qualitative analysis following the methodology of Strauss [113] and the guidelines of McDonald et al.'s [67]. I proceeded to semi-structured in-depth interviews with the goal of exhibiting new perspectives and interpretations, which could be later more thoroughly assessed with more time and manpower. All results should be understood as anecdotal and not statistically significant unless proven otherwise.

I, however, considered this as an opportunity. The lack of formal research institution allowed me to employ strategies to recruit participants that traditional studies could not. This seemed important to me because a large fraction of VRChat usage takes place in private worlds, which makes it quite hard to approach the more expert users in their "natural habitat".

I specifically wanted to focus on them because they constitute the core of VRChat user-base, and they have quite naturally the most experience with HMD and desktop coexisting. Their insightful answers as the study progressed comforted me in this approach.

The study took place over 2 weeks in August 2024, during which I interviewed 10 participants. They were recruited using the following seed methods:

- Participant was my friend
- Participant was in a public world
- Participant was streaming their gameplay on the platform *twitch.tv*

On top of which I asked each participant to introduce me to one of their friends. I prioritized breadth of exploration, and tried to penetrate as deep as possible into the social graph. The list of participants can be found in Appendix A. Appendix B contains the skeleton of questions that served as a base for the semi-structured interviews.

First impressions

One point that should be noted quite early is the level of technical knowledge and self awareness of the VRChat population. The portion of the population that is familiar with the game engine and technical terminology is quite impressive.

This echoes the findings of the Social VR, which survey measured that 75% of respondents use VRChat more than once a week [84] and more than 80% have customized their avatar themselves. In fact, 63% of the respondents even have dedicated hardware for body tracking.

We are therefore dealing with a pretty big community of well-informed expert users, which is pretty clear when talking to them. This introduces some bias in the responses, but it also is a fascinating phenomenon which well warrants further research.

Of course, the platform is also home to newer and more casual users, but they were not the focus of this investigation as I assumed people with a long history on the platform had more to say about the topic.

I did not measure personal space and interpersonal distance, but they seemed to me way smaller in my expert groups than reported in the literature in section 3.4.2. In particular, SVR does not have natural constraints on how close people can get to each other. I saw many people using extremely short personal and intimate distances, frequently overlapping their virtual bodies. This warrants dedicated follow-up investigations.

Desktop vs HMD usage

During the discussions, some recurring points were corroborated by several participants.

- It is quite frequent to see people start out on Desktop and transitioning to HMD later with expertise. This is notably the case of participants #4 and #5.
- Motion sickness is a phenomenon that constrains some people to desktops, such as participant #2. Price is another factor that makes some people stick to desktop over HMD.
- A very common scenario for people using HMD is to stick to desktop when they are working on avatar modifications or hardware customization. Desktop makes for much faster and easier platform for tests and iteration. This workflow was explicitly described by participants #1, #3, #4, #6, #7.
- There is still a lot more friction in HMD usage over desktop even for expert users, resulting in users frequently using desktop "to chat when I'm too lazy to connect with headset" (#1). For some, like #5, it seems to be the main usage mode of the platform.
- We won't dive into this point, but there is a perceived correlation between some types of HMD hardware and overall VRChat expertise. Some types of headsets are perceived as representative of more casual and younger users.

Problems of desktop and HMD cohabitation

The main takeaway of this study is that every single participant, when asked "Do you recall experiencing communication problems between people with different devices?" replied "no" in first instance. This provide quite strong evidence leaning towards the refutation of hypothesis **H0**, in accordance to the observations of Christoffer Stockselius [111]:

H0: *The difference of affordances between desktop and HMD users will cause communication problems.*

However, the semi-structured format of the interview allowed me to dive deeper and exhibit some relevant insights:

- The people I talked to have no trouble distinguishing desktop and HMD users based on limb and body motion. All of them reported seeing people with both kind of hardware very frequently, though none of them was able to give me any kind of quantification.
- Some participants noticed specifically the differences in affordances, but said they "did not matter". For instance, #7, #5 and #1 noticed that "desktop users have wider field-of-view so they come closer", echoing our conclusions from chapter 3.5. Even though nobody I talked to saw this as a nuisance, #6 reminded me that it might be a problem for people with strong sense of phantom touch.
- Some experiences in VRChat are not suitable for desktop users (games, physics-based buttons), but users know this and self-regulate accordingly.

All of this leads me to the conclusion that although the effects studied in literature are present and active, to the point of being reported by the users themselves, they might not translate into communication problems as one could imagine. I hypothesize that this is due to compensatory mechanisms and confounding factors that I will discuss in the next sections.

Hardware segregation

One possible reason for the lack of communication problems could have been segregation. Recall **H0.1**

H0.1: *HMD users and desktop users will tend to congregate into segregated communities defined by hardware affordances.*

This hypothesis is seriously challenged by the large proportion of users who use both HMD and desktop. However, probing the topic in the interviews revealed some surprising insights.

Participant #8 admitted a certain fear of desktop users, but only in public worlds. "It's often people who just download the desktop client for an evening to come troll and bully people". #8 added that an HMD can be seen as a sign of dedication. #7 goes in the same direction when saying "desktop players who see VRChat as a video game and not a social experience can be really disruptive".

This lead me to considering the question not from the perspective of hardware, but from the perspective of social dynamics, with hardware being a confounding factor only loosely correlated to the observed social groups. Instead of hardware-based segregation, #6 puts the split between "serious" and "non-serious" players, roughly echoing the beginner/expert distinction mentioned above.

Cultural adaptations

I do not have enough data to explore in any significant way the actual taxonomy of VRChat users beyond the vague draft gestured here, but whether they are called "experts", "regulars" or "serious", it seems clear that the participants I interviewed are part of a population that do not consider the difference of affordances in hardware as an issue for communication. I concluded my observations by focusing on this question to try and understand why that is the case.

From the works presented in section 3.3.2, I noticed that the non verbal cues used in VRChat can be quite different from those used in real life (headpat, cheek squeezing, etc...). #6 told me that they might even vary between friends groups and countries, much like for instance the amount of kisses on the cheeks famously differs between french regions. This brings me to the following conclusion:

I hypothesize that VRChat users have developed a distinct culture that includes specific non verbal cues that act as a language. Differences in hardware would then be negligible compared to differences in mastery of this culture. This is because this culture includes by construction both desktop and HMD players and has evolved mechanisms to mitigate affordance heterogeneity.

It is especially salient in the case of mute players (people who don't talk), a frequent phenomenon on VRChat [21,63]. They communicate exclusively through non verbal cues, so welcoming them on a platform with such diverse affordances requires high adaptability. #7 pointed out that mute players often exaggerate their body language to be understood. On the other side, #7 also gave the example of a workaround that desktop players use to gain finer control of their avatar's arms by grabbing an item, which allows them to still partake in non verbal interactions that would require arm tracking. It seems that those compensatory workarounds evolved naturally over time.

It would appear that over the years, VRChat's core users have developed a common language of non verbal cues and adaptive mechanisms to execute it. This echoes the observations of Maloney et al. [63] and Kukshinov et al. [59]. Although this remains to be confirmed by further study by people specialized in ethnography, this suggests that the heterogeneous affordances of different hardware might be overcome quite naturally given enough time for a common culture to evolve.

4.3 A resulting conceptual framework

In light of this study, we can revisit the conclusions from the scientific literature and propose a new way to conceptualize the coexistence of HMD and desktop users.

Tentative taxonomy of VRChat users

Instead of a simple dichotomy between HMD and desktop users, it seems more accurate to locate the fault line in VRChat population along the dimension of expertise. From the interviews and observations, I drew a tentative taxonomy that remains to be confirmed by more significant in depth studies:

- **Casual users**
  - On destkop: since the desktop client has the lowest barrier of entry, it is frequently used by people to occasionally check out the platform or for trolling purposes.
  - On HMD: in recent years, Meta released some relatively cheap HMD with low-end specs, making VR a lot more accessible. These devices are pretty limiting though, so they are used mostly by children or very occasional users.
- **Expert/regular users**
  - Stuck on desktop for financial or motion-sickness related reasons.
  - On HMD. This is probably a transitory step, as it appears people tend to transition quickly to the next bullet point.
  - Using a mix of HMD and desktop, to various frequency. This appears to be the majority of expert users. The reason is that as time goes, expert users naturally tend towards wanting more customization and performances out of their setup, acting as a driver for this hybridisation:
    * They will tend to get sophisticated hardware that requires phases of testing and calibration on desktop.
    * They will also tend to participate in the creator economy of VRChat: worlds and avatars are user generated, and most users find themselves at least tweaking their avatar [84]. Iterative development and testing requires them to use desktop.
    * Finally, the lower friction of the desktop client is appreciated by users who just want to talk.

A paradigm of expertise and habituation

In my observations, I could not help but note that VR in general and VRChat in particular can have quite a steep learning curve. My brain was not very efficient at switching instantly between two different 3D environments, and on top of this there were many social conventions to understand and adopt (as mentioned in section 4.2).

The taxonomy outlined above makes the user progression from casual to expert a core aspect of this research. From my interviews, it appears that there is an implicit natural funnel of engagement driven by social factors. User start out with simple hardware and standard software, and grow technical expertise as they spend time on the platform with other expert users. They will want more customization, better hardware for a more pleasant experience, but also for better non-verbal communication.

We can therefore understand the observed lack of communication problems between heterogeneous devices as the result of the learning process of new users. I hypothesize that this process is accompanied by an evolution of the user's sense of space, which has repercussions on personal space and proxemics. I further identify the following as potential drivers of this phenomenon:

- **Social mimesis**: the user's sense of space might evolve seeing other people interact in VR. Typically, they might begin considering it normal for some people to stand closer than in real life, or pass through each other's body.
- **Exposure to non-colliding entities**: frequent exposure to objects and avatars that do not have colliders (i.e. that one can simply "pass through") is likely to have an effect on the user's sense of space, in particular with respect to self-perception.
- **Experience with varied hardware**: expert users are likely to have experience with both desktop and HMD. This knowledge of various "point of view" might widen their conception of space.
- **Natural selection**: a selection pressure makes it so only people with a special tolerance for heterogeneous devices remain in this environment, the other ones giving up on the platform altogether.

These different points warrant further research. They will be partly explored in section 5 where I discuss my work at Albyon, but this thesis remains exploratory and superficial, and I do hope that all of these pathways receive further exploration from the scientific community.

Non-verbal communication and trackers

As noted above, if heterogeneous devices do not seem to cause major communication issue, they are a core part of the conversion funnel between casual and expert users. Indeed, one of the classic VRChat user journey is for people with simple HMD to invest in more sophisticated hardware to enrich their non-verbal communication cues. This puts a spotlight on a much neglected area of research.

Indeed, research (most notably chapter 3.3.2) tends to consider HMD as a single category, but recent technological advances paint a more complex picture that needs proper investigation. More and more peripherals come in to bridge the gap in non-verbal cues between VR and real world communication, as well as potentially augment users beyond what is possible in mere reality. Any study of socialization in VR needs to pay close attention to those factors:

- **Finger tracking:** Some recent VR controllers include rough finger tracking (hand open/close), but more precise solutions allow tracking for distinct phalanges.
- **Eye tracking:** Direction of gaze is an important non-verbal cue, so it's no surprise to see the growing inclusion of tracking cameras directly in HMD.
- **Face tracking:** In the same way, it is now possible to have one's avatar reflect one's facial expression directly.
- **Limb tracking:** It is now possible to track with high precision the motion of arms, legs, head, hands... It is not uncommon to see people use "full body tracking".
- **In-room tracking:** Tracking the position of the body allows for more degrees of freedom as the avatar reproduces rotations and translations of the body.
- **Brain-computer interfacing:** Some people even use brain waves sensors to control aspects of their avatars.

The specific technical solutions are outside the scope of this thesis. What should be noted, however, is that the virtual environment is home to people with a wide spectrum of non-verbal communication modalities, not simply a binary. This exposure to diversity is bound to make users more aware and tolerant when it comes to non-verbal communication differences.

Chapter 5

Prototyping work

The ethnographic study of VRChat showed us that it is possible to see fruitful coexistence of HMD and desktop users. As discussed before, it appears that this equilibrium has naturally evolved over time. However, the exact mechanisms and driving factors of this adaptation remain to be studied in more details. Furthermore, ”wait and see” is hardly a reasonable actionable advice we can give to creators of SVR environments.

During my internship at Albyon, I had the opportunity to work on a project on Meta’s social VR platform, Horizon Worlds [72]. It centered around the creation of two different game worlds, with a specific focus on heterogeneous devices. One of the desired outcomes of this R&D project was to try and extract guidelines and recommendations to favor this osmosis phenomenon in order to facilitate the coexistence of heterogeneous devices.

The context of the internship and our limited access to the Horizon Worlds platform did not allow me to run quantitatively significant observations, but I did get the chance to leverage each world for exploratory qualitative experimentation.

5.1 Rocket Launch Room

The first prototype was dedicated to exploring the basics of the behavior of users with heterogeneous devices in SVR. It took place at the same time as the ethnographic study of chapter 4 and therefore did not leverage the conclusions of this study. This work took place in a R&D program that was also exploring in parallel the themes of motion and activities in the context of 180° videos in VR, so the product and some of the tested hypotheses also reflect these themes.

5.1.1 Description of developed game world

The main artifact produced during this phase was a world (an environment) hosted on the platform Meta Horizon World [72]. It can be described as a video-watching room with light gameplay elements.

It is a big control room with one wall removed, opening onto a 180° video screen (i.e. a half-sphere) which displays rocket parts being hauled to a launch pad. Appendix C contains screenshots of the player point of view (figure C.1) and of the developer point of view (figure C.2).

The room was designed to offer several viewing areas for the video: the center of the room is the main viewing zone, it is surrounded by balconies and desks that have a clear view too. Near the back of the room, hallways go down to secluded caves that offer a more private viewing experience.

The video itself lasts 12 minutes, at the end of which the rocket can be seen to take off. The video was chosen to be low-engagement by design. As the video progresses, users have the possibility to play a mini-game: they must fix the various pieces of equipment that are breaking down around them.

Fixing is very simple, the player just needs to stand in front of the damaged equipment, where a visual effect will let them know the fix is underway. However, we designed three different difficulty levels based on the distance between the visual stimulus (flames or screen ping) and the damaged equipment. Table C in appendix C recapitulates the details.

5.1.2 User testing and experiment protocol

This test was run with the help of my colleagues. We ran the first test of this world in our office with 6 participants taken among the staff. They were familiar with VR, but not with the project or the platform Horizon Worlds.

We had a perfect split between desktop and HMD (3/3). The platform did not allow us to test an unpublished world with the desktop client which is still in development, as a result the test subjects on desktop had to use the world editor preview feature and ignore the interface. Furthermore, all the desktop accounts had to be whitelisted for this experimental feature, so the users could not operate on their own accounts.

To even things out, the HMD users were also operating on throwaway guest accounts. This introduced significant communication difficulties as nobody knew who was who. However, we'd like to point out that this is the way it would happen in an actual virtual world, so this caveat compensates for the bias of familiarity that comes with our participants knowing each other.

Another major point of note is that the participants were all co-located in the same room. We chose this consciously to make up for potential microphone difficulties on the part of the people using the desktop editor version of the software which is not meant for this, so that people could still talk even if the client did not allow it. However, it did introduce some frustration as audio was doubled for some participants.

Users played through the entirety of the video (12 minutes) and then filled the questionnaire, followed by a group debriefing session with the participants, and another one with only the organizing staff.

The conclusions and results are detailed in the following section. All the metrics reported are using a 1-5 Likert scale. The full questionnaire can be seen in appendix D.

The observed hypotheses were as follow:

- ***H1.0*** *(VIDEO) The background video is a positive factor for user engagement.*
- ***H1.1*** *(ONBOARDING) A few in-game panels repeated everywhere are enough for players to understand the tasks.*
- ***H1.2*** *(ACTIVITIES) The presence of side-activities is positive for (this) low-engagement video.*
- ***H1.3*** *(DESKTOP EXPLORATION) Desktop users are more prone to exploring the world fully.*

- ***H1.4*** *(DESKTOP MOVEMENT) Desktop users move more within viewing areas.*
- ***H1.5*** *(DESKTOP EFFICACY) Desktop users have a more global view and will be more efficient on the hardest tasks.*
- ***H1.6*** *(HETEROGENEITY CONFLICT) HMD users will report desktop users violating their personal space more than other headset users. It corresponds to* ***H0*** *from section 3.5.*
- ***H1.7*** *(HETEROGENEITY SEGREGATION) Desktop users and HMD users will tend to form separate groups. It corresponds to* ***H0.1*** *from section 3.5.*

5.1.3 Observations and conclusions

The test ran without major crashes or halting issues, but the overall atmosphere was very chaotic. We faced a few technical issues due to poor synchronization of the video between different users, but this did not impact the experience. Testers rated enjoyability as "medium" (2.67/5) and reported a lot of misunderstanding. In spite of this, the setting was well understood, and the players spent time with the various game activities. Overall, it was reminiscent of the kind of ludic chaos seen sometimes on multiplayer events, which was unexpected to us considering how slow paced the video was. Players described the activity with various degrees of seriousness, ranging from "NASA scientists" to "Despicable me minions". Players praised the environment design of the world which implicitly guided the gaze towards the video while giving the impression of a "window to the outside". Another interesting quote was that it "gave the impression that the room exists in a much bigger world", to which people replied that it created feelings of "openness" and "vastness". The world seemed pretty full even though only 6 people were in it. We hypothesize that motion, running into people, the alternation between seeing familiar and unfamiliar people might all be important factors in the perception of fullness of a world. This requires more investigation.

Methodological learnings

Testers reported that the protocol was too abrupt: they were thrown into the activity without time to get familiar with Horizon Worlds. Although this mitigated bias linked to their relationships and VR familiarity, it would have been much more pleasant for them to give them access to the tool and the world for some time before the experiment. We recommend leaving the users in a neutral space in VR for a while to get familiar with the platform before any test.

VR testing is logistically ambitious. A lot of things need to be set up, and each participant should ideally be attended to individually. A lot of testers reported frustration with the way the experiment was run. We've identified two main problems with our protocol:

- Organizer to participant ratio.
- Synchronization: the activity had a start time shared for all participants, which meant that they all needed to be ready at the same time, and forced most people to wait in case of any trouble.

180° video

Our assumptions about the video itself were confirmed by the testers, which will be useful to draw conclusions about activities in the context of a low engagement video. One of our players did not notice the video, some thought it was a background loop without progression. Only two participants could recall the full events depicted. People reported paying little attention to the video (2.33/5). This is expected, as the video is extremely slow and low action. This was confirmed by the survey rating unanimously the video as very low engagement content (1.4/5).

As for

H1.0 *(VIDEO) The background video is a positive factor for user engagement.*

Our results seem to validate it, with a caveat. Despite the passivity of the video, two users rated highly (4/5) the benefit of the video compared to a world without it. They correspond to the participants who could recall the full events depicted. The others, however, rated poorly (1.3/5) the importance of the video. This cohort had a marginally lower enjoyment of the whole experience (2.5/5) than the people enjoying the video (3/5). This confirms that a video, even a boring one, can be a positive factor, but is not necessarily the case. We hypothesized that this might be related to how much the storytelling/worldbuilding is important to the player, as opposed to players who care more about other aspects, but further investigation is needed to confirm this.

Activities

The hypothesis

H1.1 *(ONBOARDING) A few in-game panels repeated everywhere are enough for players to understand the tasks.*

is disproved for this specific test: only half the players reported understanding the easiest activity. Our panels were only enough for some players to understand the task.

Interestingly, two players understood the hard task and not the medium. It is due to the fact that the hard task was semantically similar to the easy task: the difficulty was in the distance between the signal and the fix.

H1.2 *(ACTIVITIES) The presence of side-activities is positive for (this) low-engagement video.*

on the other hand, seems validated, although further more controlled A/B testing would be needed for a significant result.

- The activities were globally rated as medium for enjoyment (2.6/5).
- The low engagement video was confirmed to not impact the activities (1.3/5), and the activities were confirmed to have a mild distracting effect on the video (3/5). This is exactly the kind of score we want: not too disturbing, but disturbing enough

to be a motivating factor driving users to the tasks. This strikes us as the right balance. In particular, nobody reported wanting less activities.

- The activities were rated as a medium factor for driving immersion (3/5), slightly more than the video. In particular, one tester strongly favored activities over video (5/5 vs 1/5). This underlines the importance of providing varied content for different profiles of players.

- One thing that observers noticed was that the activities turned a pretty empty world with a relatively boring video into a frantic collaborative game, completely transforming the virtual world from one extreme to another. As an R&D experiment, we did not aim at creating a production-ready engaging experience, but it turns out that our mediocre setup was sufficient to create a lot of emergent fun. In fact, people were active until the very end, and we had to forcibly take them out of the world.

Heterogeneous devices

The questionnaire result was very interesting when split along the variable of "device used".

- Although overall enjoyment is remarkably similar (2.6/5 average for both groups), the HMD cohort had much more polarized results (standard deviation of 1.5 versus 0.5). We can conclude that HMD is a lot more polarizing, with people strongly favoring it and people strongly disliking it (possibly for motion sickness related reasons).

- To the question "was the world adapted to the device you were using", both groups also answered similarly (3/5 for HMD, 3.3 for desktop).

- However, when asked if they thought the other group had more enjoyment, everyone agreed that HMD was the best way to experience this world: (to "do you feel like the other device is better for enjoyment, HMD group rated 2/5 and desktop 3.3/5).

This corroborates the research literature where HMD are known to be superior to desktop for immersion, embodiment and presence. We can conclude that these factors are key factors for this world.

Coming back to the question of activities and videos. Both cohorts considered HMD to be the best condition for both activities and videos:

- To "was your device the best for ...?", the HMD cohort rated 3.5/5 for activities and 3/5 for videos, whereas the desktop cohort rated 1.6/5 and 2/5 respectively.

- HMD users rated higher than desktop viewers the contribution to immersion of activities (3.3/5 vs 2.6/5) and video to a lesser degree (2.5/5 vs 2.3/5). This confirms the correlation between immersion and enjoyment for HMD usage with the activities in this world.

- We were also able to confirm that the HMD condition was more impacted by the activities disturbing the video viewing experience (3.3/5 vs 2.6/5), which confirms that our activities will have a stronger pull for HMD users than desktop.

- Some people in HMD nonetheless reported that the tunnel vision effect to alleviate motion sickness had an adverse effect on the video: they were not able to view the video in the background because they kept moving. This highlights an interesting conflict between video and motion sickness measures when the experience requires people to move.

- Finally, and perhaps most interestingly, conversations with the participants exhibited a major difference in understanding the affordances of objects in the world.

  - It was much easier for desktop users to figure out what objects were interactible, as they are highlighted and a popup effect invites them to press a button.
  - However, desktop users were very quick to give up on some tasks, assuming that they required gestures they couldn't perform. They were very quick to assume any failure was related to the lacking affordances of their device. In fact, this effect counteracted the previous one, as it seems that desktop users had more trouble fixing tasks than the HMD cohort.
  - Finally, it was easier for HMD users to miss visual effects and feedback since they can only see in front of their avatar. This highlights the need for careful design of visual language, different between the two kinds of devices.

The main takeaway here is that every part of the experience, especially item props, needs to have a great UX flow, distinct for each device. The devices being so different (pressing with the hand vs pressing a keyboard key), it is extremely important to make sure that each group can understand every object clearly. Otherwise, the variability in devices will muddy the picture and increase misunderstanding. In our case, it resulted in desktop users giving up quickly on activities, but this kind of phenomenon could lead to more severe miscommunications.

Considerations of desktop users

We expected to see most different behavior between desktop and HMD with regards to motion, since the literature reports that desktop users with a more global viewpoint (third person, flying camera) would have a tendency to move more. However, the 2 players who explored the world fully in our experiment were both HMD users. We hypothesized that this is due to the fact that the basement caves were pretty hard to find. The heightened spatial awareness of immersed HMD players helped them find small hidden hallways. Furthermore, it seemed to our observers that desktop people tended to go towards more open rooms and higher viewpoints, a point that should be investigated further.

In any case, this offers very little and very noisy data which tends to invalidate our hypotheses **H1.3** (DESKTOP EXPLORATION), **H1.4** (DESKTOP MOVEMENT) and **H1.5** (DESKTOP EFFICACY) but this doesn't seem significant enough to draw any kind of conclusion. Instead, it highlights confounding factors in our world and our experimental setup that we need to control for in order to get clearer results.

However, we would like to point out that the desktop users did seem to exhibit more leadership, in keeping with the finding of Olin et al. [87]. Desktop users were the only ones we observed coordinating and trying to organize the tasks around the activities for maximal efficiency. This is probably related to their more global viewpoint.

Social dynamics and communication

Interestingly, not all people were impacted consciously by the difference of devices. Overall, it seems to have relatively little impact during the experience (rated 2.3/5), with half the participants reporting not having noticed any difference.

At first glance, this would tend to invalidate the basic hypothesis **H0**, like the VRChat ethnographic study of Chapter 4. However, taking a closer look at what little impact it had is also very insightful.

- It appears that destkop users seem more likely to notice (and potentially suffer from) the asymmetry in devices. They reported noticing the difference more (2.6/5 vs 2/5), and were much more vocal about it in the freeform questions. Several points stand out from these questions:
  - Gestures: desktop users do not have as much flexibility with gestures as headset users. They felt this difference pretty strongly, with comments like "we were the group who couldn't raise our arms".
  - Real-world awareness: Due to being less immersed, the desktop users were more aware of the mediation related to the device, and were more likely to comment on immersion-breaking elements (lag, stutter in motion, etc. . .).
- Desktop users were also more likely to report coordination issues due to the difference in devices (3.3/5 for desktop vs 2.5/5 for HMD).

We hypothesize that this conclusion is not a generic truth but relative to the experience in the world. Our experience was perceived as valuing highly gestures, which were perceived as key to progression. We believe that this conclusion would be reversed if the gameplay favored an element for which desktop users were advantaged, like a highly exploratory task. In fact, one of the participants noted that it was "unfair" that desktop users were running faster.

This leads us to believe that the kind of experience present in a world will greatly impact the social dynamics between headset users and desktop users. The potential alignment of a perceived goal with the affordance asymmetry of devices creates a feeling of unfairness that can exacerbate segregation. This insight requires further research.

H1.6 *(HETEROGENEITY CONFLICT) HMD users will report desktop users violating their personal space more than other headset users.*

is clearly disproved in the context of our experiment, as people did not feel their personal space invaded by anyone (1.2/5). HMD users were only marginally more aware of their personal space.

We attribute this to the frantic and dynamic character of our experience, which raises the question of personal space in motion.

This could also be a bias in our user group who were familiar with VR. A player pointed out that it may be because people with HMD experience are quite used to bumping into 3D shapes without colliders, so personal space might be less important to them.

H1.7 *(HETEROGENEITY SEGREGATION) Desktop users and HMD users will tend to form separate groups.*

on the other hand got some marginal support. To the question “did you prefer spending time with people using the same device as you”, the HMD cohort’s reply was leaning yes (3.5/5) while the desktop users were of an opposite mind (1.6/5). This seems contradictory to the fact that the HMD cohort declared noticing very little difference between the devices. Therefore, we can hypothesize that this phenomenon is relatively unconscious. Furthermore we also believe this phenomenon to be related to non-verbal cues and gesture affordances, which is what HMD users with higher degree of immersion would pick up on intuitively.

Summary takeaway

Despite many caveats, this experiment was extremely informative and helpful. It concurs with my VRChat ethnographic study of Chapter 4 to provide further evidence to disprove **H0**, while highlighting the role of non-verbal cues in social dynamics between heterogeneous devices. It also made clear the need to carefully design a specific user experience adapted to the affordances of each supported device.

5.2 Influence of objects on personal space perception

My internship concluded with a final prototype that I could work on using the lessons learned from the first prototype (section 5.1) and the ethnographic study (chapter 4). After all this exploration, I took this opportunity to explore in more depth and details one specific angle of the conceptual framework detailed in section 4.3.

Section 4.3 identifies a list of potential drivers of the adaptation of users to a virtual environment with heterogeneous devices. Since my company works from the point of view of the worldbuilder, I chose to focus on objects (non-colliding entities), which is one of the actionable levers at the disposition of the world creators. It is also one of the salient points differing between heterogeneous devices that the previous prototype picked up on (section 5.1.3).

As a reminder, the underlying reasoning is that as users interact with objects that do not have hitbox or collision (that is to say object that pass through their virtual bodies), they could unconsciously associate this flexibility with their own virtual body, which would result in a more flexible conception of personal space (and in fine less problems related to heterogeneous device affordances). This presented me with a well defined concrete hypothesis to explore:

H2*: A new user will have a more adaptable personal space after having been confronted with non-colliding objects (i.e. will be less bothered by people entering their personal space).*

In the limited time I had, I was able to conduct a pilot study that showed promising results but also underlined severe limitations in my testing protocol.

5.2.1 Experimental setup

For this study, I could not turn to my colleagues or their usual network of testers because I needed participants unfamiliar with VR. Observing the influence of objects on the habituation curve of VR beginners is challenging because every person will only learn the technology once, and they will each have their own learning speed.

I decided against a between-subject protocol, since it would require hazardous comparisons between the learning curve of different people. Instead, I opted for a within-subject protocol, alternating different conditions for the same subject. This unfortunately means that a carryover effect might happen between the different condition phases. To mitigate this, I decided to alternate between conditions several times, in the hope of seeing a repeated effect on the learning speed.

I also took special care in designing the user questionnaire. To limit carryover effect, the user questionnaire included many decoy questions, as well as a generic baseline ("Do you feel familiar with VR?") to assess the overall habituation to the technology. The actual question about social distance was phased as "Do you feel like your interlocutor is standing too close?". The full questionnaire can be seen in appendix F. All the questions were answered on a 1-10 Likert scale, with a special instruction inviting the user to use the whole scale for maximum expressiveness (otherwise we wouldn't expect VR familiarity to change within one session).

The test took place on a Horizon World environment designed specifically for it based on the tutorial worlds provided by Meta. Pictures can be seen in appendix E. The environment was split in four zones, two of which had no objects at all. The two others had numerous objects lying around, for decoration (hats, swords...) or for mini-games (axe throwing, musical instrument...). An important difference between Horizon World and VRChat is that Horizon World users actually have collisions with objects around their hands, but the experimental setup tried to limit these to a minimum.

Finally, this study was aiming at measuring social distance. As such, I played the role of the interlocutor, trying to maintain a fixed distance between the participant and me that would be unequivocally shorter than real-life social distance (75cm in-game). I was using a desktop computer, while the participant was immersed in HMD. Screenshots representing it can be seen in figure E.2.

A pilot study was run in early October during the last week of my internship with two participants recruited among my friends. Each phase lasted 3 minutes, followed by a free-roam phase where the participant was allowed to explore the world indiscriminately. The questionnaire was used after each phase, with question order randomized. It was followed by an unstructured debriefing interview. One of the participant could not go through the full experiment because of motion sickness.

5.2.2 Observations and results

First of all, it is very important to preface any analysis by a disclaimer. This pilot study is a qualitative exploration and no result should be considered significant until proven properly.

We can validate the assumptions underlying the choice of the controlled social distance. All participants rated initially the interlocutor as way too close (9.5/10) and repeatedly mentioned this point in the interview section. However, this distance was bigger than the ones observed in the ethnographic study of chapter 4. This hints at it being a good choice for such experiments.

Interestingly, one participant under the impression that I was also using a HMD said "I would have been more forgiving if I had understood you'd be using a desktop computer". This shows that even beginners can be pretty forgiving and intuit the different affordances of heterogeneous devices. This also highlights the need to help users understand which devices are used by the users around them. There is almost no support for this on major SVR platforms, but it is crucial in understanding behavior and intent.

I pushed the discussion further on this question, and the participant told me that they were intuitively projecting their experience onto me. If I was wearing the same device, how could I bear standing so close? This tends to show that humans consider proxemics to be universal in first approximation. This reinforces the hypothesis of social mimesis. We can suppose that after such an initial shock due to a discrepancy of proxemics expectations, the user could renegotiate their conception of personal space to fit the default of the community.

It also turned out that the participant was assigning intentionally to this perceived breakage of proxemics norms. They felt that I was "aggressive" and doing this "on purpose". This is a good reminder of the kind of harms that can arise from heterogeneous device capabilities, and reminds us that we have yet to evaluate H0 in the context of beginners to VR. Furthermore, this also hints at the possibility that spatial positioning might be perceived as more purposeful and intentional in VR, a field of inquiry that needs more investigation.

One of the parameters I wanted to check with this pilot study was the length of the protocol. The choice of 3 minutes per phase was inspired by studies like [1,87]. It would make for a relatively short session (around 15 minutes total), both in order to avoid boredom (the task was unsupervised conversation) but most importantly motion sickness (which is especially prevalent in new users of VR).

However, during the interview section, both participants reported that they found the session too short. They highlighted that they barely had time to figure out the virtual world and the device before it was over, cautioning me against considering any result as more than noise. One of them recommended a phase of familiarization with the device and the platform before the actual start of the experiment. This raises an interesting question, however, since we're precisely interested in measuring the very beginning of the learning curve. There is probably literature on this specific question that I am unfamiliar with.

Survey results hint that despite the reported skepticism, this short experiment may be enough to observe some effects. We can see a sharp drop in motion sickness reported between the first and the second survey, especially for Participant 1 (see figure 5.1). We can also observe marginal variations of the other scores. This somewhat validates our time scale, but the on-boarding remains an open question.

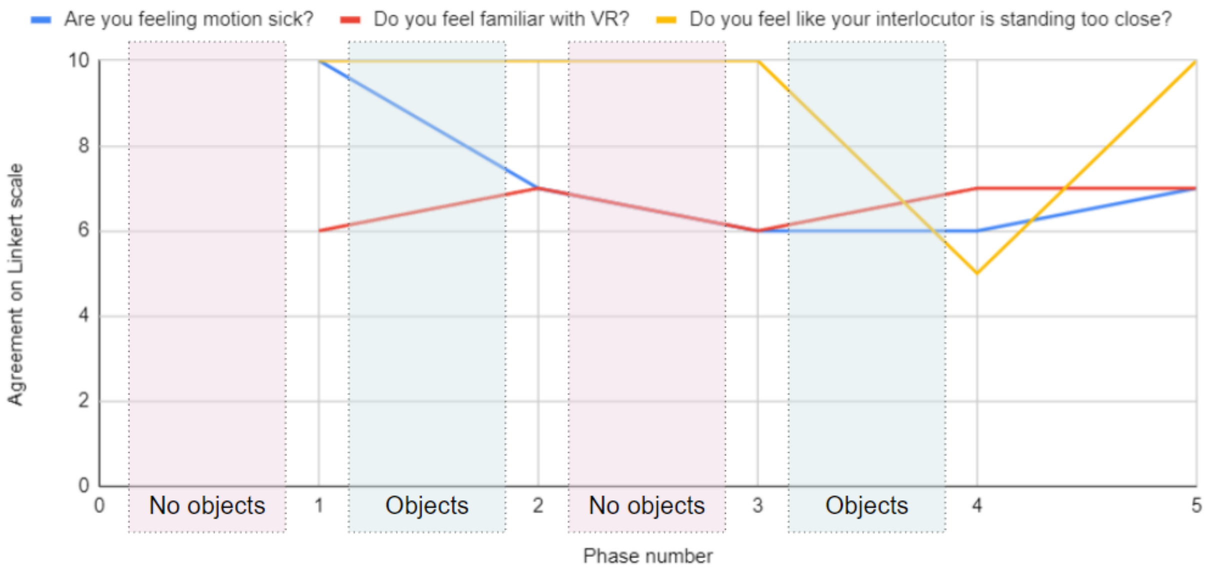

Figure 5.1: Evolution of Linkert scale answers over time for participant 1

Diving deeper in the questionnaire results, we can observe on figure 5.1 a slight uptick of VR familiarity after each phase involving objects. This leads us to the consideration that it might be possible to decorrelate and study separately the influence of object manipulation on VR familiarity and the influence of VR familiarity on proxemics. It is entirely possible that familiarity is the main factor at play and objects just act as a common catalyst for VR familiarity and proxemics. Further research will need to pay careful attention to this detail by controlling for the effect of objects manipulation on VR familiarity.

This, combined with the sharp drop to the question "Do you feel like your interlocutor is standing too close?" after the second object phase (from 10 to 5) brings some support to hypothesis H2. Interestingly, the first object phase did not result in a similar change. One possible reason that could explain this (ad hoc) is the fact that objects in the second phase were more interactive than objects in the first phase (musical instruments vs weapon props). They also required finer motion (such as fingers strumming guitar strings for instance). This makes clear the fact that further study in this direction will have to consider interactivity of the objects as an important dimension.

Although the participants, in the interviews, were begging for more time and saying things like "it was way too fast, you should ask participants to come once a week for a month", I do think the results show that an effect on very short time scales is not out of the question. Although a proper longitudinal study would of course be helpful, especially to properly deal with motion sickness, I believe a lot can still be learned by paying attention to the first minutes of a session.

From the point of view of user experience, it seems crucial to help new users navigate as fast as possible the steep learning curve of SVR. In real life, users are not patient enough to stick with a product for months before it becomes properly usable. This study might be embryonic, but it really highlights the need for SVR to have a carefully designed onboarding experience as soon as the first minutes. Objects would be paramount in this experience to help the new user get familiar with their new virtual body and its capabilities.

This is especially crucial if SVR is this user's first experience of VR. We could imagine, for instance, a single player zone to help the user get familiar with VR with computer-controlled bots, before letting them loose in a virtual world with different proxemics. In fact, proxemics probably need some dedicated explanations and tutorials. A simple warning could already go a long way to help users prepare psychologically for a world where people will regularly bump into them. The increase of heterogeneous devices in SVR will only make this kind of situation more and more likely, and there is a crucial gap in the way SVR platforms address proxemics. Who knows how many people already gave up on SVR because of this?

Chapter 6

Conclusions and future perspectives

The popularity of experiences relying on heterogeneous devices is growing. We've seen this trend in social virtual reality, but also in games or in other scenarii like therapies or classes. These asymmetric systems usually have much lower adoption cost than fully immersive ones, and allow to leverage specific capabilities of diverse hardware.

However, due to the recency of the technical solutions underlying them, the research about them is in its infancy. In particular, the impact of such systems on human behavior, social dynamics and proxemics is still poorly understood. This is most clear in the onboarding experiences of contemporary social virtual reality platform, where no mention is made of proxemics.

Yet, the review of the state of the art of human-computer interaction research of chapter 3 is clear on one point: the diverse affordances of heterogeneous devices can't help but cause imbalances in spacial cognition and group dynamics. It is a major challenge for the design of social virtual spaces, which at first glance could seem insurmountable.

Our brief ethnographic study of VRChat in chapter 4, complemented by the first prototype of Horizon World presented in section 5.1, demonstrate that it is possible for stable systems to exist, but it is not clear how such a balance can be reached purposefully. Our second prototype reminds us that it is a challenge especially hard for beginners to VR.

We've discussed several hypotheses to explain how users manage to develop a conception of social space flexible enough to accommodate heterogeneous devices: imitation through social mimesis, natural selection keeping only the most flexible users, experience with a diverse range of hardware, or exposure to non-colliding entities. The pilot study of the prototype in section 5.2 shows a marginal effect of last point, and discusses methodological insights for future investigations.

What stands out of this thesis is the need for further research in digital ethnography, especially towards the group of expert social virtual reality users. Our observations show that the naive picture we have of SVR is overly simplistic: users actually span over a wide specter of different affordances (through variety of body trackers) impacting non-verbal communication in a variety of ways. For instance, we can already observe the emergence of adaptive mechanisms and new non-verbal cues. It seems that the immersiveness of social virtual reality drastically accentuates the development of an emergent culture that was observed in early computer digital worlds like Second Life.

Many open questions remain, providing perspectives for future research. For instance, I did not have the time to investigate the relationship with proprioception (which is bound to be very informative when it comes to the impact of VR on spatial cognition), or with the perspectives opened in humanities by the field of cybernetics. Further research will indeed benefit from being trans-disciplinary. It would also be interesting to try and observe if the change of proxemics that we notice in VR impact the users in the real world. This thesis is only a first step towards the study of the impact of heterogeneous devices on digital proxemics. I hope that it paves the way for more research into this fascinating new world at the border between the virtual and the real.

Appendix A

Ethnographic Study Participants

#	Meeting circumstances	Physical country	Device(s) used	VRChat experience
1	friend of the author	France	Desktop + HMD	1 year
2	public world	Japan	Desktop	1 week
3	twitch streamer	United Kingdom	Desktop + HMD	4 years
4	twitch streamer	America	Desktop + HMD	Not communicated
5	friend of a participant	France	Desktop + HMD	2.5 years
6	friend of a participant	France	Desktop + HMD	6 years
7	friend of a participant	Denmark	Desktop + HMD	4 years
8	friend of a participant	Hungary	Desktop + HMD	”years”
9	friend of a participant	Not communicated	HMD	Not communicated
10	friend of a participant	Germany	HMD	Not communicated

Appendix B

Ethnographic Study Questions

This is the questionnaire that served as the basis for the semi-structured interviews presented in chapter 4.

1. What country is your physical body in?
2. How many hours do you spend on VRChat weekly?
3. How long have you been using VRChat?
4. How often do you use PC vs HMD headset?
5. How often do you see people using the other type of device?
6. When do you think a certain device is better as opposed to the other?
7. Do you recall experiencing communication problems between people with different devices?
8. Do you recall elements in the world that made those problems better or worse?
9. Do you recall any kind of segregation or bullying related to different devices?

Appendix C

Prototype 1 - Rocket Launch Room

Figure C.1: Player perspective of the world
Overview (left), Burning computers (center), Pinging hologram (right)

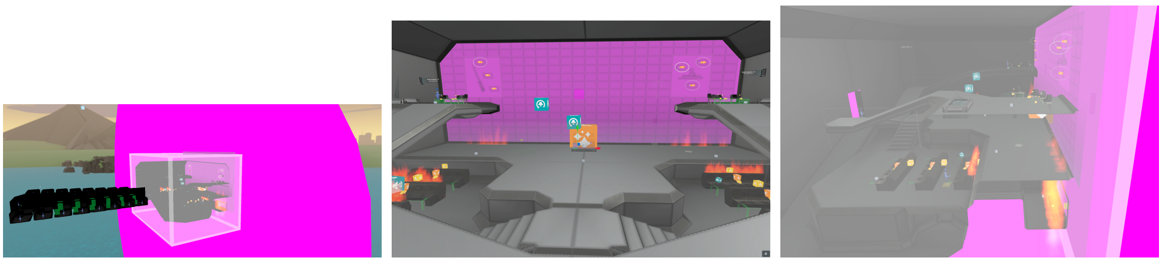

Figure C.2: Developer view of the world
Overview (left), Face view (center), Side cut (right)
The pink is a half-sphere representing the 180° video player

Level	Quantity	Task	Stimulus-fix distance
Easy	2*3	Computer catches fire	Co-located
Medium	2*3	Part of the top of the giant screen breaks	Same zone
Hard	2	Bottom of the giant screen is on fire	Different room

Table C.1: Breakdown of task difficulties

Appendix D

Prototype 1 - User questionnaire

Introduction

- Please tell us your name: ________________________
- What device were you using?
  - VR headset
  - Desktop
- How would you describe the world's activities? ________________________
- How much did you enjoy the experience?
  - Not at all 1 2 3 4 5 Very much

Video

- How long did you pay attention to the video in the background, overall?
  - Not at all 1 2 3 4 5 Very much
- In your opinion, how much did the video contribute to the experience, compared to a world without it?
  - Not at all 1 2 3 4 5 Very much
- How engaging did you find the video content?
  - Not at all 1 2 3 4 5 Very much
- Describe the content of the video in your own words. ________________________

Activities

- Did you understand how to fix the burning computers in the main room?
  - Not at all 1 2 3 4 5 Very much
- Did you understand how to fix the pinging errors at the top of the screen?
  - Not at all 1 2 3 4 5 Very much

- Did you understand how to fix the flames at the bottom of the screen?
  - Not at all 1 2 3 4 5 Very much
- Did you find the activities enjoyable?
  - Not at all 1 2 3 4 5 Very much
- Did you feel that the activities prevented you from watching the video?
  - Not at all 1 2 3 4 5 Very much
- Did you feel that the video prevented you from participating in the activities?
  - Not at all 1 2 3 4 5 Very much
- Do you think the activities helped you feel immersed in the experience?
  - Not at all 1 2 3 4 5 Very much

Different devices

This section focuses on the interaction between people on desktop and people wearing VR headsets.
SAME device refers to the one you were using (desktop or headset), *OTHER device* refers to the other.

- Did you feel any difference in behavior between people using different devices?
  - Not at all 1 2 3 4 5 Very much
- Describe all the factors that helped you distinguish between people with the SAME device and people with OTHER devices. ________________________
- Was your device the best (according to you) to participate in activities?
  - Not at all 1 2 3 4 5 Very much
- Was your device the best (according to you) to watch the video?
  - Not at all 1 2 3 4 5 Very much
- Did you feel like the world was adapted to the device you were using?
  - Not at all 1 2 3 4 5 Very much
- Were there challenges you had to face because of your device? ________________________
- Do you feel like people with the OTHER device moved more?
  - Not at all 1 2 3 4 5 Very much
- Do you feel like people with the OTHER device enjoyed the world more?
  - Not at all 1 2 3 4 5 Very much

- Did you notice more things about the difference between VR headset and desktop?

  ____

Different devices (2)
This section is centered around the interactions between different devices.

- Did the fact that people had different devices make it difficult to coordinate?
  - Not at all 1 2 3 4 5 Very much
- Did you feel like the people with VR headsets invaded your personal space?
  - Not at all 1 2 3 4 5 Very much
- Did you feel like the people with desktop computers invaded your personal space?
  - Not at all 1 2 3 4 5 Very much
- Did you prefer spending time with people sharing the SAME device as you?
  - Not at all 1 2 3 4 5 Very much

Appendix E

Prototype 2 - Medieval fair

Figure E.1: Aerial representation of the world, adapted from the "Medieval World" template
Red zones contain no objects, blue zones contain many virtual objects that have no collision logic.

Figure E.2: Social distance maintained during the observation.
Researcher point of view (left), Participant perspective (right)

Appendix F

Prototype 2 - User questionnaire

The following questions are asked after each phase, in randomized order. Users were invited to use the scale fully.

- Are you feeling motion sick?
  - Not at all 1 2 3 4 5 6 7 8 9 10 Very much
- Is the zone enjoyable?
  - Not at all 1 2 3 4 5 6 7 8 9 10 Very much
- Is the zone beautiful?
  - Not at all 1 2 3 4 5 6 7 8 9 10 Very much
- Is the zone welcoming?
  - Not at all 1 2 3 4 5 6 7 8 9 10 Very much
- Is the decoration too crowded/oppressive?
  - Not at all 1 2 3 4 5 6 7 8 9 10 Very much
- Do you feel like your interlocutor is standing too close
  - Not at all 1 2 3 4 5 6 7 8 9 10 Very much
- Are you bored?
  - Not at all 1 2 3 4 5 6 7 8 9 10 Very much
- Do you feel familiar with VR?
  - Not at all 1 2 3 4 5 6 7 8 9 10 Very much

Bibliography

[1] Nadine Aburumman, Marco Gillies, Jamie A Ward, and Antonia F de C Hamilton. Nonverbal communication in virtual reality: Nodding as a social signal in virtual interactions. *International Journal of Human-Computer Studies*, 164:102819, 2022.

[2] Albyon. Albyon studios, an immersive media group company. `https://albyon.io/`, 2019.

[3] Rasmus Asshoff. Welcome to vrchat: an ethnographic study on embodiment and immersion in virtual reality, 2022.

[4] Wagner James Au. Vrchat layoffs and metaverse challenges: An analysis, 2024. Accessed: 2024-08-30.

[5] Mehmet Berkman, Güven Çatak, and Mıstık Eremektar. Comparison of vr and desktop game user experience in a puzzle game: "keep talking and nobody explodes". *AJIT-e: Academic Journal of Information Technology*, 11(42):180–204, 2020.

[6] David Beyea, Rabindra Ratan, Yiming Lei, Hanjie Liu, Gabriel E Hales, and Chaeyun Lim. A new meta-analysis of the proteus effect: Studies in vr find stronger effect sizes. *Presence: Virtual and augmented reality*, 31:189–202, 2022.

[7] Lindsay Blackwell, Nicole Ellison, Natasha Elliott-Deflo, and Raz Schwartz. Harassment in social virtual reality: Challenges for platform governance. *Proceedings of the ACM on Human-Computer Interaction*, 3(CSCW):1–25, 2019.

[8] Blizzard Entertainment. World of warcraft. Video Game, 2004.

[9] Andrea Bönsch, Sina Radke, Heiko Overath, Laura M Asché, Jonathan Wendt, Tom Vierjahn, Ute Habel, and Torsten W Kuhlen. Social vr: How personal space is affected by virtual agents' emotions. In *2018 IEEE conference on virtual reality and 3D user interfaces (VR)*, pages 199–206. IEEE, 2018.

[10] Frederik Brudy, Christian Holz, Roman Rädle, Chi-Jui Wu, Steven Houben, Clemens Nylandsted Klokmose, and Nicolai Marquardt. Cross-device taxonomy: Survey, opportunities and challenges of interactions spanning across multiple devices. In *Proceedings of the 2019 chi conference on human factors in computing systems*, pages 1–28, 2019.

[11] Lauren E Buck, John J Rieser, Gayathri Narasimham, and Bobby Bodenheimer. Interpersonal affordances and social dynamics in collaborative immersive virtual environments: Passing together through apertures. *IEEE transactions on visualization and computer graphics*, 25(5):2123–2133, 2019.

[12] Kent Bye. Discussion about vrchat layoffs & paths to profitability with four community members. https://voicesofvr.com/1394-discussion-about-vrchat-layoffs-paths-to-profitability-with-four-community-members/, 2023. Podcast Episode 1394, Voices of VR, Accessed: 2024-08-30.

[13] Shi Cao, Krithika Nandakumar, Raiju Babu, and Benjamin Thompson. Game play in virtual reality driving simulation involving head-mounted display and comparison to desktop display. *Virtual Reality*, 24(3):503–513, 2020.

[14] Bing Chen, Yunqing Wang, and Lianghui Wang. The effects of virtual reality-assisted language learning: A meta-analysis. *Sustainability*, 14(6):3147, 2022.

[15] Jiadong Chen, Chenghao Gu, Jiayi Zhang, Zhankun Liu, and Shin 'ichi Konomi. Sensing the intentions to speak in vr group discussions. *Sensors*, 24(2), 2024.

[16] Zubin Choudhary, Matthew Gottsacker, Kangsoo Kim, Ryan Schubert, Jeanine Stefanucci, Gerd Bruder, and Gregory F Welch. Revisiting distance perception with scaled embodied cues in social virtual reality. In *2021 IEEE Virtual Reality and 3D User Interfaces (VR)*, pages 788–797. IEEE, 2021.

[17] Sue Cobb, Luke Beardon, Richard Eastgate, Tony Glover, Steven Kerr, Helen Neale, Sarah Parsons, Steve Benford, Eileen Hopkins, Peter Mitchell, et al. Applied virtual environments to support learning of social interaction skills in users with asperger's syndrome. *Digital Creativity*, 13(1):11–22, 2002.

[18] Noah Coomer, Joshua Ladd, and Betsy Williams. Virtual exploration: Seated versus standing. In *VISIGRAPP (1: GRAPP)*, pages 264–272, 2018.

[19] James J Cummings and Jeremy N Bailenson. How immersive is enough? a meta-analysis of the effect of immersive technology on user presence. *Media psychology*, 19(2):272–309, 2016.

[20] Mark H Davis. Measuring individual differences in empathy: Evidence for a multidimensional approach. *Journal of personality and social psychology*, 44(1):113, 1983.

[21] Derafog. What's the deal with vrchat mutes? https://www.youtube.com/watch?v=uucE1xd72lM, January 2023. Accessed: 2024-09-02.

[22] William C Diehl and Esther Prins. Unintended outcomes in second life: Intercultural literacy and cultural identity in a virtual world. *Language and intercultural communication*, 8(2):101–118, 2008.

[23] Weihua Dong, Tianyu Yang, Hua Liao, and Liqiu Meng. How does map use differ in virtual reality and desktop-based environments? *International Journal of Digital Earth*, 13(12):1484–1503, 2020.

[24] Denise Doyle. Empathy vr and the beanotherlab collective. *Virtual Creativity*, 10(2):191–200, 2020.

[25] Daniel Enriquez. *Investigating Asymmetric Collaboration and Interaction in Immersive Environments*. PhD thesis, Virginia Tech, 2024.

[26] Barrett Ens, Joel Lanir, Anthony Tang, Scott Bateman, Gun Lee, Thammathip Piumsomboon, and Mark Billinghurst. Revisiting collaboration through mixed reality: The evolution of groupware. *International Journal of Human-Computer Studies*, 131:81–98, 2019.

[27] Leighton Evans and Michal Rzeszewski. Hermeneutic relations in vr: Immersion, embodiment, presence and hci in vr gaming. In *International Conference on Human-Computer Interaction*, pages 23–38. Springer, 2020.

[28] Facebook, Inc. Introducing meta: A social technology company. `https://about.fb.com/news/2021/10/facebook-company-is-now-meta/`, October 2021.

[29] Jingchao Fang, Victoria Chang, Ge Gao, and Hao-Chuan Wang. Social interactions in virtual reality: What cues do people use most and how. In *Companion publication of the 2021 conference on computer supported cooperative work and social computing*, pages 49–52, 2021.

[30] Yan Feng, Dorine C Duives, and Serge P Hoogendoorn. Wayfinding behaviour in a multi-level building: A comparative study of hmd vr and desktop vr. *Advanced Engineering Informatics*, 51:101475, 2022.

[31] Liviu A Fodor, Carmen D Cotet, Pim Cuijpers, Ștefan Szamoskozi, Daniel David, and Ioana A Cristea. The effectiveness of virtual reality based interventions for symptoms of anxiety and depression: A meta-analysis. *Scientific reports*, 8(1):10323, 2018.

[32] Jann Philipp Freiwald, Sünje Gollek, and Frank Steinicke. Vr invite: A project-independent smartphone app for vr observation and interactivity. In *Human-Computer Interaction–INTERACT 2021: 18th IFIP TC 13 International Conference, Bari, Italy, August 30–September 3, 2021, Proceedings, Part I 18*, pages 352–370. Springer, 2021.

[33] Kexue Fu, Yixin Chen, Jiaxun Cao, Xin Tong, and RAY LC. " i am a mirror dweller": Probing the unique strategies users take to communicate in the context of mirrors in social virtual reality. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*, pages 1–19, 2023.

[34] Alberto Gallace and Matteo Girondini. Social touch in virtual reality. *Current Opinion in Behavioral Sciences*, 43:249–254, 2022.

[35] Jan Gugenheimer, Evgeny Stemasov, Julian Frommel, and Enrico Rukzio. Sharevr: Enabling co-located experiences for virtual reality between hmd and non-hmd users. In *Proceedings of the 2017 CHI conference on human factors in computing systems*, pages 4021–4033, 2017.

[36] Edward T Hall. The hidden dimension. *Garden City*, 1966.

[37] Bing Han and Fernanda Leite. Measuring the impact of immersive virtual reality on construction design review applications: Head-mounted display versus desktop monitor. *Journal of Construction Engineering and Management*, 147(6):04021042, 2021.

[38] Eugy Han and Jeremy N Bailenson. Social interaction in vr. In *Oxford Research Encyclopedia of Communication*. 2024.

[39] Eugy Han, Cyan DeVeaux, Mark Roman Miller, Gabriella M Harari, Jeffrey T Hancock, Nilam Ram, and Jeremy N Bailenson. Alone together, together alone: The effects of social context on nonverbal behavior in virtual reality. *PRESENCE: Virtual and Augmented Reality*, 33:425–451, 2024.

[40] Heiko Hecht, Robin Welsch, Jana Viehoff, and Matthew R Longo. The shape of personal space. *Acta psychologica*, 193:113–122, 2019.

[41] Daniel Hepperle and Matthias Wölfel. Similarities and differences between immersive virtual reality, real world, and computer screens: a systematic scoping review in human behavior studies. *Multimodal Technologies and Interaction*, 7(6):56, 2023.

[42] Stewart Hicks. How reality tv houses are built to break people, 2024. Accessed: 2024-11-15.

[43] Bill Hillier and Julienne Hanson. *The social logic of space.* Cambridge university press, 1989.

[44] Kevin Hogan and Ron Stubbs. *Can't get through: eight barriers to communication.* Pelican Publishing Company, 2003.

[45] Matt C Howard, Melissa B Gutworth, and Rick R Jacobs. A meta-analysis of virtual reality training programs. *Computers in Human Behavior*, 121:106808, 2021.

[46] Ming-Wei Hsu, Te-Yen Wu, Yu-Chian Wu, Yu-An Chen, Yu-Chih Lin, and Pin-Sung Ku. Party animals: Creating immersive gaming experience for physically co-present vr and non-vr players. In *Proceedings of the 2017 CHI Conference Extended Abstracts on Human Factors in Computing Systems*, pages 222–225, 2017.

[47] Sebastian Hubenschmid, Johannes Zagermann, Daniel Fink, Jonathan Wieland, Tiare Feuchtner, and Harald Reiterer. Towards asynchronous hybrid user interfaces for cross-reality interaction. 2021.

[48] Tina Iachini, Yann Coello, Francesca Frassinetti, Vincenzo Paolo Senese, Francesco Galante, and Gennaro Ruggiero. Peripersonal and interpersonal space in virtual and real environments: Effects of gender and age. *Journal of Environmental Psychology*, 45:154–164, 2016.

[49] Hikaru Ibayashi, Yuta Sugiura, Daisuke Sakamoto, Natsuki Miyata, Mitsunori Tada, Takashi Okuma, Takeshi Kurata, Masaaki Mochimaru, and Takeo Igarashi. Dollhouse vr: a multi-view, multi-user collaborative design workspace with vr technology. In *SIGGRAPH Asia 2015 emerging technologies*, pages 1–2. 2015.

[50] Stella Jacob. Safe dans le métavers: Enjeux, limites et perspectives de la modération et des dispositifs de signalement dans les espaces sociaux numériques. Master's thesis, Institut de Design et d'Innovation, 2021.

[51] Marcel Jonas, Steven Said, Daniel Yu, Chris Aiello, Nicholas Furlo, and Douglas Zytko. Towards a taxonomy of social vr application design. In *Extended abstracts of the annual symposium on computer-human interaction in play companion extended abstracts*, pages 437–444, 2019.

[52] Sukran Karaosmanoglu, Katja Rogers, Dennis Wolf, Enrico Rukzio, Frank Steinicke, and Lennart E Nacke. Feels like team spirit: Biometric and strategic interdependence in asymmetric multiplayer vr games. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*, pages 1–15, 2021.

[53] Yuki Kato, Kazuya Nagamachi, Maki Sugimoto, Masahiko Inami, and Michiaki Kitazaki. Vrchat ni okeru giji shintai sesshoku ga koukando to komyunikeeshon nanido ni oyobosu kouka ni kansuru chousa kenkyuu. *Nihon Baachuaru Riariti Gakkai Ronbunshi*, 26(1):22–31, 2021.

[54] Andrea Shinyoung Kim. *Virtual Worldmaking: A Phantasmal Media Approach to VRChat*. PhD thesis, Massachusetts Institute of Technology, 2021.

[55] Ikhwan Kim and Junghan Sung. New proxemics in new space: proxemics in vr. *Virtual Reality*, 28(2):85, 2024.

[56] Jongyong Kim, Jonghoon Song, Woong Seo, Insung Ihm, Seung-Hyun Yoon, and Sanghun Park. Xr framework for collaborating remote heterogeneous devices. In *2020 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW)*, pages 586–587. IEEE, 2020.

[57] Pawel Kiper, Andrzej Szczudlik, Michela Agostini, Jozef Opara, Roman Nowobilski, Laura Ventura, Paolo Tonin, and Andrea Turolla. Virtual reality for upper limb rehabilitation in subacute and chronic stroke: a randomized controlled trial. *Archives of physical medicine and rehabilitation*, 99(5):834–842, 2018.

[58] Christina Koutra. Is 'second life'taking over 'real life'? sociability and social interaction within the 'second life'environment: An electronic ethnographic study. *Computing and Information Systems*, 18(2):2–17, 2014.

[59] Eugene Kukshinov, Daniel Harley, Kata Szita, Reza Hadi Mogavi, Cayley Macarthur, and Lennart E Nacke. Disembodied, asocial, and unreal: How users reinterpret designed affordances of social vr. In *Proceedings of the 2024 ACM Designing Interactive Systems Conference*, pages 1914–1925, 2024.

[60] Bhone Myint Kyaw, Nakul Saxena, Pawel Posadzki, Jitka Vseteckova, Charoula Konstantia Nikolaou, Pradeep Paul George, Ushashree Divakar, Italo Masiello, Andrzej A Kononowicz, Nabil Zary, et al. Virtual reality for health professions education: systematic review and meta-analysis by the digital health education collaboration. *Journal of medical Internet research*, 21(1):e12959, 2019.

[61] Ben Lang. Vrchat records largest numbers yet amid coronavirus pandemic. https://www.roadtovr.com/vrchat-record-users-coronavirus/, 2020. Interview, Accessed: 2024-08-30.

[62] Linden Lab. Second life. Online Virtual World, 2003.

[63] Divine Maloney, Guo Freeman, and Donghee Yvette Wohn. "talking without a voice" understanding non-verbal communication in social virtual reality. *Proceedings of the ACM on Human-Computer Interaction*, 4(CSCW2):1–25, 2020.

[64] Maozhu Mao. How people social in vr: A behavior mapping study in virtual environments. 2021.

[65] Stefan Marks and David White. Multi-device collaboration in virtual environments. In *Proceedings of the 2020 4th International Conference on Virtual and Augmented Reality Simulations*, pages 35–38, 2020.

[66] Michael McCready, Alexandra Covaci, and Luma Tabbaa. Labxscape: A prototype for enhancing player experience in cross-reality gameplay. In *2022 IEEE International Symposium on Mixed and Augmented Reality Adjunct (ISMAR-Adjunct)*, pages 183–187. IEEE, 2022.

[67] Nora McDonald, Sarita Schoenebeck, and Andrea Forte. Reliability and inter-rater reliability in qualitative research: Norms and guidelines for cscw and hci practice. *Proceedings of the ACM on human-computer interaction*, 3(CSCW):1–23, 2019.

[68] Joshua McVeigh-Schultz, Anya Kolesnichenko, and Katherine Isbister. Shaping pro-social interaction in vr: an emerging design framework. In *Proceedings of the 2019 CHI conference on human factors in computing systems*, pages 1–12, 2019.

[69] Joshua McVeigh-Schultz, Anya Osborne, Max Kreminski, Sean Fernandes, Sabrina Fielder, Victor Li, Tym Lang, and Katherine Isbister. Social superpowers in social vr. In *Social VR: A New Medium for Remote Communication & Collaboration 2021 Workshop in 2021 ACM CHI Virtual Conference on Human Factors in Computing Systems*, pages 1–5, 2021.

[70] Beatriz Mello, Robin Welsch, Marissa Christien Verbokkem, Pascal Knierim, and Martin Johannes Dechant. Navigating the virtual gaze: Social anxiety's role in vr proxemics. In *Proceedings of the CHI Conference on Human Factors in Computing Systems*, pages 1–15, 2024.

[71] Zahira Merchant, Ernest T Goetz, Lauren Cifuentes, Wendy Keeney-Kennicutt, and Trina J Davis. Effectiveness of virtual reality-based instruction on students' learning outcomes in k-12 and higher education: A meta-analysis. *Computers & education*, 70:29–40, 2014.

[72] Meta Platforms, Inc. Horizon worlds. Online Virtual Reality Platform, 2021.

[73] Meta Quest blog. Meta horizon worlds begins expansion to mobile and web. `https://www.meta.com/en-gb/blog/quest/horizon-worlds-web-mobile-social-vr-free/`, September 2023.

[74] Paul Milgram and Fumio Kishino. A taxonomy of mixed reality visual displays. *IEICE TRANSACTIONS on Information and Systems*, 77(12):1321–1329, 1994.

[75] Pedro Monteiro, Guilherme Gonçalves, Hugo Coelho, Miguel Melo, and Maximino Bessa. Hands-free interaction in immersive virtual reality: A systematic review. *IEEE Transactions on Visualization and Computer Graphics*, 27(5):2702–2713, 2021.

[76] Kyle Moore and Marcus Carter. It is not an island it's a world: Fortnite and "worldness". In *Proceedings of the 54th Hawaii International Conference on System Sciences*, Hawaii, USA, 2021. Hawaii International Conference on System Sciences.

[77] Robert Moore, E Cabell Hankinson Gathman, and Nicolas Ducheneaut. From 3d space to third place: The social life of small virtual spaces. *Human Organization*, 68(2):230–240, 2009.

[78] David Moreno-Lumbreras, Roberto Minelli, Andrea Villaverde, Jesus M Gonzalez-Barahona, and Michele Lanza. Codecity: A comparison of on-screen and virtual reality. *Information and Software Technology*, 153:107064, 2023.

[79] David Moreno-Lumbreras, Gregorio Robles, Daniel Izquierdo-Cortázar, and Jesus M Gonzalez-Barahona. Software development metrics: to vr or not to vr. *Empirical Software Engineering*, 29(2):42, 2024.

[80] Fares Moustafa and Anthony Steed. A longitudinal study of small group interaction in social virtual reality. In *Proceedings of the 24th ACM symposium on virtual reality software and technology*, pages 1–10, 2018.

[81] Bonnie Nardi and Justin Harris. Strangers and friends: Collaborative play in world of warcraft. In *Proceedings of the 2006 20th anniversary conference on Computer supported cooperative work*, pages 149–158, 2006.

[82] Rebecca Coates Nee and David M Dozier. Second screen effects: Linking multiscreen media use to television engagement and incidental learning. *Convergence*, 23(2):214–226, 2017.

[83] Chan Nel. Social vr lifestyle survey 2021. https://medium.com/@nemchan$_{n}el/social - vr - lifestyle - survey - 2021 - 5914c35ec063, 2021. Accessed: 2024 - 08 - 30.$

[84] Chan Nel. Social vr lifestyle survey 2023. https://medium.com/@nemchan$_{n}el/social - vr - lifestyle - survey - 2023 - a61d1c4cae33, 2023. Accessed: 2024 - 08 - 30.$

[85] Chan Nel. Vrchat breaks records with 92,000 simultaneous users. https://medium.com/@nemchan$_{n}el/vrchat - breaks - records - with - 92 - 000 - simultaneous - users - 9464a33f3561, 2023. Accessed: 2024 - 08 - 30.$

[86] A Nikolaou, A Schwabe, and H Boomgaarden. Changing social attitudes with virtual reality: a systematic review and meta-analysis. *Annals of the International Communication Association*, 46(1):30–61, 2022.

[87] Patrick Aggergaard Olin, Ahmad Mohammad Issa, Tiare Feuchtner, and Kaj Grønbæk. Designing for heterogeneous cross-device collaboration and social interaction in virtual reality. In *Proceedings of the 32nd Australian Conference on Human-Computer Interaction*, pages 112–127, 2020.

[88] Allison Ong. *Immersive Realities: An Expansion Pack for Landscape Architecture*. PhD thesis, 2018.

[89] Kaitlyn M Ouverson and Stephen B Gilbert. A composite framework of co-located asymmetric virtual reality. *Proceedings of the ACM on Human-Computer Interaction*, 5(CSCW1):1–20, 2021.

[90] Kaitlyn Michelle Ouverson, Carsten Scherber, Emily Oldham, and Stephen B Gilbert. What does "asymmetric vr" mean? a directed content analysis of co-located use of vr by users on reddit. *Frontiers in Virtual Reality*, 2:765881, 2021.

[91] Perkins Coie LLP. Xr industry insider 2021 ar/vr/xr survey results. `https://www.perkinscoie.com/insights/publication/xr-industry-insider-2021-arvrxr-survey-results`, November 2021.

[92] Simon T Perrault, Eric Lecolinet, Yoann Pascal Bourse, Shengdong Zhao, and Yves Guiard. Physical loci: Leveraging spatial, object and semantic memory for command selection. In *Proceedings of the 33rd annual acm conference on human factors in computing systems*, pages 299–308, 2015.

[93] Irene Petrizzo, Kyriaki Mikellidou, Savvas Avraam, Marios Avraamides, and Roberto Arrighi. Reshaping the peripersonal space in virtual reality. *Scientific Reports*, 14(1):2438, 2024.

[94] Playtracker Team. How insight works. https://playtracker.net/info/how$_{i}nsight_{w}orks/, 2024. Accessed : 2024 - 08 - 30$.

[95] Doros Polydorou. The last play: A transmedia installation. In *International Symposium on Electronic Art*, 2022.

[96] Rabindra Ratan, David Beyea, Benjamin J Li, and Luis Graciano. Avatar characteristics induce users' behavioral conformity with small-to-medium effect sizes: a meta-analysis of the proteus effect. *Media Psychology*, 23(5):651–675, 2020.

[97] Katja Rogers, Sukran Karaosmanoglu, Dennis Wolf, Frank Steinicke, and Lennart E Nacke. A best-fit framework and systematic review of asymmetric gameplay in multiplayer virtual reality games. *Frontiers in Virtual Reality*, 2:694660, 2021.

[98] Daniel Roth, Constantin Klelnbeck, Tobias Feigl, Christopher Mutschler, and Marc Erich Latoschik. Beyond replication: Augmenting social behaviors in multi-user virtual realities. In *2018 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*, pages 215–222. IEEE, 2018.

[99] Sebastian Rutkowski, Pawel Kiper, Luisa Cacciante, Blazej Cieslik, Justyna Mazurek, Andrea Turolla, Joanna Szczepanska-Gieracha, et al. Use of virtual reality-based training in different fields of rehabilitation: A systematic review and meta-analysis. *Journal of Rehabilitation Medicine*, 52(11):1–16, 2020.

[100] David Saffo, Andrea Batch, Cody Dunne, and Niklas Elmqvist. Through their eyes and in their shoes: Providing group awareness during collaboration across virtual reality and desktop platforms. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*, pages 1–15, 2023.

[101] Rossella Salerno. From "la dérive" to virtual reality: Representing the urban walk between reality, senses and imagination. *Experiential Walks for Urban Design: Revealing, Representing, and Activating the Sensory Environment*, pages 163–174, 2021.

[102] Gustavo Saposnik, Mindy Levin, and Stroke Outcome Research Canada (SORCan) Working Group. Virtual reality in stroke rehabilitation: a meta-analysis and implications for clinicians. *Stroke*, 42(5):1380–1386, 2011.

[103] Kelsea Schulenberg, Guo Freeman, Lingyuan Li, and Bart Knijnenburg. Does who you are or appear to be matter?: Understanding identity-based harassment in social vr through the lens of (mis) perceived identity revelation. *Proceedings of the ACM on humancomputer interaction*, 2024.

[104] Ketaki Shriram and Raz Schwartz. All are welcome: Using vr ethnography to explore harassment behavior in immersive social virtual reality. In *2017 IEEE Virtual Reality (VR)*, pages 225–226. IEEE, 2017.

[105] Mel Slater, Domna Banakou, Alejandro Beacco, Jaime Gallego, Francisco Macia-Varela, and Ramon Oliva. A separate reality: An update on place illusion and plausibility in virtual reality. *Frontiers in virtual reality*, 3:914392, 2022.

[106] Mel Slater, Amela Sadagic, Martin Usoh, and Ralph Schroeder. Small-group behavior in a virtual and real environment: A comparative study. *Presence*, 9(1):37–51, 2000.

[107] Steven Spielberg, Alan Silvestri, Zak Penn, Ernest Cline, and Donald De Line. *Ready player one*. Warner Bros USA, 2018.

[108] Priyanka Srivastava, Anurag Rimzhim, Palash Vijay, Shruti Singh, and Sushil Chandra. Desktop vr is better than non-ambulatory hmd vr for spatial learning. *Frontiers in Robotics and AI*, 6:50, 2019.

[109] SteamDB. Steamdb charts - vr games. https://steamdb.info/charts/?tagid=21978, 2024. Accessed: 2024-08-30.

[110] Jean-François Stich. A review of workplace stress in the virtual office. *Intelligent Buildings International*, 12(3):208–220, 2020.

[111] Christoffer Stockselius. Social interaction in virtual reality: Users' experience of social interaction in the game vrchat, 2023.

[112] Straszfilms. Here's why vrchat is the new (old) internet. YouTube, 2020. Accessed: 2024-10-02.

[113] Anselm Strauss. *Qualitative analysis for social scientists*. Cambridge university press, 1987.

[114] Alistair G Sutcliffe and K Deol Kaur. Evaluating the usability of virtual reality user interfaces. *Behaviour & Information Technology*, 19(6):415–426, 2000.

[115] Ivan E Sutherland. A head-mounted three dimensional display. In *Proceedings of the December 9-11, 1968, fall joint computer conference, part I*, pages 757–764, 1968.

[116] Philipp Sykownik, Linda Graf, Christoph Zils, and Maic Masuch. The most social platform ever? a survey about activities & motives of social vr users. In *2021 IEEE Virtual Reality and 3D User Interfaces (VR)*, pages 546–554. IEEE, 2021.

[117] Styliani Sylaiou, Maria Chountasi, and Elena Lagoudi. Towards a digital age psychogeography and the hybrid flâneur. *Interartive: a platform for contemporary art and thought. Walking Art/Walking Aesthetics Special Issue. Available at: https://walkingart. interartive. org/2018/12/towards-a-digital-age (Accessed: 10 April 2022)*, 2018.

[118] Theresa Jean Tanenbaum, Nazely Hartoonian, and Jeffrey Bryan. " how do i make this thing smile?" an inventory of expressive nonverbal communication in commercial social virtual reality platforms. In *Proceedings of the 2020 CHI conference on human factors in computing systems*, pages 1–13, 2020.

[119] Md Tahsin Tausif. *Understanding the Effects of Tablet-based Virtual Reality (VR) Viewing Systems for an Inclusive, Cross-device Virtual Environment.* PhD thesis, Virginia Tech, 2022.

[120] Yusra Tehreem, Sofia Garcia Fracaro, Timothy Gallagher, Ryo Toyoda, Kristel Bernaerts, Jarka Glassey, Fernando Russo Abegão, Sven Wachsmuth, Michael Wilk, and Thies Pfeiffer. May i remain seated: A pilot study on the impact of reducing room-scale trainings to seated conditions for long procedural virtual reality trainings. In *2022 8th international conference on virtual reality (ICVR)*, pages 62–71. IEEE, 2022.

[121] Jason Tham, Ann Hill Duin, Laura Gee, Nathan Ernst, Bilal Abdelqader, and Megan McGrath. Understanding virtual reality: Presence, embodiment, and professional practice. *IEEE Transactions on Professional Communication*, 61(2):178–195, 2018.

[122] Mattia Thibault and Mila Bujic. " mirror dwellers": Social vr, identity and internet culture. In *DIGRA*, 2022.

[123] Wai Tong, Meng Xia, Kam Kwai Wong, Doug A Bowman, Ting-Chuen Pong, Huamin Qu, and Yalong Yang. Towards an understanding of distributed asymmetric collaborative visualization on problem-solving. In *2023 IEEE Conference Virtual Reality and 3D User Interfaces (VR)*, pages 387–397. IEEE, 2023.

[124] Wei-Te Tsai, I-Jui Lee, and Chien-Hsu Chen. Inclusion of third-person perspective in cave-like immersive 3d virtual reality role-playing games for social reciprocity training of children with an autism spectrum disorder. *Universal Access in the Information Society*, 20(2):375–389, 2021.

[125] Reddit User. Are desktop users appreciated? https://www.reddit.com/r/VRchat/comments/10i8srd/ar
$2024-09-24$.

[126] Reddit User. How many of you seen desktop users in vr chat?
https://www.reddit.com/r/VRchat/comments/1arussa/how$_{m}any_{o}f_{y}ou_{s}een_{d}esktop_{u}sers_{i}n_{v}r_{c}hat/$, 2023
$2024-09-24$.

[127] Reddit User. I'm a desktop user and i have this stigma that...
https://www.reddit.com/r/VRchat/comments/s5v8en/im$_{a}d_{e}sktop_{u}ser_{a}nd_{i}h_{a}ve_{t}his_{s}tigma_{t}hat/$, 2023.A
$2024-09-24$.

[128] Reddit User. Why are some people mean to desktops?
https://www.reddit.com/r/VRchat/comments/18msjln/why$_{a}re_{s}ome_{p}eople_{m}ean_{t}o_{d}esktops/$, 2023.$Acce$
$2024-09-24$.

[129] Nesse van der Meer, Vivian van der Werf, Willem-Paul Brinkman, and Marcus Specht. Virtual reality and collaborative learning: A systematic literature review. *Frontiers in Virtual Reality*, 4:1159905, 2023.

[130] Witold Jerzy van Ratingen. Loitering with intent. 2017.

[131] Sara Ventura, Laura Badenes-Ribera, Rocio Herrero, Ausias Cebolla, Laura Galiana, and Rosa Baños. Virtual reality as a medium to elicit empathy: A meta-analysis. *Cyberpsychology, Behavior, and Social Networking*, 23(10):667–676, 2020.

[132] Sara Ventura, Eleonora Brivio, Giuseppe Riva, and Rosa M Baños. Immersive versus non-immersive experience: Exploring the feasibility of memory assessment through 360 technology. *Frontiers in psychology*, 10:2509, 2019.

[133] Hélder do Espírito Santo Vicente. Virtual reality as a mitigation for remote work issues: Advantages, disadvantages and challenges. Master's thesis, ISCTE - University Institute of Lisbon, 2023.

[134] VRChat Community Metrics. Vrchat community metrics. https://metrics.vrchat.community/, 2024. Accessed: 2024-08-30.

[135] VRChat Inc. Vrchat. Online Virtual Reality Platform, 2014.

[136] VRChat Team. Mobile alpha now live! https://hello.vrchat.com/blog/mobile-alpha-now-live, 2023. Accessed: 2024-08-30.

[137] Xiaoying Wei, Xiaofu Jin, and Mingming Fan. Communication in immersive social virtual reality: A systematic review of 10 years' studies. In *Proceedings of the Tenth International Symposium of Chinese CHI*, pages 27–37, 2022.

[138] Stanley G Weinbaum. *Pygmalion's spectacles*. Simon and Schuster, 2016.

[139] Laurie M Wilcox, Robert S Allison, Samuel Elfassy, and Cynthia Grelik. Personal space in virtual reality. *ACM Transactions on Applied Perception (TAP)*, 3(4):412–428, 2006.

[140] Julie Williamson, Jie Li, Vinoba Vinayagamoorthy, David A Shamma, and Pablo Cesar. Proxemics and social interactions in an instrumented virtual reality workshop. In *Proceedings of the 2021 CHI conference on human factors in computing systems*, pages 1–13, 2021.

[141] Julie R Williamson, Joseph O'Hagan, John Alexis Guerra-Gomez, John H Williamson, Pablo Cesar, and David A Shamma. Digital proxemics: Designing social and collaborative interaction in virtual environments. In *Proceedings of the 2022 CHI conference on human factors in computing systems*, pages 1–12, 2022.

[142] Amal Yassien, Passant ElAgroudy, Elhassan Makled, and Slim Abdennadher. A design space for social presence in vr. In *Proceedings of the 11th Nordic Conference on Human-Computer Interaction: Shaping Experiences, Shaping Society*, pages 1–12, 2020.

[143] Andrew Yoshimura and Christoph W Borst. Evaluation of headset-based viewing and desktop-based viewing of remote lectures in a social vr platform. In *Proceedings of the 26th ACM Symposium on Virtual Reality Software and Technology*, pages 1–3, 2020.

[144] Fengyuan Zhu, Mauricio Sousa, Ludwig Sidenmark, and Tovi Grossman. Phoneinvr: An evaluation of spatial anchoring and interaction techniques for smartphone usage in virtual reality. In *Proceedings of the CHI Conference on Human Factors in Computing Systems*, pages 1–16, 2024.

[145] Daniel Zielasko and Bernhard E Riecke. To sit or not to sit in vr: Analyzing influences and (dis) advantages of posture and embodied interaction. *Computers*, 10(6):73, 2021.