



**HAL**  
open science

# Toward a Characterization of Perceptual Biases in Mixed Reality: A Study of Factors Inducing Distance Misperception

Etienne Peillard

► **To cite this version:**

Etienne Peillard. Toward a Characterization of Perceptual Biases in Mixed Reality: A Study of Factors Inducing Distance Misperception. Graphics [cs.GR]. École centrale de Nantes, 2020. English. NNT: 2020ECDN0030 . tel-03162186

**HAL Id: tel-03162186**

**<https://theses.hal.science/tel-03162186v1>**

Submitted on 8 Mar 2021

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# THÈSE DE DOCTORAT DE

L'ÉCOLE CENTRALE DE NANTES

ÉCOLE DOCTORALE N° 602  
*Sciences pour l'Ingénieur*  
Spécialité : Architecture et études urbaines

Par

**Étienne PEILLARD**

**Vers une caractérisation des biais perceptifs en réalité mixte :  
une étude de facteurs altérant la perception des distances**

Thèse présentée et soutenue à Nantes, le 24 novembre 2020

Unités de recherche : UMR 1563, Laboratoire Ambiances, Architectures, Urbanités (AAU)

## **Rapporteurs avant soutenance :**

Victoria INTERRANTE Professeure, University of Minnesota (États-Unis)  
Jean-Louis VERCHER Directeur de recherche, Aix-Marseille Université

## **Composition du Jury :**

Présidente :	Valérie GYSELINCK	Directrice de recherche, Université Gustave Eiffel, Versailles
Examineurs :	Sarah CREEM-REGEHR	Professeure, University of Utah (États-Unis)
	Anatole LÉCUYER	Directeur de recherche, Inria Rennes
Dir. de thèse :	Guillaume MOREAU	Professeur des universités, École Centrale de Nantes
Co-enc. de thèse :	Ferran ARGELAGUET	Chargé de recherche, Inria Rennes
Co-enc. de thèse :	Jean-Marie NORMAND	Maître de conférences, École Centrale de Nantes



# ACKNOWLEDGEMENTS

---

First, I would like to express my sincere gratitude to my thesis jury, Valérie, Sarah, Vicky, Jean-Louis, for taking the time to read my work and attend to my defense. It has been a pleasure and an honor to be able to present my work and defend my thesis in front of them. Their insightful comments and encouragement, as well as the exciting questions and discussions we had, will allow me to widen my research from different perspectives and I am very grateful for that.

Je voudrais ensuite bien sûr remercier mes encadrants pour leur immense bienveillance. Ils m'ont permis de conduire ce travail en confiance, avec l'assurance de leur indéfectible soutien. Je remercie en particulier Jean-Marie pour sa présence quotidienne et sa patience légendaire, Guillaume grâce à qui j'ai pu entrer et m'épanouir dans le monde de la recherche, Ferran pour sa prévenance et sa vigilance, et enfin Anatole pour sa foulditude d'idées et ses encouragements constants.

Évidemment, je voudrais aussi remercier tous les gens que j'ai pu côtoyer au travail durant tout ce temps. Deux équipes comme autant de familles où partager travail et plaisir, stress et réussite. Je remercie tous les membres de l'équipe Hybrid pour leur chaleur et leur entrain. Merci à Rebecca pour sa présence facétieuse, merci à Diane pour son indulgente surveillance, merci bien sûr à Florian et Nathalie pour toute l'aide et le soutien qu'ils apportent au quotidien. Je veux remercier également tous mes collègues Nantais dont l'exemplarité et le dévouement forcent chaque jour mon admiration. Merci à Myriam, Vincent, Jean-Yves pour ces moments partagés; et merci également à Léa et Martin pour leur enthousiasme et leur bonne humeur.

More generally, I would like to thank all those who participated in the elaboration of this work: my supervisors and my colleagues, but also the researchers with whom I was able to collaborate. In particular, I would like to thank Yuta for welcoming me to his laboratory and offering me this wonderful opportunity for cooperation. 伊藤先生、ありがとうございました!

It also seems important to me to deeply thank all the participants who took part in my studies. Without them, this work would not have been possible.

Merci également à mes amis qui m'ont supporté, en particulier durant ces dernières années. Merci à Arthur, Roland, Félix, Rémi avec qui j'ai tant partagé et qui continuent de m'apporter beaucoup. Merci également à Victorien et Hervé pour les instants et les joies partagées qui m'ont tellement nourri.

Enfin, il me reste à remercier ma famille, et en particulier mes parents, qui de l'autre bout du monde continue à m'apporter tout leur soutien.

Encore une fois, merci à tous.

\* \*  
\*

# TABLE OF CONTENTS

---

<b>Introduction</b>	<b>15</b>
<b>I Theoretical Background: Realities and Perceptions</b>	<b>23</b>
<b>1 From sensation to perception; visual perception as an example, distance evaluation as a metric</b>	<b>25</b>
1.1 The origin of perception . . . . .	26
1.1.1 The Perceptual Process . . . . .	26
1.1.2 The perception system's purposes: compensating and interpreting .	28
1.2 Forewords about the human visual system . . . . .	30
1.3 The specific case of distance evaluation and depth perception . . . . .	31
1.3.1 Depth cues and cue theory . . . . .	32
1.3.2 Distance fields . . . . .	37
1.3.3 Global theories of visual space in real environments . . . . .	37
1.4 Measuring perception . . . . .	40
1.4.1 Usual perception evaluation techniques . . . . .	40
1.4.2 Distance perception evaluation techniques . . . . .	42
<b>2 Mixed Reality - How to render a virtual world and blend it with reality</b>	<b>47</b>
2.1 Definitions . . . . .	47
2.1.1 Real and Virtual . . . . .	48
2.1.2 Virtual Reality . . . . .	48
2.1.3 Augmented Reality . . . . .	49
2.1.4 Continuum, Mixed Reality and other notions . . . . .	49
2.2 Devices . . . . .	50
2.2.1 Near-eye displays . . . . .	51
2.2.2 Handheld Devices . . . . .	55
2.2.3 Distant displays . . . . .	55
2.3 Mixed reality and perception, limitations and challenges . . . . .	56

## TABLE OF CONTENTS

---

2.3.1	The vergence-accommodation conflict . . . . .	57
2.3.2	X-ray vision . . . . .	58
<b>3</b>	<b>Overview of key studies in distance perception in Augmented Reality</b>	<b>61</b>
3.1	Depth cues . . . . .	63
3.1.1	Binocular disparity and convergence . . . . .	63
3.1.2	Monocular studies . . . . .	66
3.1.3	Accommodation . . . . .	67
3.1.4	Motion parallax . . . . .	68
3.1.5	Peripheral vision . . . . .	69
3.1.6	Brightness . . . . .	69
3.1.7	Shadows . . . . .	69
3.1.8	Height of the object . . . . .	70
3.2	Intrinsic factors . . . . .	71
3.2.1	Gender . . . . .	71
3.2.2	Age . . . . .	72
3.2.3	Avatar . . . . .	72
3.2.4	Learning and Feedback . . . . .	73
3.2.5	Emotions . . . . .	73
3.3	Device and rendering parameters . . . . .	74
3.3.1	Latency . . . . .	74
3.3.2	Graphics quality and realistic rendering . . . . .	74
3.3.3	Blurring . . . . .	75
3.3.4	Screen Size . . . . .	76
3.3.5	Resolution . . . . .	77
3.3.6	Handheld specific factors . . . . .	77
3.3.7	Video See-Through displays . . . . .	78
3.3.8	HMD specific issues . . . . .	79
3.4	Environment and object related factors . . . . .	80
3.4.1	Environments . . . . .	80
3.4.2	Occluder and x-ray vision . . . . .	81
3.4.3	Auxiliary augmentations . . . . .	82
3.4.4	Object appearance and position . . . . .	83
3.5	Besides depth: exocentric distance perception . . . . .	84

3.5.1	Exocentric distance perception in real life . . . . .	84
3.5.2	Exocentric distance perception in VR . . . . .	85
3.5.3	Exocentric Distance Perception in AR . . . . .	86
3.6	Conclusion: Are all perceptual biases in MR explained by a specific factor effect? . . . . .	87
<b>II Research Contributions</b>		<b>91</b>
<b>4</b>	<b>Intrinsic Biases: the Anisotropy of Distance Perception</b>	<b>93</b>
4.1	Main experiments . . . . .	94
4.1.1	Experiment 1: Anisotropy of distance perception . . . . .	95
4.1.2	Experiment 2: Influence of the viewing angle . . . . .	103
4.1.3	Provisional Conclusions . . . . .	107
4.2	Control experiments . . . . .	108
4.2.1	Experiment 3: Disambiguating head and eye directions effects . . .	108
4.2.2	Experiment 4: Does the quality of the virtual environment matter? . . .	111
4.3	Discussion . . . . .	114
4.4	Conclusion . . . . .	116
<b>5</b>	<b>Transfer of Perceptual Bias: From Egocentric to Exocentric?</b>	<b>119</b>
5.1	User Experiment . . . . .	120
5.1.1	Participants . . . . .	120
5.1.2	Experimental Apparatus . . . . .	120
5.1.3	Procedure . . . . .	124
5.1.4	Experimental Design . . . . .	125
5.1.5	Statistical Analysis . . . . .	126
5.2	Results . . . . .	127
5.3	Discussion . . . . .	130
5.4	Conclusion . . . . .	132
<b>6</b>	<b>The Impact of Depth Cues: Shadows and Accommodation Distance</b>	<b>135</b>
6.1	Impact of shadow shape and consistency on depth perception . . . . .	136
6.1.1	Research Hypotheses . . . . .	136
6.1.2	Experimental design . . . . .	137



TABLE OF CONTENTS

---

6.1.3 Experiment 1 - Studying the influence of the shape of virtual objects' shadows on distance perception in OST AR . . . . . 141

6.1.4 Experiment 2 - Studying the influence of lighting misalignment on distance perception in OST AR . . . . . 151

6.1.5 Discussion . . . . . 159

6.1.6 Conclusion . . . . . 161

6.2 Studying the influence of accommodation distance through the use of Retinal Projection Displays . . . . . 161

6.2.1 Methods . . . . . 162

6.2.2 Results . . . . . 170

6.2.3 Discussion . . . . . 171

6.2.4 Conclusion . . . . . 174

**III Prospects and Discussion 177**

**7 Prospects – How to Compensate Perceptual Biases? The Use of Interaction Techniques 179**

7.1 Theoretical Background and Research Hypotheses . . . . . 180

7.2 Experimental Protocol Proposal . . . . . 181

7.2.1 Apparatus . . . . . 181

7.2.2 Calibration . . . . . 182

7.2.3 Task and Procedure . . . . . 183

7.2.4 Experimental Design . . . . . 185

7.3 Discussion . . . . . 186

**8 Discussion 189**

8.1 Space perception and Mixed Reality - Current state of knowledge and research challenges . . . . . 189

8.2 Contributions of this Thesis . . . . . 191

8.3 Limitations and Future Works . . . . . 193

8.3.1 Experimental variations . . . . . 193

8.3.2 Perspectives on conducted experiments . . . . . 196

8.3.3 Long term perspectives . . . . . 199

**Conclusion 203**

<b>IV Appendix</b>	<b>205</b>
A Résumé long en français	207
B Author's publications	215
Bibliography	217

# LIST OF FIGURES

---

1.1	Overview of the perceptual process . . . . .	27
1.2	Experiments showing flaws in the visual perception system. . . . .	29
1.3	Inner structure of the eye (from Goldstein and Brockmole, 2017) . . . . .	30
1.4	Depth cues classification from Howard, 2002. . . . .	32
1.5	Depth cues classification from Goldstein and Brockmole, 2017. . . . .	32
1.6	Depth cues classification proposed. . . . .	33
1.7	Example of main depth cues. . . . .	34
1.8	Effective range of depth cues . . . . .	38
1.9	Relative depth cues strength . . . . .	39
1.10	Different distance evaluation techniques. . . . .	43
2.1	Classification of Mixed Reality (MR) devices . . . . .	51
2.2	Example of commercial Head Mounted Displays (HMDs) . . . . .	52
2.3	Possible optical configurations of Retinal Projection Displays (RPDs). . . . .	54
2.4	Examples of distant displays . . . . .	56
2.5	The Vergence-Accommodation conflict . . . . .	57
2.6	Examples of visualization techniques using x-ray vision . . . . .	58
3.1	A Sankey diagram of distance perception studies in AR . . . . .	65
4.1	Exp 1 - Experimental setup . . . . .	96
4.2	Exp 1.1-1.3 - Virtual stimulus . . . . .	96
4.3	Exp 1 - Calibration procedure . . . . .	97
4.4	Exp 1.1,3,4 - Stimulus configuration . . . . .	99
4.5	Exp. 1.1 - Psychometric curves . . . . .	101
4.6	Exp. 1.2 - Stimulus configuration . . . . .	103
4.7	Exp. 1.2 - Psychometric curves . . . . .	106
4.8	Exp. 1.2 - PSE for each angle condition . . . . .	107
4.9	Exp. 1.3 - Psychometric curves . . . . .	110
4.10	Exp. 1.4 - Virtual stimulus . . . . .	112

---

4.11	Exp. 1.4 - Psychometric curves . . . . .	113
5.1	Exp. 2 - Apparatus . . . . .	121
5.2	Exp. 2 - Answering protocol . . . . .	122
5.3	Exp. 2 - Stimuli conditions . . . . .	122
5.4	Exp. 2 - Calibration markers . . . . .	123
5.5	Exp. 2 - Computer vision for answer recording . . . . .	124
5.6	Exp. 2 - Relative distance estimation errors across conditions . . . . .	128
5.7	Exp. 2 - Relative distance estimation errors considering dominant eye . . .	129
5.8	Exp. 2 - Individual results . . . . .	129
6.1	Exp. 3.1-2 - Overview . . . . .	138
6.2	Exp. 3.1-2 - Experimental procedure . . . . .	141
6.3	Exp. 3.1 - Shadows conditions . . . . .	143
6.4	Exp. 3.1 - Relative error per distance . . . . .	145
6.5	Exp. 3.1 - Relative error per shadow type . . . . .	146
6.6	Exp. 3.1 - Completion time per distance . . . . .	147
6.7	Exp. 3.1 - Completion time per shadow type . . . . .	148
6.8	Exp. 3.1 - Individual results . . . . .	148
6.9	Exp. 3.2 - Lighting conditions . . . . .	152
6.10	Exp. 3.2 - Relative errors per distance . . . . .	153
6.11	Exp. 3.2 - Relative errors per lighting displacement . . . . .	154
6.12	Exp. 3.2 - Completion time per distance . . . . .	155
6.13	Exp. 3.2 - Completion time per lighting displacement . . . . .	156
6.14	Exp. 3.2 - Individual results . . . . .	156
6.15	Exp. 4 - Overview of the experimental setup . . . . .	163
6.16	Exp. 4 - User point of view and close-ups . . . . .	164
6.17	Exp. 4 - Target positioning system . . . . .	164
6.18	Exp. 4 - Custom binocular RPD headset . . . . .	165
6.19	Exp. 4 - Global results . . . . .	171
6.20	Exp. 4 - Individual data . . . . .	172
6.21	Exp. 4 - Precision results . . . . .	173
7.1	Exp 5 - Interaction device . . . . .	182
7.2	Exp 5 - Calibration markers . . . . .	183

LIST OF FIGURES

---

7.3 Exp 5 - Interaction techniques . . . . . 184

# LIST OF TABLES

---

1.1	Different depth cues and their properties . . . . .	38
3.1	Overview of characteristics of distance perception studies in Augmented Reality . . . . .	64
4.1	Exp 1.1,3,4 - Summary of experimental variables . . . . .	99
4.2	Exp. 1.2 - Summary of experimental variables . . . . .	104
4.3	Exp. 1.2 - Fitting values for psychometric curves . . . . .	105
5.1	Exp. 2 - Summary of experimental variables . . . . .	125
6.1	Exp 3.1 - Individual mean relative errors by shadow type . . . . .	149
6.2	Exp 3.1 - Individual mean relative errors by distance . . . . .	149
6.3	Exp 3.1 - Completion time . . . . .	150
6.4	Exp 3.2 - Individual mean relative errors by light position . . . . .	157
6.5	Exp 3.2 - Individual mean relative errors by distance . . . . .	157
6.6	Exp 3.2 - Completion time . . . . .	158
6.7	Exp. 4 - Summary of experimental variables . . . . .	167
7.1	Exp. 5 - Summary of experimental variables . . . . .	185



# INTRODUCTION

---

“Virtual Reality”, “Augmented Reality”, “Mixed Reality”; these concepts already convey in their formulation the oppositions characterizing them. They seem to point to a reality that is not the actual reality, or even a reality that does not quite manage to be real. In essence, they demonstrate the inherent purpose of these technologies to substitute, modify and enrich reality. Starting from the actual reality, multiple digital devices show the user a different reality by adding or substituting virtual elements. It is interesting to note that the naming of these concepts adds an adjective to “reality” that replaces or completes it. These computer generated realities seek to achieve actual reality, to integrate virtual objects or environments into the perceptual environment of the observers, but do not perfectly achieve this level of realism. In addition, perception is a complex process involving many elements. Even in reality, what is and what is perceived are two different things. There are differences between reality and perception, but these differences are even greater when reality is not the actual reality but a modified reality. Thus, the observer is not fooled; what he perceives is not the “actual reality” but a digitally modified version of the real world.

In 1965, Sutherland wrote the following about the “Ultimate Display” (Sutherland, 1965):

“The ultimate display would, of course, be a room within which the computer can control the existence of matter. A chair displayed in such a room would be good enough to sit in. Handcuffs displayed in such a room would be confining, and a bullet displayed in such a room would be fatal. With appropriate programming such a display could literally be the Wonderland into which Alice walked.”

In this example, we can recognize an ideal “virtual reality” device, but it is defined as a “display”, i.e. a device capable of giving something to perceive. Before that, Sutherland also wrote that “the task of the display is to serve as a looking-glass into the mathematical wonderland constructed in computer memory”. We can see here an expression of a direct issue, where the aim of the device would be to show an environment, without doubting about its “reality”. To make the analogy with the Turing’s test, what could be called the “Sutherland’s test” would be validated for virtual reality devices able to simulate virtual



elements indistinguishable from real ones. This device would be capable of displaying reality in its simplest expression. This phenomenon is already seen in virtual environments, where participants react as if the environment were real (Slater, 2009). Interestingly, this is achieved even if the Virtual Environments (VEs) do not look “real” at all. However, there is still much work to be done to achieve such a state permanently.

Subsequently, it was only in the 1980s that the term Virtual Reality (VR) was introduced by Jaron Lanier (Kurt, 2007). This shift from “display” to “VR” was gradual and many VR devices began to emerge while the term had not yet become widespread. Could it be said that, as technology evolves, it became clear that the reality presented by these “displays” remains far from the “actual” reality? And that it then became necessary to characterize this new form of imperfect reality? What then led to these ambivalent terms of virtual/augmented/mixed realities? It is, however, the thesis that I wish to defend here, that the maturation of mixed realities only further underlines their differences with the actual reality and that this constant friction between reality and computed simulation is constitutive of these notions, right down to their denominations. In the same way as the “Uncanny Valley”, where the too close resemblance of an android with a human underlines its differences, the proximity with the actual reality of mixed reality also underlines its weaknesses.

Thus, this dissonance between real and virtual can be noticed at every level of observation: biases specific to this Mixed Reality appear, whether they are due to physical differences between real and virtual, to perceptual differences or even to biases internal to individuals. Reality and its perception are hard to define. What is reality? What makes us feel/perceive that something is real? Why do we perceive reality as we do? How do we learn to perceive reality? How and why perceptual biases arise? Interactions of virtual content with these perceptual open questions are even more puzzling. When faced with virtual content, are new biases created? Do we apply these biases to a new “reality”? Is it that, since we do not have the same cues for understanding the environment, ambiguity arises and then a perceptual mismatch appears? Is that virtual element carry their own biases? These questions are extremely broad and my hope is to bring some elements of answer in this thesis.

To reduce the scope of our research, the most notable difference between real and virtual lies in the technical limitations of the devices used to show this virtual reality. If what is shown to the users differs from reality, then they will perceive this virtual reality differently from a real situation. The paradox here is that it is impossible to provide a

simulated environment that is an exact replica of reality. The computation time may be extremely low, it will never be zero; colors and details may be extremely faithful, they will never be strictly identical. The current objective is therefore to reduce these defects as much as possible, and to achieve a system's realism below the perceptibility threshold of the human sensory system. Thus, with a sufficiently precise system, the observer would not be able to distinguish the real from the virtual. However, the human perceptual system is extremely precise and complex. For example, the human visual system includes both sensors (the eyes), the optic nerves and the brain that processes information. However, this system works as a whole and it is impossible to dissociate the specific action of each of these elements. While some elements are theoretically below the observer's perception threshold, they nevertheless have effects on perception. For example, an imperceptible modification of the field of view can affect motion sickness (Fernandes & Feiner, 2016), or an image displayed during a theoretically imperceptible flash (also known as subliminal image) influences decision making (Huang et al., 2012). In short, biases induced by the human perceptual system may appear, even if the simulation is sufficiently accurate to be confused with reality. In this respect, they constitute a different category of biases from the technical ones mentioned above. Finally, it should also be noted that since perception is an internal cognitive process, many elements specific to the observer could modify his perception. Thus, even considering that it would be possible to generate a virtual environment identical to a real environment, the fact that observers are aware of the virtual nature of this environment can generate an inner bias, impacting perception. However, these biases are always intrinsically linked. Perception is a complex process that involves the different levels described above, but also implies a permanent interaction between these different levels. Thus, it may not be possible to define a complete model of perception, and to systematically specify the biases due to MR. These biases may be intertwined in essence, as is perception.

From this premise derives the approach that I have chosen to follow in my work. In order to shed light on the origin of certain biases, I have chosen to use a per-factor approach as precise as possible. Each of the experiments carried out aims to focus on a specific bias but also to identify the more general links with the theory of perception. The ideal objective would be to highlight the various external perceptual biases (linked to the devices used) and then to identify more complex biases such as those internal to individuals and thus answer the following question: "Does the fact of being aware of the virtual nature of the environment, without being able to distinguish it from reality,

influences the perception of users?”. However, the resolution of this problem goes far beyond the scope of this work. In this thesis, I will attempt to present, highlight and propose explanations in a systematic way; on an attempt to explain some specific biases that appear in VR or in Augmented Reality (AR).

The elements provided previously demonstrate the complexity of the study of visual perception and the challenges added when considering Mixed Reality. As such, we propose to focus on one of the translations of perception: distance perception. Indeed, visual perception is considered as the first form of perception, and the spatialization of objects has therefore been extensively studied. Some authors even showed that any element, even moral ones, can be associated with a spatial position by the mind (Casasanto, 2009; Davidson, 1992). Thus, the study of distance perception is a metric that is especially adapted to the general study of perception. This is therefore the research angle on which we focus in this thesis.

However, beyond the precise measurement of these perceptual biases, emphasis will be placed on the origin of these biases and their potential existence beyond the experimental field where they are observed. Indeed, the biases revealed during strictly controlled experiments are precisely revealed thanks to the strict and sober character of the environments used. These experiments lack an ecological validity and do not really reflect a realistic situation of using virtual or augmented reality. It therefore seems essential to me to also take into account the differences potentially induced by this experimental choice.

## Outline

This manuscript is composed of 8 chapters that are summarized as follows.

### **Part I – Theoretical Background: Realities and Perceptions**

This part focuses on the previous work on perception and Mixed Reality.

#### **Chapter 1 – From sensation to perception; visual perception as an example, distance evaluation as a metric**

The goal of the first chapter is to present an overview of related work on perception and specifically distance perception. Starting from the origin of perception studies, we present the perceptual process, which transforms sensations into perception. We also present the

challenges of the measure of perception. This chapter also provides a detailed presentation of depth cues theory and general theories on perceptual space.

## **Chapter 2 – Mixed Reality - How to render a virtual world and blend it with reality**

The second chapter shows the theoretical background of Virtual Reality (VR) and Augmented Reality (AR). First it introduces the field of VR and AR and their main concepts. Second, it presents the different techniques and devices allowing to achieve Mixed Reality, underlining the stakes of these elements from the point of view of perception. Finally, it focuses on two main challenges of MR: the Vergence-Accommodation Conflict (VAC) and the x-ray vision.

## **Chapter 3 – Overview of key studies in distance perception in Augmented Reality**

The third chapter reviews the state for the art of distance perception in AR. It is structured around the list of potential biases that can occur in AR characterized as follows: (1) depth cues related biases, (2) intrinsic biases,(3) rendering induced biases and (4) environment and object specific biases. The specific case of exocentric distance perception is also reviewed.

## **Part II – Research contributions**

In this part, we present four series of experiments highlighting several biases in MR. For each experiment, background and research hypotheses are presented, then the protocol is detailed and finally results are reported and discussed.

## **Chapter 4 – Intrinsic Biases: the Anisotropy of Distance Perception**

The fourth chapter presents a study of an intrinsic bias in Virtual Reality. Observers are asked to perform an estimation of egocentric distances in an otherwise blank environment. This configuration makes it possible to highlight the anisotropy of the perceptive system, in the absence of any other element or reference.

## **Chapter 5 – Transfer of Perceptual Biases: From Egocentric to Exocentric?**

In the fifth chapter, we further study the potential transfer of a perceptual bias. Indeed, the contraction of egocentric distances is a notable fact of visual perception in Augmented Reality, it is thus interesting to observe the potential shift of this effect towards the exocentric space. This chapter reports a user experiment evaluating exocentric distance perception for virtual and/or real objects in a MR context.

## **Chapter 6 – The impact of depth cues on egocentric distance perception in Augmented Reality**

The sixth chapter questions in more detail some specificities of visual perception of distances through two important elements responsible of a potential bias: anchoring and the accommodation-vergence conflict. First, anchoring is studied under the light of shadows, evaluating the effect of shape and position of shadows on depth perception in the first section of this chapter. Second, the accommodation cue is tackled through the technological point of view. A study is thus designed using a focus-free device: a Retinal Projection Display (RPD). A study using a Optical See-Through (OST) HMD and a RPD is presented in the second section of this chapter.

## **Part III – Prospects and Discussion**

### **Chapter 7 – Prospects - How to compensate Perceptual Biases? The use of interaction techniques**

The seventh chapter investigates the possibility of influencing these biases not by strictly perceptual means but by using dedicated interaction techniques in order to correct, in a relevant way, the space perception accuracy in augmented environments. First, it presents the challenges of interaction in the perceptual process and its potential impact on perception in AR. Then, it details an experimental protocol to evaluate the effect of different interaction techniques on distance perception in AR.

### **Chapter 8 – Discussion**

The eighth chapter concludes the manuscript with a general discussion. First, it recalls the main challenges of space perception and MR. Second, it summarizes the contributions

---

of this thesis. Finally, it provides a general discussion of the issues, biases and potential improvements provided by the different studies and present potential perspectives.



PART I

# Theoretical Background: Realities and Perceptions

---





Most elephants are usually big, but this one, built at Nantes (France) by the company "La Machine" is twice as big as an ordinary elephant. In this chapter, we present the different elements allowing perception and we presents more in depth the theories of distance perception.

# FROM SENSATION TO PERCEPTION; VISUAL PERCEPTION AS AN EXAMPLE, DISTANCE EVALUATION AS A METRIC

---

## Contents

---

<b>1.1 The origin of perception . . . . .</b>	<b>26</b>
1.1.1 The Perceptual Process . . . . .	26
1.1.2 The perception system's purposes: compensating and interpreting	28
<b>1.2 Forewords about the human visual system . . . . .</b>	<b>30</b>
<b>1.3 The specific case of distance evaluation and depth perception</b>	<b>31</b>
1.3.1 Depth cues and cue theory . . . . .	32
1.3.2 Distance fields . . . . .	37
1.3.3 Global theories of visual space in real environments . . . . .	37
<b>1.4 Measuring perception . . . . .</b>	<b>40</b>
1.4.1 Usual perception evaluation techniques . . . . .	40
1.4.2 Distance perception evaluation techniques . . . . .	42

---

Perception is a complex process that is still not fully explained. This is evidenced by the numerous works studying vision, from the earliest theories (Berkeley, 1709) to recent work on the neurological component of visual perception (Epstein & Baker, 2019; Niketeghad & Pouratian, 2019). The number of senses involved, the sensory sensors used by the body and the physical and chemical processes are numerous. However, a common thread can be pulled together, regardless of which sense is considered. With a global approach, the process of perception can thus usually be divided into a succession of steps common to all senses.

First, this global perceptual process is detailed. Then, this chapter reports several vision's related concepts. Finally, it focuses on distance perception as the main metric to

evaluate perception and provides a detailed presentation of distance perception theory.

## 1.1 The origin of perception

The principles underlying human perception have been widely studied for many centuries (Berkeley, 1709). The initial theory of the internal representation of space in the form of a mind map (S. Kaplan, 1973) was later challenged by a topological approach of space (Kuipers, 1983), rejecting the idea of a metric representation in favor of a semantic representation of objects in relation to each other. However, these theories have also proved to be too restrictive. The most recent works agree on a mixed, more ecological conceptualization of perception. They focus on describing how different types of information, resulting from the interaction of the individual with his environment, can contribute to the construction of a mental representation (Tversky, 1993). This approach is the one underlined by Goldstein and Brockmole (2017) in their model of perception and in particular of the perceptual process. It is on this theory and related concepts that we will base the remainder of this chapter.

### 1.1.1 The Perceptual Process

This process allows external stimuli to be transformed into a conscious experience, and eventually action. It can be broken down into several relatively independent steps, which are detailed below. However, it is important to note at this point that although relatively independent, these steps are not simply unidirectional links in a chain. There are constant exchanges between the different levels of perceptual information processing, and the organization and segmentation presented are a simplified version of the overall perceptual process. Most of the examples in this section will focus on vision, but it should be noted that this process is not limited to vision alone and is common to all human senses.

The perceptual process can be described as follows, and is presented in Figure 1.1:

- stimuli (distal & proximal),
- physiological response (receptors & neuronal processing),
- perception & recognition,
- action.

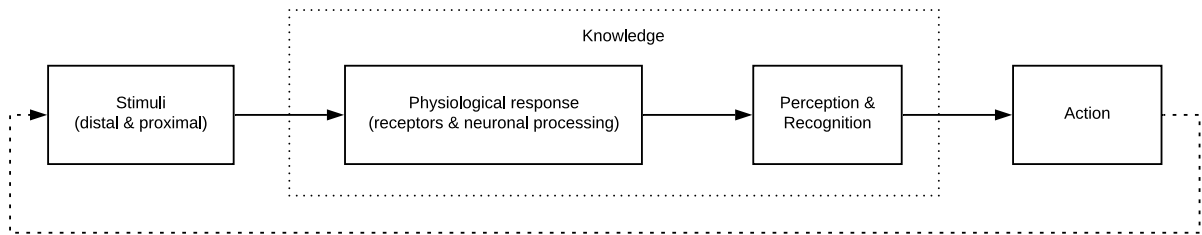


Figure 1.1 – Overview of the perceptual process, adapted from Goldstein and Brockmole, 2017.

**Stimuli.** They are the first step of perception. They consist of all the elements external to the observer’s body that work together to create what will be perceived. There are two types of stimuli: distal stimuli and proximal stimuli. **Distal stimuli** include all elements that could be perceived and are “distant” from the observer. For example, a tree, a bird song or the sun. They are not directly perceived (the tree, like the sun, does not come into contact with the eye that sees them) but are at the origin of the perception reaction. They are then transformed to finally become **proximal stimuli**. This second category includes the final stimuli that are then detected by the human body’s sensors: light image projected on the retina, vibration of the air in contact with the eardrum, etc.

**Physiological response.** Once the stimulus hits the sensory **receptor**, it is transformed, through a series of electro-chemical reactions that depend on each receptor, into an electrical signal. It is then routed through the neural system, which transmits and transforms the signal. At this point, the reaction becomes internal to the individual. However, this is still an electrical impulse and not perception. The signals are transmitted and modified by nerves and then **pre-processed** by neurons, but this part of the perceptual process is unconscious and solely physical.

**Perception & Recognition.** At this step, electrical signals are transformed into conscious experience. All the signals are combined to give a mental representation of what is perceived. At this stage perception and recognition are two distinct yet interrelated processes. The notion of perception reflects the fact that the subject is conscious of perceiving something, which is therefore different from recognizing it. Some brain pathologies sometimes can provoke a defect of one of these two functions. Patients are then conscious of seeing certain things but unable to recognize them.

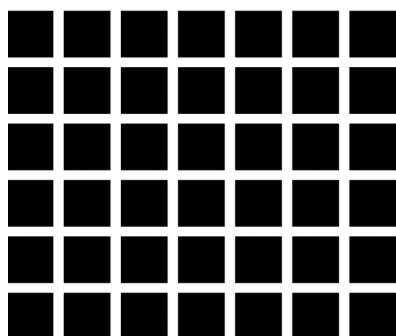
The fundamental role of **knowledge** at this stage should also be underlined. Indeed,

the conscious experience that results from perception is not only the sum of the electrical signals coming from the receptors but also the whole of the observer’s past experiences and even his cognitive state.

**Action.** Action can be seen as the last step in the process of perception and is one of the main outcomes of perception. While its integration as a part of the perceptual process is debatable, some authors suggest that it must be associated in it because action used to be a major goal of perceptual processing for animals (Milner et al., 1995). However, the interconnection between perception and action must be stressed. Action makes it possible to adapt one’s point of view, one’s way of perceiving as well as reacting to what is perceived. The links between perception and interaction will thus be explained further on.

### 1.1.2 The perception system’s purposes: compensating and interpreting

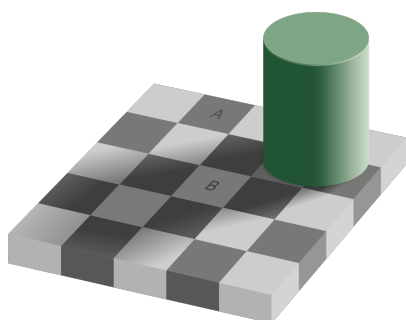
Although we have detailed the perceptual process systemically here, it must be stressed that none of these steps can be considered “simple” or “perfect”. Imperfections, unexpected behaviors or simply ambiguous results can be observed at each step. Firstly, the sensors of the human body are limited: the eyes have a limited field of view and depth of field, they are not able to distinguish some nuances of colour, the ears have a limited frequency range, the olfactory receptors may be saturated after too long an exposure, etc. But it does not stop there. The next steps, perception and recognition, internal to the individual, can also be biased. Cognitive state, experience, a priori are all factors likely to blur perception. The possibilities of interaction can even be responsible for significant perceptual biases (Franklin & Tversky, 1990). Thus, the cognitive process that enables perception is by nature limited by its inputs and must constantly adapt to compensate for its shortcomings. This adaptation constitutes a formidable capacity but is also at the origin of many defects of perception. While it is useful to be able to interpret a deficiency, it can also alter perception by causing over-interpretation. Many illusions and perceptual biases have their origins not in a failure of the sensory system but in an erroneous interpretation of the entire scene (Bartz et al., 2008).



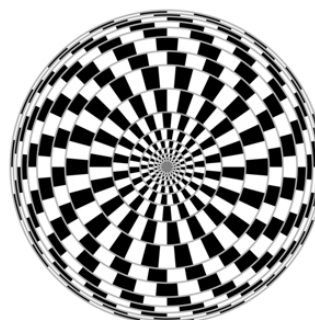
(a) The Hermann's grid. Contrast enhancement in the visual system create gray points at the intersections which vanish when looked directly.



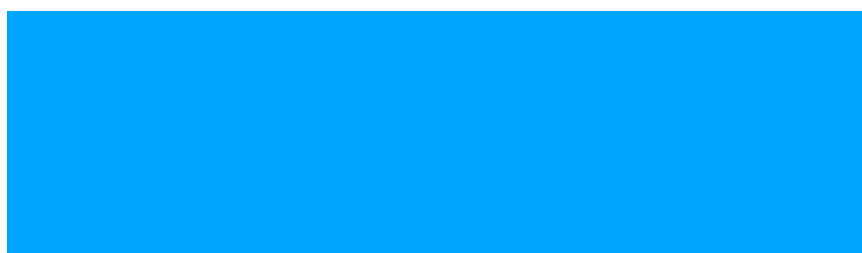
(b) The blind spot. When closing the left eye, staring at the cross with the right eye and slowly moving the head towards and away from the screen, the dot disappears. This happens when the image falls in the "blind spot" of the retina, where the optic nerve starts and no receptors are present.



(c) The squares A and B have the same color but the brain interpret one as darker than the other because of the shadow.



(d) Concentric gray circles appear to be spirals, from Kitaoka, 1998



(e) Blue field entoptic phenomenon. Tiny bright dots appear in the visual field when looking to a bright blue light. The speed of the dots varies in sync with the pulse; they briefly accelerate at each heartbeat. They are created by an accumulation of red blood cells behind the white blood cells moving in the capillaries in front of the retina of the eye. The brain usually edit out those dots under normal vision but since the blue light is especially absorbed by the red cells, some artifacts remain visible when the eye is exposed to bright blue light.

Figure 1.2 – Experiments showing flaws in the visual perception system.

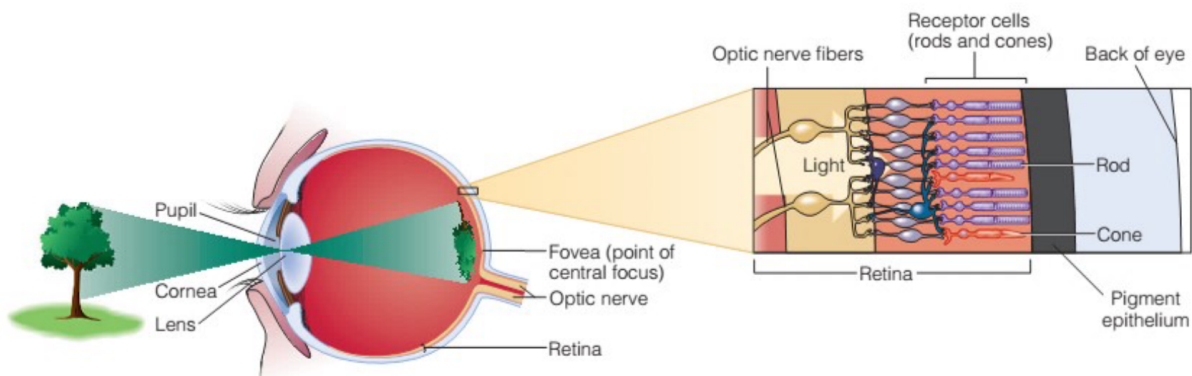


Figure 1.3 – Inner structure of the eye (from Goldstein and Brockmole, 2017)

## 1.2 Forewords about the human visual system

Vision is the main sense of humans. This sense is processed by the visual cortex, the largest system in the human brain and is provided by the *visual system*. The visual system comprises the sensory organ (the eye) and the part of the central nervous system (the optical nerve and the brain) which gives organisms the ability to process visual detail. It detects and interprets information from visible light to build a representation of the surrounding environment.

At the beginning of the visual system lies the eye. As represented in Figure 1.3, the light enters the eye via the pupil, passing through the cornea. Then, the lens focuses the light on the retina. The light is absorbed and activates the receptor cells (cones and rods) of the retina. The cones, of three different types, allow to perceive the different colors and are more sensitive to contrast. The rods, more numerous but less precise, are mostly used for low luminosity vision. When excited, the receptor cells generate an electric impulse which is transmitted to the brain by the optical nerve.

Because of its design, the image formed by the eye must be projected onto the retina to be seen clearly. Thus, to see objects placed at different distances, the eye has to adjust the focal length of the lens to form the image on the retina. At rest, the image is sharp to infinity. The ciliary muscles need to contract to see closer objects sharply.

It has to be noted that the horizontal field of view (FoV) of a still eye is only  $160^\circ$  (Denion et al., 2014). The movements of the eye and the head, when combining the FoV of both eyes, increase this value to  $300^\circ$ . Moreover, the eye has an especially sharp vision in the center of the retina (the fovea), which makes these movements even more important. Indeed, the fovea, which is 0.01% of the retina is represented by a large area in the visual

cortex (8-10% of the cortical map's area, Goldstein and Brockmole, 2017).

### 1.3 The specific case of distance evaluation and depth perception

The topic of visual perception and in particular the perception of distances has been the subject of a very large number of studies for several centuries. In fact, visual perception is considered to be the first sense of human beings and the perception of distances – i.e. the ability to evaluate the position of an object in space, relative to another object or to the observer position – makes it possible to measure and quantify this perception. Numerous levels of detail are considered in this field, from the biological phenomenology of perception (Hubel & Wiesel, 1965) up to models seeking to provide a global model of all the factors influencing the perception of distances in visual space (Koenderink et al., 2008).

In the perceptual space, the three dimensions are *up-down*, *sideways* and *towards-away*. This last dimension informs the brain about depth and is known to be the one giving the less accurate information (Ware, 2008). The majority of studies presented in this work aims to evaluate the distance perception of the observers, from their own point of view. That is to say the distance between an object and their own body, defined as *egocentric distance*. Some elements may extend to other dimensions, such as the perception of distances between objects, called *exocentric distances*, and will be discussed further in this thesis.

Moreover, in this thesis we will use various names for similar but still different concepts.

- **Distance perception:** the outcome from the perceptual process, applied to distances; it includes distance evaluation but is also part of some other perception, such as space perception.
- **Distance evaluation:** externalization of distance perception, by word or gesture.
- **Depth perception:** distance perception along the depth axis. It includes egocentric distance perception but also exocentric distances in the depth axis or relief perception.
- **Space perception** includes “perception of the spatial attributes of objects, their size, shape, stability, motility, and their distance and directional locations in reference to each other and to the perceiving subject” (Carr, 1935).



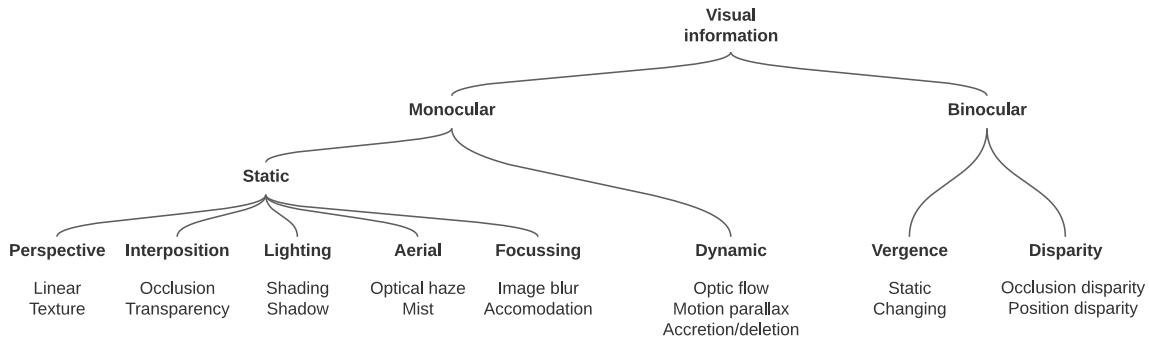


Figure 1.4 – Depth cues classification from Howard, 2002.

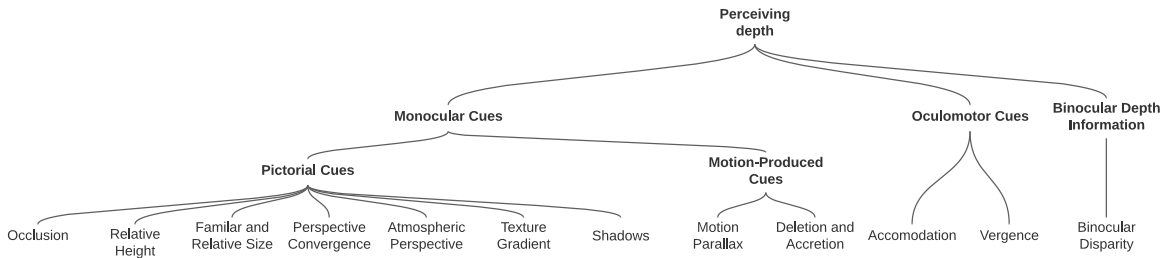


Figure 1.5 – Depth cues classification from Goldstein and Brockmole, 2017.

Thus, we will focus on more general elements of depth perception, without dwelling on the physical and physiological upstream elements. We will first introduce common definitions to categorize distance perception and then present some context elements related to the “visual space”.

### 1.3.1 Depth cues and cue theory

The human is able to analyse the 2D images given by its eyes to get a 3D representation of his environment. This is reported to be achieved by using *depth cues*. Categorizing these depth cues is a difficult task as they are usually closely related and similar in nature. Several authors have proposed slightly different classifications (Goldstein & Brockmole, 2017; Howard, 2002). We present here a typology adapted from Goldstein’s classification (Goldstein & Brockmole, 2017) but simplifying it into only three categories. Our classification is presented in Figure 1.6. Refer to Figure 1.4 and Figure 1.5 for the classifications of the authors mentioned above.

Depths cues can be classified in three categories: oculomotor, pictorial and motion

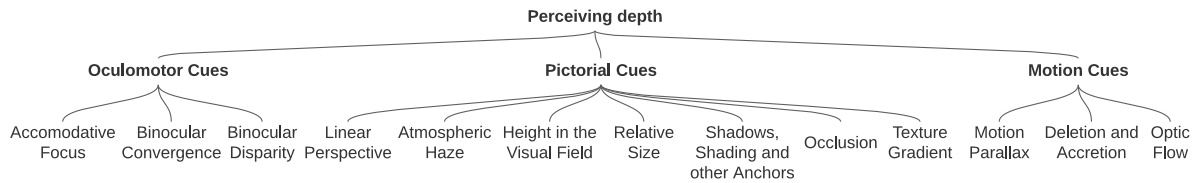


Figure 1.6 – Depth cues classification proposed.

cues.

**Oculomotor cues** refer to the cues provided by the oculomotor system of the eyes and the fact that the human being has two eyes. This category includes:

- **Binocular disparity** takes advantage of the difference in the points of view and the overlap of the visual field of each eye. It allows, by comparing the position of an object on one side with respect to the other, to have an estimation of the distance. For example, an object placed on the right in the field of vision of the right eye and on the left when seen by the left eye will thus be further away than an object in the center of each eye.
- **Binocular convergence** corresponds to the rotational movement of the eyes, i.e. the angle at which the eye muscles must position the eyes to look at an object. Somatosensory receptors in the muscles of the eye allow for continuous feedback of this angular position. Interestingly, the resting position of the eyes is divergent. When the eye muscles relax, the eyes align with the axis of the orbits and are therefore slightly exotropic, i.e. turning outwards.
- **Accommodative focus** results from the movements of the eye muscles which must compress the crystalline lens in order to clearly project the image of the object on the retina. When these muscles are completely relaxed, the objects placed at infinity (beyond 6 meters for the human eye (Koretz & Handelman, 1988)) are sharp. Although this cue gives an absolute indication of distance, it is much more effective in discriminating distance in an ordinal manner.

**Pictorial cues** are sources of depth information that can be depicted in a picture (Goldstein, 2005). This category includes:

- **Atmospheric haze**. The further away the objects are, the greater the air between the observer and the object. Light rays must therefore pass through more air and hence more airborne particles (dust, water droplets, air pollution, etc.). They are



(a) Linear perspective



(b) Distance from the horizon



(c) With no familiar object visible, the distance information provided by the size of the objects is only relative ...



(d) ... while the distance information becomes absolute when familiar elements are present.



(e) Atmospheric haze



(f) Texture gradient

Figure 1.7 – Example of main depth cues.

then deviated by these elements which causes a less sharp image with a slight blue tint.

- **Linear perspective** or **perspective convergence**. Looking at parallel rectilinear elements we can see that they seem to come closer and converge towards a point. This cue is particularly useful in urban environments with many rectilinear elements.
- **Height in the visual field** also provide depth information, in particular for open views. The closer the elements are placed to the horizon line, the farther from the observer they will be perceived.
- **Relative size** makes it possible to interpret larger objects as being closer. The size of an object is inversely proportional to the distance from the viewer.
- **Occlusion** and transparency allow the observer to see the order of the objects. When one overlaps the other then it is perceived as closer. However, this depth cue is only ordinal and does not give an absolute indication of distance.
- **Shadows, shading and other anchors**. Shading refers to the shadows that appear on the object due to its shape and orientation in relation to the light source. It makes it possible to interpret the shape of an object and thus to better understand its position. Shadowing designates shadows cast by the object on another surface. They allow to have more elements of reference, called anchors.
- **Texture gradient**. When elements are repeated many times over a large area, they appear as a texture whose patterns narrows with distance. Distant elements are seen smaller, but unlike the depth cue of relative size, it is the whole texture that serves as a reference. The texture contrast seems to diminish with distance.

**Motion cues** take advantage of the continuity of observation during a displacement to infer distance information from the speed and relative positions of the elements seen. This category includes:

- **Motion parallax**. As the observer moves and fix a still object (or when objects move and the observer stays still), the closest elements will move more than the farthest elements. This makes it possible to estimate the distance between the observer and the object under consideration. This depth cue is not only relative but also allows an absolute estimate of the distance. This cue is massively used in animated movies to simulate depth (Nawrot & Stockert, 2010).
- **Deletion and accretion** are the dynamic counterpart of the occlusion and an ordinal approach to the parallax motion. As the observer moves, some objects are

occluded by others. The more an object is subject to accretion/deletion, the more it is perceived as distant (G. A. Kaplan, 1969).

- We can also mention **optic flow**, which translates all the visual elements generated by a movement. It includes parallax motion but also other types of movement. For example, a rotation of the head or a displacement in the direction of the gaze generates an optic flow but does not involve motion parallax. This cue is a major asset used by animals for visual perception (Eckmeier et al., 2013; Srinivasan et al., 1996; Wallace, 1959).

As summarized in Table 1.1, in addition to type (oculomotor/pictorial/motion) and effective distance range (see subsection 1.3.2), depth cues can also provide an absolute or relative distance information.

Moreover, while the depth cue theory explains high-level depth information, some specificities of the human vision such as stereoscopic depth perception of local surfaces provide low-level interpretation. For example, random-dot stereograms demonstrate that stereopsis based on a simple point-to-point comparison of the two eyes' images can provide a depth sensation and does not require the presence of monocularly visible forms or contours (Collett, 1985; Nakayama, 1996; Ramachandran & Cavanagh, 1985; Takeichi et al., 1992).

In the real world, several of those cues are then combined to evaluate distances. Interaction and fusion models have been proposed to specify the weight of each cues (Landy et al., 1995). Many models were presented, including more or less depth cues and with more or less strict fusion criteria. Overall, no model seems to be clearly dominant. The two main obstacles seem to be the strong and non-linear interactions between cues coupled with the highly individual nature of the perception, as highlighted in the previous section. However, those models suggest that the *consistency* between cues play a great role. Consistency between cues is a much more impacting factor than the individual effect of each cue (Fischer, 2000; Hubel & Wiesel, 1965; Landy et al., 1991).

Finally, a last factor arises: the specific distance tendency (SDT) (Gogel & Tietz, 1973). Indeed, this inner bias is defined as a tendency for the observers to perceive objects at a relatively near distance when presented without any depth cue. In such a configuration, observers tend to perceive objects at about 2 meters (Gogel, 1969).

### 1.3.2 Distance fields

Cutting (2003) proposed a categorization of spaces based on the relative strength of depth cues according to distances. (See Fig. 1.9 for this relative weights plotted.) This induced three different spaces: *near-*, *medium-* and *far-field*, also reported in Vishton and Cutting (1995) as *personal*, *action* and *vista* space. As the second naming suggests, these spaces are also related to human action.

**Personal space** is defined as the zone immediately surrounding the observer, within arm's reach and slightly beyond. Cutting defined it within 1.5 meters, which is the working space for a stationary individual. At that range, five cues are the most important: occlusion, retinal disparity, relative size, convergence and accommodation, in that order.

**Action space** is defined as the space of an individual's public action. He can move quickly within this space, talk without difficulty and throw objects. This space induces also a different ranking of depth cues. The five most, sorted by weight are occlusion, height in the visual field, binocular disparity, motion perspective and relative size. This field extends to about 30 meters.

**Vista space** extends beyond these 30 meters, until the visual field limit, a space which can be seen but not reached quickly. At these distances, the four more important cues are the usually called *pictorial cues*: occlusion, height in the visual field, relative size and aerial perspective.

As explained in the previous part, distance fields are associated with the evolution of the weight of depth cues according to the distance but also actions that can be done in these spaces. On the one hand, they are defined by the depth cues, but they also make it possible to predict and interpret the evolution of perception as a function of distance (See Figure 1.9). As we have previously detailed, action and perception are strongly linked, and the fact that distance fields allow for actions of different nature also induces differences in perception.

### 1.3.3 Global theories of visual space in real environments

Beyond the specific aspect of perception and a factorial modeling of perception, some authors seek on the contrary a more holistic approach to the modeling of perception. The notion of **visual space**, i.e. the global perception observers have of the space based on their visual perception, has been largely investigated in real environments. Depth perception in real environments is covered by a very wide corpus of studies dating back

Table 1.1 – The different depth cues and their properties, from Kytö et al., 2013.

Depth cue	Cue type: (M)onocular (B)inocular (O)culomotor	Effective distance range: (P)ersonal (A)ction (V)ista	Type of distance: (A)bsolute (R)elative
Occlusion	M	P, A, V	R
Angular declination	M	A, V	A, R
Relative size	M	P, A, V	R
Texture density	M	P, A, V	R
Brightness	M	P, A	R
Aerial perspective	M	V	R
Binocular disparities	B	P, A	R
Accommodation	M, O	P	A
Convergence	B, O	P	A
Motion parallax	M	P, A	A, R

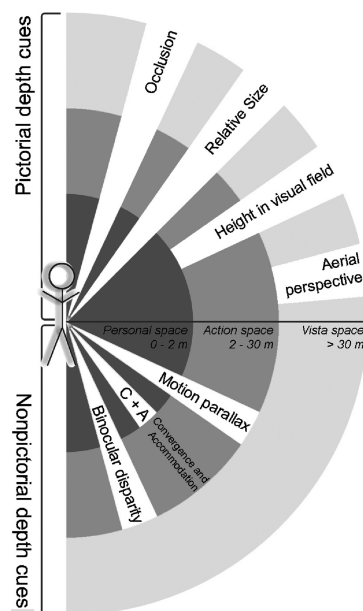


Figure 1.8 – Effective range of depth cues, from Renner et al. (2013).

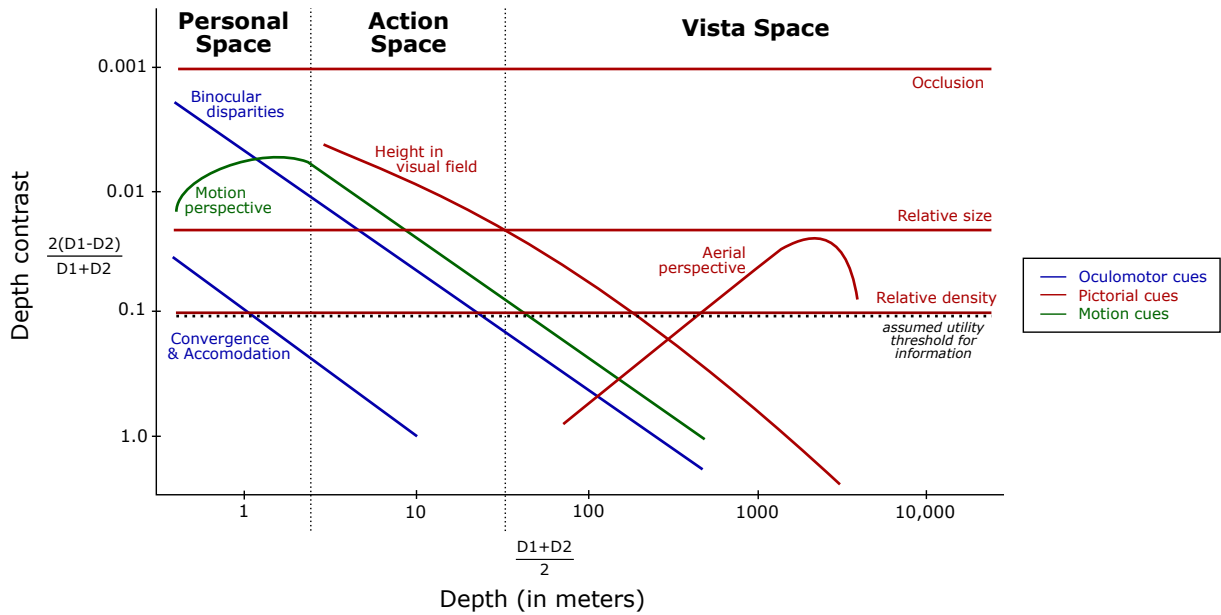


Figure 1.9 – Relative depth cues strength, adapted from Vishton and Cutting (1995). These functions delimit three types of spaces, determined by different depth cues’ weight.

to the 18<sup>th</sup> century, providing both experimental studies and complex numerical models, as detailed thereafter. Many of those studies report that people commit systematic errors when asked to estimate geometrical properties of their environment from a fixed point of view (Battro et al., 1978; Battro et al., 1976; Wagner, 1985).

Several mathematical models have been proposed to explain these differences between the visual space and the real physical space. All these models seek to propose a global theory of transformation between the real space and the visual space measured through perceptual experiments, and thus to find the mathematical shape that best describes the visual space. J. J. Gibson (1950) stated that the visual space is Euclidean. Angell (1974), Daniels (1974) and Reid (1764) suggested it was spherical. a. a. Blank (1953) and A. A. Blank (1959, 1978), Indow (1967, 1979) and Luneburg (1950) and Luneburg (1947) hinted it was hyperbolic. W. C. Hoffman (1966, 1980) proposed that it reflects a Lie algebra group. Then, Wagner (1985) suggested affine-transformed Euclidean models and another observer-based model. To date, there is no consensus yet on this topic. Koenderink et al. (2008) even suggested that the visual space cannot be described by any Riemannian structure.

Thus, most perceptual models lean on the notion of space’s intrinsic curvature. This notion is defined by Gauss (1828) and used by many authors to measure the deformation



of the perceived visual field. They used various techniques to evaluate this curvature, such as alley building (Indow, 1967), exocentric pointing (Koenderink et al., 2000), or distance, angle and area estimations (Wagner, 1985). The studies conducted on that specific parameter state that the curvature of the visual space might not be constant (Indow, 1991). Moreover, Koenderink et al. (2000) suggested that the curvature changes from elliptic in near space to hyperbolic in far space, and then becomes parabolic at very large distances. Koenderink et al. (2002) also evaluated the curvature of frontoparallels: observers were asked to place a radio-commanded vehicle at the bisected point of a linear segment. They found that frontoparallels were significantly concave toward the observers. Interestingly, these results were the opposite of those obtained in a similar open environment, but with a different task (Koenderink et al., 2000).

It therefore appears that the visual space cannot be easily modeled and that current models fail to completely explain it. Furthermore, the inconsistency of distance perception is not only specific to the visual field. The haptic space, i.e. the space perceived through the active exploration of surfaces and objects by a moving subject, also appears to be non-Euclidean (Kappers & Koenderink, 1999) and the haptic perception of spatial properties of objects seems distorted and geometrically inconsistent (Fasse et al., 2000). Even if the in-depth explanation of other perception channels is beyond the scope of this study, it illustrates the global inconsistency of human space perception.

## 1.4 Measuring perception

### 1.4.1 Usual perception evaluation techniques

As detailed previously, perception is a complex and mostly internal process. As such, it is difficult to define what “measuring” means for such a process. For the entire internal part, physical models can often be established. Translations into a functional model of the structure of certain areas of the brain are also beginning to emerge. However, since the core of perception lies in the creation of perceptual experiences, it remains very difficult to measure it precisely. In addition, its variability (inter-personal but also for a single subject) makes the measurement of perception quite difficult. The methods used take advantage of considering the perception mechanism as a whole and primarily evaluate the final outcome of this process.

Therefore, several techniques can be described:

- **Thresholds** are the simplest element to evaluate since the expected answer is binary. People are asked either to compare two stimuli or to assess the presence or absence of a stimulus to evaluate the smallest noticeable stimulus. Methods used to assess these thresholds are usually referred as “classical psychophysical methods” (Fechner et al., 1966).
- **Magnitude estimation** consists in asking a person to evaluate one object or any stimulus criteria on an arbitrary scale. Given a standard reference, they are asked to evaluate a given stimulus. This can be performed with an arbitrary scale but also with more standard scales, such as the metric system. This estimation allows to go beyond the dichotomy of threshold evaluation and enables a continuous measurement of the perception of specific features.
- **Recognition testing** more specifically evaluates the perception at the recognition state. Subjects are asked to recognize certain elements, with or without guidance. This allows on the one hand to test some basic functions in the case of treatment of pathologies but also, for healthy subjects, to evaluate the performance and accuracy of recognition, for example by limiting the available features or in a limited time.
- **Reaction time** is another metric that can be used to evaluate perception. This is rather a roundabout form of metric but allows for example to evaluate the performance of recognition or attention.
- **Phenomenological report.** Only qualitative, this form of metric consists of asking subjects to describe what they perceive. Although this does not provide precise measures of perceptual elements, it does allow more complex elements to be highlighted and then analyzed. Again, the internal nature of perception must be emphasized. Asking observers to express their perception in their own words allows to raise effects that could have been overlooked.
- **Physical tasks or judgments.** This last evaluation method takes advantage that action is part of the perceptual process to use it as a metric. Reaching, navigating, pointing, driving, etc. are all actions that can be used as metrics to assess the perceptions behind them.

After this general presentation of the main categories of evaluation techniques, we will now consider the particular case of the evaluation of the perception of distances.

## 1.4.2 Distance perception evaluation techniques

When it comes to perception a major problem usually arises: the perceived distance of an object by an observer is a subjective cognitive data which cannot directly be measured. Whichever is the chosen reporting technique, the *distance judgement* given by the observer induces but not perfectly represents the perceived distance.

However, many studies suggested various techniques to report judged distance. In this section, we will focus in particular on techniques that measure perceived distances and that can be applied in MR. Other techniques exist, for example to measure more globally visual biases such as alignment or angular deviations (used for example to answer the questions mentioned in the previous paragraph) but they will not be presented here. Loomis and Knapp (2003) presented a review of most of the distance evaluation techniques, and recently Swan et al. (2015) provided a more in depth analysis of some of these for AR.

Perceived distance evaluation techniques can be defined in several groups: verbal report, open-loop action based tasks, perceptual matching and indirect techniques.

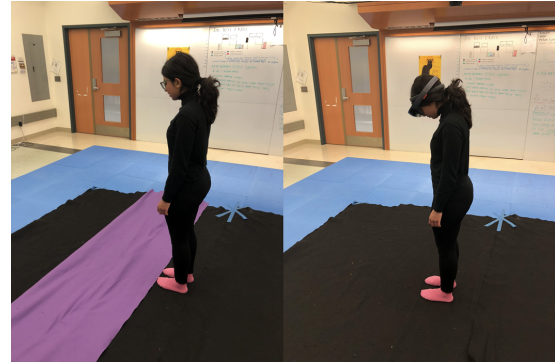
**Verbal report** is the more common technique to evaluate perceived distance. Observers have to verbally estimate the distance to an object in whatever unit they are familiar with (Foley, 1977; E. J. Gibson & Bergman, 1954; E. J. Gibson et al., 1955; Gogel, 1976, 1981; Gogel & Tietz, 1973; Mershon et al., 1977), or using multiples of a given referent (called *magnitude estimation* in psychophysical studies) (Baird & Biersdorf, 1967; Galanter & Galanter, 1973; M. Teghtsoonian & Teghtsoonian, 1970; R. Teghtsoonian & Teghtsoonian, 1970a, 1970b, 1978). This technique has been widely used in real environments (up to 9km, Da Silva, 1985) but indicates systematically compressed perceived distances and can be biased by cognitive knowledge (Loomis & Philbeck, 2008). Using the same technique but with an indirect metric, observers can also verbally estimated the size of familiar objects (Loomis & Knapp, 2003) which makes it possible to deduce the distance they estimate to be from the object. However, this method does not provide an absolute measure of estimated distance because the relationship between perceived size and estimated distance is not accurate for human perception.

**Open-loop action based tasks.** Every technique that use an observer's action to indicate a distance falls into this category. Most of the time, when these techniques are used, the observer looks at a stimulus and then the stimulus disappears and they must recreate the distance they perceived without seeing the original object. This category includes:

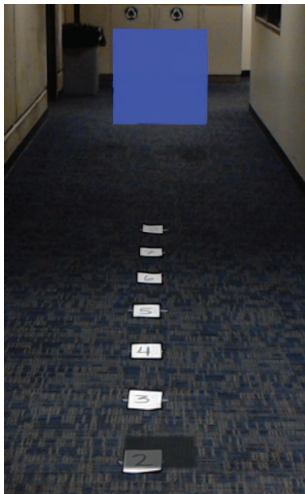
- *Visually directed walking* where, after seeing an object, subjects must walk to the



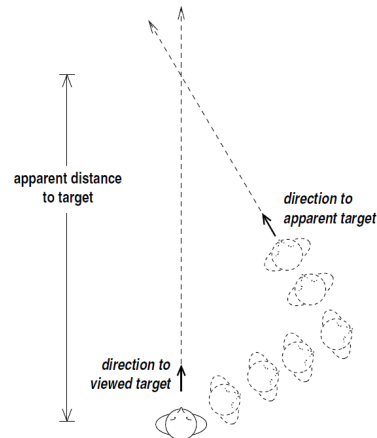
(a) Matching and reaching task from Swan et al. (2015).



(b) Gap affordance estimation from Pointon et al. (2018).



(c) Matching task from Diaz et al. (2017).



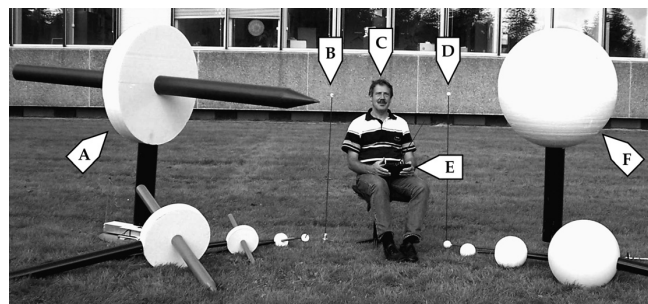
(d) Triangulation by walking task from Thompson et al. (2011).



(e) Bisection task from Swan et al. (2017).



(f) Grasping task from Al-Kalbani et al. (2019).



(g) Exocentric pointing task from Koenderink et al. (2008).

Figure 1.10 – Different distance evaluation techniques.

location of the object (Creem-Regehr et al., 2005; Knapp & Loomis, 2004; Loomis & Knapp, 2003; Messing & Durgin, 2005; Rieser et al., 1995; Willemsen et al., 2004; Wul et al., 2004). This technique has been widely used in many environments and is reported to be accurate up to 20 m in real environments (Loomis & Knapp, 2003), with little systematic bias (Waller & Richardson, 2008) and improving performance with training (Mohler et al., 2006; Richardson & Waller, 2005).

- *Imagined visually directed walking*, where participants estimate how long it would take to walk to targets (Plumert et al., 2005).
- *Triangulation by walking*, where subjects start walking in an oblique direction from the direction of a previously viewed target. Then, on the experimenter signal, they turn and take several steps towards where they perceived the previously viewed target to be (Loomis & Knapp, 2003; Thompson et al., 2004; Willemsen et al., 2004).
- *Open-loop pointing tasks*, where participants indicate distance with a finger or manipulate a slider that is hidden from their view (Mon-Williams & Tresilian, 2000; Swan et al., 2015). This technique can also be used for measuring derivative visual space’s indicators such as orientation (Koenderink et al., 2000).
- *Throwing*, where participants have to throw (blindly or not) an object toward the target previously seen (Peer & Ponto, 2016, 2017; Sahm et al., 2005; Thomson, 1983).

**Perceptual matching** procedures consist in closed-loop action-based tasks (Bingham et al., 2001; Prablanc et al., 1979). Participants are asked to align a cursor with the target. The distinctive aspect of these techniques is that the participant can compare the position of the target and the cursor at any time. Thus, some authors point out that the nature of the task can be changed from an evaluation task of a perceived distance to simply minimizing the disparity between the target and cursor. Then, these techniques would fail to measure *absolute distance perception* (Bingham & Pagano, 1998; Prablanc et al., 1979). However, this technique is still widely used to compare different targets and cursors, for example in AR (Ellis & Menges, 1998; McCandless et al., 2000; Wul et al., 2004).

**Indirect techniques** include all techniques that do not directly measure distance by asking the user to recreate it but use a derived metric. The verbal estimation of the size of objects mentioned earlier may also arguably fall into this category.

This category also includes:

- *Forced-choice tasks*. Participants are asked to choose a position among several (Livingston et al., 2003), or to perform an ordinal task by choosing for example the closest one among several stimuli (Rolland, Gibson, et al., 1995).
- “*Can I pass?*” *techniques* which rely on affordance and the body perception of participants. They are asked to look toward two poles and to assess if they can pass through without touching them (Geuss et al., 2010; Walker et al., 2012). Similar techniques can be found with other tasks relying on affordance, such as crossing a gap (Pointon et al., 2018; Wu et al., 2019).
- *Bisection*. Participants are asked to point or place an object at the midpoint between two objects. This technique is close to the matching technique presented before but does not suffer from the alignment bias underlined. However, distances given by bisected intervals are reported as compressed by some studies (Gilinsky, 1951), accurate by others (Bodenheimer et al., 2007; Da Silva, 1985; Purdy & Gibson, 1955; Rieser et al., 1990), or even expanded and environment dependent (Vishton & Cutting, 1995).

Overall, some general outlines can be drawn from all these different techniques. First, one of the major issues of reporting distances and also one of the theory to explain the variability of distance reporting is that distance judgments are influenced by the amount of energy observers anticipate to expend in order to reach the item to which distances are judged (Decety et al., 1989; Proffitt, 2006). This so called effort-based theory helps to explain some perception results and enriches attempts to create models of perception. However, like all the theories on perception, it is still neither fully understood nor fully accepted. Then, the methods used to obtain distance judgments have been reported to have a major effect on the accuracy of those judgments (Elliott, 1987; Rieser et al., 1990; Thomson, 1983; Witmer & Kline, 1998). As such, since the techniques affect the results, both by their nature and by the fact that they need an action from the observers, perception studies cannot be carried out without this multitude of techniques. A combination of these techniques could therefore be needed in order to compensate for the biases they may cause.



Mixed reality allows to render virtual elements in real. However, this is not without effects on reality. In this chapter, we present the definitions and stakes of mixed reality, as well as the technologies allowing to realize this mixed reality.

# MIXED REALITY - HOW TO RENDER A VIRTUAL WORLD AND BLEND IT WITH REALITY

---

## Contents

---

<b>2.1</b>	<b>Definitions . . . . .</b>	<b>47</b>
2.1.1	Real and Virtual . . . . .	48
2.1.2	Virtual Reality . . . . .	48
2.1.3	Augmented Reality . . . . .	49
2.1.4	Continuum, Mixed Reality and other notions . . . . .	49
<b>2.2</b>	<b>Devices . . . . .</b>	<b>50</b>
2.2.1	Near-eye displays . . . . .	51
2.2.2	Handheld Devices . . . . .	55
2.2.3	Distant displays . . . . .	55
<b>2.3</b>	<b>Mixed reality and perception, limitations and challenges . .</b>	<b>56</b>
2.3.1	The vergence-accommodation conflict . . . . .	57
2.3.2	X-ray vision . . . . .	58

---

## 2.1 Definitions

The definitions and scope of so called Virtual, Augmented or Mixed Reality (MR) are numerous and changing. The purpose of this section is then to elicit the main principles underlined by these notions, in order to highlight the specificities related to perception.



### 2.1.1 Real and Virtual

First of all, and because these concepts will be widely discussed afterwards, let us first take some time to define these two notions. We will use the definitions given by Milgram and Kishino (1994):

- **Real** objects are any objects that have an actual physical existence.
- **Virtual** objects are objects that exist in essence or effect, but not formally or actually.

### 2.1.2 Virtual Reality

We will not discuss here the premises of what can be seen as VR, as they can be found in descriptions such as Plato’s myth of the cave (Plato, 400) or Giovanni Fontana’s magic lantern (Fontana, 1420). We will focus on Virtual Reality in its current use, a term popularized by Jaron Lanier in the 1980s (Machover & Tice, 1994).

For the remainder of this work, we will use the following definition, stated as *technical definition* in Fuchs and Moreau (2003):

Virtual Reality is a technical and scientific area making use of computer science and behavioral interfaces in order to simulate 3D entities behavior in a virtual world that interact in real time among themselves and with the user in pseudo-natural immersion through sensory-motor channels.

However, it is also interesting to mention another, so-called *functional* definition from the same book:

Virtual reality will allow the user to extract himself from the physical reality in order to virtually change the time, place and/or type of interaction: interaction with an environment simulating reality or interaction with an imaginary or symbolic world.

This definition is quite close to one of those given by Lanier (2017):

The substitution of the interface between a person and the physical environment with an interface to a simulated environment.

Moreover, I cannot resist quoting a few other definitions, from Jaron Lanier’s book, which are related to concepts that will be detailed later in this work. One recalls the fundamental notion of perception, while underlining the difference between real and virtual, and the potential contribution of one to the other:

A coarser, simulated reality fosters appreciation of the depth of physical reality in comparison. As VR progresses in the future, human perception will be nurtured by it and will learn to find ever more depth in physical reality.

And this last definition, which underlines the contribution of the brain in the process of creation and perception, essential to the perception of a coherent environment:

Technology that rallies the brain to fill in the blanks and cover over the mistakes of a simulator, in order to make a simulated reality seem better than it ought to.

Overall, some common elements of these definitions can be highlighted. First, the fact that “virtual” reality is defined in opposition to “actual” reality; as a mirror or as a way of substituting for it. In addition, one also observes the major place given to interaction in these definitions. We will discuss later in this work the effects of interaction through the prism of perception in VR, but it is worth emphasizing that interaction lies in most of the VR’s definition.

### 2.1.3 Augmented Reality

Concerning AR, we will use the definition of Azuma (1997), revised in Azuma et al. (2001) and supported by the book from Schmalstieg and Hollerer (2016). This definition describes AR as:

- [...] any system that has the following three characteristics:
- combines real and virtual objects in a real environment;
  - runs interactively, and in real time; and
  - registers (aligns) real and virtual objects with each other.

This definition can also be compared with that of VR. Both emphasize the importance of interaction, but they are at odds in their purpose. While VR seeks to substitute the virtual for the real, AR seeks to mix the two in the best possible way. However, does this difference in purpose between the two concepts imply a difference in nature? Are these two notions really so different?

### 2.1.4 Continuum, Mixed Reality and other notions

Rather than two separate concepts, Milgram et al. (1994) proposed the vision of a continuum between reality and virtual reality, which they named Mixed Reality (MR). It would contain all the concepts and techniques that mix real and virtual, excluding solely

real or virtual techniques. Technologies are then placed on this continuum according to their proportion of real/virtual elements. AR corresponds here to any system which, in a real environment, adds virtual elements. We can also cite *Augmented Virtuality* which, mirrored to AR on this continuum, corresponds to the addition of real elements in virtual environments (Milgram et al., 1994).

**Does VR belongs to MR?** The term “Mixed Reality (MR)” is sometimes preferred over “augmented reality” because it is considered to be broader and more versatile (Schmalstieg & Hollerer, 2016). Therefore, MR is both defined as a continuum including AR but not VR, but also as a more generic word to refer to any kind of application mixing real and virtual. Moreover, in its motivation VR appears to be in opposition with every shape of the concept of MR which seeks to mix real and virtual content, while VR seeks to substitute the real for the virtual. However, it does not fully succeed to. Indeed, VR always includes interaction with a human (who is undoubtedly real). For example, a CAVE is a VR device but the real body of the user is still visible. This necessity of interaction implies that VR cannot be fully virtual as the interaction always has a real origin. As such, and despite its purpose, VR does not succeed to be solely virtual. Since the Ultimate Display (Sutherland, 1965), which would be able to fully substitute real for virtual is not ready yet, a “complete” Virtual Reality cannot be achieved yet. As such, we believe that for now, and as long as a real human would interact with virtual elements in VR, Virtual Reality has to be considered as a Mixed Reality technology.

## 2.2 Devices

The devices used for AR purposes can be classified in three categories, according to the distances between the device, the observer and the object (Bimber & Raskar, 2005; Schmalstieg & Hollerer, 2016): near eye displays, handheld devices and projectors. In order to underline the perceptual challenges associated with these categories and be able to compare with VR devices, we will complement these categories with VR devices. The main difference is that the “projectors” category will be extended in order to include all planar devices, such as screens. See Figure 2.1 for an overview of this classification.

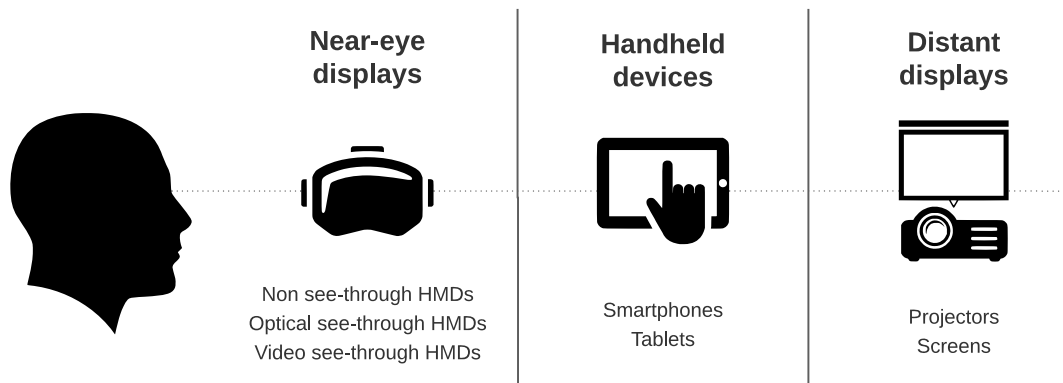


Figure 2.1 – Classification of MR devices, adapted from Schmalstieg and Hollerer (2016).

### 2.2.1 Near-eye displays

Near-eye displays are devices that “place the image directly in front of the user’s eye using one or two small screens” (Bowman et al., 2004). Nowadays near-eye displays are generally Head Mounted Displays (HMDs) (LaViola Jr et al., 2017). Wearing a device on the head involves complex designs in order to make the device unobtrusive and comfortable (Rolland & Cakmakci, 2009). Even if the first one were not tracked in position, today’s every modern HMDs have in-built tracking systems. There are three categories of HMDs, depending on if and how the real environment is visible: Non see-through HMDs, video see-through HMDs and optical see-through HMDs.

**Non see-through HMDs** In VR, near-eye displays are usually referred as “non see-through HMDs” or simply “HMDs”. Those devices are designed to make the user unable to see the real world and then provide a complete immersion in the VE. Nowadays, several major companies have developed their own non see-through HMD like HTC, Oculus, Valve or Samsung (see Figure 2.4). These devices use either two small LCD (or OLED) screens or a smartphone combined with optical lenses to enable the user to see the rendered image. Usually, the devices that use smartphones are limited by the smartphone’s performance and it is common to have more latency and a lower refresh rate.

**Video See-Through (VST) HMDs** These HMDs use cameras placed in front of the user’s eye in order to provide the visual feedback of the real world. Images from real and virtual environment are then displayed by the HMD in the same way as for non see-through HMDs. The Video See-Through (VST) technique allows for a good superimposition of the



Figure 2.2 – Example of commercial HMDs: the HTC Vive (left), the Microsoft HoloLens 2 (center) and the HP VR1000 (right)

real and the augmented layers, since they are both displayed by the same screen. However, the intrinsic specificities of the cameras used to provide visual feedback, such as field of view or resolution limitation and optical distortion, may induce viewing issues.

**Optical See-Through (OST) HMDs** These HMDs mostly use semi-transparent mirrors to superimpose the augmentation and the direct view of the environment. The main advantage of this technology is that it provides a direct view of the environment, with sharp and quick view but induce more constraints on the augmentations' rendering (e.g. latency, colors and depth of field). Indeed, the augmentations will suffer from the constraints of the displays while the direct real view will not be affected. Then, the differences between real and virtual rendering will be even more visible.

**More details on accommodation issues** The more common Optical See-Through (OST) HMDs are devices which use fixed lenses and screens to provide the virtual images (Schmalstieg & Hollerer, 2016). However, this induces a fixed accommodation cue.

In order to overcome this phenomenon, some AR devices propose other techniques to display virtual images, which can be categorized as follows:

1. The most common of those devices are *varifocal displays*. Images provided by these devices are displayed at a certain depth but the focal plane is movable. Varifocal displays can be implemented in several ways, each having some limitations, e.g. FoV, bulkiness or eyebox size (Akşit et al., 2017; Dunn et al., 2017; Hu & Hua, 2014; S. Liu et al., 2008).
2. *Multifocal devices* use multiplanar volumetric projection displays based on a stack of laminated planes. This technique can create a more plausible accommodation depth

cue and, if dense enough, can provide a continuous and accurate depth cue (Hua & Javidi, 2014; Rolland et al., 2000). This technique is used in the Magic Leap AR headset<sup>1</sup> to provide a better depth sensation. However, this device only provides two distinct focal planes.

3. Other devices try to reproduce the exact light-field which would have been created by the virtual images with *holographic rendering*. These devices are able to provide a three dimensions image, with a per-pixel focal control. If the final image is precise enough, the depth cues provided by the virtual object are the same as the ones which would have been created by a real object (H. J. Kim et al., 2015; Maimone et al., 2017).

All these technologies are more accurate than standard fixed-focal OST displays. However, they still suffer from other major drawbacks and limitations, such as a limited resolution, a high space requirement and high cost. Overall, studies presenting those technologies usually focus on a specific rendering technique and are never compared to each other. Furthermore, these studies focus on technological challenges but the techniques they are presenting remain untested regarding their impact on perception. Then, another technology can be presented; a type of displays that does not have a fixed focal plane since they do not have any focal plane: Retinal Projection Displays (RPDs).

**What are Retinal Projection Displays?** In 1980, the Scanning Laser Ophthalmoscope (SLO) was developed by Webb et al. (Webb et al., 1980). This device was used to provide retina images using a scanning light beam. This first prototype of retinal scanning device was then iteratively improved along the years to include a modulation of the light beam to eventually project a whole image in the retina. As a result, the light beam stimulates the retina like any ordinary light ray seen by the eye, allowing the device to act as a virtual display.

Retinal scanning technology relies on the Maxwellian view, cf. Figure 2.3a. This specific configuration of an optical display forces all the light beams from the visual source to converge at the center of the pupil. This point is the center of the optical system of the eye and the light beams are supposed thin enough not to be affected by the eye lens. Thus, the thin light beam directly hits the retina and forms a very small and sharp point of light.

---

1. <https://www.magicleap.com/>

2. <https://www.qdlaser.com/en/applications/eyewear/>

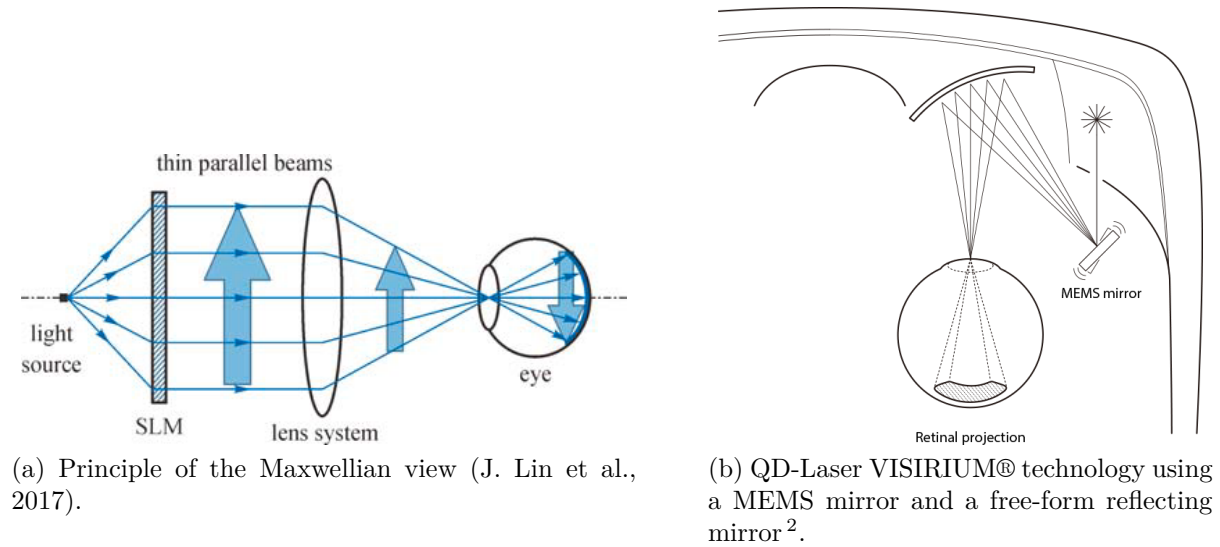


Figure 2.3 – Possible optical configurations of RPDs.

Enforcing Maxwellian view on an optical display can be achieved in several ways. The two main ones are: (1) using a Spatial Light Modulator (SLM) to generate a grid of thin parallel beams that converge toward the eye pupil using a lens system and (2) using a Micro-Electro-Mechanical System (MEMS) scanning mirror to reflect a laser beam toward a reflecting free-form mirror which lastly focuses the light in the eye pupil (see Figure 2.3a&b.). While the first approach is mainly used for rendering only virtual images (Dewald et al., 2016), the second one is more suitable for AR purposes since the free-form mirror can be semi-transparent to make both real and virtual elements visible. Several other Maxwellian-based configurations and more devices layout have also been proposed in the literature and, for more details, the reader is referred to Lin et al.’s survey (J. Lin et al., 2017).

Thanks to their design based on the principle of Maxwellian view, RPDs are able to render an image that is always sharp, regardless of the focal accommodation of the eye. In addition, recent technical improvements related to MEMS mirrors allow the emergence of more compact devices that could be used for AR headsets. These devices therefore have the potential to replace the current OST HMDs if their specificities allow them a better depth perception for users.

## 2.2.2 Handheld Devices

This category includes phones and tablets, as well as some more unusual devices like autostereoscopic tablets<sup>3</sup>. Those devices have the characteristic to use a camera to capture the video stream of the real environment and then display it on a screen; that is to say from a displaced point of view. This specific paradigm has been detailed in (Kruijff et al., 2010) but can also be compared with some studies conducted on pictures and even paintings. This particular effect appends the more common issue of the limited FoV for such devices. The tracking part of this tools are mostly image-based. Thus, in the past decade, new tracking techniques have been developed, and computational power of handheld devices has increased (Dey et al., 2012). This technology being mature enough to provide proper AR, more experiments could be conducted with these devices. However, since they have to be carried by hand, the interactions with the virtual content remain limited to indirect interactions.

## 2.2.3 Distant displays

We choose here to regroup all those devices in one category but we can distinguish different types of devices.

**Projectors and screens for VR** This category includes all the devices based on a fixed plane to provide virtual images. This encompasses wide or surrounding displays such as immersive screens or CAVEs (Cruz-Neira et al., 1993), but also smaller displays such as workbench displays. These devices render stereoscopic images on planar surfaces and the point of view of the observer is usually tracked to provide a coherent rendering. One of the key advantages of these technologies is the ability to immerse the user in a virtual world while keeping the possibility to see its own body.

**Projectors for AR** Two types of projectors can be distinguished here, characterized by their uses. Indeed, projectors can be used to do Spatial Augmented Reality (SAR) (Bimber et al., 2005) which directly projects virtual elements onto a physical real scene. If the projected image is not stereoscopic, the image is not dependent on the point of view and allows several people to see and interact at the same time with the same augmentation.

---

3. Autostereoscopic displays have the ability to provide stereoscopic images without needing addition tools such as specific glasses. The reader is referred to the survey from Holliman et al. (2011) for a complete review of autostereoscopic techniques.



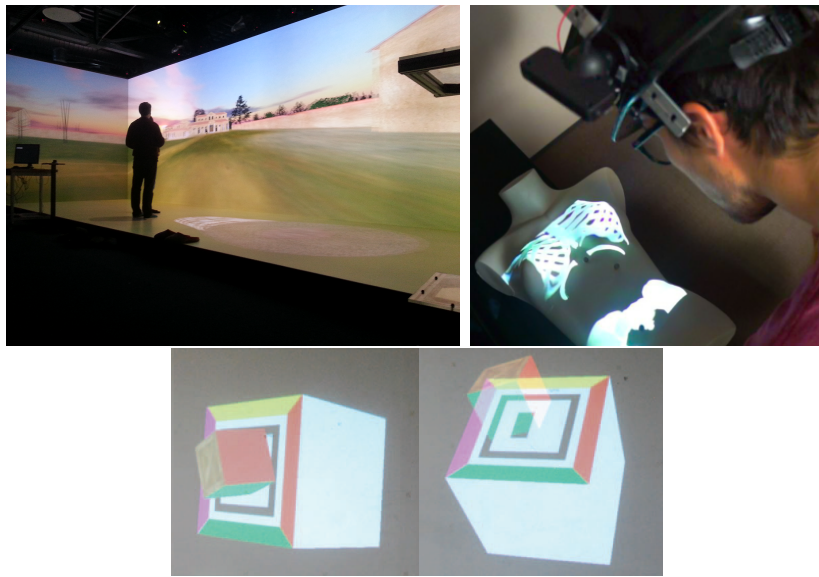


Figure 2.4 – Examples of distant displays: the Immersia CAVE-like display from Inria Rennes-Bretagne Atlantique and IRISA (top left), MoSART (Mobile Spatial Augmented Reality for 3D Interaction With Tangible Objects), from Cortes et al. (2018) (top right), VST/OST replication with projectors, from Plopski et al. (2016) (bottom).

Projectors can also be used to display images on semi-transparent glasses. This case is similar to the screen condition described for VR but allows to see the real scene superimposed with the augmentations.

Some researchers have also used projectors to mimic OST and VST conditions (Plopski et al., 2016).

## 2.3 Mixed reality and perception, limitations and challenges

Due to all technical specificities of MR displays cited previously, the perception of mixed environments can be different from the perception of reality, and even different according to the device used. In this part, we will focus more in depth on some issues specific to MR. One is related to the technical limitation of the devices and the functioning of the human perceptive system, the Vergence-Accommodation Conflict (VAC), while the other issue is related to the display capacity of MR, which goes beyond the usual rendering of objects: the X-ray vision.

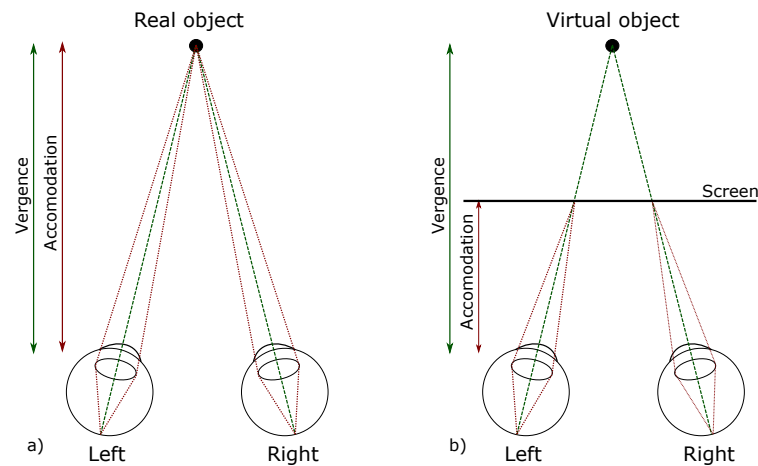


Figure 2.5 – The Vergence-Accommodation conflict: a) natural vision versus b) screen based-vision.

### 2.3.1 The vergence-accommodation conflict

As detailed in subsection 1.3.1, accommodation and vergence are two important cues for depth perception (Azuma, 1997). Under normal circumstances, these cues are coherent: they both provide the same depth information when looking at a specific distant object. The human vision system is familiar with this coherence and acts accordingly. When the eyes converge at a certain distance, they accommodate simultaneously to see the image sharp at this distance (Fincham & Walton, 1957; Martens & Ogle, 1959) (see Figure 2.5.a).

In AR OST headsets as for stereoscopic screens, the distance of vergence and the distance of accommodation are different as shown in Figure 2.5.b. To perceive the image sharp, the eyes have to focus at a fixed distance (at the screen position for screen-based devices, and at the focal distance of the lenses used for OST HMDs). However, they have to converge at another distance, which corresponds to the intended distance of the virtual object. This phenomenon is known as the Vergence-Accommodation Conflict (VAC) (D. M. Hoffman et al., 2008; Julesz, 1971)

This effect has been extensively studied in real, virtual and augmented reality and is known to be responsible for inducing visual fatigue and in a more general way to hinder visual performance (D. M. Hoffman et al., 2008). However, the VAC may be even more impacting in AR OST devices since an augmented environment contains both real objects, which provide coherent depth cues, and virtual objects, which do not.



Figure 2.6 – Examples of visualization techniques using x-ray vision: virtual hole, from State et al. (1994) (left), edge-based representation, from Avery et al. (2009) (center), and saliency-based representation, from Sandor et al. (2010) (right).

### 2.3.2 X-ray vision

Showing to the user occluded information that would not be visible in the real world is called “x-ray vision” (Feiner et al., 1995). This possibility is specific to AR applications. Different visualization techniques have been experimented to provide an accurate representation of this unusual vision. Moreover, this ability has been used to many applications in different fields such as medical, architectural, inspection or military applications.

Some of the first applications that used this so-called x-ray vision were providing a symbolic representation of the occluded objects, while other ones displayed a “virtual hole” in the real object (State et al., 1994). Subsequent studies experimented other representations which provide a more realistic view of the occluded part such as edge-based (Avery et al., 2009), saliency-based (Sandor et al., 2010) or melted visualization (Dey et al., 2010). In their review, Dey and Sandor (2014) advocate that these representations for occluded objects have to be considered as a suite of visualizations in order to complement each other in different situations.

However, the usual depth cue of occlusion is absent from this representation which deeply impact the ability of observers to perceive distances. To understand this kind of representation, the human perceptual system has to interpret both the relation between the occluded object with its surroundings and the relation with the viewer’s point of view. Moreover, the occlusion depth cue is stated as one of the most important, in particular to order objects along the depth axis (see subsection 1.3.1). Breaking the coherence with the other cues might induce a major impact on depth perception.





Today, Augmented Reality is used in a wide variety of fields, including entertainment, prototyping, training, etc. However, the distance perception biases in Augmented Reality remain major challenges for the adoption of this technology. In this chapter, we propose a detailed presentation of studies showing the impact of different factors on perception in Augmented Reality.

# OVERVIEW OF KEY STUDIES IN DISTANCE PERCEPTION IN AUGMENTED REALITY

---

## Contents

---

<b>3.1</b>	<b>Depth cues</b>	<b>63</b>
3.1.1	Binocular disparity and convergence	63
3.1.2	Monocular studies	66
3.1.3	Accommodation	67
3.1.4	Motion parallax	68
3.1.5	Peripheral vision	69
3.1.6	Brightness	69
3.1.7	Shadows	69
3.1.8	Height of the object	70
<b>3.2</b>	<b>Intrinsic factors</b>	<b>71</b>
3.2.1	Gender	71
3.2.2	Age	72
3.2.3	Avatar	72
3.2.4	Learning and Feedback	73
3.2.5	Emotions	73
<b>3.3</b>	<b>Device and rendering parameters</b>	<b>74</b>
3.3.1	Latency	74
3.3.2	Graphics quality and realistic rendering	74
3.3.3	Blurring	75
3.3.4	Screen Size	76
3.3.5	Resolution	77
3.3.6	Handheld specific factors	77
3.3.7	Video See-Through displays	78

3.3.8	HMD specific issues . . . . .	79
<b>3.4</b>	<b>Environment and object related factors . . . . .</b>	<b>80</b>
3.4.1	Environments . . . . .	80
3.4.2	Occluder and x-ray vision . . . . .	81
3.4.3	Auxiliary augmentations . . . . .	82
3.4.4	Object appearance and position . . . . .	83
<b>3.5</b>	<b>Besides depth: exocentric distance perception . . . . .</b>	<b>84</b>
3.5.1	Exocentric distance perception in real life . . . . .	84
3.5.2	Exocentric distance perception in VR . . . . .	85
3.5.3	Exocentric Distance Perception in AR . . . . .	86
<b>3.6</b>	<b>Conclusion: Are all perceptual biases in MR explained by a specific factor effect? . . . . .</b>	<b>87</b>

---

In this section, we present the related work linked to distance perception in MR. Overall, we will focus on studies evaluating this perception in AR, compared with real objects or environments in order to elicit the potential factors of the perceptual biases in MR. However, we must already acknowledge that, to provide a precise evaluation of each factor’s influence, the studies have to tackle each factor, with several objects and environment type (real/virtual) and for different distances, since we already underlined the evaluation of the weight of depth cues according to the distance. Considering every potential factor and interaction between them represents a tremendous number of potential studies, and most of them are still open research questions. As such, we will focus in this section on AR studies, but real and VR works will also be presented when needed to provide a more complete overview of the potential effect of every factors on distance perception.

## Foreword on distance underestimation

Before going further, we must mention the most famous result of distance perception in MR: distance underestimation. Primarily from VR studies and then extended to AR, this result is one of the most documented result in perception in MR. Distance estimation is about 94% accurate, but in a Virtual Environment, it is only about 80% accurate (C. J. Lin & Woldegiorgis, 2015). Egocentric depth is underestimated when objects are viewed on the ground plane, at near- to medium-field distances, and the VR environment is presented in a HMD<sup>1</sup>. However, no real consensus exists on the origin of this effect. A lot

---

1. Bruder et al. (2015), Creem-Regehr et al. (2015), Creem-Regehr et al. (2005), Knapp and Loomis (2004), Lampton et al. (1995), Li et al. (2014), Loomis and Knapp (2003), Messing and Durgin (2005),

of potential factors explaining it have been studied but none of them fully explains this phenomenon. The underestimation of distances in VR is more likely to be caused by a combination of technical limitations as well as inner biases from the observers. It has to be mentioned that this effect is greatly reduced in modern HMDs even if some differences remain regarding their distance compression depending on the experimental conditions (Buck et al., 2018; Kelly et al., 2017).

This result was also found for AR (Swan et al., 2007; Swan et al., 2006) but less studies tackled this specific aspect. Moreover, while it makes sense to evaluate distances in a virtual environment, doing so in an “augmented environment” is equivalent to evaluating distances in a real environment or with virtual objects. So there are a lot of possible configurations. As such, we decided to structure the following work not by trying to strictly characterize the perception in AR but by detailing the different parameters likely to influence it. These parameters can be specific to AR but also more widely shared for other types of environments. We also provide the Table 3.1 which intends to report all the published studies (as of July 2020) on distance perception in AR and the key elements of these studies, as well as a Sankey diagram showing the main characteristics of the reported studies.

To organize the works presented below, potential factors on distance perception will be categorized as follows: depth cues, intrinsic factors, technical and device related factors, and environments and objects related factors.

## 3.1 Depth cues

As we extensively presented in subsection 1.3.1, the human perceptual system is primarily relying on several depth cues in order to evaluate distances. As such, many perceptual studies focused on the evaluation of the effect of each of these depth cues.

### 3.1.1 Binocular disparity and convergence

Binocular disparity and convergence are simulated in virtual rendering devices with two distinct images displayed on each eye. This factor is usually considered as the most

---

Plumert et al. (2005), Rolland, Gibson, et al. (1995), Rousset et al. (2015), Thompson et al. (2004), Willemsen et al. (2004, 2009), Witmer and Kline (1998), and Witmer and Sadowski (1998), Wul et al. (2004)



Table 3.1 – Overview of characteristics of distance perception studies in Augmented Reality

Reference	Device	Occluder	Field	Reporting technique	Accommodation	Background	Comparison with real	Head movements
Ellis and Bacher, 1994	OST HMD	Yes	Personal	Matching	Close	Blank / Plain	No	No
Ellis et al., 1995	OST HMD	Yes	Personal	Matching	Close	Blank / Plain	No	Limited
Rolland, Gibson, et al., 1995	OST HMD	No	Personal	Forced choice	Far/Collimated	Blank / Plain	Yes	?
Ellis and Menges, 1997	OST HMD	No	Personal	Matching	Close	Blank / Plain	No	Limited
Ellis and Menges, 1998	OST HMD	Yes	Personal	Matching	Far/Collimated	Blank / Plain	No	Limited
	OST HMD	Yes	Personal	Matching	Close	Blank / Plain	No	Limited
	OST HMD	Yes	Personal	Matching	Far/Collimated	Blank / Plain	No	Limited
	OST HMD	No	Personal	Matching	Close	Blank / Plain	No	Limited
	VST HMD	No	Personal	Pointing	?	Wall / Table	Yes	Yes
McCandless et al., 2000	OST HMD	No	Personal	Matching	Close	Blank / Plain	No	No
Mon-Williams and Tresilian, 2000	Other	No	Personal	Reaching	Coherent	Blank / Plain	No	No
Furnanski et al., 2002	Other	Yes	Personal	Forced choice	Coherent	Wall / Table	No	Yes
Rolland et al., 2002	OST HMD	No	Personal	Forced choice	Close	Blank / Plain	No	No
Livingston et al., 2003	OST HMD	Virtual	Vista	Forced choice	Close	Outdoor	No	Yes
Sugano et al., 2003	VST HMD	No	Vista	Ordering	?	Wall / Table	No	Yes
Jerome and Whitner, 2005	OST HMD	No	Personal	Matching	?	Indoor (room)	Yes	Yes
	OST HMD	No	Action	Verbal	?	Indoor (room)	Yes	Yes
Livingston et al., 2005	OST HMD	No	Action	Verbal	Close	Indoor (room)	Yes	Yes
Wither and Hollerer, 2005	VST HMD	No	Action	Matching	Close	Indoor (corridor)	Yes	Yes
Swan et al., 2007	OST HMD	No	Vista	Blind walking	-	Outdoor	No	Yes
	OST HMD	No	Action	Blind walking	Close	Indoor (room)	Yes	No
	OST HMD	No	Action	Verbal	Close	Indoor (room)	Yes	No
	OST HMD	Virtual	Action	Blind walking	?	Indoor (room)	No	No
Jones et al., 2008	OST HMD	No	Action	Blind walking	?	Indoor (corridor)	Yes	Yes
Livingston, Ai, and Decker, 2009	OST HMD	No	Personal	Forced choice	Far/Collimated	Indoor (corridor)	Yes	Yes
Livingston, Ai, Swan, et al., 2009	OST HMD	No	Personal	Forced choice	Far/Collimated	Indoor (corridor)	Yes	Yes
	OST HMD	No	Action	Blind walking	?	Wall / Table	No	No
Greechkin et al., 2010	OST HMD	No	Action	Matching	Far/Collimated	Indoor (corridor)	No	Yes
	OST HMD	No	Action	Blind walking	Far/Collimated	Indoor (corridor)	Yes	Yes
	OST HMD	No	Action	Blind walking	Far/Collimated	Indoor (corridor)	Yes	Yes
	OST HMD	No	Action	Timed imagined walking	Far/Collimated	Outdoor	No	Yes
	OST HMD	Yes	Vista	Blind walking	-	Indoor (corridor)	Yes	Yes
Dey et al., 2010	OST HMD	No	Personal	Matching	?	Blank / Plain	Yes	Yes
Jones et al., 2011	OST HMD	No	Personal	Reaching	?	Blank / Plain	Yes	Yes
Slugh et al., 2011	OST HMD	No	Personal	Reaching	-	Blank / Plain	Yes	Yes
	Handheld	Yes	Vista	Verbal	-	Outdoor	No	Yes
	Handheld	Yes	Vista	Verbal	-	Outdoor	No	Yes
	Handheld	Yes	Vista	Verbal	-	Outdoor	No	Yes
Dey et al., 2012	VST HMD	Yes	Action	Verbal	?	Wall / Table	No	Yes
	Handheld	Yes	Action	Verbal	-	Blank / Plain	No	Yes
Kytö et al., 2013	Handheld	No	Personal	Matching	-	Blank / Plain	No	Yes
Berning et al., 2014	VST HMD	No	Action	Matching	Far/Collimated	Indoor (room)	No	No
Kytö et al., 2014	Projector	No	Personal	Sketching	-	Blank / Plain	No	No
C. J. Lin et al., 2015	Projector	Yes	Action	Forced choice	-	Indoor (room)	No	No
Long et al., 2015	OST HMD	No	Personal	Reaching	Collimated	Blank / Plain	Yes	No
Swan et al., 2015	OST HMD	No	Personal	Matching	Collimated	Blank / Plain	Yes	No
	OST HMD	No	Personal	Pointing	Close	?	No	Yes
Medeiros et al., 2016	VST HMD	No	Personal	Pointing	Close	?	No	Yes
	Projector	Virtual	Personal	Matching	Coherent	Blank / Plain	No	Yes
Plopski et al., 2016	Projector	No	Action	Matching	Close	Indoor (corridor)	No	Yes
S. Schmidt et al., 2016	OST HMD	No	Action	Matching	Close	Indoor (corridor)	-	Yes
Diaz et al., 2017	OST HMD	No	Action	Midpoint	Consistent	Blank / Plain	Yes	Yes
Slugh et al., 2018	OST HMD	No	Personal	Matching	Far/Collimated	Blank / Plain	Yes	No
	OST HMD	No	Personal	Matching	Close	Blank / Plain	Yes	No
Swan et al., 2017	Handheld	No	Action	Midpoint	-	Outdoor	No	Yes
	Projector	Yes	Action	Matching	-	Indoor (corridor)	No	Yes
S. Schmidt et al., 2017	Other	Yes	Action	Matching	-	Blank / Plain	No	Yes
Lislie et al., 2018	OST HMD	No	Action	Affordance	Other	Indoor (room)	No	Yes
Pointon et al., 2018	OST HMD	No	Personal	Affordance	Close	Indoor (room)	Yes	Yes
Rosales et al., 2019	OST HMD	No	Action	Blind walking	Close	Indoor (room)	No	Yes
Heinrich et al., 2019	Projector	No	Personal	Ordering	Close	Indoor (room)	No	Yes
Wu et al., 2019	OST HMD	No	Personal	Affordance	Close	Wall / Table	No	Yes
J. M. Liu et al., 2020	Handheld	No	Vista	Verbal	-	Indoor (room)	Yes	Yes
Ping et al., 2020	OST HMD	No	Action	Matching	Close	Indoor (corridor)	Yes	Yes
Gagnon et al., 2020	OST HMD	No	Action	Affordance	Close	Indoor (room)	No	Yes

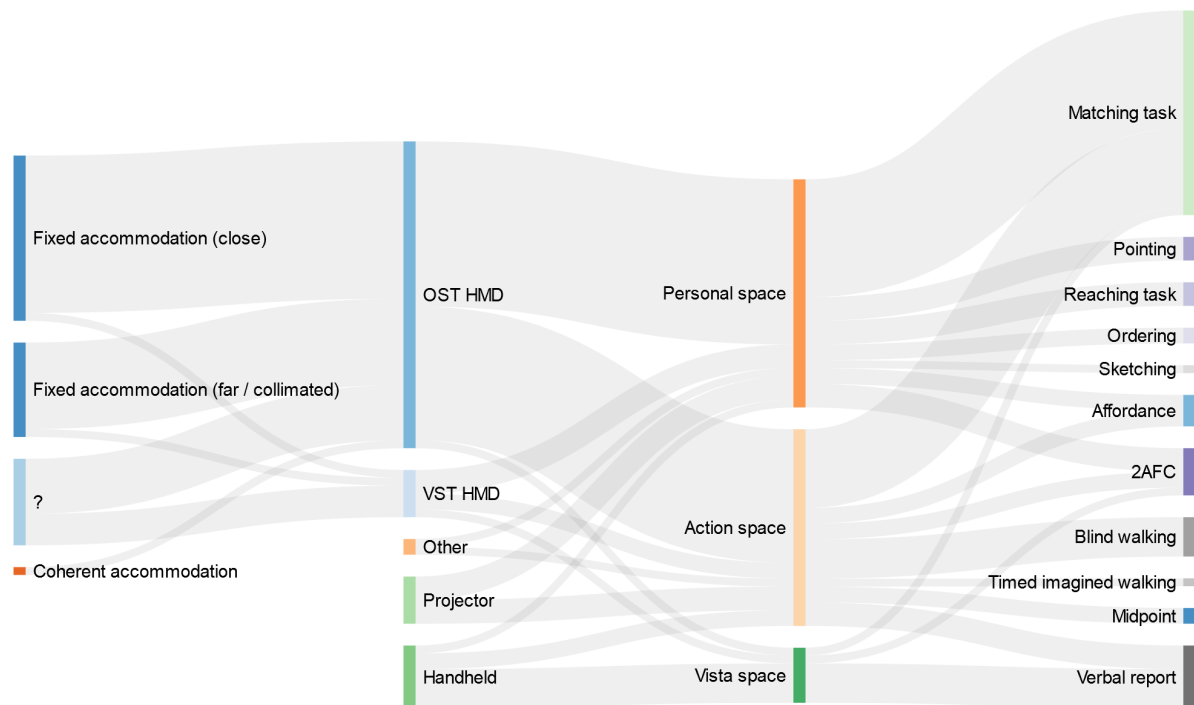


Figure 3.1 – A Sankey diagram of distance perception studies in AR, reporting the accommodation distances and the type of devices used, the distance space evaluated and the reporting technique used.

important oculomotor factor for depth estimation in MR (see Figure 1.9). All the other pictorial cues are usually present in every rendering device, but the specificities of MR devices is the ability to provide these so-called “3D images”, which are usually stereoscopic.

The isolated effect of disparity on depth estimation has been extensively tackled in the past in optical studies. Overall, disparity is known to be a very effective cue until a certain distance, related to the interpupillary distance (IPD) of the observer, and its accuracy decreasing exponentially with the disparity (Gooding et al., 1991; Ogle, 1953; Ritter, 1979). Interestingly, most of these results were obtained while using a projected stereoscopic screen, i.e. with an inaccurate accommodation cue, in a setup similar to a classical projected-based VR display. However, the study from Ogle (1953) was conducted with an optical setup allowing a proper control of the accommodation and found similar results.

Thus, binocular disparity seems to be recognized as the most important and effective cue for HMD-based MR devices, but also for SAR devices (S. Schmidt et al., 2016). The results for autostereoscopic handheld devices are more nuanced since stereoscopic

cues appear to be useful only if monocular depth cues are weak (Berning et al., 2014). However, the specific effect of binocular disparity seems to be mostly overlooked in recent MR perception studies. A few studies tackled this subject, such as Naceri and Chellali (2012) which evaluates the isolated effect of disparity on depth perception, comparing real and virtual targets. They found important individual differences in the virtual condition, compared with the real condition. Livingston, Ai, and Decker (2009) also found that stereo perception with AR HMDs behaves like stereo perception should in everyday vision, while the stereo threshold users were able to achieve with the AR displays was much greater than the thresholds that have been achieved by experiments with normal vision (i.e. the Just Noticeable Difference (JND) in depth was higher for virtual than for real objects). Both studies underlined the importance of other depth cues for judging depth, such as apparent size or accommodation.

### **3.1.2 Monocular studies**

In order to define the specific effect of binocular vision on depth perception, many studies were conducted with different viewing conditions, such as monocular vision. While monocular vision underlines the effect of specific cues, such as pictorial ones that we will tackle later, it also provides a different perspective and as such has a significant effect on distance perception.

Ellis and Menges (1997) explored the impact on depth judgments of different display techniques: monocular, binocular (with the same image on each eye shifted to mimic depth) and ordinary stereoscopic display, and different accommodation distances. Results showed a significant overestimation in the monocular condition and a significant effect of accommodation cue under this condition. In the VR configuration, while binocular vision was considered as a weak cue beyond a few meters, several experiments showed different results. Willemsen et al. (2008) showed similar performance in distance perception within the far distance field with stereo, binocular, and monocular viewing, while Allison et al. (2009) showed a significant underestimation of distances with monocular viewing.

While the monocular viewing might not be only responsible for this phenomenon, it is interesting to note that even a real environment is perceived as compressed when seen with a monocular VST headset (Messing & Durgin, 2005).

Overall, studies suggest that monocular viewing not only enhances the effect of all the monocular cues, but also emphasizes the potential bias caused by one of them, i.e. the height in the field of view (Rosales et al., 2019) or accommodation and background

(Ellis & Menges, 1998). Interestingly, Loomis et al. (2002) showed that whereas perception of location does not depend on whether viewing is monocular or binocular (when other distance cues are abundant), perception of shape becomes more veridical when viewing is binocular. This means that perception of shape and perception of distance are linked yet distinct perceptions.

### 3.1.3 Accommodation

Accommodation is one of the major candidates to explain distance misestimation in MR. Indeed, the vast majority of devices used to provide virtual images does not provide a proper accommodation distance. They mostly provide fixed accommodation distances: the accommodation distance of the optics for lens-based devices (which can be infinite for collimated optics), or the distance of the screen for screen-based ones. Thus, dissociating accommodation and vergence is known to lead to a biased perception toward the accommodation distance (Swenson, 1932). Moreover, wrong accommodation cues for virtual objects can even impact real objects perception (Edgar et al., 1993).

As such, Ellis and Menges (1997) showed the effect of the accommodation cue on distance perception for virtual objects in AR. They also highlighted an interesting correlation between participants' age and accommodation-driven shift in perceived distance: younger participants rely more on the accommodation cue in the monocular condition (Ellis & Menges, 1998).

Swan et al. (2015) reproduced this experiment, using various reporting techniques and providing a comparison with real objects distance estimations. They found that distances for virtual objects were consistently overestimated when compared with real estimates. They hinted that the collimated optics were responsible for this overestimation as they might causes the eyes' vergence angle to rotate outward by a constant angular amount.

They notably modified their display apparatus (Singh et al., 2018) to be able to change the focal distance of the virtual object. When comparing three different accommodation situations (1) with collimated optics, (2) with an accommodation equal to the intended vergence distance, (3) with an accommodation fixed at midpoint of the distance range of the study, results show an increasing overestimation for collimated AR and a slight underestimation when the accommodative depth cue is consistent with the actual distance. The AR consistent case showed also results as good as the midpoint case. This suggests that, with collimated optics, the incorrect accommodation cue will induce a major overestimation of distances but that this effect is not strong enough to bias distance perception

for close but still inaccurate focal distances. Interestingly, the authors also observed an interaction effect between age, brightness and accommodation. It appears that elder people, who may be less able to accommodate at different focal distances, are more accurate in the collimated case, suggesting that they are less relying on the accommodation cue.

Using Spatial Augmented Reality (SAR) devices, S. Schmidt et al. (2017) showed that increasing the distance between two projection screens (i.e. increasing the difference between the focal distances of the targets projected onto) increases the absolute error in depth perception.

Following a completely different path, some authors (Johnson et al., 2016; Konrad et al., 2016) tried to use different accommodation distances (fixed) for each eye in order to improve depth perception in VR. However, this did not succeed to improve perception nor comfort.

Finally, Lisle et al. (2018) reported that volumetric AR displays, i.e. displays able to render an image in a 3D space, outperformed a traditional fixed focal plane AR head-up display (HUD) when virtual objects are placed at more than 5 meters from the focal depth plane of the screen.

### **3.1.4 Motion parallax**

Jones et al. (2008) conducted an experiment comparing users' depth perception in real, AR and VR conditions via a blind walking protocol. They also crossed these conditions with motion parallax conditions: where observers stood still or rocked back and forth. Results showed that subjects performed veridically in the AR condition while a significant underestimation effect was reported in VR. Thus, the motion parallax factor did not affect the observers' depth perception neither in VR nor in AR. The only effect reported for motion parallax was that it made depth judgments less accurate (more underestimated) in the real environment with an HMD.

While the motion parallax cue has not been studied in other works in AR, several studies have been conducted in VR and suggested that this cue is not a major cue for distance perception in virtual environments (Luo et al., 2007; Rousset et al., 2015). It has also been reported to be weak even in real environments (Philbeck & Loomis, 1997). However, Reichelt et al. (2010) showed that motion parallax is one of the cues that contributed the most to visual discomfort in VR when not consistently provided.

Overall, motion parallax is a difficult cue to tackle since its impact is low compared to other depth cues. Its evaluation is even more difficult in AR since many studies are

conducted with fixed point of view and objects. Nevertheless, recent modern HMDs, with better tracking systems, might allow new studies dedicated to this factor.

### 3.1.5 Peripheral vision

While motion parallax is recognized as a weak cue, the overall depth cues given by optic flows, and in particular for peripheral vision, are recognized as important for distance perception. Indeed, peripheral visual information is reported as an important source of information for the calibration of movement within medium-field virtual environments (Jones et al., 2011).

However, this effect depends on the amount of information available in the FoV, and thus is directly linked with the size of this FoV (see section 3.3.8 for a complete review of the effect of the FoV).

### 3.1.6 Brightness

In real environments, brighter objects are usually perceived as closer (Ashley, 1898; Coules, 1955; Farne, 1977). This result was also found in AR experiments since Livingston, Ai, and Decker (2009) showed that dim targets are perceived farther. Singh et al. (2018) experiments also found the most accurate matches for diminished targets, which brightness is more alike those of the real targets. The authors underlined the brightness possible impact in previous studies since AR targets are usually brighter than the real ones and a brighter target may appear closer. Ping et al. (2020) also reported an effect of the color of the target on distance perception, which they hinted would influence the perceived brightness of this target. This effect was also reported by Do et al. (2020). Hence, this result is especially impactful in AR since brightness is a difficult rendering factor to reproduce, in particular in OST HMDs (Billinghurst et al., 2015).

### 3.1.7 Shadows

Sugano et al. (2003) proposed the first study using a VST HMD in personal space in order to evaluate the effect of drop shadows on distance perception in AR. They measured users' performance with an ordering task of colored spheres (from the closest to the farthest, and the highest to the lowest). The authors showed that these shadows improved distance perception in depth and also in height when presented in the monocular condition.

However, different types of shadows have to be considered. The more common ones are “drop shadows”, as in the studies previously mentioned. These shadows are non realistic ones, which are projected on the surface (usually the floor) vertically below the object. These shadows are acting as anchors and provide additional information for depth perception. On the contrary, realistic shadows, are computed according to the positions of the light sources and other environmental parameters in order to provide a more realistic rendering.

Diaz et al. (2017) provided an extensive analysis of different rendering and shadow effects in AR and their impact on depth perception. They demonstrated that drop shadows greatly reduce the overall underestimation of virtual objects in AR. The effect of realistic shadows was also significant, but to a lesser extent. Regarding the effect of drop shadows on grasping performance, Al-Kalbani et al. (2019) showed that drop shadows improve grasping performance and also significantly improve user’s depth estimation of AR objects’ position.

Moreover, the shading model can also impact the depth perception of AR objects. In particular, shading models with specular highlights could improve depth perception in AR (Díaz et al., 2017; Ping et al., 2020).

Overall, the topic of lighting and shadows remain vivid in AR perception, especially with OST HMDs since the rendering of realistic shadows with additive-only displays is still a major challenge (Sei et al., 2020).

### **3.1.8 Height of the object**

The vast majority of studies reporting distance underestimation in MR conducted their experiment with objects placed on the floor, or in the lower part of the field of view of the observers. However, height in the field of view is a major cue, especially for objects placed at greater distances. Moreover, grounding is a very important cue for visual perception (Bian et al., 2005). If targets are not grounded but perceived as “floating” mid air, they will be more likely to be perceived farther. However, the rendering of virtual elements in AR might disrupt this grounding relation and then affect distance perception.

Thus, some studies were conducted in AR to evaluate the effect of the height in the field of view for virtual targets seen in real environments in far-field. Long et al. (2015) showed an overall overestimation of virtual elements’ depth but no significant impact of height or horizontal position of the augmentation. In their study, Swan et al. (2006) showed that distances for targets in the lower field are more likely to be underestimated,

while this underestimation decreases with practice and finally reaches the upper field distance estimation after 7-8 repetitions.

Rosales et al. (2019) confirmed this result in the action space, showing that distances are judged to be farther for targets off the ground versus on the ground. They also reported that participants perceived off-ground objects farther in a monocular condition, compared to a stereo viewing condition. However, the experiment showed just the opposite for on-ground objects. While results for off-ground objects was in line with previous studies on monocular vs binocular perception (see subsection 3.1.2), results for the on-ground condition are more surprising and underline the potential unexpected interactions which can occur in mixed environments.

Moreover, Pointon et al. (2018) compared distance estimation using affordance-based techniques. They showed that while on-floor gap crossing results were underestimated, the affordances for passing through an aperture were as precise as in the real world. They suggested that this difference relies on the height of the objects (on the floor vs extending off the ground).

## **3.2 Intrinsic factors**

In this section, we will discuss what I will call “intrinsic” factors. This includes all the elements related to the person, his physical characteristics, his experiences and knowledge, as well as his representation (avatar) and his emotions.

### **3.2.1 Gender**

The specific case of gender has been scarcely tackled for distance perception in AR. Jerome and Witmer (2005) proposed an experiment involving depth matching and verbal report of AR objects. They found a significant underestimation of AR targets compared with real targets. This bias is due to an interaction of the participants’ gender on AR depth perception, showing that women are more likely to underestimate distances in AR than men. However, this study is the only one highlighting the gender-related differences in distance perception, and was not designed to do so. Moreover, it has been proven that the perception of distances varies according to the gender of the observer (Linn & Petersen, 1986). This question is therefore of major interest in MR since recent studies underlined the overall underrepresentation of female participants in VR research, leading



to potential biases (Peck et al., 2020).

### 3.2.2 Age

Overall, with aging, some perceptive sensors become less efficient. For vision for example, elder people tend to suffer from presbyopia, which reduces the efficiency of the accommodation depth cue (Glasser, Campbell, et al., 1998). As such, in AR, Ellis and Menges (1998) showed that elder participants are less relying in the accommodation cue to assess distances in monoscopic vision. Singh et al. (2018) reproduced these results, showing that, with collimated optics, older observers showed less overestimation than younger observers. However, distance estimation performance of elder people are reported to be at least as precise and sometimes more precise than of younger people (Bian & Andersen, 2013). As such, the weight of the different depth cues can be affected by the age of the observer, in order to preserve the quality of distance perception. Aging will also affect the perception of other senses related to space perception, such as body orientation (Nestmann et al., 2020). This is especially important in AR since we already underlined all the potential inconsistency between depth cues and more studies are needed to elicit all the potential factors influenced by aging on distance perception in MR.

### 3.2.3 Avatar

In VR, avatars are often used as a reference for evaluation tasks. As such, seeing an avatar is known for improving depth perception (Mohler et al., 2008), even at third-person view (Mohler et al., 2010).

However, when it comes to AR, the notion of avatar might be difficult to define. Indeed, the user does not normally see an avatar in AR but directly their own body. There is therefore, to the best of my knowledge, no study on this question in AR. However, two situations deserve to be studied in my opinion from this point of view. (1) When the body is used as a support for an augmentation. In this case, the body and potentially its perception can be modified, which could impact the perception of other objects. (2) Similarly, in the same way that avatars can be modified in VR, they could be added in AR by taking advantage of display devices with occlusion. Direct body modification could then have an impact on perception. One could even extend this idea to other concepts more distant from perception, such as presence, but this would take us far beyond the scope of this work.

### 3.2.4 Learning and Feedback

*Feedback* is characterized as sensory information that indicates something about the actual state of a desired action. It can be either intrinsic or extrinsic. *Intrinsic feedback* is the sensory information which directly comes from the interaction with the world (e.g. vestibular input, optic flow, tactile feedback), while *extrinsic feedback* relies on external sources (e.g. visual or verbal correction, other objects moving) (R. A. Schmidt & Wrisberg, 2008).

As knowledge takes a major place in the perceptual process (see subsection 1.1.1), so do learning, repetitions and feedback in distance perception. However, Swan et al. (2006) is the only study reporting a learning effect in distance perception in AR. They showed that virtual objects' distances are underestimated when placed at ground level, but that underestimation decreases and achieved veridical distances after 7-8 repetitions.

However, this increase of distance estimation performance as been tackled in VR and two main results arises: (1) Feedback is able to improve distance perception in VR to match real perception, and can even affect distance estimation in reality afterwards (Mohler et al., 2006; Richardson & Waller, 2005). (2) The underestimation effect might disappear when observers are placed in a VR room that is an accurate model of the real room they are located in (Interrante et al., 2006).

Overall, these results from VR cannot be directly applied in AR since the targets and the environment can be of different natures (real/virtual). However, these results underline the numerous biases and effects that can be induced by learning, and the necessity to investigate this specific topic in AR. Actually, S. Schmidt et al. (2017) showed that the experience of the user in AR can affect his precision in depth estimation since the authors reported better performance for participants familiar to AR compared to less experimented participants.

### 3.2.5 Emotions

A recent study from Wu et al. (2019) tackled the issue of the impact of emotions on distance estimation. Studies in real (Jiang & Mark, 1994) and virtual (Geuss et al., 2016) have already shown the effect of emotions, and fear in particular, on space perception (please refer to the survey in Wu et al.'s paper for more details). Thus, in AR, the authors showed that participants generally underestimated their abilities to cross gaps and this underestimation increased with deeper gaps. These results can open many exciting

evaluations on the interaction between emotions and visual perception in AR, as well as potential effects of immersion and even presence.

### **3.3 Device and rendering parameters**

Rendering devices and techniques are two major sources of differences between real and virtual and therefore can cause perceptual biases. In this section, we will study the biases caused by the characteristics and shortcomings of devices and rendering techniques.

#### **3.3.1 Latency**

The JND for latency in VR environments is usually admitted to be below 17 ms (Adelstein et al., 2003). However, the situation for AR applications may be more complex. Indeed, observers are able to see both real and virtual stimuli and compare their latency. Hence, with AR displays such as OST HMDs, the real elements does not suffer from latency while virtual ones does, and this difference might impact perception.

Several studies are testing AR setups by simulating them in VR. However, results from a study by Lee et al. (2010) underlined that, even if simulating AR in VR seems to be an efficient way to evaluate the impact of latency, the inherent latency of VR system has to be taken into account. Thus, simulating AR with a VR experiment, Nabiyouni et al. (2017) showed that unequal levels of latency between real and virtual lead to a reduction in performance, even when overall latency levels are lower compared to the matched case.

Considering more specifically distance perception, McCandless et al. (2000) studied the influence of time-delayed viewing in monocular display, in conjunction with motion parallax. Users had to match a real cursor with a virtual target which stability was impaired by a time delay between the observers' head motions and the corresponding change of the object's position on the display. In this situation, results showed that depth judgment errors increased systematically with increasing distance and latency.

#### **3.3.2 Graphics quality and realistic rendering**

Does the quality of the virtual elements matter for distance evaluation? In VR, this question has been the subject of numerous studies but the final outcome remains unclear. While some studies showed that distance underestimation in VR might not be caused by the quality of rendered graphics (Thompson et al., 2004), other studies found that

quality of the rendered graphics alters the distance perception (Phillips et al., 2009). Thus, the underestimation effect might be compensated by modifying the way the graphics are rendered (Kuhl et al., 2006) and presenting a richer and more realistic virtual environment might improve distance estimation (Renner et al., 2013). Significant differences were also found when estimating distances in non-photorealistic versus high-fidelity VEs (Naceri et al., 2011; Phillips et al., 2009).

Overall, while the quality of the rendering in VR cannot be considered as the main factor to explain distance underestimation in VE, it is still possible to mitigate this misestimation using more realistic graphics. Unfortunately, this topic has not been tackled yet for AR but given previous results in VR, integration and differences in rendering between real and virtual elements may once again play a great role in this evaluation.

### **3.3.3 Blurring**

Blurring has been studied in MR in two different aspects: as a factor to explain distance misperception in MR and as a mean to mitigate depth perception errors. Indeed, while studying occlusion in AR, Fuhrmann et al. (1999) successfully used blurring to help compensate for depth perception errors.

However, blur was also suspected to be a factor of misestimation of distances in MR. Thus, two kinds of blur have to be distinguished: (1) motion blur, which occurs when a target is moving, and (2) Depth of Field (DoF) blur, which is created when an element is out of focus.

Since blur is known to convey information about the direction and speed of moving objects (Burr & Ross, 2002; Francis & Kim, 2001; Ross, 2004), similar to motion lines (H. Kim & Francis, 1998), Langbehn et al. (2016) tried to evaluate the effect of blur on distance or speed perception in action space. They found that visual blur has no impact on distance or speed perception in VR. However, the study from Moehring et al. (2009) showed that DoF blur can improve spatial perception in VR in personal space. The authors also found that, while users performed better, they did not like this blurring effect. This difference between perceived and measured effect is especially interesting since Carnegie and Rhee (2015) reported that visual blur was able to reduce visual discomfort and fatigue and mitigate the VAC.

Cidota et al. (2017) compared the effect of visual blur on distance perception in VR and AR situations. They showed that visual blur reduced perceived performance and the level of presence or engagement of the users. However, the results on measured performance are

less conclusive since visual effects seem to reduce performance in AR but improve them in VR.

Taken together, those results suggest that even if visual blur is not a primary cue for distance perception in MR, the reduction it can provide to visual discomfort, in particular for close distances evaluation, may be beneficial for distance estimation.

### **3.3.4 Screen Size**

Obviously, this effect only applies for screen-based devices, i.e. handheld devices and projected immersive systems. However, the latter are now mostly big enough to fill the entire FoV of the observers and handheld devices used to be the major devices used for AR so this section will focus only on these devices. Moreover, parallels could be drawn between the effect of screen size and the effect of the FoV for HMDs but unfortunately, no study tackled this particular topic yet. A specific review of FoV effect for HMDs can be found in section 3.3.8.

Overall, Dey et al. (2012) is the only study evaluating the specific effect of screen size of handheld devices on distance perception. They found that depth compression is reduced when using a mobile phone, while participants subjectively preferred a tablet. Moreover, the tablet provided significantly better ordinal depth perception and faster response time than the mobile phone.

Unfortunately, studies from VR are not likely to provide insight on this topic since screen based VR systems are way larger than handheld AR devices. The main result about screen size in VR is that a bigger screen allows it to be placed farther, which is beneficial for distance perception (Bruder et al., 2016). We can also mention the positive effect of screen size on the sense of presence in VR (Clemente et al., 2014; Ijsselsteijn et al., 2001; Lombard et al., 2000) but however exciting alleging a correlation between presence and distance perception, as well as the extension of presence from VR to AR might be, it would prove far beyond the scope of this work. Thus, more studies are needed to elicit the specific effect of screen size on distance perception in AR. In particular, the size of the screen may impact the availability of motion cues and the evaluation of distance for moving objects or point of view (PoV) could be especially interesting under those viewing conditions.

### 3.3.5 Resolution

In VR HMDs, resolution has been found to have a minor effect on depth perception and anti-aliasing was reported as being more important than stereoscopy in this configuration (Jaeae-Aro & Kjelldahl, 1997). Thus, no significant effect of resolution on depth perception was found in AR for usual handheld devices (phones and tablets) (Dey et al., 2012). However, for a wider range of resolutions tested, really low resolutions were found to reduce distance evaluation performance (Ryu et al., 2005).

### 3.3.6 Handheld specific factors

**2D only rendering (pictorial effect)** Since the image seen on the screen of usual handheld displays is flat, the main cues available are pictorial cues. Thus, pictures usually seem compressed, even if the light-field matches the current PoV of the observer (see below) (Cutting, 2003; Rogers, 1995). As such, a similar result is expected in AR and has been reported by Dey et al. (2012). Using a bisection task, Swan et al. (2017) reported different distance estimation, from overestimation to underestimation when evaluating distance perception at 15 and 30m. J. M. Liu et al. (2020) reported a similar underestimation of distances for real and virtual objects at 40m, while the virtual targets were perceived to be farther than the real targets at the closer distances.

**Secondary PoV** Another specificity of handheld devices is that the point of view of the scene, captured by the device's camera, is different from the point of view of the observer. As such, since the center of projection of the image differs from the observer's PoV, the image is deformed (Sedgwick, 1991). Geometric Field of View (GFOV) is a property of any graphical application, and is defined as the subtended angle between the left, right, top, and bottom ends of the view frustum originating from a virtual camera used in that application, while Display Field of View (DFOV) refers to the physical subtended angle from the eyes to the edges of the screen (Dey & Sandor, 2014).

Hence, with a constant GFOV, a smaller DFOV reduces the size of the environment, while a larger DFOV increases it (Steinicke et al., 2011). While these distortions are usually not perceived (Rogers, 1995; Vishwanath et al., 2005) they can induce perceptual biases depending of the setting and task (Rogers, 1995). To date, no specific study has been conducted to elicit the effect of the eye height in AR with handheld displays but Dey and Sandor (2014) underlined that depth perception could be improved if handheld

AR systems were able to dynamically adapt their geometric field of view to match the display field of view.

**Autostereoscopic devices** Berning et al. (2014) reported an experiment in AR with an autostereoscopic handheld device. They evaluated the participants' performance for a matching task in personal space between a real cuboid and some virtual objects in monoscopic and stereoscopic conditions. Results showed that depth judgment with this device is mostly influenced by monoscopic cues but the stereoscopic cue of this system can improve depth accuracy for challenging situations. Similar results were found previously by Mikkola et al. (2010) with an autostereoscopic display in a non-AR situation.

### 3.3.7 Video See-Through displays

Since the video stream of the environment is captured by an external camera, the image displayed in VST HMDs is not what should be seen if observed directly with the user's eyes. This effect is similar to the displacement of PoV described previously but occurs with immersive displays. Thus, Biocca and Lansing (1998) and Rolland, Biocca, et al. (1995) explored the potential bias in depth judgment created by this phenomenon for real elements seen through VST OST in personal space. They showed that VST HMDs altered the perceptual experience of users, leading to significant effects on visuo-motor performance. Thus, users tend to adapt to this sensory rearrangement in about 15 minutes. Post-experience negative effects were also observed when conducting the same task without the HMD.

Another characteristic of VST devices is the blending between virtual and real elements. Indeed, VST devices allow for a coherent occlusion between real and virtual objects. Hence, in order to compare depth judgments between OST and VST HMDs, Plopski et al. (2016) mimicked those conditions with projectors. In the VST condition, both real and virtual object were projected with the same projector, with a coherent occlusion. In the OST condition, two projectors were used simultaneously, one projecting the real object and the other the virtual one. The authors measured the performance of participants to perform an alignment task between those two objects, under each display condition. They found that users are less perceptive to rotational errors overall, and translational accuracy is worse in OST condition than in VST.

Finally, Medeiros et al. (2016) evaluated the difference in pointing performance between OST and VST HMDs. They found that participants were better in depth perception

with VST than with OST HMDs. They also performed the task quicker and reported less effort. However, the differences between the HMDs used in this study limit the generalization of these results.

Overall, those results underline the specificities of VST HMDs compared with OST HMDs where, while VST HMDs are able to provide proper occlusion, they suffer from the modification of the real environment seen through a camera. Nevertheless, it has to be mentioned that a factor which was extensively tackled with OST HMDs was left over for VST HMDs: the accommodation cue. Indeed, since VST HMDs mostly use headsets with fixed optics, the accommodation cue must be inaccurate and an evaluation of the effect of this inaccuracy would be interesting to perform.

### 3.3.8 HMD specific issues

**Weight** When dealing with underestimation of distances in VR, some authors suggested that the weight of the HMD itself might be a factor of this underestimation (Willemsen et al., 2004). While results of this study did not conclude in a significant effect of the weight of the HMD on distance estimation, this effect can also be studied in conjunction with the reduction of the FoV and the modification of the navigation process. Indeed, with a smaller FoV, observers have to move their head more in order to investigate the environment. An interaction effect might then be considered for tasks requiring more head movements.

For AR, the fact that the observer is always wearing a headset, while seeing both real and virtual elements, can affect both real and virtual perceptions. The effects of HMDs' specificities on distance perception in AR has not been specifically tackled yet and many studies remain to be conducted to elicit all the potential effects on real and virtual perception.

**Field of View** The field of view of commercial HMDs is still significantly smaller than what can be seen by the eyes. As an example, the horizontal FoV provided by the Oculus Rift S<sup>2</sup> is 90° and the FoV of the HTC Vive<sup>3</sup> is 110°. Even ultra wide HMDs such as the StarVR One<sup>4</sup> only reach 210° of horizontal FoV, which is still less than the 270° usually considered as the maximum visual field for the human eye (with eye rotation) (Denion

---

2. <https://www.oculus.com/rift-s/>

3. <https://www.vive.com/>

4. <https://www.starvr.com/>



et al., 2014).

This reduction of FoV might result in a distortion of the optic flow available to perceive the environment. Nilsson et al. (2014) showed that the walking speed is underestimated with smaller FoVs. This has been suspected to be responsible for the underestimation of distances in VR (Wul et al., 2004) but other studies just showed the opposite (Creem-Regehr et al., 2005; Knapp & Loomis, 2004). Willemsen et al. (2009) assessed that even if the combination of weight and reduction of the FoV in VR HMDs can be partially responsible for the distance misestimation in virtual environments, its effect is not sufficient to explain the overall underestimation bias in VR.

In AR, the effect of FoV may be even more complex to tackle since real and virtual have different FoV. For example, in the Microsoft HoloLens<sup>5</sup>, the augmented horizontal FoV is about 30°, but the visual field is also reduced in its upper part due to the frame of the HMD and disrupted on the side due to the borders of the screens. Overall, all the factors of occlusion and FoV reduction have to be considered to evaluate the effect of FoV in AR HMDs. As of now, the only study tackling this subject is that from Trepkowski et al. (2019), which showed that larger FoV can be beneficial for search performance in AR. Thus, much work remain to be done to evaluate the specific effects of HMDs' FoV on distance perception.

## 3.4 Environment and object related factors

Finally, the elements seen by the observer, whether they are objects or environment, virtual or real, can also influence perception. Their arrangement as well as the choices of symbolization can have an impact on the perception. This last category includes the study of all these elements.

### 3.4.1 Environments

In real environments, the type of surrounding context can influence people's judgment of distances (Lappin et al., 2006). As such, the environment might also be a strong cue to determine distances in MR, in particular in the case of outdoor AR. To do so, Livingston, Ai, Swan, et al. (2009) designed an experiment comparing egocentric distance

---

5. <https://www.microsoft.com/hololens>

perception in an indoor and an outdoor environment with an OST HMD.<sup>6</sup> They found a significant underestimation in outdoor environments and an overestimation bias for indoor environments. These results are coherent with the study of Lappin et al. (2006).

However, using a handheld device, Swan et al. (2017) presented an experiment comparing egocentric depth perception in AR in different environments (an open field, a corridor and a ice-covered lake) but did not find any significant effect of the environment on distance perception. J. M. Liu et al. (2020) neither found significant differences between estimates in indoor and outdoor environments with handheld devices. Hence, while the environment seems to play a role in distance estimation in AR, this factor might also be intertwined with other factors related to AR devices.

### 3.4.2 Occluder and x-ray vision

As detailed in subsection 2.3.2, AR allows to render images behind real elements. As such, two factors can be established as potential sources of distance estimation biases: (1) the presence, position and shape of the occluding element, (2) the rendering chosen for the usually occluded elements.

Ellis and Bucher (1994) showed that positioning an opaque physical object either at the perceived depth of the virtual image or at a position substantially in front causes the virtual image to apparently move closer to the observer. Moreover, this effect is strengthened when the opaque physical occluding object is moving, regardless of the rendering quality of the virtual object. In a subsequent study, Ellis and Menges (1998) showed that the occluder still “pushes” the perceived virtual object toward the observer but also reported a significant correlation between this displacement and an over-convergence tendency. This effect occurs independently of the occluder shape but is significantly reduced when the occluder is slotted. With a VST setup, Furmanski et al. (2002) experimented perceptual heuristics for helping participants to perceive relative positions of augmentations on a wall. They found that subjects reported virtual objects as always being in front of real objects, except when additional occlusion and motion cues were presented together. In SAR, S. Schmidt et al. (2017) also reported that the absolute error in estimated depth increases with the distance between the target and the projected surface. Taken together, those results suggest that the usual x-ray rendering induces observers to perceive objects

---

6. In fact, the authors reported the opposite in their study. However, since they were reporting the distance error of a virtual target, a positive error implies an object placed farther, i.e. perceived closer than it actually was, which induces an underestimation of distances for a virtual target.

always in front of the occluder but a sufficient difference in depth or providing more depth cues allows them to compensate for this bias.

It must be noted that uniform transparency destroys spatial relationships (Buchmann et al., 2005). As such, several techniques have been developed to render occluded content in AR, such as x-ray tunnel (Bane & Höllerer, 2004), magic lenses (Looser et al., 2007; Viega et al., 1996), "virtual hole" (State et al., 1994), alpha-blending overlay (possibly user-driven x-ray visualization window (Kalkofen et al., 2007; Mendez et al., 2006)), edge-based (Avery et al., 2009; Kalkofen et al., 2007, 2009), saliency-based (Sandor et al., 2010; Santos et al., 2016) and even mixed approach (Kalkofen et al., 2013; Zollmann et al., 2010). Thus, edge-based techniques are less legible when there are too many edges on the occluding real object, while saliency-based techniques are less legible when there are high contrasts in color or bright lighting but are more efficient to identify large items (Santos et al., 2016).

However, two issues are overlapping here: the efficient display of many superimposed elements and the precise estimation of depth for occluded objects, the latter remaining a major challenge (Mendez & Schmalstieg, 2009). Thus, Dey and Sandor (2014) reported that while visual clutter deteriorates depth judgment and reduces the visibility of occluded objects, depth judgment can still be improved by providing clear visibility of the occluded objects. Hence, Livingston et al. (2003) reported that, with appropriate rendering, such as drawing style and opacity settings, users are able to accurately interpret three layers of occluded objects, even in the absence of perspective cues.

### 3.4.3 Auxiliary augmentations

In order to improve depth perception in AR for outdoor use, Wither and Hollerer (2005) proposed several auxiliary augmentations (AAs) and representations: semitransparent planes with dropped shadows of the virtual targets, top-down view and color changes according to depth. They evaluated relative and absolute distance perception but reported non significant results for the absolute case, since the participants answers were not precise enough. Therefore, having other markers visible improved the results.

Kytö et al. (2013) explored the use of other AAs to improve depth perception in AR. With monocular or stereoscopic viewing, they asked subjects to evaluate the ordinal position of an augmented red circle on a wall, while seeing an other similar fixed target as an AA. Results showed a significant performance improvement with AAs, especially in the stereoscopic condition. In order to provide a similar experiment for farther distances, Kytö

et al. (2014) conducted an other experiment using AAs and found that depth judgments were the closest to veridical depth when binocular disparity and AAs were combined.

In fact, drop shadows (as detailed in subsection 3.1.7) can also be considered as AAs. Thus, AAs seem to be a great way to improve distance perception in AR. Heinrich et al. (2019) also reported that drop shadows (here implemented as a drop shadow with a supporting line between the object and its shadow) can improve depth perception. In the same study, they reported that providing a pseudo-chromadepth rendering, i.e. rendering the target with a color encoding its depth from blue to red, also improved the depth estimation. However, this study underlines the intrinsic bias of distance perception studies relying on AAs: the distance perception might be improved not thanks to the AA as a positive tool for depth perception but rather because they allow the observer to perform a different task. Here, the participants may rely more on the color of the target or the relative distances between the shadows, rather than estimating the depth of the target themselves.

As such, much work remains to be done to characterize the aspect of AAs that better affect distance estimation, without providing such significant changes in rendering that the depth can be inferred from a different perception mean.

#### 3.4.4 Object appearance and position

Among many aspects of object representation, orientation has been studied as a factor for distance estimation in AR. In particular, Diaz et al. (2017) evaluated the effect of “billboarding”, i.e. the fact that a 2D object is always facing the observer, on distance perception. This consistency is supposed to enforce pictorial cues and in particular the relative size cue. Hence, results showed an improvement of depth perception with this billboarding effect.

Do et al. (2020) showed an effect of color and brightness on depth perception in AR. Bright colors were perceived closer than darker ones, with bright red perceived closer than any other color. However, the authors also underlined that this effect was present for a spherical target but not for cubes or with more complex 3D objects. They suggested that these objects provide more depth cues and then are less subject to distance perception biases.

Overall, numerous studies with real and virtual objects showed that size (Smith, 1953), shape (Loomis, 2014), color (Chatterjea et al., 1974), orientation and other characteristics of objects have a significant impact on distance perception. Using a SAR configuration to

augment real targets, S. Schmidt et al. (2016) showed that projecting illusions, such as color or blur effect can modify depth perception of real objects. Thus, characterizing the interaction of these factors with mixed rendering is a major challenge for MR perception.

## **3.5 Besides depth: exocentric distance perception**

So far, we have only considered the egocentric perception of distances, also called depth perception. However, space perception includes any evaluation of distances, including that between two objects. Moreover, since all the different types of perception are related, egocentric distance perception can influence other types of perception. As such, this section will focus on detailing the previous work on exocentric perception of distance.

### **3.5.1 Exocentric distance perception in real life**

One of the first studies focusing on a comparison between egocentric and exocentric distance perception in real environments is that of Loomis et al. (1992). Authors used a matching task where users had to adjust two metallic rods placed on the observer's frontal plane so that the distance between them matched that of two target metallic rods that were placed on the depth plane (i.e. the sagittal plane, being perpendicular to the frontal one). The distance between the rods lied within the 4-12m range. The question motivating this work was whether natural environments viewed binocularly from a stationary vantage point were perceived correctly. They found that there was an underestimation of egocentric distances compared to exocentric distances. Kudoh (2005) reproduced the work from Loomis et al. (1992) using blind walking as a reporting technique and confirmed the overall overestimation of exocentric distances.

An hypothesis presented in Loomis et al. (1992) and confirmed in Foley et al. (2004) and Loomis et al. (2002) is that there is a dissociation between the perception of location and shape. Indeed, while our perception of location is accurate (up to at least 20m under full cues conditions), the perception of 2D shapes viewed under the same conditions exhibits large systematic distortions. Given these results, this hypothesis further assumes that the perception of extent and of shape is based on different neural computations from those involved in the perception of location.

Other studies in real environments (Foley, 1968; Norman et al., 1996) showed that stimuli lying in the depth (i.e. sagittal) plane are indeed underestimated compared to stimuli

placed in the frontal plane. In a later work, Foley et al. (2004) proposed a mathematical model for the relationship between perceived and physical objects (i.e. visual angles) as well as perceived distances. Nevertheless, their model showed that this relationship cannot be represented by any Euclidean or other metric geometry.

Levin and Haber (1993) confirmed that a change in the viewing angle can modify the exocentric depth perception of exactly the same stimuli. Orientation has also been shown to be an issue (Aznar-Casanova et al., 2008; Levin & Haber, 1993).

Finally, Predebon (1991) investigated the effect of familiar and unfamiliar size of objects in the vista space (between 10m and 187.5m). Results showed that the familiar/unfamiliar characteristics of stimuli do not influence depth-matching judgments, judgments of relative depth, nor judgments of the relative lengths and relative egocentric distances of extents presented in the observers' frontoparallel plane. Finally, it was shown that familiar sizes influence judgments of the relative egocentric distances of unfamiliar and familiar objects.

To summarize, the very extensive literature regarding exocentric distance perception in real environments seems to agree on the fact that exocentric distance of stimuli placed in the sagittal plane is underestimated compared to that of similarly distant objects positioned in the frontoparallel plane.

### 3.5.2 Exocentric distance perception in VR

Regarding exocentric distance perception in virtual environments, results are less systematic since they also depend on the type of VR display used in the experiments.

Waller (1999) showed an overestimation of exocentric (i.e. inter objects) distances and that this systematic error based on verbal reporting could be improved when the experimenter gave participants direct feedback after their verbal report. The experimental conditions allowed the participant to move freely and two display conditions were tested: HMD-based and desktop-based VR. Nevertheless, the VE was very simple (a huge gray cube with a regular grid superimposed on the cube's faces) and thus not fully representative of nowadays' VEs.

Conversely, Wartenberg and Wiborg (2003) allowed participants to move but also observed distance overestimation in the VR condition (within a CAVE) while the desktop conditions lead to both over- and underestimations.

Regarding size estimation, there is no complete consensus: while some authors tend to demonstrate a certain accuracy (Geuss et al., 2010; Kenyon et al., 2007), (Stefanucci

et al., 2015) showed that virtual objects are estimated as smaller than in reality. It should be noted that the setup from Stefanucci et al. (2015) used flat TV screens to display 3D objects in a “desktop VR” configuration unlike Geuss et al. (2010) and Kenyon et al. (2007) that were conducted in “immersive VR”.

Geuss et al. (2012), showed that, while wearing an HMD, exocentric distances are compressed when evaluated in the depth plane but are accurate when objects are located in the frontal plane. Based on these findings, Kelly et al. (2015) proposed an alternative hypothesis to explain these results: they were due to a combination of underestimation caused by the VE and overestimation of frontal relative to depth extents. They thus conducted two experiments in which they reproduced Geuss et al. findings in a similar VE but also demonstrated an underestimation in both frontal and depth planes when using a simpler VE composed of a single textured plane.

C. J. Lin et al. (2015) asked participants to sketch a 2D line representing the center-to-center distance between a real object and a virtual object presented using a 3D stereoscopic projector onto a screen. They showed that participants underestimated the distance and that the accuracy worsened along with an increasing distance between both objects. They also showed that aligning both objects in the vertical direction reduced this underestimation.

Finally, still using a projection-based stereoscopic environment, Woldegiorgis and Lin (2017) showed that the accuracy of a pointing task when objects lie in the frontal plane was lower (85%) compared to that of the physical world (94%). They also showed that this accuracy was affected by the egocentric distance between the participant and the frontal plane.

### 3.5.3 Exocentric Distance Perception in AR

Finally, we conclude this section with an overview of the few studies dedicated to exocentric distance perception in AR. Dey et al. (2012) showed that in hand-held AR devices, there was an underestimation (using verbal reporting) of exocentric distances and egocentric distances in the vista space (distances from 30m to 110m were considered in the paper). Nevertheless, exocentric distances were only evaluated between virtual objects as no real objects were involved.

Then, Sugano et al. (2003) showed that the presence of virtual shadows increased participants’ accuracy in a vertical ordering task using a VST HMD setup.

Finally, Pointon et al. (2018) studied space perception by asking participants if they

could pass through a virtual aperture between two poles, presented at different widths and distances. Their results showed no difference between virtual poles in AR and real ones. Using affordance evaluation of apertures in a wall, Gagnon et al. (2020) showed that judgments were more accurate near the aperture compared to a farther viewing point, despite the inability to see the entirety of the aperture within the FoV of the device. They also reported a improvement of verbal report with feedback.

From the presented literature review, and while having received substantial interest from the community both in real and virtual environments, exocentric distance perception has been relatively left aside in augmented environments. In particular, there is no study focusing on exocentric distance perception in OST AR environments. This is an important and relevant matter, especially given the recent development of high quality consumer grade OST AR headsets, such as the Microsoft HoloLens.

### **3.6 Conclusion: Are all perceptual biases in MR explained by a specific factor effect?**

All along this section we extensively presented the potential factors which can appear in MR and contribute to a distance miscalculation. However, one last question remains open: Are virtual objects in real environments always subject to distance misperception? With rendering parameters as close to reality as possible, what is the specific effect of mixed elements on distance perception?

To answer this question, we focused on studies which compared distance estimation with real and AR setups. We will focus on HMD-based AR since the rendering of handled displays is more specific due the different DFOV (see subsection 3.3.6 for more details). Unfortunately, a few studies fall in this category, probably because of the technical investment needed to accurately test distance perception in reality.

The first reported studies which compared real and virtual distance perception in AR, with OST HMDs in the personal distance field showed an overestimation of distances for virtual objects compared to real objects (Rolland, Gibson, et al., 1995; Swan et al., 2015). However, the authors hinted that this overestimation was caused by the collimated optics of the HMD used. Hence, Singh et al. (2018) used a focal-adjustable device and showed that, with a coherent accommodation cue, the distance estimation of the observers for virtual targets was as good as for real targets.

For farther distances (beyond 2 meters), most of the studies using AR OST HMDs



reported an underestimation of distance evaluation for virtual targets compared with real ones (Grechkin et al., 2010; Jerome & Witmer, 2005; Swan et al., 2007) while some other studies did not (Jones et al., 2008). Jones et al. (2011) reproduced this result and showed that, with walking based evaluation techniques, peripheral information is a very strong re-calibration cue for space perception. As such, the authors assessed that, when this cue is available, participants can succeed to recalibrate their distance perception in AR environment, explaining the lack of underestimation in recent AR studies. Indeed, if we consider the mentioned studies in detail, Grechkin et al. (2010) and Jerome and Witmer (2005) used a setup similar to VR, and Grechkin et al. even compared VR and AR and found an underestimation trend in AR similar to those found in VR. Thus, Swan et al. (2007) used a walking based evaluation technique and still reported an underestimation, while they underlined that it was significantly lower than the usual underestimation in VR. However, this peripheral recalibration cannot be the only factor responsible for the better perceptual performance in AR compared to VR. As already detailed (see section 3.3.8), the reduced FoV of HMDs might not even be responsible for the distance underestimation in VR (Creem-Regehr et al., 2005; Knapp & Loomis, 2004). Thus, Jones et al. (2008) also suggested that the better results in AR may be due to a better configuration of the AR HMDs which can calibrate virtual images with real referents.

Overall, it seems difficult to assess a specific effect of AR rendering on depth perception. The general trend shows similarity with VR biases but the ability to see the real environment as well as virtual objects seems to allow observers to obtain a better performance compared to VR. Moreover, if seeing the real environment makes observers more akin to recalibrate their perception, it is still possible that, on the contrary, the differences in rendering between real and virtual enforce some biases. It is thus possible that, as stated by Livingston et al. (2005) “whether the user ‘suspends disbelief’ of the virtual objects; such a belief would not prevent a subconscious cue from interfering with the user’s performance”.





PART II

# Research Contributions

---



Distance perception, like any form of perception, is a mainly internal process. Thus, only internal biases can be responsible for a distortion of this perception. In this chapter, we present a study aimed at highlighting the anisotropy of the perception of distance in Virtual Reality.

# INTRINSIC BIASES: THE ANISOTROPY OF DISTANCE PERCEPTION

---

## Contents

---

<b>4.1</b>	<b>Main experiments</b>	<b>94</b>
4.1.1	Experiment 1: Anisotropy of distance perception	95
4.1.2	Experiment 2: Influence of the viewing angle	103
4.1.3	Provisional Conclusions	107
<b>4.2</b>	<b>Control experiments</b>	<b>108</b>
4.2.1	Experiment 3: Disambiguating head and eye directions effects	108
4.2.2	Experiment 4: Does the quality of the virtual environment matter?	111
<b>4.3</b>	<b>Discussion</b>	<b>114</b>
<b>4.4</b>	<b>Conclusion</b>	<b>116</b>

---

We underlined the importance of knowledge for perception. This can play a great role for distance perception, via learning, re-calibration or other intrinsic factors (see section 3.2 for more details). However, most studies that reported intrinsic biases were primarily focused on other factors and incidentally found evidences of intrinsic effects. Here, we would like to focus exclusively on intrinsic factors. As such, the experiment will be held in VR in order to reduce any potential environmental effect. Moreover, we will remove every other external distance cues. Hence, the main remaining “factor” will be the observer himself.

In order to choose the precise element that will be evaluated in this first study, we would like to highlight two specif points from the previous section:

1. All the studies evaluating egocentric distance perception in MR have been conducted with objects placed in front of the participants. Then, what happens when we look to the sides?

2. Haptic space anisotropy appears to be related to the position of arms and the areas reachable by the hand. As such, if the hand movements influence the haptic perception, can head movements influence visual perception?

Numerous studies use reality as a baseline for perception to provide theories and guidelines about depth perception. This subject has been broadly covered for real environments, although not fully understood yet (Loomis & Knapp, 2003). Most of these studies tackle the phenomenology of depth perception and the perceptual illusions which can arise when combining distance, orientation and speed perception. In particular, several models for the visual space (i.e. the shape of the space perceived with vision) have been proposed. They underline the effects of the observer’s point of view and orientation on egocentric distance perception (Koenderink et al., 2008).

However, in VR studies regarding egocentric distance perception, the potential influence of the user’s head orientation has not been tackled yet. What is more, in such studies the VE is generally oriented and the participant faces a specific direction. The *forward* direction is specified, and objects are always presented in front of the participant. Then, what happens when the objects are placed on the sides? Could it alter perceived distances, as suggested by visual field theories in real environments?

The purpose of the present study is to evaluate egocentric distance perception to targets to the side in VR. During our experiments, participants had to choose between two virtual spheres which one was the farthest. Systematically, and sequentially, one virtual sphere was placed on the side, and another one was placed in front. For each trial, they had to compare the distance between them and each sphere (egocentric distances) and select the farthest one. Thus, we assessed if the perceived distance to a virtual object on the side differs from that of an object in front. Hence, we studied the potential anisotropy of egocentric distance perception in immersive VEs.

## 4.1 Main experiments

In this section, we describe the first two psychophysical experiments conducted in this study.

### 4.1.1 Experiment 1: Anisotropy of distance perception

This experiment aimed at verifying our research hypothesis **H1**: there is a difference of egocentric distance perception between a virtual object placed in front and one placed on the side.

#### 4.1.1.1 Participants

This experiment involved 21 participants (all males). They were aged from 20 to 43 (mean=24.3, SD=4.9). Participants were students or members of the laboratory and naive about the purpose of the experiment. They were not paid for their participation. All participants had normal or corrected-to-normal vision and 16 out of 21 of them had already used an HMD before.

#### 4.1.1.2 Experimental Apparatus

Participants were seated on a fixed chair in front of a table and a fixed joystick was used for answering (cf. Figure 4.1). They were asked to stay straight, their back resting on the back of the chair and both hands grasping the joystick. The goal was to fix the torso orientation and to keep a steady position while still being able to move their head freely. Participants wore headphones that were used to vocally announce in which direction the next stimulus would appear. Participants could answer using the joystick by tilting it in the direction of the chosen sphere and by pressing the trigger button to validate.

The VE was displayed in a HTC Vive Head Mounted Display (HMD) which provides a resolution of  $1200 \times 1080$  per eye with a horizontal FoV of approximately  $110^\circ$ . The HMD is self-tracked and provides participants' head positions and orientations with a refresh rate of 90 Hz (Niehorster et al., 2017). In order to provide a correct perspective and stereoscopic rendering on the HMD, the IPDs of each participant was measured before the experiment with a ruler tape.

The VE (depicted in Figure 4.2) was designed to provide a limited amount of visual information. It consisted in an empty space with a white sky, a brown floor and a blurred horizon. The stimuli used were red spheres of 10 cm radius. The reference sphere was placed at 3 meters from the participant head's center, at eye level. The only virtual light provided was a white ambient light. No reflection or shadows were visible in the VE.





Figure 4.1 – Experimental setup used in all our experiments. Participants were seated on a chair, wore a HTC Vive HMD and headphones. They grabbed a joystick, which was fixated to the table, with both hands.

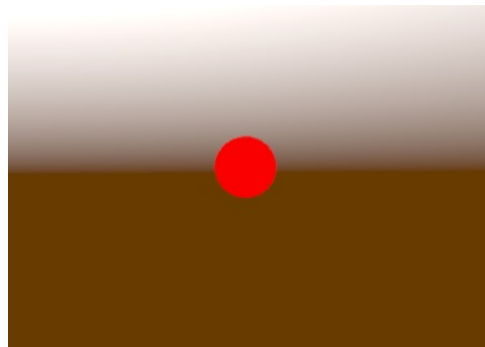


Figure 4.2 – Virtual stimulus used in Experiments 1, 2 and 3 from the participant's PoV.

#### 4.1.1.3 Head Position Calibration

Due to the nature of the experiment (objects displayed at different angles with respect to the user), a potential bias on distance estimation could be introduced by the position of the user. Indeed, the center defined by the HMD is some centimeters ahead from the pivot point of the head, which can lead to a miscalculation of the egocentric distance as depicted in Figure 4.3. Thus, the following calibration procedure was designed to calibrate the center of the experiment.

First, participants were asked to look to the right and to the left several times. The resulting head positions were distributed along circular arcs, and their centers were used to compute the head's rotation center via an optimization method. Finally, a validation phase followed the calibration to ensure that the computed head's center matched the

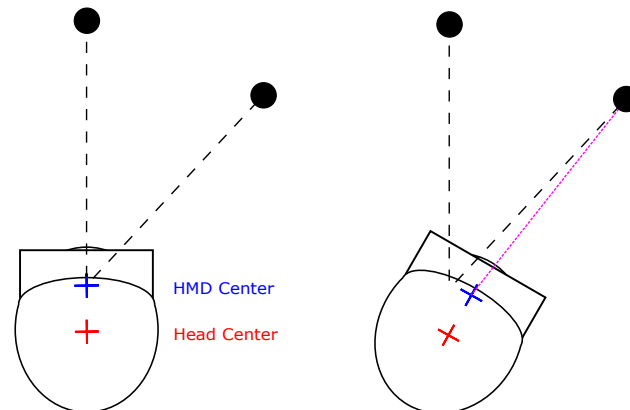


Figure 4.3 – Calibration procedure. Without calibration, the positions of the head and of the HMD centers are distinct. When the spheres are placed at the same distance from the HMD center (dotted lines), the egocentric distance to the sphere placed on the side could be perceived closer because of centers mismatch.

real head center position (a difference of less than 3 cm) during the experiment. After carefully executing this calibration, red spheres covered 46 pixels wide on each of the HMD’s screens at the reference distance and the retinal projection of the spheres was identical in size when they were placed at the same distance.

#### 4.1.1.4 Experimental Procedure

Participants started by filling out a short form containing written instructions about the experiment. After verbal explanations, they carried out the calibration procedure as described in subsection 4.1.1.3. Then, they performed 10 trials during which they could get used to the experimental procedure. Finally, the participants were presented with the set of 120 trials. The procedure for each trial was as follows.

In order to assess egocentric distance perception, a two-alternative forced choice (2AFC) protocol was considered. Participants were presented with two successive stimuli and they had to determine the one which answered the following question: “Which object is the farthest one?”. The stimuli were spheres placed at different positions and angles with respect to the participant’s position (see subsection 4.1.1.6 for more details), in which one of them was always a reference stimulus (a sphere placed always at the same distance). The spheres were displayed using a red flat shading rendering.

At the beginning of each trial a vocal instruction indicated in which direction the participant should look for the current stimulus (front, left or right). Once in the field of

view of the participant, the stimulus stayed visible during 2.5 s before disappearing. After a short delay of 0.5 s the second stimulus was presented following the same procedure. When the second sphere disappeared, the participant provided the answer to the question “Which object is the farthest one?” by tilting the joystick to indicate which sphere seemed farther from him and validated with the trigger.

The sky of the VE was used to provide visual information about different phases of the experiment by changing color when the participant could answer or rest. While it was white when the spheres were visible, it turned yellow to indicate that the participant could answer and changed to green between each repetition. This was achieved to inform the participant that his answer was effectively recorded without giving localized references (as would a text panel).

The average total time per participant, including pre/post-questionnaires, instructions, experiment, breaks, and debriefing, was 30 min. After each break, the calibration detailed in subsection 4.1.1.3 was performed. Participants wore the HMD for approximately 20 min.

#### 4.1.1.5 Pilot experiment

A pilot experiment was conducted in order to validate the protocol and determine several parameters for the stimuli configurations (distance, size, angle). As described in subsection 4.1.1.2, participants were seeing two red spheres at different positions and had to choose which one was the farthest.

This pilot experiment involved 12 naive participants, and results validated the chosen distances for the visual stimuli to obtain proper psychometric curves. The results of this pilot study already showed a shift (or perceptual bias) in the ordinal estimation according to the position of the sphere (in front or on the side). This shift did not seem to differ for the two angles tested (45° and 90°).

#### 4.1.1.6 Experimental Design

We followed a  $2 \times 2 \times 5 \times 2$  full-factorial within-subjects design:

- **C1** The position of the reference sphere, which could be in front or on the side (but always displayed 3m away from the participant’s head center).
- **C2** The side of the “on the side” sphere, which could be on the right or on the left.

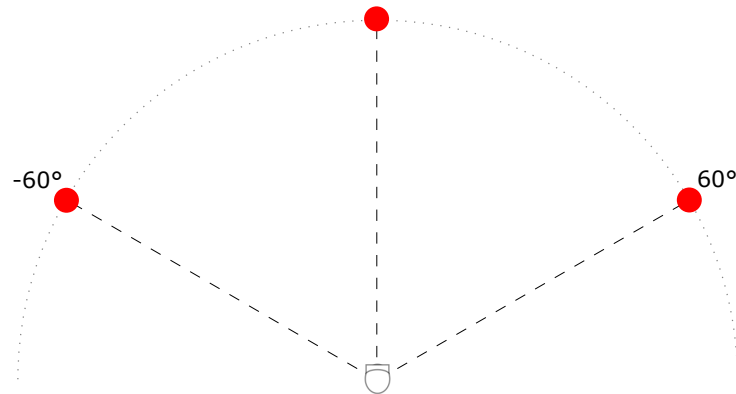


Figure 4.4 – Possible positions of the spheres in Experiments 1, 3 and 4. Only one sphere is visible at a time. They are displayed in pairs, one in front and one on the side, sequentially in a random order.

Table 4.1 – Independent and Dependent Variables of Experiments 1, 3 and 4

Independent Variables			
	<i>observers</i>	21	(random variable)
<b>C1</b>	<i>reference position</i>	2	In front On the side
<b>C2</b>	<i>side</i>	2	Right Left
<b>C3</b>	<i>distance</i>	5	-10%, -5%, 0%, 5%, 10%
<b>C4</b>	<i>first sphere</i>	2	Reference Comparison
	<i>repetition</i>	3	1, 2, 3
Dependent Variables			
	<i>chosen sphere</i>		Reference or comparison sphere chosen

- **C3** The relative distance between the reference and the comparison spheres. Five possible values were chosen after the results of a pilot experiment:  $-0.3$  m,  $-0.15$  m,  $0$  m,  $+0.15$  m,  $+0.3$  m compared to the reference position ( $3$  m).
- **C4** The order of presentation of stimuli (reference first, comparison first), was counterbalanced to minimize order effects (Ziemer et al., 2009). Every couple of spheres was presented in every possible order (side first/front first, reference first/comparison first).

The angle defined by the two stimuli and the head center was fixed to  $60^\circ$ . This experiment was more focused on the identification of a potential bias rather than measuring the effect with respect to the viewing angle.

In summary, participants were presented with 120 trials, divided in 3 blocks of 40 trials in a different randomized order for each block. Each block of 40 trials presented a set of couples of spheres made of: 2 reference positions (**C1**)  $\times$  2 sides (**C2**)  $\times$  5 relative positions (**C3**)  $\times$  2 presentation orders (**C4**).

#### 4.1.1.7 Statistical Analysis

The choices of participants (reference perceived closer) were analyzed using a logistic regression model with a logistic link function. The model included the independent variables as fixed effects and the *participant* factor as a random effect. Tukey post-hoc tests adapted for logistic regression models were employed. For the sake of simplicity, the main effects of factors involved in interactions are reported but they are not discussed.

The statistical analysis was carried out using the R software, with the *glm* function and the *anova.glm* function to analyze the deviance of the logistic regression model, and the *glht* function of the *multcomp* package for pairwise comparisons.

Moreover, when an independent variable showed a strong bias on the users accuracy for distance estimation, we were able to further analyze its effect by fitting the following psychometric curve to the data for each level of the variable ( $\beta$  provides the value in meters of the Point of Subjective Equality (PSE)):

$$f(x) = \frac{1}{1 + e^{-\alpha(x-\beta)}} \quad (4.1)$$

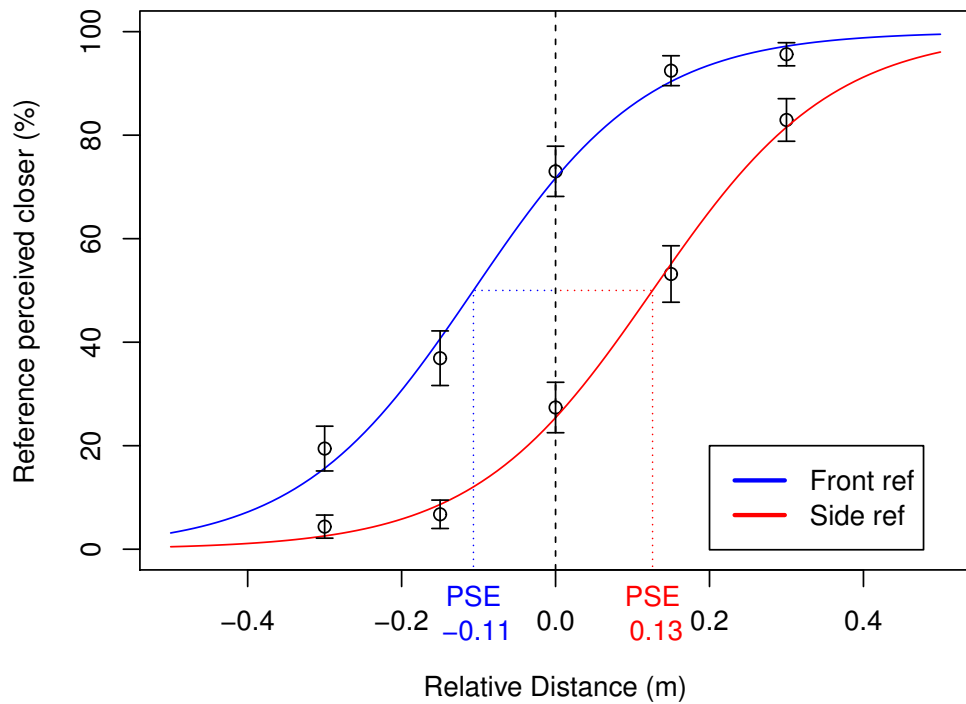


Figure 4.5 – Experiment 1 - Psychometric curves. For each possible relative distance between the reference and the comparison sphere we plot the percentage and the standard error of answers stating that the reference is closer. The red (resp. blue) curve shows the psychometric curve with a “on the side” (resp. “in front”) reference sphere. The corresponding PSE is displayed for each condition.

#### 4.1.1.8 Results

The analysis of deviance of the logistic regression model (reference perceived closer vs **C1**, **C2**, **C3** and **C4**) showed a main effect for **C1** ( $\chi^2 = 318.75; p < 0.001$ ), **C3** ( $\chi^2 = 868.30; p < 0.001$ ) and **C4** ( $\chi^2 = 30.39; p < 0.001$ ). We also observed an interaction effect between **C3** and **C4** ( $\chi^2 = 39.16; p < 0.001$ ). In contrast, there was no main effect for the presentation side **C2** ( $\chi^2 = 0.18; p = 0.66$ ). Regarding **C1**, post-hoc tests showed that when the reference sphere was presented in front of the participant, it was more likely to be considered as closer ( $z = 4.07; p < 0.001$ ), see Figure 4.5.

Post-hoc tests for the interaction effect (**C3** and **C4**) showed that when the comparison stimulus was placed at  $-0.3$  m and  $-0.15$  m and the reference stimulus was presented second, the reference is perceived closer than if it was presented first ( $-0.3$ m:  $z = -4.80; p < 0.001$ ,  $-0.15$ m:  $z = -5.216; p < 0.001$ ). This effect was not visible for the other levels of **C3** (all *n.s.*).

Due to the significance of **C1**, we further analyzed the ability to discriminate the farthest sphere by fitting the psychometric curves (see Figure 4.5). In particular, we split the analysis between the discrimination when the reference stimulus was in front and on the side. When comparing *Side* and *Center* positions, the obtained coefficients were  $\alpha = 8.5428$  ( $SE = 0.5401$ ) and  $\beta = 0.1260$  ( $SE = 0.0077$ ) for the reference on the side and  $\alpha = 8.7196$  ( $SE = 0.8198$ ) and  $\beta = -0.1067$  ( $SE = 0.0114$ ) for the reference at the center.

#### 4.1.1.9 Discussion

The presented results showed first a significant order effect. Under some distance conditions, the last presented object is perceived closer. Thus, the main observed effect is the perceptual bias which leads the observer to perceive an object presented in front of him closer than an object located at  $60^\circ$  on the side which validates our first hypothesis **H1**. Interestingly, the side (left or right) on which the object is presented does not significantly affect the results, it is always perceived as being farther away.

In order to explore this phenomenon, we conducted another experiment to further investigate the potential influence of the angle formed by the “in front” and the “side” virtual objects.

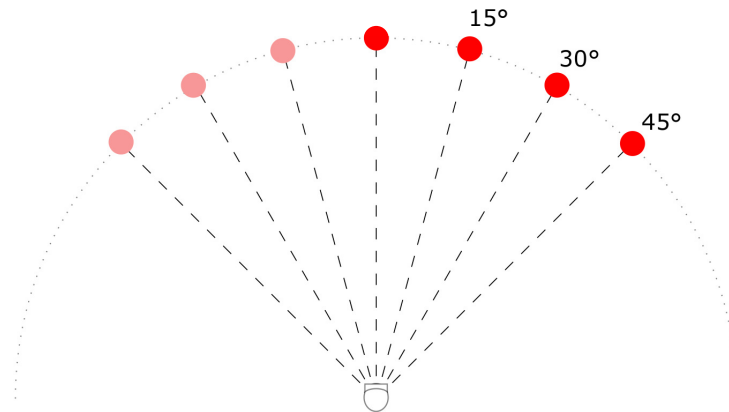


Figure 4.6 – Configuration of the displayed spheres for Experiment 2. Only one sphere is visible at a time. They are displayed by pairs of two, one in front and one on the side, sequentially, in a randomized order. The side (left or right) is not a parameter of the experimental design and is chosen randomly.

## 4.1.2 Experiment 2: Influence of the viewing angle

While the first experiment showed that distance perception in VR was anisotropic, it did not assess how this anisotropy changed with respect to the angle defined by the “in front” and “on the side” stimuli. Thus, this second experiment aimed to test our second hypothesis **H2**: the anisotropy of egocentric distance perception in VR is correlated with this angle. It followed the same apparatus and response protocol as the first experiment (see subsection 4.1.1.4), but the experimental design was modified to assess changes in the estimated distance for a discrete number of angle values (15°, 30° and 45°).

### 4.1.2.1 Participants

This experiment involved 16 participants (2 females, 14 males). They were aged from 18 to 56 (mean=29.4, SD=10.4). The participants were students or members of the laboratory, naive about the purpose of the experiment. They were not paid for their participation. All our participants had normal or corrected-to-normal vision and 13 out of 16 already took part in an experiment or demo involving HMDs before. None of them took part in any of the previous experiments.

### 4.1.2.2 Experimental design

We used a  $2 \times 3 \times 5 \times 2$  full-factorial within-subjects design:



Table 4.2 – Independent and Dependent Variables of Experiment 2.

<b>Independent Variables</b>			
	<i>observers</i>	21	(random variable)
<b>C1</b>	<i>reference position</i>	2	In front On the side
<b>C2</b>	<i>angle</i>	3	15°, 30°, 45°
<b>C3</b>	<i>distance</i>	5	-10%, -5%, 0%, 5%, 10%
<b>C4</b>	<i>first sphere</i>	2	Reference Comparison
	<i>repetition</i>	3	1, 2, 3
<b>Dependent Variables</b>			
	<i>chosen sphere</i>	Reference or comparison sphere chosen	

- **C1** is the position of the reference sphere, which can be in front or on the side (but always displayed 3m away from the participant’s head center).
- **C2** is the angle between the reference and comparison spheres. The three possible values chosen were 15°, 30° and 45°.

We decided to focus on smaller angles in order to determine the angle from which the effect appears. These values can also be completed by the 60° case studied in the first experiment.

- **C3** is the relative distance between the reference and the comparison spheres. Five possible values were chosen after preliminary testings:  $-0.3$  m,  $-0.15$  m, 0 m,  $+0.15$  m,  $+0.3$  m compared to the reference position.
- **C4** is the order of presentation of the two spheres, which was counterbalanced to minimize order effect.

Since the *side* factor was found not significant in the previous experiment, we chose to remove it from our experimental factors to reduce the number of trials. However, the *on the side* sphere was randomly placed right or left to avoid displaying the *on the side* stimuli always on the same side. This could have generated discomfort due to the fact that the subjects would have been turning their head always to the same side. In addition this helped the participant to keep facing front during the experiment.

In summary, participants were presented with 180 trials, divided in 3 blocks of 60 trials in a different randomized order for each block. Each block of 60 trials presented a set of couples of spheres made of: 2 reference positions (**C1**)  $\times$  3 angles (**C2**)  $\times$  5 relative

Table 4.3 – Experiment 2 – Fitting values for psychometric curves

angle	Ref. sphere in front		Ref. sphere on the side	
	$\alpha$ (SE)	$\beta$ (SE)	$\alpha$ (SE)	$\beta$ (SE)
15°	10.23 (0.72)	0.0180 (0.0076)	9.33 (0.29)	−0.0016 (0.0038)
30°	8.83 (0.69)	0.0607 (0.0096)	8.82 (0.86)	−0.0440 (0.0121)
45°	7.84 (0.84)	0.1232 (0.0139)	6.54 (0.60)	−0.1325 (0.0134)

positions (**C3**)  $\times$  2 presentation orders (**C4**).

#### 4.1.2.3 Results

The logistic regression analysis showed a main effect for **C1** ( $\chi^2 = 115.42; p < 0.001$ ), **C3** ( $\chi^2 = 1182.11; p < 0.001$ ) and **C4** ( $\chi^2 = 16.78; p < 0.001$ ), no effect was found for **C2** ( $\chi^2 = 0.77; p = 0.68$ ). In addition, two main interaction effects were found, the first being between **C1** and **C2** ( $\chi^2 = 52.69; p < 0.001$ ) and the second being between **C2** and **C3** ( $\chi^2 = 33.42; p < 0.001$ ). Thus, we will only discuss the interaction effects and the main effect of **C4**.

First, regarding the interaction effect between **C1** and **C2**, post-hoc tests showed that differences between the two levels of **C1** increased as the the angle (**C2**) increased. Precisely, post-hoc tests showed that while the effect of **C1** is not significant when the angle is 15° ( $z = 0.721; p = 0.98$ ), the effect of **C1** becomes significant for 30° ( $z = 4.274; p < 0.001$ ) and 45° ( $z = 9.711; p < 0.001$ ). This effect is also visible in the psychometric curves computed for each level of **C2** (see Table 4.3). The greater the angle, the greater the difference between PSE values.

Second, regarding the interaction effect between **C2** and **C3**, post-hoc tests did not provide conclusive results. Nevertheless, they suggest that participants' accuracy was slightly higher for the 15° and 30° conditions compared to the 45° condition. Finally, post-hoc tests for **C4** did not show any significant effect ( $z = 0.46; p = 0.65$ ).

#### 4.1.2.4 Discussion

The main effect on **C1**, as in the first experiment, validates our hypothesis **H1**.

The main finding of this experiment is the interaction effect between **C1** and **C2** highlighting the influence of the *angle* factor on the observed bias, i.e. distance overestimation of the *on the side* stimulus. This observation validates our hypothesis **H2**.

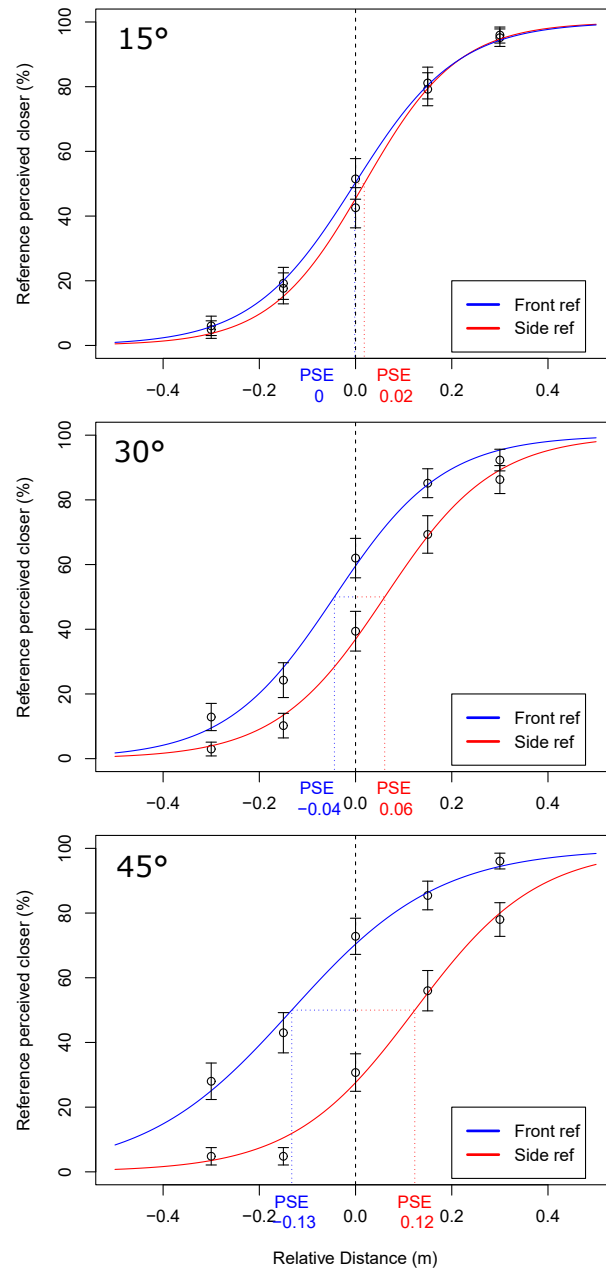


Figure 4.7 – Experiment 2 – Psychometric curves at different angular conditions (15° top, 30° middle, 45° bottom). For each possible relative distance between the reference and the comparison sphere we plot the percentage and the standard error of answers stating that the reference is closer. The red (resp. blue) curve shows the psychometric curve with a “on the side” (resp. “in front”) reference sphere. The corresponding PSE is displayed for each condition.

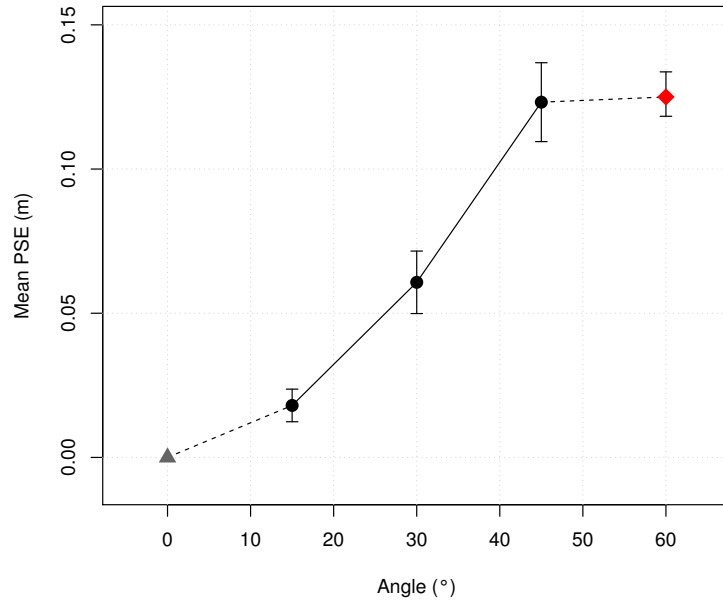


Figure 4.8 – PSE for each angle condition for Experiment 2, in black. The standard error was computed from the psychometric curves fitting. The red dot is from Experiment 1.

While the effect is non-significant for an angle of  $15^\circ$  (see Figure 4.7), the overestimation is significant for  $30^\circ$  and  $45^\circ$ . Interestingly this relationship seems to follow a non-linear behavior (see Figure 4.8) at least until  $45^\circ$ . For comparison purposes, the point at  $60^\circ$  in Figure 4.8 is the one obtained in the first experiment. The  $60^\circ$  condition was added for reference and suggests that the overestimation effect might reach a step around  $45^\circ$ .

### 4.1.3 Provisional Conclusions

In this section we presented a new perceptual bias leading people to perceive virtual objects displayed on their side as being farther away than virtual objects displayed in front of them. We showed that this overestimation is significant for angles of at least  $30^\circ$  and that the relationship between overestimation and angle does not follow a linear relationship.

Regarding the potential effects which could cause the observed bias, we suggested that the observers could prefer the “forward” direction in a forced choice when the task is too difficult. This non-perceptual but judgmental bias could explain the first experiment’s results. However, the second experiment shows an increase in bias along with the angle which indicates a perceptual bias linked with head orientation. Moreover, we conducted a

side experiment with 3 people (not reported here) and a reverse protocol (observers had to choose the closest sphere) and this experiment showed the same results of anisotropic distance estimation.

In our first experiment we observed that the *in front* stimulus always appeared on the center of the field of view (area of maximal vision acuity) while the *side* stimulus could appear outside of the area of maximal vision acuity. This occurred when participants did not rotate their head enough to put the *side* stimulus at the center of their FoV. This effect was observed due to the search behavior while looking for the object on the side and did not happen for the in front object since *center* is their resting position.

This was validated by analyzing participants' head orientation records taken during Experiment 1. Most of them turned their head  $5^\circ$  to  $10^\circ$  less than the correct angle when looking on the side.

Since an object which is not in the central vision provides deteriorated depth cues and that a deteriorated vision can lead to a bias in distance estimation (Ashley, 1898; Coules, 1955; Farne, 1977) we assumed that this difference in visual stimulus could imply the observed bias.

Another question that naturally arises regarding this surprising bias and the ecological validity of the VEs used in our experiments. Would the same effect be observed in a VE with more depth cues?

Based on these observations, we designed two control experiments aiming at studying (i) the potential influence of the position of the virtual object in the FoV and (ii) the influence of additional depth cues in VEs on the observed effect.

## 4.2 Control experiments

The following experiments were designed to rule out potential biases that could influence the results of our experiments.

### 4.2.1 Experiment 3: Disambiguating head and eye directions effects

This experiment was conducted in order to assess our hypothesis **H3**: the position of the stimulus in the field of view is responsible for the anisotropy of the distance estimation of virtual objects. In particular, it was designed to remove the potential bias that could

be introduced if the stimuli are not presented close to the center of the user's field of view. Thus, the protocol ensured that the stimuli were always presented as close as possible to the center of the user's field of view. The described experiment shared the same apparatus, calibration procedure, experimental design and procedure (see subsection 4.1.1.2) with the first experiment.

#### **4.2.1.1 Participants**

This experiment involved 15 participants (5 females, 10 males). They were aged from 20 to 51 (mean=31.3, SD=10.1). The participants were students or members of the laboratory, naive about the purpose of the experiment. They were not paid for their participation. All our participants had normal or corrected-to-normal vision and 10 out of 15 already had used an HMD before. None of them took part in any of the previous experiments.

#### **4.2.1.2 Experimental Procedure**

The procedure was identical to the one used in the first experiment, the only change consisted on how the user was guided in order to orient their head to face the desired stimulus. Since stimuli were always visible in the first experiment, voice commands (left, right, center), were sufficient to guide users towards the desired stimulus. On the contrary, in this experiment the visual stimulus was only displayed when the participant's head was correctly aligned ( $\pm 3^\circ$ ) with the desired position.

We discarded the use of visual cues as they could introduce additional biases. Instead, we focused on the use of auditory cues in order to guide the user. Stereo headphones were used to help the user aligning his head in the direction of the next stimulus using a stereo beeping sound. The sound was stronger on the side corresponding to the direction in which participants had to turn their head. In addition, to help participants aligning their head, the frequency of the sound increased when they approached the correct orientation. When the head was aligned correctly ( $\pm 3^\circ$ ), the stimulus appeared. As in the previous experiment, the user had to look at two different stimuli, one placed in front and another on the side. Each stimulus was presented during 2.5 s before disappearing. Once the first stimulus disappears, after a delay of 0.5 s the audio alignment procedure was repeated in order to guide the user to the second stimulus. When the second stimulus disappeared, the participant had to determine which stimuli was farther away using the joystick.

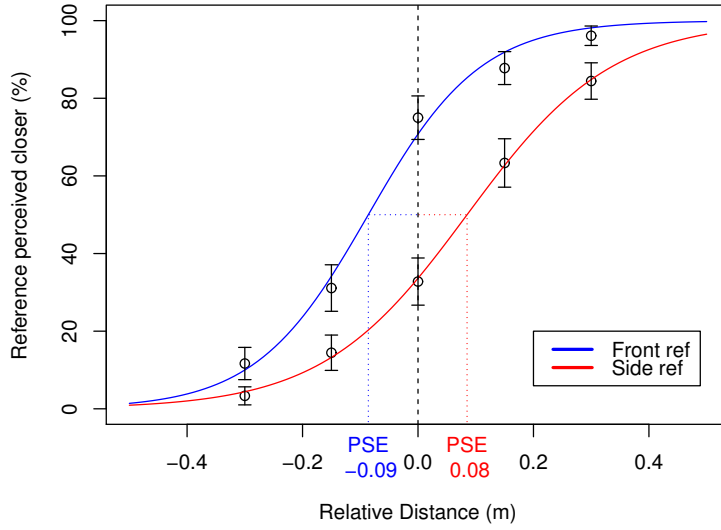


Figure 4.9 – Experiment 3 – Psychometric curves. For each possible relative distance between the reference and the comparison sphere we plot the percentage and the standard error of answers stating that the reference is closer. The red (resp. blue) curve shows the psychometric curve with a “on the side” (resp. “in front”) reference sphere. The corresponding PSE is displayed for each condition.

The average total time per participant, including pre-questionnaires, instructions, experiment, breaks, post-questionnaires, and debriefing, was 40 minutes. Participants wore the HMD for approximately 25–30 minutes.

#### 4.2.1.3 Results

As described in subsection 4.1.1.7, we used the same logistic regression analysis as in the first experiment. The analysis showed a main effect for **C1** ( $\chi^2 = 126.51; p < 0.001$ ), **C3** ( $\chi^2 = 696.64; p < 0.001$ ) and **C4** ( $\chi^2 = 13.05; p < 0.001$ ). No significant effect was found for **C2** ( $\chi^2 = 1.60; p = 0.2$ ) and no interaction effect was found.

Regarding **C1**, as in the first experiment, post-hoc tests showed that when the reference sphere was presented in front of the participant it was more likely to be considered as closer ( $z = 2.586; p < 0.01$ ). Post-hoc tests also showed significant differences for the different levels of **C3** showing that the choices of the participants were strongly dependent on the distance between the reference and the comparison stimulus. Finally, the post-hoc test did not show any significant difference for the levels of **C4** ( $z = 0.82; p = 0.415$ ).

Due to the significance of **C1**, we further analyzed the choosing rate of the reference sphere by fitting psychometric curves as stated in subsection 4.1.1.7. When comparing

*Side* and *Center* positions, the obtained coefficients were  $\alpha = 8.000$  ( $SE = 0.2330$ ) and  $\beta = 0.085$  ( $SE = 0.0038$ ) for the reference on the side and  $\alpha = 10.276$  ( $SE = 1.1624$ ) and  $\beta = -0.086$  ( $SE = 0.0122$ ) for the reference in front.

#### 4.2.1.4 Comparison with Experiment 1

The only difference between experiments 1 and 3 was the exploration behaviour of the user while searching for the stimuli. In order to assess its potential impacts on the results, we decided to compare the results of both experiments with the addition of a between-subjects factor *directed* with two levels (visual and acoustic).

The linear regression analysis showed that the *directed* factor was not significant ( $\chi^2 = 0.16; p = 0.68$ ). Regarding the other factors, no changes were observed. As previously, **C1** ( $\chi^2 = 429.59; p < 0.001$ ), **C3** ( $\chi^2 = 1547.52; p < 0.001$ ) and **C4** ( $\chi^2 = 44.20; p < 0.001$ ) had significant effects on the result while **C2** did not ( $\chi^2 = 0.16; p = 0.69$ ).

#### 4.2.1.5 Discussion

As in Experiment 1, the side where the object is presented does not affect the result. However, unlike experiment one, there is no significant order effect in this experiment. Nevertheless, in the first experiment the order effect was weaker than the other ones, so it might be present but too weak to be significant in this experiment.

Our research hypothesis **H3** was to explain the difference in distance perception of the object on the side by its peripheral position in the visual field. Thus, in this experiment, we carefully ensured that the observer always saw the sphere in the center of his FoV. However, we observed the same systematic bias as in the first experiment. Moreover, when considering the two experiments as a between-subjects design, results showed no statistically significant difference. As a result, we failed to validate our hypothesis **H3** and concluded that the horizontal position of the object in the observer's FoV is not responsible for the effect observed.

## 4.2.2 Experiment 4: Does the quality of the virtual environment matter?

Experiment 4 aimed at assessing the potential influence of depth cues on the bias observed in Experiments 1, 2 and 3. Indeed, we hypothesized that **H4**: this perceptual bias will still appear in VE with more depth cues. In the three previous experiments, due



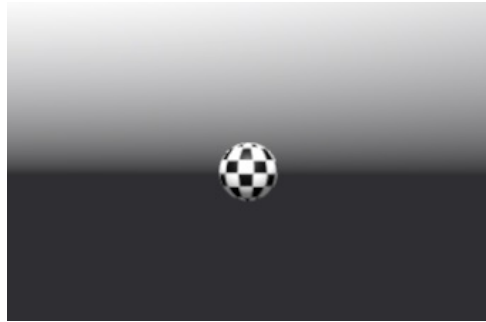


Figure 4.10 – Virtual stimulus used in Experiment 4 from the participant’s PoV. The VE is designed to provide more depth cues than the previous one (see Figure 4.2).

to the flat shading rendering, users perceived virtual spheres as 2D disks. On the contrary, in this fourth experiment the sphere was textured using a black and white checkerboard texture and a point light was placed above the head of the user (see Figure 4.10). This rendering configuration provides additional depth information in the virtual environment. Indeed, each square of the checkerboard provides anchors for the stereoscopic perception and the point light induces shading on the sphere, highlighting its 3D shape.

The experiment followed the same exact protocol and experimental design as Experiment 1 (see subsection 4.1.2).

#### 4.2.2.1 Participants

The experiment involved 16 participants (9 females, 7 males), aged from 18 to 52 (mean=31.9, SD=10.6). The participants were students or members of the laboratory naive about the purpose of the experiment. They were not paid for their participation. All our participants had normal or corrected-to-normal vision and 11 out of 16 had already used a HMD before. None of them took part in any of the previous experiments.

#### 4.2.2.2 Results

The logistic regression analysis showed a main effect for the reference position **C1** ( $\chi^2 = 105.75; p < 0.001$ ), the distance **C3** ( $\chi^2 = 734.48; p < 0.001$ ) and the order **C4** ( $\chi^2 = 16.52; p < 0.001$ ). In contrast, no significant effects were observed for the side **C2** ( $\chi^2 = 0.24; p = 0.62$ ). Interaction effects were non-significant. Post-hoc tests for **C1** showed that the stimulus was perceived closer when it was in the front and that recognition accuracy increases with the difference in distance ( $z = 3.101; p < 0.01$ ). For **C3**, post-hoc tests showed that except for comparisons +0.3 m vs +0.15 m ( $z = 2.182; p = 0.18$ ) and

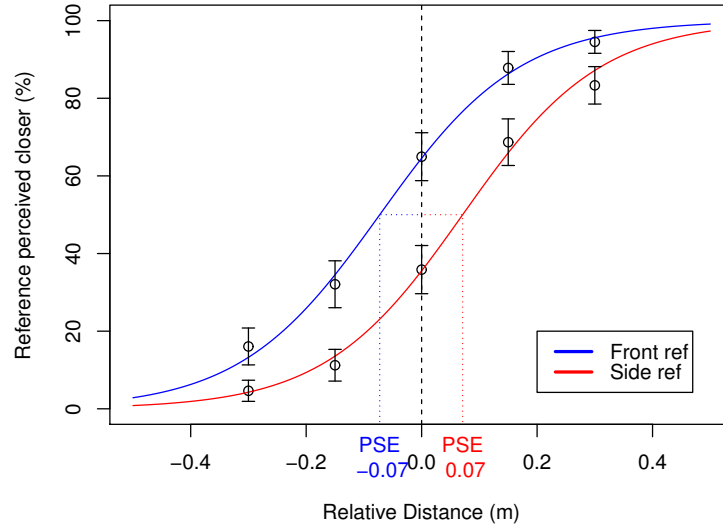


Figure 4.11 – Experiment 4 – Psychometric curves. For each possible relative distance between the reference and the comparison sphere we plot the percentage and the standard error of answers stating that the reference is closer. The red (resp. blue) curve shows the psychometric curve with a “on the side” (resp. “in front”) reference sphere. The corresponding PSE is displayed for each condition.

$-0.15$  m vs  $-0.3$  m ( $z = 2.17$ ;  $p = 0.18$ ) all other pairwise differences were significant (all  $p < 0.01$ ).

Due **C1**’s significance, we fitted a psychometric curve for each level of **C1**. When the reference was placed on the side the obtained coefficients were  $\alpha = 8.376$  ( $SE = 0.6619$ ) and  $\beta = 0.0711$  ( $SE = 0.0101$ ) and when the reference was placed in front the coefficients were  $\alpha = 8.260$  ( $SE = 0.515$ ) and  $\beta = -0.0725$  ( $SE = 0.0080$ ).

#### 4.2.2.3 Discussion

Results obtained in this experiment showed the same bias as the ones observed in Experiment 1, validating our hypothesis **H4**. But, if we compare the PSE values we can observe that the difference was smaller than the one obtained in the first experiment. This leads us to believe that the effect observed and detailed in the present section is not linked to the specific task but indeed illustrates a bias in perceived egocentric distances in VR.

### 4.3 Discussion

The results of our studies show a novel bias in distance perception in virtual environments: the perception of egocentric distances in VR is anisotropic. Virtual objects located on the sides are perceived to be farther away than objects in front. The first experiment showed that an object placed on the side is perceived 5% farther than an object placed in front. The work of Koenderink et al. (2002), conducted in a real environment, underlines that the *forward* direction is specific and leads to a deformation of the visual space. Thus, with our second experiment we aimed to characterize the observed bias by assessing distance estimation while varying the exploration angle. The results of the second experiment show that the overestimation effect on the side increases along with the angle. The greater the angle between the *in front* and the *on the side* stimuli, the more the observer tends to overestimate the distance of the stimuli on the side. Combined with the results of the first experiment, this effect seems to reach a plateau at 45°.

However, VR suffers from specific limitations which are not present in real environments. One of the most important regarding vision is the relative size of the FoV compared with ordinary real vision, and potential image distortions introduced by HMDs' optical systems. We conducted a third experiment to evaluate the potential impact of HMDs' optical systems specificities on the observed bias. We forced participants to look at specific directions in order to ensure that the visual stimulus was always displayed and viewed at the center of the HMD's FoV. With such an experimental protocol, the previously discussed issues were minimized. In this case, the results did not differ from those obtained in the first experiment. As a conclusion, we can state that the ocular particularity of side vision in VR does not seem to be responsible for the observed anisotropy of distance perception. Then, the effect observed in the two first experiments could not be due to different exploration behaviours or hardware limitations, we hypothesized that the motor action of the head was a good candidate to explain this bias.

The increase in overestimation as related to the head rotation amplitude seems to verify the effort-based theory of depth perception. Indeed, according to this theory, distance judgments seem to be influenced by the amount of energy observers anticipate to expend (Decety et al., 1989; Proffitt, 2006). In our experiments, the *in front* object is directly seen by the observer, while the *on the side* object requires an action (head movement) to be seen. This additional movement from the resting position could modify distance perception and lead to an overestimation. This assumption requires subsequent

studies to be evaluated. In particular, additional movements or forces could be added during the experiment to assess the effort-linked dimension of space perception under these conditions.

Finally, the last experiment we conducted, aimed at evaluating the impact of the VE on the observed bias. Indeed, in the three first experiments the VE mainly provided binocular depth cues which could have made the task more difficult, facilitating the appearance of the perceptual bias. In contrast, in our fourth experiment, the visual stimulus provided additional graphical depth cues, lighting, and increased contrast. Here again, results show that, even with more salient depth cues, the same perceptual bias is observed even if it is slightly weaker than in the first experiment. Since the environment provided more depth cues, the ordering task was probably easier. Participants were thus more likely to choose the correct sphere, decreasing a bit the appearance of the perceptual bias.

To analyze this study with respect to those conducted in real environments, our series of experiments can be compared with the frontoparallel curvature experiments of Koenderink et al. (2002). In this paper, they determined that frontoparallels are perceived as being curved, with the concave side facing the observer. In their experiments, when participants had to place objects to a given distance to define a frontoparallel, the objects in front of them were placed further away than the correct position, which implies that they were perceived as being closer. This observation is coherent with our results, suggesting that the bias we observe is not dependent on the visual stimulus but on the distance perception assessment (e.g. motor actions) and on the directionality of the virtual environment. However, since experimental conditions were not the same in their experiments and in ours, other studies, at different distances, are needed to further characterize this bias.

Last, the observed bias could be considered to draw new guidelines for the design of VR applications. When a VR application deals with realism, in architectural applications for example, it could be used to better match the space perception of the observer, and/or to provide an improved experience. The VE could also be oriented in order to be perceived longer in some direction. Moreover, VR applications which rely on highly oriented VEs, such as industrial accessibility tests, driving simulations, or some specific immersive games, could also consider this new perceptual bias to better specify their scenarios and their 3D contents.

## 4.4 Conclusion

In this chapter, we presented a study that evaluated observers' distance perception in VR environments, with minimal depth and environmental cue provided. This configuration aimed at eliciting observers' inner biases which could influence depth perception in VR.

We conducted a series of four experiments involving a total of 85 participants to study differences in distance estimation when comparing objects placed in front of the observer with objects placed on his side. In all experiments a two-alternative forced choice (2AFC) standardized psychophysical protocol was employed, in which the main task was to determine the stimulus that seemed to be the farthest one.

The first experiment showed that the orientation of virtual stimuli with respect to the observer introduces a distance perception bias: objects placed on the sides are systematically perceived as farther away than objects placed in front. This effect was confirmed with a second experiment, and was found to increase along with the angle of the object on the side. The two following experiments showed that this effect appears to be independent of both the position of the object in the field of view as well as the visual quality of the virtual scene.

Overall, this series of experiments highlights what can be called an “intrinsic bias” since all the other extrinsic explanations have been ruled out. Given the results of those experiments, the “forward” deformation of the visual space is likely to be explained by an inner anisotropy, an intrinsic tendency for people to perceive distances differently depending on the angle of view.





Like egocentric distance perception, other aspects of distance perception can also be biased. In this chapter, we propose an experiment to evaluate the potential bias of exocentric perception, i.e. between two objects, in Augmented Reality.

# TRANSFER OF PERCEPTUAL BIAS: FROM EGOCENTRIC TO EXOCENTRIC?

---

## Contents

---

<b>5.1</b>	<b>User Experiment . . . . .</b>	<b>120</b>
5.1.1	Participants . . . . .	120
5.1.2	Experimental Apparatus . . . . .	120
5.1.3	Procedure . . . . .	124
5.1.4	Experimental Design . . . . .	125
5.1.5	Statistical Analysis . . . . .	126
<b>5.2</b>	<b>Results . . . . .</b>	<b>127</b>
<b>5.3</b>	<b>Discussion . . . . .</b>	<b>130</b>
<b>5.4</b>	<b>Conclusion . . . . .</b>	<b>132</b>

---

In the chapter 3, we presented the vast corpus of distance perception studies in MR. Let us remind one of the major results of this field: the underestimation of egocentric distances in AR when compared to real estimates (see also section 3.6). While this bias can be mitigated via learning and feedback, this underestimation trend seems constitutive of distance perception in AR, in particular for OST HMDs.

However, while exocentric and egocentric distance perception can be related, the visual space remains non-linear and as such, distance perception can vary with the orientation of the distances considered. Hence, is the egocentric perception bias also observed in AR for exocentric distance perception?

To elicit this bias, we conducted a user study estimating distance perception between two objects lying in a frontoparallel plane in the medium-field perceptual space.



## 5.1 User Experiment

In this section we present a user experiment aimed at the evaluation exocentric distance perception in OST AR. To do so, we designed an experiment where participants had to evaluate distances between objects (spheres placed in a plane perpendicular to the users' view) in the following configurations: (i) two real spheres, (ii) two virtual spheres displayed in AR and (iii/iv) one real sphere and one virtual sphere. This last condition is in fact two-fold, since the real sphere can be displayed on the left hand side or on the right hand side of the participant. In the following, we detail the experimental design, the apparatus, as well as the protocol.

### 5.1.1 Participants

This experiment involved 29 participants (21 male and 8 female), aged from 19 to 54 years (mean=25.0, SD=7.2). Among them, 27 were right-handed, one was left-handed and one reported to be ambidextrous. The eye dominance was assessed using the method proposed by Durand and Gould (1910) in which participants were asked to look at a distant object through a hole made with their two hands. 9 participants had a dominant left eye and 20 had a dominant right eye. Participants were students of the university or members of the laboratory and naive about the purpose of the experiment. They were not paid for their participation. All participants had normal or corrected-to-normal vision and 23 out of 29 had already used a HMD before but none of them had used an OST AR HMD before.

### 5.1.2 Experimental Apparatus

The real stimuli were orange 3D-printed spheres with a radius of 20 mm. They were placed on a bench located 2.1 meters away from the participant (see Figure 5.1 left). This bench was composed of two symmetric rails, one for each sphere. A moving trolley was placed on each rail and controlled by a stepper motor. Given the mechanical characteristics of the motor and of the transmission system, spheres could be positioned with a precision of 0.2 mm. The control software of the bench run on a Raspberry Pi and was written in Python.

The placing mechanism was completely hidden by a black wooden structure. In addition, a *moving panel* was cut inside and could be toggled on and off in order to reveal or

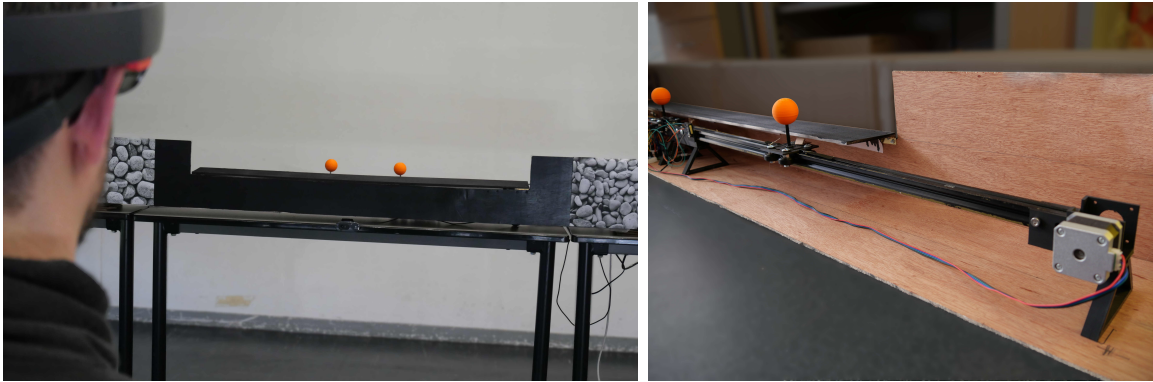


Figure 5.1 – Left, bench displaying two real spheres. The hinge-actuated moving panel, opened here, could be automatically opened/closed to reveal/hide the visual stimuli. Right, one of the two rails of the bench, seen from behind. An orange sphere is attached on top of a trolley that can slide on the rail. The trolley is moved by a stepper motor through a belt. The other half of the bench is symmetrical.

hide the visual stimuli. Figure 5.1 (center) allows to better appreciate the mechanism in which the moving panel is opened and the two real spheres are visible. All the other parts of the bench were always hidden during the experiment. AR markers were placed at each end of the panel (see Figure 5.4) which were used to calibrate the position of the bench in the VE.

The virtual stimuli were spheres displayed on an OST AR headset, the Microsoft HoloLens. The HoloLens has a FoV of  $30^\circ \times 17.5^\circ$  with a resolution of  $1268 \times 720$  px per eye. It contains an inertial measurement unit (IMU) and four sensors (a depth camera, a RGB camera, a four-microphone array, and an ambient light sensor) used to map the 3D environment in real-time. The HoloLens also supports voice input, basic gesture recognition, and head tracking. We selected the HoloLens mainly due to the quality and robustness of its tracking. The AR scenes were built using Unity3D v2018.3.0f2. The spheres were displayed using the standard Unity shader, they were lit with a single frontal directional light and without texture. This provided a coherent ambient and specular lightning (with no visible shadow since the light was frontal).

The HMD was calibrated before the experiment for each participant using the HoloLens' built-in calibration procedure. At the beginning of the experiment, the location of the bench was determined using the markers placed on it that were detected using the Vuforia Engine v8.0.10<sup>1</sup>. This position was registered and remained fixed during the whole

1. <https://developer.vuforia.com/>



Figure 5.2 – Participants could provide the perceived exocentric distance by placing two sliding spheres. After the participants placed the spheres the system automatically took a picture of both spheres which was used to measure the distance between both spheres.



Figure 5.3 – Three out of the four possible display conditions, from top to bottom: *real-real*, *real-virtual* (note that *virtual-real* is also part of the experimental design) and *virtual-virtual*.

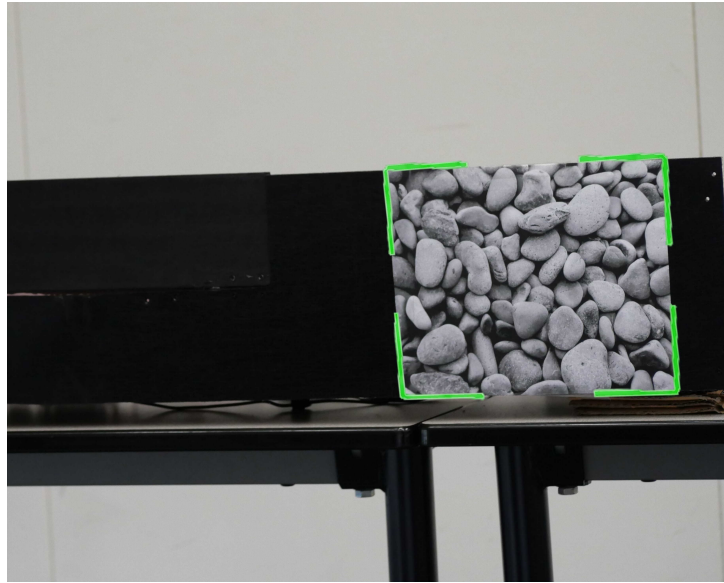


Figure 5.4 – An image used as a Vuforia marker, viewed from inside the HoloLens. The green corners represent the detection of the marker by the system and allowed users to assess a correct alignment during the calibration phase.

experiment.

In order to ensure the correct alignment between real and AR content, before each run we asked participants to visually assess that the real and the virtual spheres displayed and placed at the same position were superimposed. If there were any mismatch between the two spheres, the application was reset. This never occurred except when participants took off the headset during the break, which leads us to believe that the calibration was robust enough for our purpose.

Participants were seated on a fixed chair in front of a table. They were asked to sit straight and with their back resting on the back of the chair during the whole experiment. Two spheres, identical to those placed on the bench, were fixed on sliding support on a rail in front of them, see Figure 5.1 left. Participants could answer by sliding these *answering spheres*. The distance between the two answering spheres was recorded by a camera placed under the bench, in front of the participant. To compute the distance between both spheres, the camera took a picture and the image obtained was filtered to detect the spheres' centroid positions. To do so, the picture was first converted in the LAB color system to get rid of the luminosity component and then thresholded to only keep the specific orange color of the sphere (see Figure 5.5). The whole procedure was implemented using OpenCV 3.2.0. The camera was calibrated using OpenCV's standard

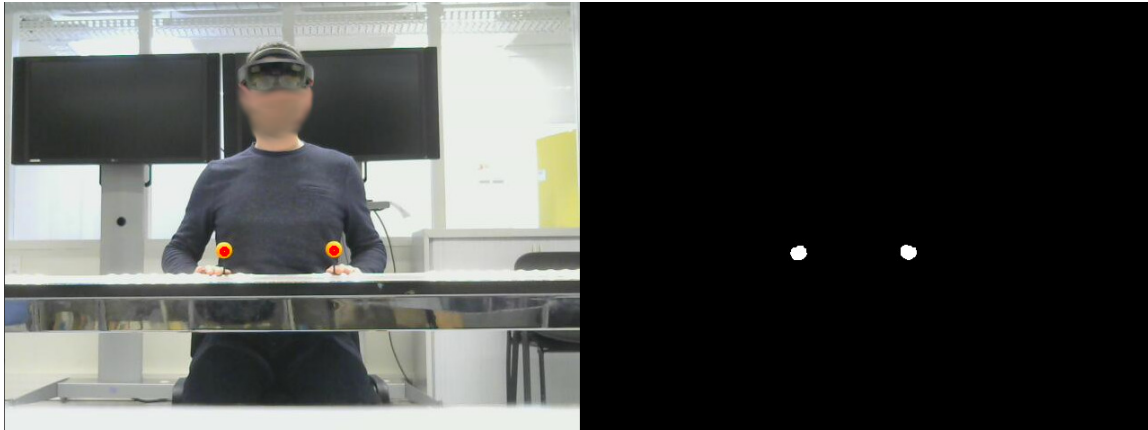


Figure 5.5 – Detection of the spheres’ positions with the camera. Left: original image from the camera. Right: Image obtained after applying a color threshold in the LAB color system.

calibration method at the beginning of the experiment by placing the spheres at a reference known distance. The resolution of the camera was  $1280 \times 720$  px which corresponds to a precision of 2 mm on the recorded answer.

### 5.1.3 Procedure

Participants started by reading and filling out a short form containing written instructions about the experiment. After verbal explanations, they carried out the in-built calibration procedure of the HoloLens. Then, participants performed 10 trials in order to get used to the experimental procedure. Finally, participants had to perform two blocks of 48 trials each (see subsection 5.1.4). Participants were allowed to have a break between the two blocks. The procedure for each trial was as follows.

At the beginning of each trial, the moving panel of the occluder was closed and the moving spheres in front of the participant were side by side, regrouped at the center of the rail. The moving panel opened, revealing zero, one or two real spheres, according to the trial condition. Simultaneously, the virtual spheres (if relevant for the trial) were displayed through the HMD. There were two distinct spheres visible at a time. The spheres were visible for five seconds. After that time, the moving panel closed hiding the physical spheres, and, if present, the virtual spheres disappeared.

The participant then had five seconds to move the spheres in front of him to reproduce the same distance as the distance between the spheres he just saw. During this time, a virtual indicator informed the participant of the time left to answer. Once this time was

Table 5.1 – Summary of Independent and Dependent Variables.

<b>Independent Variables</b>			
	<i>Observers</i>	29	(random variable)
<b>C1</b>	<i>Distance (cm)</i>	8	25, 30, 35, 40, 45, 50, 55, 60
<b>C2</b>	<i>Type</i>	4	real-real (RR), virtual-virtual (VV) real-virtual (RV) , virtual-real (VR)
	<i>Repetition</i>	2	1, 2
<b>Dependent Variables</b>			
<b>D1</b>	<i>Reporting</i>	Distance between the two spheres positioned by the participant	

up, the camera recorded the distance between the two spheres. If the participant was still adjusting the distance after five seconds this response was removed from the analysis. We believe this did not have a strong impact on the results as this happened rarely (less than 0.005% of the trials).

The participant had then five seconds to bring the sphere in front of him at their initial position while the spheres of the bench, behind the occluder, was moving to prepare the next trial.

The proposed protocol was designed to ensure that (1) all participants looked at the stimuli the same amount of time, (2) that the stimuli was placed at an optimal viewing distance to minimize the potential accommodation and vergence bias and (3) that the stimuli was visible simultaneously by both eyes without head motions even when the distance between them was the maximum. This design encouraged participants to exhibit a similar exploration behavior (static), thus potentially reducing variability.

In total, the average time per participant, including pre-questionnaires, instructions, calibration, experiment, breaks, post-questionnaire and debriefing, was 45 min. Participants wore the HMD for approximately 30 min.

#### 5.1.4 Experimental Design

We used a mixed-model within-subjects design with the following conditions:

- **C1**: The distance between the two stimuli. Eight possible values were chosen: 25 cm, 30 cm, 35 cm, 40 cm, 45 cm, 50 cm, 55 cm, 60 cm.
- **C2**: The type for each stimulus, which could be both real, both virtual, or one real and one virtual (*RR*, *VV*, *RV*, *VR*). In the *one real/one virtual* condition, in order

to minimize potential biases due to the placement of the “real” sphere (placed on the left or on the right), the placement of the real sphere was counterbalanced. Therefore, for the real-real and virtual-virtual conditions, the number of trials was also doubled to ensure the same number of repetitions for all **C2** conditions. We were expecting that the effect of the side was not significant, but we included it in the design for completeness.

The distance between the participant and the stimuli (2.1 m) was chosen to be in the comfort zone of the HoloLens<sup>2</sup>. This specific distance should limit potential biases due to a mismatching focal distance between the displayed virtual objects and the real objects. In addition, since the horizontal augmented FoV of the HoloLens is 30°, **C1** was chosen to ensure that both spheres could be simultaneously on the user’s field of view.

In summary, participants were presented with 96 trials: 8 distances (**C1**) × 4 stimuli (**C2**) × 2 (for asymmetrical cases) × 2 blocks. For each block, the order of **C1** and **C2** were fully randomized to minimize potential order effects. The only dependent variable was the reported exocentric distance **D1**.

Following previous research, and based on our experimental design, our main research hypotheses were the following:

- **H1** Exocentric distance perception is the most accurate for the real/real condition.
- **H2** Exocentric distance perception for the virtual/virtual is underestimated when compared to the real/real condition.
- **H3** Exocentric distance perception for the real/virtual and virtual/real conditions is in-between the both real/real and virtual/virtual conditions.

### 5.1.5 Statistical Analysis

The reported distance was analyzed using general linear mixed model. The model included the independent variables **C1** and **C2** as fixed effects and the *participant* factor as a random effect. Tukey post-hoc tests ( $\alpha > 0.05$ ) were done when needed using the Bonferroni correction for multiple comparisons. The degrees of freedom were corrected using the Greenhouse Geisser correction method when the sphericity assumption was violated. The statistical analysis was carried out using the R software using the function *aov\_ez* of the *afex* package. The functions *emmeans* (*multcomp* package) and *glht* (*multcomp* package) were used for pairwise comparisons.

---

2. <https://docs.microsoft.com/en-us/windows/mixed-reality/comfort>

The statistical analysis only considered the relative distance estimates, computed as the percentage between the real distance between stimuli and the estimated distance from participants.

## 5.2 Results

The general linear mixed model of **C1** and **C2** versus the relative distance estimation showed a significant effect for the placement of the stimuli, **C2** ( $F(1.99, 51.63) = 8.89; p < 0.001; \eta_p^2 = 0.25$ ), but not for the distance factor **C1** ( $F(2.66, 69.08) = 2.16; p = 0.11; \eta_p^2 = 0.08$ ), neither for the interaction term **C1**  $\times$  **C2** ( $F(10.68, 277.56) = 1.49; p = 0.14; \eta_p^2 = 0.05$ ). In order to further study the effect of **C2**, we conducted post-hoc tests to detect significant pairwise differences. Results showed three significant pairwise differences, see Figure 5.6, namely:

- stimuli were perceived as significantly closer in the *VV* condition compared to the *RR* condition ( $p = 0.017$ ),
- stimuli were perceived as significantly closer in the *RV* condition compared to the *VR* condition ( $p = 0.014$ ),
- stimuli were perceived as significantly closer in the *VV* condition compared to the *VR* condition ( $p < 0.001$ ).

The significant pairwise difference *RR* vs. *VV* did not support **H1** but it supported **H2**, the estimates for the *RR* condition were less accurate, but estimates for the *VV* condition were significantly smaller. In addition to the main differences among visual conditions, we observed a consistent and noticeable overestimation estimates for all visual conditions, above 25% in overall (cf. Figure 5.6). In addition, the non-significant main effect from **C1** showed that the similar relative estimates were observed no matter the distance between stimuli.

Regarding **H3**, results were inconclusive as they showed an asymmetry for the *RV* and *VR* conditions. Our results show that the stimuli on the *VR* condition were perceived as significantly farther away than stimuli of the *RV* as well as the *VV* condition. We did not expect that the position of the virtual sphere would play such an important role, and initially assumed that both *RV* and *VR* would give rise to comparable results.

In order to further study the asymmetry of both conditions with one real and one virtual object, we conducted a statistical analysis integrating the information of participants'



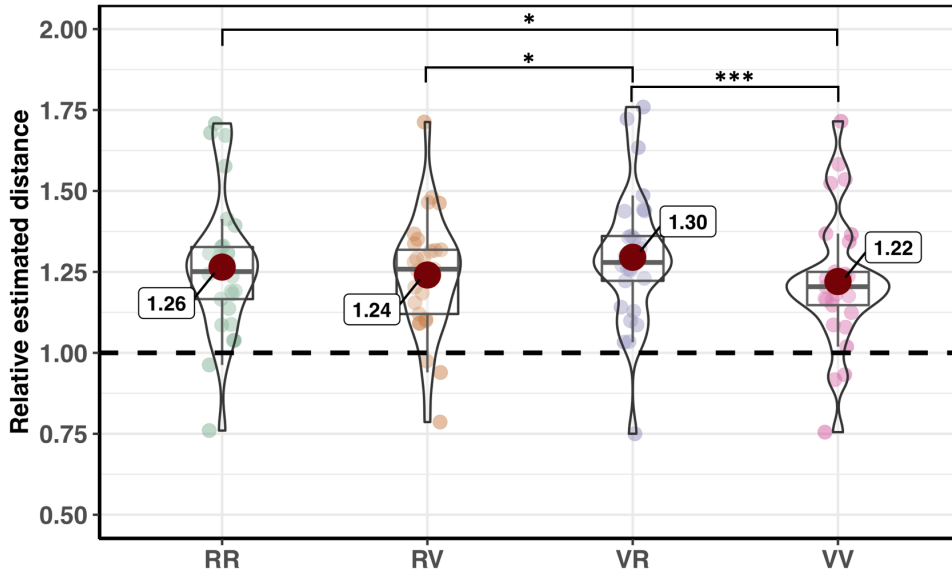


Figure 5.6 – Combined box and violin plot of the relative distance estimation errors averaged for all distances across the four visual conditions: *real-real*, *real-virtual*, *virtual-real* and *virtual-virtual*. The red dot represents the mean value.

eye dominance into account and only considering *RV* and *VR* trials. We thus grouped the trials into two groups (**C3**):

- *Vdom*: trials for which the virtual object is placed on the same side as the participant’s dominant eye.
- *Rdom*: trials for which the real object is placed on the same side as the participant’s dominant eye.

The general linear mixed model of **C1** and **C3** versus the relative estimated distance. The results only showed a main significant effect for **C3** ( $F(1, 27) = 6.47; p = 0.02; \eta_p^2 = 0.19$ ), there was no effect for **C1** ( $F(3.57, 96.30) = 2.09; p = 0.10; \eta_p^2 = 0.07$ ) nor the interaction term **C1**  $\times$  **C3** ( $F(4.78, 129.09) = 0.43; p = 0.82; \eta_p^2 = 0.02$ ). Pairwise post-hoc tests showed that *Vdom* estimates were significantly smaller than *Rdom* estimates ( $p = 0.017$ ). Figure 5.7 shows the summary of the results.

In addition, we further analyzed the data for each participant, comparing the estimate of the average relative distance for each condition. Results showed that 78% of the participants perceived *Vdom* as smaller than *Rdom* and 88% of the participants perceived *VV* as smaller than *RR*.

Finally, we also explored the potential inter-user variability (see Figure 5.8), which

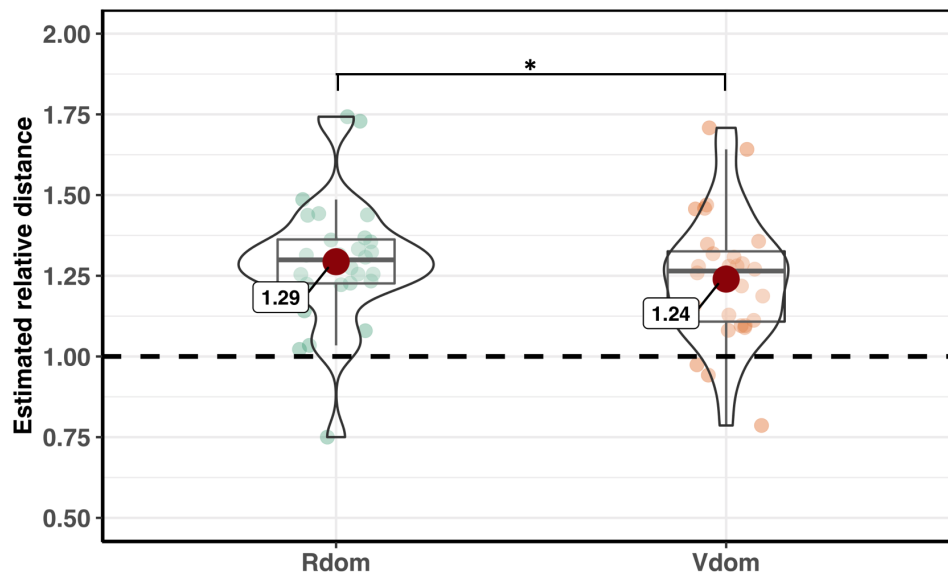


Figure 5.7 – Combined box and violin plot of the relative distance estimation errors for the *Vdom* and *Rdom* trial groups averaged for all distances and conditions.

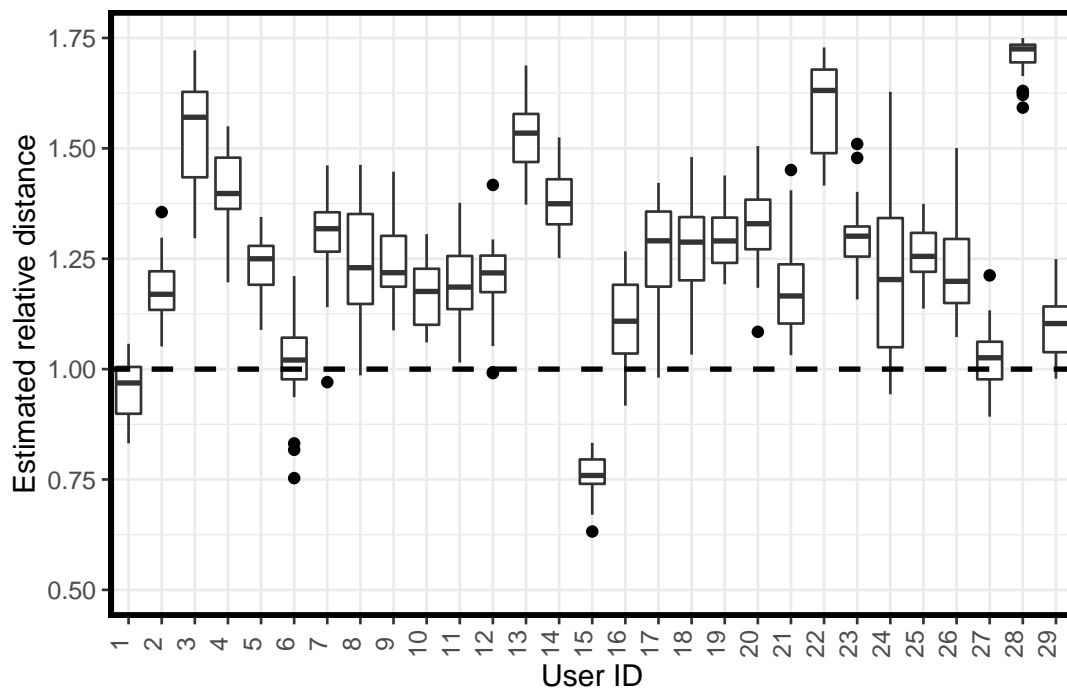


Figure 5.8 – Box plot of the relative distance estimation for each user averaged for all distances and conditions.

showed a relatively strong user variability. For example, the average overestimation for users 3, 13, 22 and 28 was greater than 50% and user 15 had an underestimation of 25% while users 1, 6 and 27 provided estimates close to the real distances.

### 5.3 Discussion

Results of this study show that the distance between two objects presented in OST AR in the perceptual middle-field are overall overestimated (by around 25.5%) across all four conditions where both objects can be either: *real-real*, *real-virtual*, *virtual-real* and *virtual-virtual*.

When further analyzing those results, we also notice that two virtual objects in OST AR are perceived as significantly closer than two real objects at the same positions. For the farthest distance (over 50 cm between both objects), this effect is equivalent to an underestimation of 4% for the *both virtual* case when compared with the *both real* layout. This effect is in line with previous results of studies showing an underestimation of exocentric distances both in AR (Dey et al., 2012) and in VR (Kelly et al., 2015), but here we observed it for smaller distances.

One possible explanation of this phenomenon can be linked to the underestimation of egocentric distances. Under the tested condition, if the virtual spheres are perceived closer than the real ones, and since their positions in the visual field is the same, the resulting exocentric distance is perceived smaller.

Even if the VE was carefully designed to provide the same lighting, the limitations of the OST display induced a different representation between the virtual and the real spheres. Therefore, there were clear, noticeable differences in lighting and shading according to the nature (real or virtual) of the sphere (Ashley, 1898). This difference in rendering between real and virtual spheres could thus lead to a bias in perceived distances.

The global overestimation of distances for all the tested conditions is not a usual result. In our case, exocentric distances are reported as 25.5% larger on average. This result is unusual since the majority of studies report a rather accurate exocentric distance perception in the frontoparallel plane, see (Geuss et al., 2012). However, these results can be linked with the chosen distances between objects as well as the reporting procedure chosen in our experiment. Here, participants had to move two real objects in front of them. Yet, the proprioceptive distance perception is known to be non-homogeneous (Fasse et al., 2000). Since the reporting task involves hand movement, the answer could be biased by

this effect.

Finally, the most surprising result in our study corresponds to the asymmetrical results from the *real-virtual/virtual-real* setups. These results are more difficult to interpret. While we expected for these two layouts to be symmetrical, this was not the case. We also hypothesized (see **H3**) that the *RV* and *VR* conditions would lead to results in-between those of the *RR* and *VV* conditions. This result is unexpected since previous results, see Lin et al. (C. J. Lin et al., 2015), showed a constant underestimation in the *one virtual/one real* situation.

We thus tried to further analyze this result to understand what could lead to this asymmetry of error estimations based on the position of the virtual stimulus. Since our experimental setup was designed to be completely symmetrical, this effect of the side was not expected and we suggested that it relies on some internal lateralization of the participants. To do so, we grouped trials based on the position of the stimulus in relation with the participants' dominant eye of the subject. Our results showed a significant effect between *Vdom* (virtual object on the side of the dominant eye) and *Rdom* (real object on the side of the dominant eye) with an overestimation in the *Rdom* case of around 5% (see Figure 5.7). Therefore, the results are closer to the *real-real* condition when the real object is on the side of the dominant eye, and are closer to the *virtual-virtual* condition when the virtual object is on the side of the dominant eye. As a first explanation of this eye dominance, we supposed that the observer could have focus more on the object on the same side as its dominant eye and consider the other object as symmetrical.

Hand dominance could also be a potential factor explaining this result, yet, only one subject was left-handed, which shows that the effect happened already with a right handed population, ruling out this potential bias.

Another possible explanation for this phenomenon could rely on luminance. Indeed, luminance has been reported as being a strong factor for depth estimation for real objects (Ashley, 1898; Coules, 1955; Farne, 1977) or more recently in AR (Singh et al., 2018). More saturated and brighter objects tend to be perceived as nearer than less saturated and bright objects. Then, a difference of illumination between the screens of the HoloLens could lead to a depth mismatch of the objects and then a perceived exocentric distance biased. Therefore, we measured the respective luminance of the screens by displaying a full screen white rectangle on the HoloLens and taking pictures of the screens using a camera (Point Grey Flea3 FL3-U3-32S2C). We then calculated the average brightness of the central part of the screens (where the sphere was displayed). We found that the right

screen was brighter than the left one with a contrast value of 14%. However, since the experiment was designed in way that both spheres were always visible by both eyes at the same time, the potential effect of luminance in that case should also be counterbalanced. Moreover, the HoloLens being an OST HMD, the relative illumination of the augmented objects is background-dependent and can vary because of the head movements of the observer.

As of now, we cannot provide a specific explanation for this phenomenon and further studies are required to better analyze it.

## 5.4 Conclusion

In this section, we presented an experiment aiming at evaluating the effect of AR on exocentric distance estimation. This experiment was rooted in the more general background of egocentric distance underestimation in the action/vista space in AR.

We reported a user study (n=29) estimating distance perception between two objects lying in a frontoparallel plane at 2.1m from the observer (i.e. in the medium-field perceptual space). Four conditions were tested in our study: real objects on the left and on the right of the participant (called *real-real*), virtual objects on both sides (*virtual-virtual*), a real object on the left and a virtual one on the right (*real-virtual*) and finally a virtual object on the left and a real object on the right (*virtual-real*). Participants had to reproduce the distance between the objects by spreading two identical real objects presented in front of them.

Results showed that two virtual objects were perceived closer one to another than two real objects, i.e. the exocentric distance was underestimated in AR compared to reality. We also reported a non-symmetrical behavior in the mixed configuration (one real object and one virtual object), which suggests that the relationship between real and virtual objects with respect to the user might be affected by other external factors.





As for reality, depth cues are used by the human perceptual system in Augmented Reality to evaluate distances. In this chapter, we present two experiments focusing on the evaluation of the effect of shadows and accommodation on distance perception.

# THE IMPACT OF DEPTH CUES: SHADOWS AND ACCOMMODATION DISTANCE

---

## Contents

---

<b>6.1 Impact of shadow shape and consistency on depth perception</b>	<b>136</b>
6.1.1 Research Hypotheses . . . . .	136
6.1.2 Experimental design . . . . .	137
6.1.3 Experiment 1 - Studying the influence of the shape of virtual objects' shadows on distance perception in OST AR . . . . .	141
6.1.4 Experiment 2 - Studying the influence of lighting misalignment on distance perception in OST AR . . . . .	151
6.1.5 Discussion . . . . .	159
6.1.6 Conclusion . . . . .	161
<b>6.2 Studying the influence of accommodation distance through the use of Retinal Projection Displays . . . . .</b>	<b>161</b>
6.2.1 Methods . . . . .	162
6.2.2 Results . . . . .	170
6.2.3 Discussion . . . . .	171
6.2.4 Conclusion . . . . .	174

---

Factors that can impact distance perception in AR and explain its misestimation are numerous, but a vast majority of these factors relies on the differences between the depth cues provided by real and virtual elements. Technical limitations or rendering specificities of AR devices imply that virtual objects are displayed differently from real objects. This leads to inaccurate depth cues and thus potentially biases depth estimation in AR. In this section, we will tackle the effect of two specific depth cues in AR: (1) shadows and (2)



accommodation.

## 6.1 Impact of shadow shape and consistency on depth perception

Shadows are often referred as “anchoring cues”. They allow the observer to get information of an object based on the the position of the lights and the shape of its shadow. When closer to other objects, or even when *on* the object, they can help the observer to correlate the object’s position with the position of the shadow on the surface, e.g. an object placed near the floor will be perceived as more “grounded” when displayed with a shadow below. This depth cue is especially important in AR since it can provide a supplementary virtual augmentation (the virtual shadow) of a virtual object in a real environment.

However, in AR, the computation of appropriate shadows is a major technical challenge. Indeed, to provide realistic shadows, regarding both their shapes and positions, requires the computation of all the positions and illumination parameters of all the objects and light sources, each of which could be real or virtual. Such a computation remains far from the rendering capacities of AR devices and current devices use approximated rendering to provide appropriate enough rendering with respect to their computational power.

In this section, we will present a study aiming at evaluating the effect of shadows on distance perception in AR, (1) using shadows of different rendering quality and (2) with lights’ positions coherent or not with the real scene. This work was conducted in collaboration with Yuan Gao, PhD student at the Beijin Institute of Technology.

### 6.1.1 Research Hypotheses

The influence of virtual objects shadows as well as lighting (which can be particularly difficult to accurately reproduce in an AR environment) on distance perception remains mainly unresolved. Based on previous research, we propose the four following hypotheses:

- **H1** Participants underestimate egocentric distances in OST AR.
- **H2** Virtual objects shadows improve distance perception accuracy in OST AR.
- **H3** Shadows with different realism levels have different influences on distance perception in OST AR.

- **H4** Lighting misalignment between the real and the virtual environments negatively impacts distance perception in OST AR.

The purpose of the present study is to explore issues related to shadows in AR and their influence on depth perception. We decided to focus on an AR scene displayed in the medium- field distance, which covers a wide scope of AR applications.

First, we explore the accuracy of distance evaluation in an OST AR scene as well as the influence of virtual shadows on distance perception. In a second step, we study the impact of lights' misalignment on depth perception in OST AR. Two experiments based on an OST-HMD were designed to tackle these issues.

Participants were required to match a virtual object with a real line on the floor in each trial. Distance judgment accuracy for each trial was recorded and analyzed.

## 6.1.2 Experimental design

Two experiments were designed to verify our hypotheses. Since both experiments have very similar experimental procedures, tasks and apparatus, we decided to regroup their description in the following.

### 6.1.2.1 Apparatus

Both experiments were conducted on a Microsoft HoloLens. It has a  $30 \text{ deg} \times 17.5 \text{ deg}$  FoV with a resolution of  $1268 \times 720 \text{ pixels}$  per eye. It contains an inertial measurement unit (IMU) and four sensors (a depth camera, a RGB camera, a four-microphone array, and an ambient light sensor) used to map the 3D environment in real-time. The HoloLens also supports voice input, basic gesture recognition, and head tracking. In addition, the HoloLens displays are fixed at an optical distance approximately 2.0m away from the user, which means that users have to accommodate near 2.0m to maintain a clear image. In other words, placing virtual augmentations around this distance can minimize the vergence-accommodation conflict. We selected the HoloLens mainly due to the quality and robustness of its tracking. In addition, a Bluetooth mouse was adopted as an input mechanism for our task (see subsection 6.1.2.2). The AR scenes were built using Unity3D v2017.3.1f1 and the HoloLens Toolkit<sup>1</sup>.

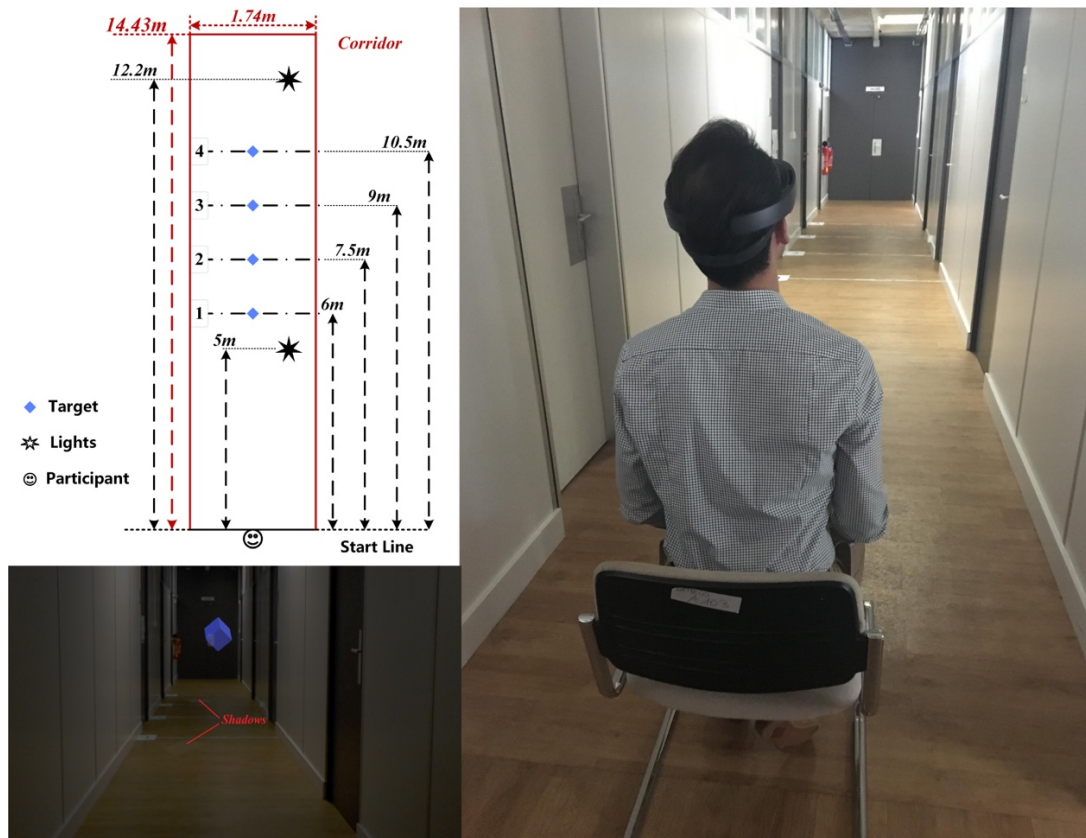


Figure 6.1 – Figure 1: Top left: Top view of the position of the different targets and two real lights. Bottom left: View from the HoloLens. Right: Overall view of a participant in the corridor

### 6.1.2.2 Experimental Environment and Tasks

The experiments were conducted in a corridor which is 14.43m long, 1.74m wide and 3.15m high. The corridor is an enclosed space and there were two lights on the ceiling for lighting, each light consisting of four fluorescent tubes. A virtual cube floated in the air at a height of 1.25m from the ground (see Figure 6.1) and rotated at a constant speed of 12 deg/s to enhance participants' stereoscopic sensation. In order to facilitate the task the cube is oriented such that one of its vertex points to the floor (see Figure 6.1). Participants saw the virtual cube through the HoloLens. Due to the HoloLens' limited FoV, it was necessary to guarantee that subjects had a low visual horizon so that they could see virtual shadows on the floor. In order to do so, subjects were seated on a chair located on a reference line at the end of the corridor (see Figure 6.1). In addition, four white lines were drawn on the floor (at 6m, 7.5m, 9m and 10.5m from the subjects position) to serve as references for four different targets distances. Four A4 (210 × 297mm) pieces of paper with numbers from 1 to 4 printed on them served as a way to recognize the targets (named as Target 1, Target 2, Target 3, Target 4 from nearest to furthest from the subjects position), (see Figure 6.1). The two real lights were located at 5m and 12.20m from the subjects' position and at 3.15m high.

Three main factors were considered when determining this range for our targets: (i) the shadows of the virtual objects had to be visible by the participants, (ii) virtual objects should be displayed at a distance allowing natural interaction and (iii) virtual objects should be within a comfort zone of the HoloLens.

Regarding (i), as shown in Figure 6.1, the two real lights were located at 5m and 12.20m from the subjects' position at a height of 3.15m. If the virtual cube was too close to the participants, its shadows could be displayed behind them. On the other hand, if the virtual cube was too far away from the participants, its shadow would be invisible (i.e. too far away and too small). Preliminary tests allowed us to determine that the target distance range of 6m to 10.5m could guarantee shadows of the virtual cube to be displayed on the ground within a visible area. This range also corresponds to a common distance range for human daily interaction which is highly suitable for most AR applications (cf. criterion (ii)). Finally, although the distance range of 6m to 10.5m exceeds the optimal distance (1.25m to 5m) for placing virtual objects in the HoloLens, this distance range is still within a relative comfort zone for HoloLens, and will not be uncomfortable<sup>2</sup> for

---

1. <https://github.com/Microsoft/MixedRealityToolkit-Unity/releases?after=2017.4.0.0-rc>  
2. <https://docs.microsoft.com/en-us/windows/mixed-reality/comfort>

users (criterion (iii)).

Each experiment consisted of a fixed number of trials. Participants saw a virtual cube floating in the air with or without a shadow. They could move the cube forward and backward along the axis of the corridor by scrolling the wheel of the Bluetooth mouse. The minimal step of the wheel induced a 5cm translation of the cube. Participants had to match the lowest vertex of the cube with a target line on the ground specified in each trial (see Figure 6.1 bottom left). Target lines were announced vocally to participants directly using the HoloLens built-in speakers at the beginning of each trial. Participants needed to confirm the target line to make sure they picked the correct one. Participants confirmed the final location of the cube by clicking on the mouse’s wheel. We recorded the final distance of the cube for each trial and loaded the next trial automatically. In order to guarantee that participants understood the procedure of the experiment correctly, each of them went through a pre-experiment practice session which included 10 trials of the same task.

### **6.1.2.3 Experimental Procedure**

Both experiments were within-subject studies and followed the same procedure. Each experiment lasted for 20-25 minutes. The procedure of the experiment is depicted in Figure 6.2 and consisted of five steps: (1) Screening; (2) Introduction and subject informed consent letter; (3) Calibration; (4) Practice session; (5) Formal experiment. Screening consisted of making sure that participants had normal or corrected to normal vision as well as normal color vision and normal stereoscopic vision. In order to do so, all the participants performed visual function tests including the visual acuity test, the stereo acuity test and the Ishihara test (Hardy et al., 1947).

In addition, the subjects experience with VR and AR was also recorded through a subjective questionnaire before the experiment. After that, the experimenter introduced the experiment content to the subjects, without mentioning the goals, and gave them the Subject Informed Consent Letter. This letter tells subjects that the experiment design and procedure were in accordance with the Declaration of Helsinki. Subjects were also informed that they could leave the experiment at any time or if they felt uncomfortable without the need of getting permission from the experimenter nor to justify themselves. Finally, they signed the letter.

Participants sat on a chair and were helped to put on the HoloLens. A calibration procedure was carried out to guarantee that the trajectory of the virtual cube matched

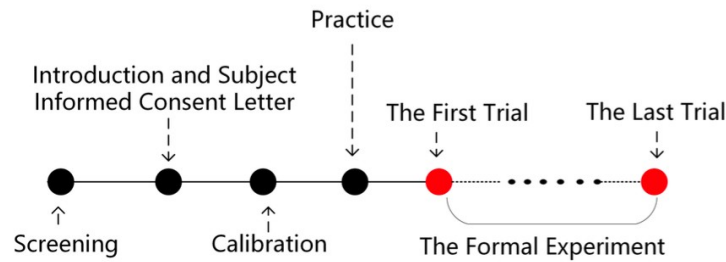


Figure 6.2 – Experimental procedure

the axis of the corridor. The calibration procedure is as follows: participants saw three dots arranged in a line in the HoloLens’ FoV. They had to align the dots with the middle axis of the corridor and make sure this line was perpendicular to the corridor’s start line (see Figure 6.1, top left). After the calibration, the practice session, consisting of 10 trials and during which the experimenter explained the task, loaded automatically. Once the practice session was over, the formal experiment started. The order of the trials was randomized for each participant. In both experiments, the task was identical to that of the practice session which is to match the position of the virtual cube with a specified target on the floor. At the beginning of each trial, a vocal announcement in the HoloLens built-in speakers specified the target.

Objective measures were recorded: signed distance errors, absolute distance errors and completion time. The signed distance error represents the distance between the center of the final position of the virtual cube (aligned with the lowest vertex) and the target along the Z-axis (i.e. the axis of the corridor). A signed distance error  $> 0$  means that it is an egocentric underestimation (i.e. the cube is behind the target line). A signed distance error  $< 0$  means that it is an egocentric overestimation (i.e. the cube is in front of the target line). The absolute distance error is the absolute value of the signed distance error.

### 6.1.3 Experiment 1 - Studying the influence of the shape of virtual objects’ shadows on distance perception in OST AR

The main purpose of the first experiment is to verify: **H1** “Participants underestimate egocentric distances in OST AR”, **H2** “Virtual objects’ shadows improve distance perception accuracy in OST AR”, and **H3** “Shadows with different realism levels have different influences on distance perception in OST AR”.

A matching task under different conditions was designed to verify these hypotheses. For more details on the task, please refer to subsection 6.1.2.2. Participants accomplished the tasks under the conditions of presence or absence of virtual objects' shadows in the AR scene. In addition, when virtual shadows were present, three different levels of shadow realism were studied. Hypotheses **H1**, **H2** and **H3** were then assessed by comparing distance errors under the aforementioned conditions.

### 6.1.3.1 Experimental Conditions

The conditions that we want to manipulate in this experiment were (i) the presence (Shadow ON) or absence (Shadow OFF) of virtual objects' shadows in the AR scene and the impact of the level of realism of virtual objects' shadows. This second condition was decomposed into three sub-conditions, namely:

- “Round” shadows (see Figure 6.3 top right): they have an incorrect shape and have sharp crispy edges. In the following, we refer to them as *round shadows*.
- “Hard-edge” shadows (see Figure 6.3 middle right): they have a correct shape but have sharp crispy edges. In the following, we refer to them as *hard-edge shadows*.
- “Soft-edge” shadows (see Figure 6.3 bottom right): the more realistic shadows which shape is correct and which fade off at the edges to represent a penumbra area. In the following, we refer to them as *soft-edge shadows*.

When the shadows are ON, we display two shadows (see Figure 6.3) on the ground in a coherent way: the ground plane being detected during the HoloLens' calibration procedure; the virtual shadows lie on a virtual plane that exactly corresponds to the real ground plane.

To implement our shadows we rely on Unity3D's internal shadow implementation using the “hard shadows” and “soft shadows” options as well as the “Ultra” quality settings for our *hard-edge* and *soft-edge* shadows respectively. While Unity's “hard shadows” are generated using the *very high* resolution parameter, the “soft shadows” are produced in *low resolution*. This creates a blurrier shadow and increases the difference between the two shadow renderings. Regarding the “round shadows” we used the same settings as for the “soft-edge shadows” for which we used a sphere to cast the shadows instead of the cube. The sphere has a diameter equals to that of the cube's side. Participants still see the virtual cube, the sphere was only used as the geometry to cast shadows. More information

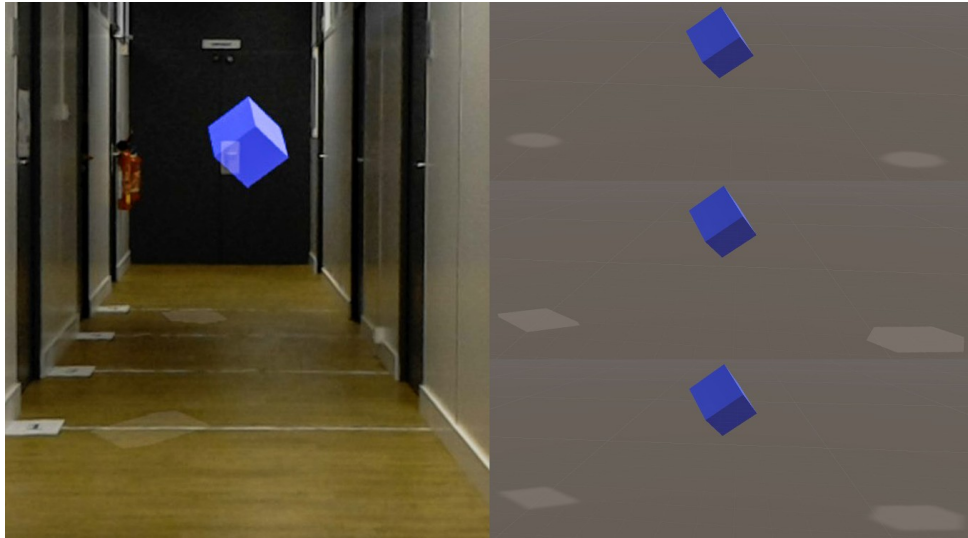


Figure 6.3 – Left: Participant’s viewpoint within the HoloLens in the “hard-shadow” condition. Right: The different shadow conditions, from top to bottom: “Round”, “Hard-Edge” and “Soft-Edge” shadows.

about Unity3D’s shadows is available online<sup>3</sup>.

We therefore had 4 shadows conditions: (i) Shadows OFF: no shadows were displayed in the AR scene; (ii) “Round” Shadows ON; (iii) “Hard-Edge” Shadows ON; (iv) “Soft-Edge” Shadows ON.

The shadows that would have been cast by the two lights of the room are computed in real-time in Unity3D and are consistent with the size and positions of the real lights in the corridor thus the shape, size and location of the virtual shadows vary according to the position of the virtual cube in the corridor.

Regarding virtual shadows and OST-HMDs, one should take into account the following limitation: due to the transparency of the screens, it is impossible to display black in OST-HMDs. Two options were then possible regarding the display of virtual shadows in our AR scene: (i) artificially increase the brightness of the whole AR scene, except for the shadows or (ii) increase the brightness of the shadows only.

Both options present advantages and drawbacks. Option (i) would allow for more realistic shadows since they would appear darker than option (ii). Nevertheless, given the HoloLens’ relatively small FoV, option (i) would give rise to a brighter window within the user’s perspective. Indeed, the screens (i.e. where it is possible to render in the HoloLens) would seem brighter than the rest of the user’s view, e.g. the peripheral view. Option

3. <https://docs.unity3d.com/Manual/ShadowOverview.html>



(ii) on the contrary would keep a coherent brightness between the screens and the “non renderable” part of the HoloLens’ lenses. As a consequence, we chose option (ii), that is to display our virtual shadows in a slightly brighter way.

The experiment was carried out at four different target distances: 6m, 7.5m, 9m and 10.5m. Moreover, the largest distance used in our studies (10.5m) guarantees that participants were still within a relative comfort zone for the HoloLens<sup>4</sup>.

960 trials (20 observers  $\times$  4 types of shadow  $\times$  4 distances  $\times$  3 repetitions) were conducted in the experiment.

### 6.1.3.2 Participants

A total of 22 participants passed the screening step and took part in this experiment (11 males and 11 females). They were between 21 and 31 years of age ( $M = 24.91$ ,  $SD = 2.72$ ). Results of the subjective questionnaires show that 7 participants had no experience with VR or AR, 12 participants had little experience (used glsVR or AR 1-5 times), and 3 participants were experienced users ( $> 5$  times). Each participant completed 48 trials.

### 6.1.3.3 Data Analysis and Results

In order to conduct our analysis, and based on our measurements, we used as dependent variable the relative error of egocentric distance estimation. This is computed as follows:

$$E_{relEgo} = (d_{cube} - d_{target})/d_{target}$$

where  $d_{cube}$  is the distance of the cube estimated by the participant and  $d_{target}$  is the real distance of the target.

We removed 5 participants from the initial dataset: 2 of them misunderstood instructions and thus did not conduct the tasks in a right way, another 2 subjects had abnormally high standard deviations (more than 3 times the standard deviations of other subjects). The last one was removed because of its large number of outliers (12 among 48 trials). Data from the remaining 17 participants were kept for our statistical analysis, using IBM SPSS Statistics 24.0 with a 0.05 significance level.

We first conducted Shapiro-Wilk tests to assess data normality for different levels of factors, which includes four types of shadows (OFF, Round, Hard-edge, Soft-edge) and

---

4. <https://docs.microsoft.com/en-us/windows/mixed-reality/comfort>

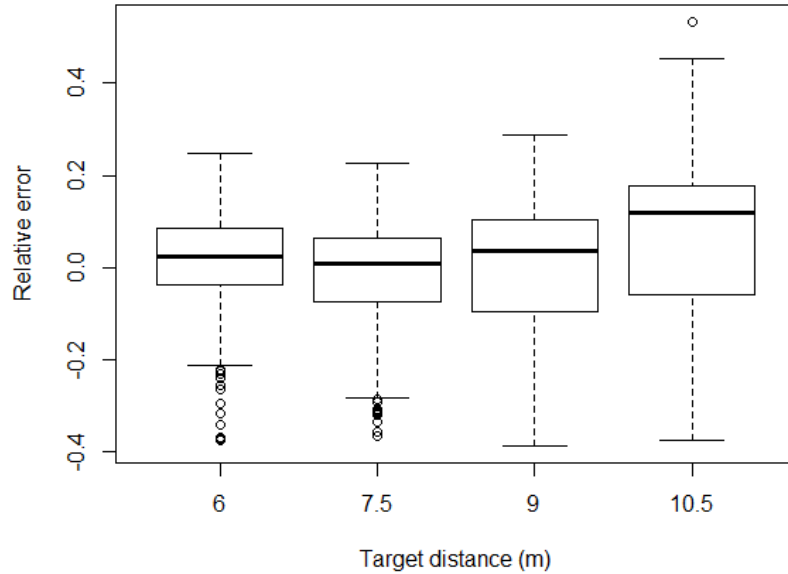


Figure 6.4 – Experiment 1. Relative errors in distance judgments, averaged for all participants grouped per distance condition.

four distances (6m, 7.5m, 9m, 10.5m). The results show that the data does not distributed normally: OFF ( $W = 0.970$ ;  $p < 0.001$ ), Round ( $W = 0.981$ ;  $p = 0.008$ ), Hard-edge ( $W = 0.984$ ;  $p = 0.022$ ), Soft-edge ( $W = 0.976$ ;  $p = 0.002$ ); 6m ( $W = 0.924$ ;  $p < 0.001$ ), 7.5m ( $W = 0.945$ ;  $p < 0.001$ ), 9m ( $W = 0.960$ ;  $p < 0.001$ ), 10.5m ( $W = 0.975$ ;  $p = 0.001$ ).

Thus, we decided to conduct non-parametric tests on the dataset. First, a Kruskal-Wallis H Test was applied. Results showed, as expected, a significant effect of distance ( $H = 52.631$ ;  $p < 0.001$ ) confirming that judgment errors increase along with the distance between the cube and participants, see Figure 6.4. We followed by computing post-hoc tests: all pairwise comparisons with adjusted  $p$ -values using the Bonferroni correction for multiple tests. Significant effects were found between (i) 6m and 10.5m ( $Z = -5.227$ ;  $p < 0.001$ ), (ii) 7.5m and 10.5m ( $Z = -6.951$ ;  $p < 0.001$ ), (iii) 9m and 10.5m ( $Z = -4.442$ ;  $p < 0.001$ ).

As for the effects of shadows, unexpected to us, the Kruskal-Wallis H test showed a non-significant effect of shadow type ( $H = 7.446$ ;  $p = 0.059$ ), see Figure 6.5. This was unexpected since the presence of shadows has been reported to increase accuracy of distance judgments in AR (see e.g. Diaz et al. (2017)). However, since the Kruskal-Wallis'  $p$  value is close to the significance level used in our analysis, we computed Mann-Whitney U Tests to compare data of each different shadow pair.

Results showed that there is significant difference between the Shadow OFF condition

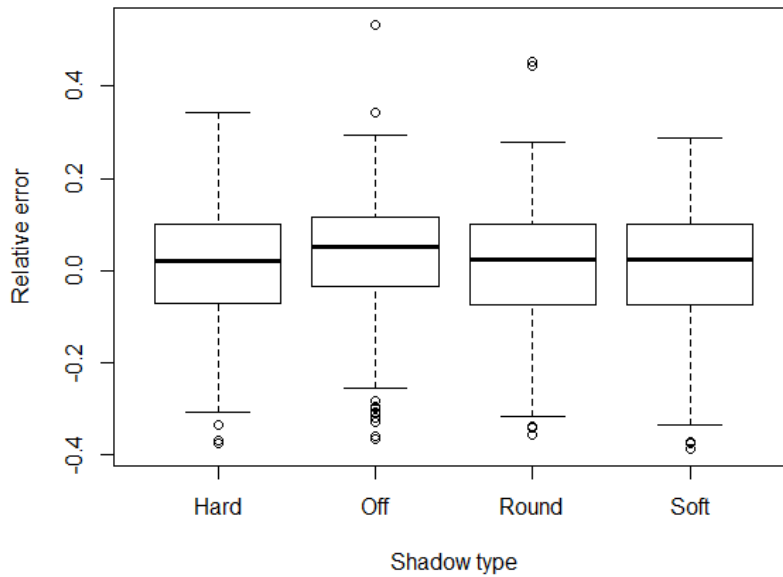


Figure 6.5 – Experiment 1. Relative errors in distance judgments, averaged for all participants grouped per shadow type.

and other conditions (i) OFF and Round-edge ( $Z = -2.064$ ;  $p = 0.039$ ), (ii) OFF and Hard-edge ( $Z = -2.435$ ;  $p = 0.015$ ), (iii) OFF and Soft-edge ( $Z = -2.106$ ;  $p = 0.035$ ), see Figure 6.5. Moreover, results showed a non-significant effect of the three Shadow ON conditions.

In addition, the completion time (i.e. the time it took for each participant to complete the task) has been analyzed for each trial. First, Shapiro-Wilk tests were applied to assess data normality for different levels of factors. Results show that the data is not normally distributed for any of the factors: OFF ( $W = 0.841$ ;  $p < 0.001$ ), Round ( $W = 0.752$ ;  $p < 0.001$ ), Hard-edge ( $W = 0.840$ ;  $p < 0.001$ ), Soft-edge ( $W = 0.869$ ;  $p < 0.001$ ); 6m ( $W = 0.905$ ;  $p < 0.001$ ), 7.5m ( $W = 0.851$ ;  $p < 0.001$ ), 9m ( $W = 0.840$ ;  $p < 0.001$ ), 10.5m ( $W = 0.770$ ;  $p < 0.001$ ).

Then, a non-parametric Kruskal-Wallis H Test was applied to the data of completion time. Results show a significant effect of distance ( $H = 67.520$ ;  $p < 0.001$ ) proving that completion time increases along with the distance between the cube and participants, see Figure 6.6. We followed by computing post-hoc tests: all pairwise comparisons with adjusted  $p$ -values using the Bonferroni correction for multiple tests. Significant effects were found between (i) 6m and 9m ( $Z = -4.488$ ;  $p < 0.001$ ), (ii) 6m and 10.5m ( $Z = -7.078$ ;  $p < 0.001$ ), (iii) 7.5m and 9m ( $Z = -3.886$ ;  $p = 0.001$ ), (iv) 7.5m and 10.5m ( $Z = -6.515$ ;  $p < 0.001$ ).

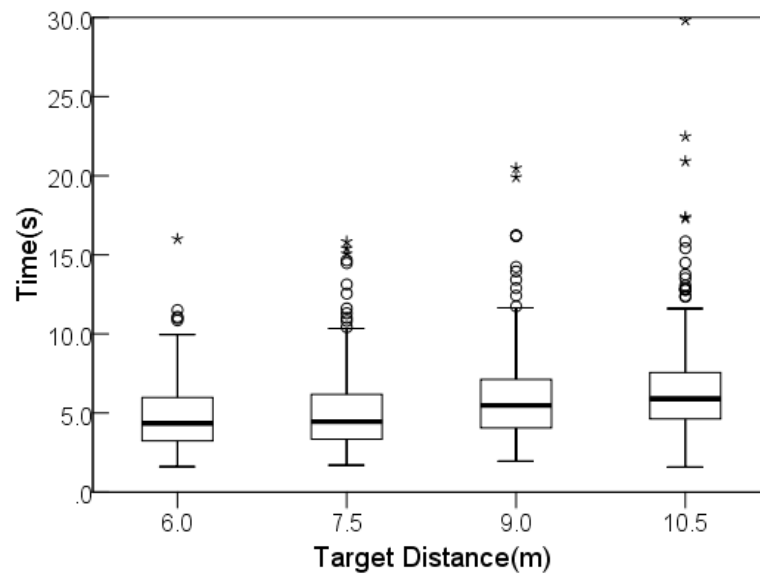


Figure 6.6 – Experiment 1. Completion time in distance judgments, averaged for all participants grouped per target distance.

As for the effects of shadows, the Kruskal-Wallis H test showed a non-significant effect of shadow type ( $H = 1.904$ ;  $p = 0.593$ ), see Figure 6.7. Afterwards, we computed Mann-Whitney U Tests to compare data of each different shadow pair. However, no significant effect was found in the results.

From the results, it can be seen that there is a trend towards underestimation but the high variance levels of our results could not confirm this bias. In order to further analyze this effect, we studied results for each participant individually.

Participants' individual data were observed to confirm the stability of the above findings. As shown in Figure 6.8, 9 out of 17 subjects have positive relative errors, 5 out of 17 subjects have negative relative errors, and 3 out of 17 subjects have both positive and negative relative errors. In other words, the number of subjects who underestimate distances in AR accounts for the largest proportion in experiment 1. This result is consistent with the general trend of the data.

In addition, the mean relative error of each subject under different shadow types was computed. Results show that 11 out of the 17 subjects have higher distances judging accuracy under at least one of the shadows ON conditions compare with the shadows OFF condition (see Table 6.1).

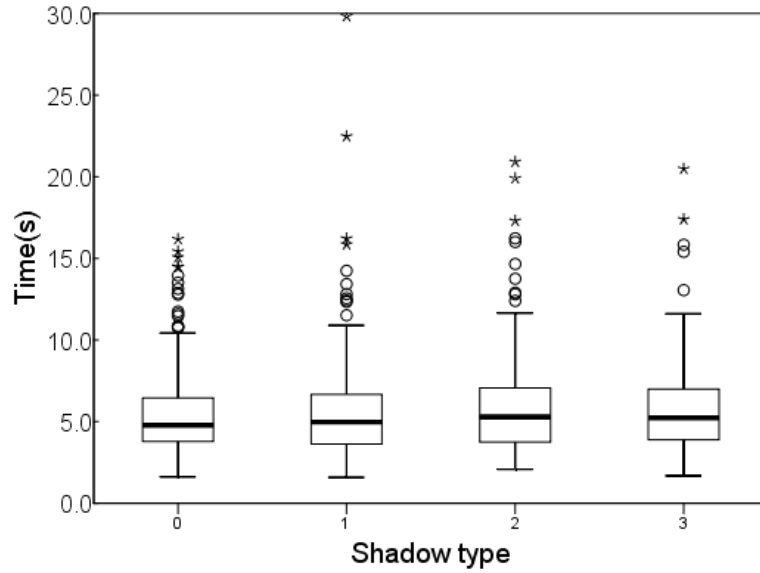


Figure 6.7 – Experiment 1. Completion time in distance judgments, averaged for all participants grouped per shadow type

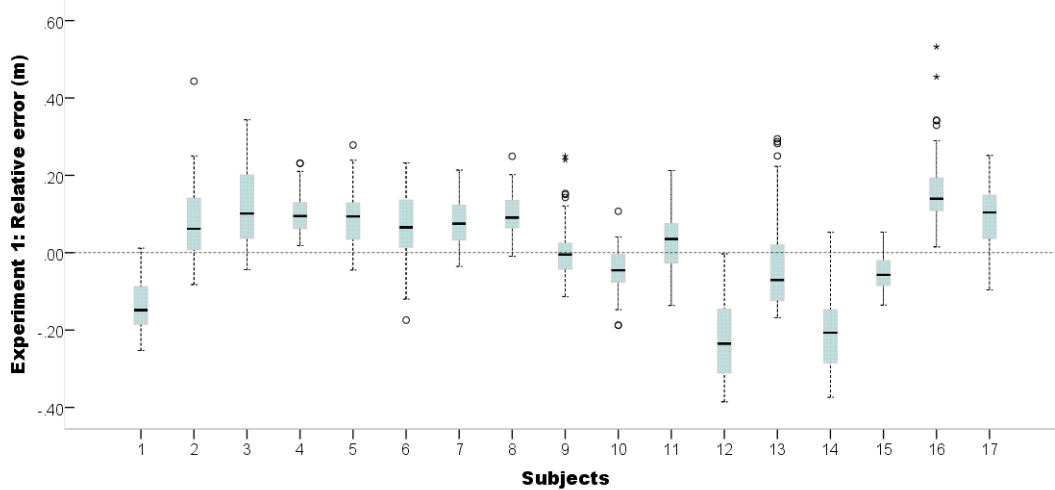


Figure 6.8 – Experiment 1. Individual relative errors in distance judgments.

Table 6.1 – Individual mean relative errors in distance judgments under different shadow types.

Shadows / Subject ID	1	2	3	4	5	6	7	8	9
Off	-.115	.076	.128	.099	.107	.086	.073	.114	.014
Round	-.143	.100	.115	.108	.118	.035	.075	.091	.008
Hard	-.144	.063	.123	.091	.072	.060	.080	.084	.011
Soft	-.144	.060	.118	.108	.075	.076	.081	.095	-.012
Shadows / Subject ID	10	11	12	13	14	15	16	17	-
Off	-.030	.049	-.225	.164	-.177	-.058	.199	.091	-
Round	-.028	.014	-.227	-.085	-.208	-.055	.147	.093	-
Hard	-.062	.019	-.220	-.087	-.226	-.036	.156	.076	-
Soft	-.057	.027	-.235	-.086	-.211	-.054	.156	.112	-

Table 6.2 – Individual mean relative errors in distance judgments under different targets' distances.

Distances / Subject ID	1	2	3	4	5	6	7	8	9
6m	-.184	-.002	.039	.091	.052	.017	.087	.123	-.004
7.5m	-.195	.049	.055	.058	.078	.007	.014	.059	-.039
9m	-.113	.083	.142	.079	.098	.067	.056	.061	-.002
10.5m	-.054	.170	.249	.178	.144	.165	.152	.142	.066
Distances / Subject ID	10	11	12	13	14	15	16	17	-
6m	-.014	.073	-.219	.019	-.216	-.044	.133	.032	-
7.5m	-.065	.007	-.233	-.031	-.223	-.064	.106	.067	-
9m	-.027	-.016	-.239	-.031	-.212	-.074	.123	.114	-
10.5m	-.072	.045	-.216	-.050	-.171	-.022	.298	.159	-

On the other hand, the mean relative error of each subject under different targets' distances was computed. Results show that 14 out of 17 subjects have the highest mean relative errors under distance 10.5m (see Table 6.2).

As for the results of the completion time, we also computed the individual data for each subject. Due to our statistical analysis only found significant distances' effect on completion time, we computed the mean completion time for each subject under different targets' distances (see Table 6.3). The result shows 14 out of 17 subjects have longer mean completion time under the distance of 9m and 10.5m than under the distance of 6m and 7.5m. This result is coherent with the global statistical analysis.

#### 6.1.3.4 Discussion

As mentioned in subsection 3.4.3, many previous studies showed that shadows are very important for human's visual perception, and that they represent an important depth cue

Table 6.3 – Individual mean completion time in distance judgments under different targets’ distances.

Distances / Subject ID	1	2	3	4	5	6	7	8	9
6m	4.184	4.901	5.012	4.472	5.568	3.604	6.043	5.062	5.634
7.5m	5.299	5.319	6.373	4.768	6.433	3.762	6.465	4.854	4.620
9m	5.300	6.176	5.467	5.922	6.857	4.182	8.763	5.794	6.148
10.5m	6.099	5.599	6.455	6.461	7.126	5.036	8.668	6.966	5.610
Distances / Subject ID	10	11	12	13	14	15	16	17	-
6m	3.504	3.534	6.894	3.423	4.242	8.830	3.982	2.807	-
7.5m	3.176	3.125	7.970	3.854	4.028	11.473	4.193	2.557	-
9m	4.624	4.685	8.540	3.857	3.797	13.706	5.510	3.497	-
10.5m	4.227	5.177	10.292	7.202	4.425	15.265	6.338	4.113	-

(Hubona et al., 1999; Mamassian et al., 1998). While it has been showed that shadows can help people confirm objects’ position in a virtual scene (Thompson et al., 2011; Yonas et al., 1978), research on shadows’ influence on depth perception in AR is still limited, even more so in OST AR (Diaz et al., 2017).

The present experiment aimed at investigating the influence of different levels of realism shadows on distance perception in OST AR. First of all, we wanted to test whether the presence or absence of virtual objects’ shadows would improve distance perception accuracy in OST AR (hypothesis **H2**). Our results show that while distance perception is better in the shadows ON condition, no significant difference was found between the three shadows conditions. Participants’ individual data were also observed to confirm this phenomenon. Results show that for most subjects (11 out of 17), their individual data were coherent with this result. We believe it is a very interesting result, since it would tend to show that shadow type does not to play an important role on distance judgment in OST-HMDs, disproving our hypothesis **H3**. This is a meaningful result since application developers could use “cheap” non-realistic shadows, like our round shadows, instead of costly ones, especially since most OST-HMDs use standalone computers with limited CPUs/GPUs.

Finally, regarding hypothesis **H1** (participants underestimate distances in OST AR), we can confirm a trend towards underestimation but high variance levels of our results could not confirm this bias. The relative errors for each individual were also observed, and we found that most of the participants (9 out of 17) underestimate distance in AR, and 3 out of 17 subjects have both positive and negative relative errors. In addition, our results also confirm that errors on perceived distances increase along with the distance

between the virtual object and participants.

As for the completion time of each trial, although the initial position of the virtual cube is random, the statistical analysis reveals that completion time of trials increases significantly along with the distance, however, the shadow type has no significant effect on the completion time of trials. It could be inferred that the participants need to spend more time on matching task when the targets' distances were farther away.

#### 6.1.4 Experiment 2 - Studying the influence of lighting misalignment on distance perception in OST AR

This second experiment aims at assessing hypotheses **H1**: “Participants underestimate egocentric distances in OST AR” and **H4**: “Lighting misalignment between the real and the virtual environments impact negatively depth perception in OST AR”. Similar to experiment 1, details about the apparatus and the task have been given in subsection 6.1.2.2.

##### 6.1.4.1 Experimental conditions

In order to select the type of virtual shadow used in this experiment, we used results from the first experiment. Since we did not want to induce a bias in depth perception, we chose to rely on a realistic type of shadow, namely the *hard-edge* shadow. Thus, all of the trials in this experiment were performed with the shadow ON condition and using the *hard-edge shadow* setting, see Figure 6.3 middle right.

We selected three virtual lights' positions, see Figure 6.9: (i)  $L0m$ : coherent lighting conditions where the positions of virtual lights coincide with the real lights' positions, (ii)  $L+3m$ : virtual lights were positioned 3 meters behind the real lights' positions along the Z-axis, and (iii)  $L-3m$ : virtual lights were positioned 3 meters before the real lights' positions along the Z-axis.

In addition, we also conducted the experiment under four different target distances, as in the first experiment, namely 6m, 7.5m, 9m and 10.5m. All the other settings of the trials as well as the task are similar with those of the first experiment. A total of 756 trials (twenty-one observers  $\times$  three lights' positions  $\times$  four distances three repetitions) were conducted in this experiment.



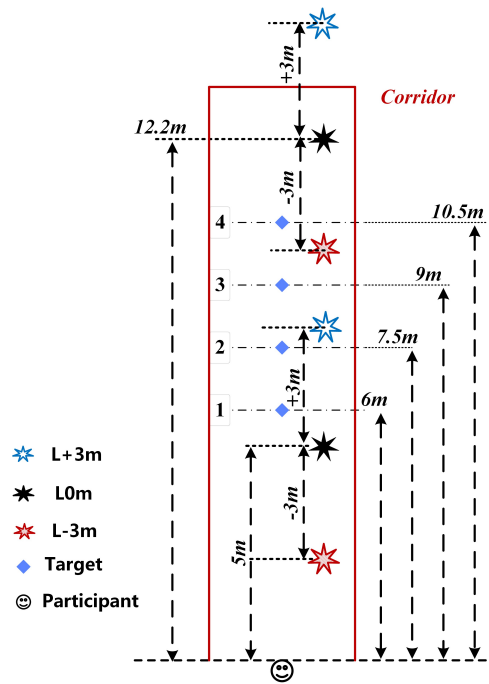


Figure 6.9 – Experiment 2. Lighting conditions.

### 6.1.4.2 Participants

A total of 21 participants (all different from experiment 1) passed the screening step and took part in this experiment (11 males and 10 females). Participants were between 21 and 29 years of age ( $M = 25.9$ ,  $SD = 3.19$ ). Results of the subjective questionnaires showed that 11 participants had no previous experience with VR or AR, 9 participants had little experience (1-5 times) with VR or AR, and 1 participant had extensive experience (> 5 times) with VR or AR. Each participant completed a total of 36 trials.

### 6.1.4.3 Data Analysis and Results

Like experiment 1, the dependent variable of our statistical analysis is the relative error between the cube and the floor reference target.

From the initial dataset, 3 participants were removed: 2 because of a standard deviation more than 5 times higher than the standard deviation of the other participants and one because of an abnormal mean (being more than three times the standard deviation smaller than the global mean). Here again the analysis was performed using IBM SPSS Statistics 24.0 with a 0.05 significance level.

The final data of different light position levels and distance levels was processed by

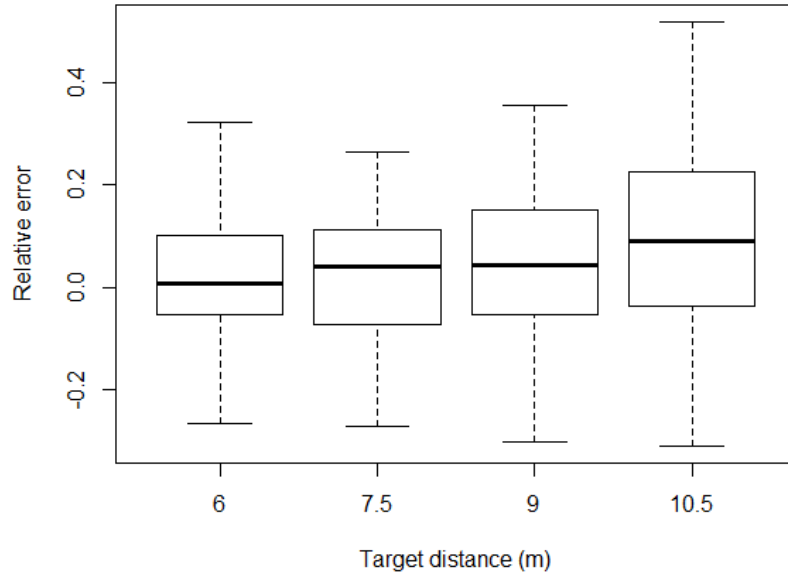


Figure 6.10 – Experiment 2. Relative errors in distance judgments, averaged for all participants grouped per distance condition.

a Shapiro-Wilk test respectively. The results show that some levels of data were not distributed normally: L0 ( $W = 0.978$ ,  $p = 0.002$ ), 7.5m ( $W = 0.978$ ,  $p = 0.011$ ). Thus, non-parametric tests were applied to all of the data.

As in the previous experiment, we first computed a Kruskal-Wallis H Test. Results showed a significant effect of distance ( $H = 22.610$ ;  $p < 0.001$ ) and lights' position ( $H = 15.961$ ;  $p < 0.001$ ).

We thus followed with post-hoc tests: all pairwise comparisons with adjusted  $p$  values using the Bonferroni correction for multiple tests. Results found significant differences between (i) 6m and 10.5m ( $Z = -4.277$ ,  $p < 0.001$ ), (ii) 7.5m and 10.5m ( $Z = -3.873$ ,  $p = 0.001$ ). Other distances groups were not significantly different.

As for lights' positions: pairwise comparisons revealed significant difference between groups (i) L+3m and L-3m ( $W = -3.947$ ,  $p < 0.001$ ) and (ii) L+3m and L0m ( $W = -2.511$ ,  $p = 0.036$ ). Note that no significant difference was found between the L0m and L-3m groups.

Figure 6.10 presents the mean relative errors for all participants under four distance conditions. It can be observed that the mean relative error increases along with the distance of the virtual object.

Regarding hypothesis **H1**, our results are not conclusive. There seems to be a tendency toward a slight underestimation ( $mean = 0.046m$ ), but the high variance levels do not

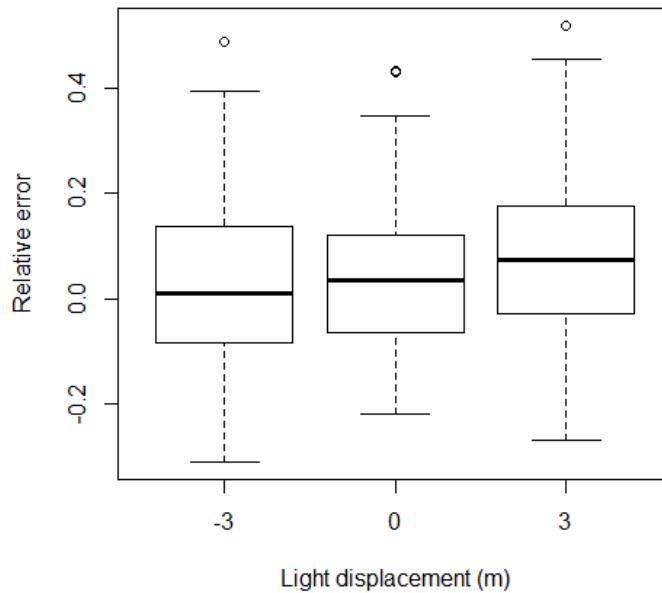


Figure 6.11 – Experiment 2. Relative errors in distance judgments, averaged for all participants grouped per lighting placement.

allow us to conclude on a clear distance underestimation in OST AR no matter the lighting misalignment condition.

Fig. 11 presents the mean relative errors for all participants in distance judgments grouped per lighting placement. The mean relative error has the highest value under the condition L+3m. The condition L0m has a mean relative error slightly higher than that compared with the condition L-3m, however, the condition L0m also corresponds to a smaller standard deviation.

In the same way, the completion time for each trial was analyzed. Results of the Shapiro-Wilk test show that data were not distributed normally: L-3m ( $W = 0.902$ ,  $p < 0.001$ ), L0m ( $W = 0.827$ ,  $p < 0.001$ ), L+3m ( $W = 0.879$ ,  $p < 0.001$ ), 6m ( $W = 0.838$ ,  $p < 0.001$ ), 7.5m ( $W = 0.795$ ,  $p < 0.001$ ), 9m ( $W = 0.874$ ,  $p < 0.001$ ), 10.5m ( $W = 0.909$ ,  $p < 0.001$ ). Thus, a Kruskal-Wallis H Test was applied to all of the data. Results showed a significant effect of distance ( $H = 41.523$ ;  $p < 0.001$ ), see Fig. 12, and non-significant effect of lights' position ( $H = 0.468$ ;  $p = 0.791$ ), see Figure 6.13.

As post-hoc tests for the effect of distance we computed all pairwise comparisons with adjusted  $p$  values using the Bonferroni correction for multiple tests. Results found significant differences between (i) 6m and 9m ( $Z = -3.902$ ,  $p = 0.001$ ), (ii) 6m and 10.5m ( $Z = -5.618$ ,  $p < 0.001$ ), (iii) 7.5m and 9m ( $Z = -3.091$ ,  $p = 0.002$ ), (iv) 7.5m and 10.5m

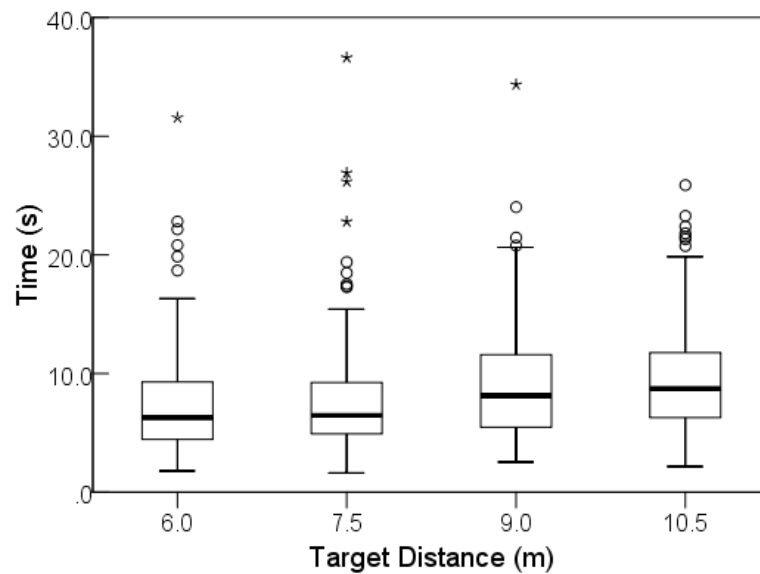


Figure 6.12 – Experiment 2. Completion time in distance judgments, averaged for all participants grouped per target distance

( $Z = -4.807$ ,  $p < 0.001$ ). Other distances groups were not significantly different.

Then, the Mann-Whitney U Test was used to check the differences between different levels of lights position, however no significant result was found.

Similar to experiment 1, we computed the participants' individual data to confirm the stability of our findings. As shown in Figure 6.13, 8 subjects have positive relative errors, 4 subjects have negative relative errors, and 6 of them have both positive and negative relative errors. In other words, the number of subjects who underestimate distances in AR accounts for the largest proportion in experiment 2. This result is consistent with the general trend of the data.

In order to reinforce the findings about the lights placements' effects on distance relative errors, we computed for each subject the mean relative error under different light placements (see Table 6.4). As mentioned above, results show that L+3m corresponds to a significant higher relative error than L0m, while L-3m corresponds to smaller relative error than L0m. From Table 6.4, we can see that for 9 out of the 18 subjects, results were coherent with the general trend. In addition, three subjects (Subject 3, Subject 13, and Subject 16) show the highest distance judging accuracy under the condition L0m. A Shapiro-Wilk test was adopted to assess data normality for relative error of these three

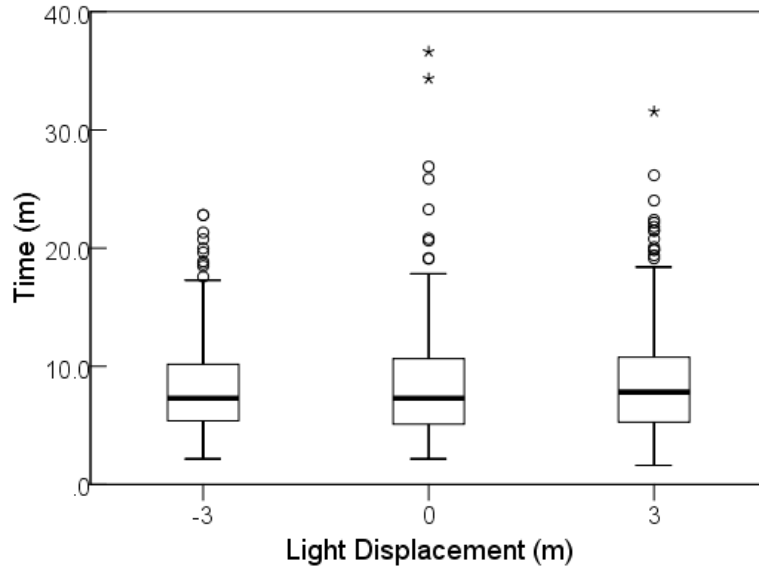


Figure 6.13 – Experiment 2. Completion time in distance judgments, averaged for all participants grouped per lighting placement.

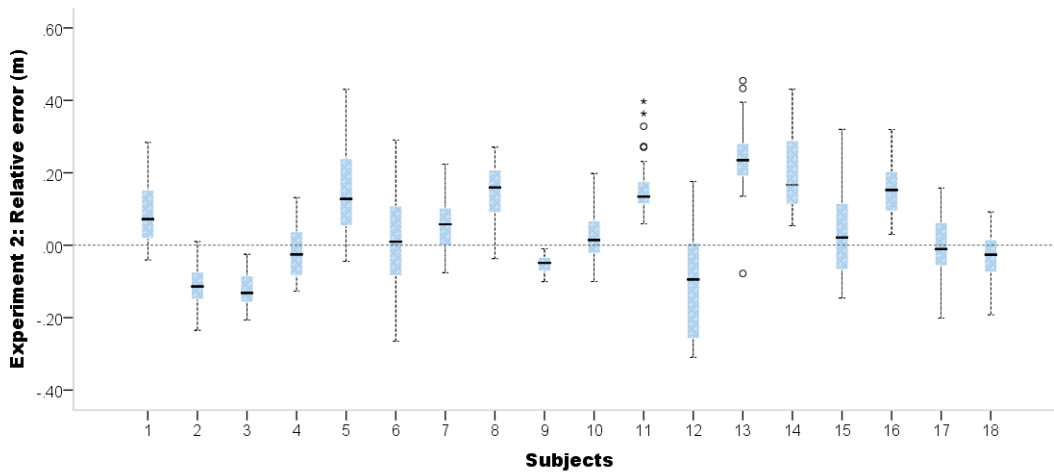


Figure 6.14 – Experiment 2. Subjects individual relative errors.

Table 6.4 – Individual mean relative errors under different lights placements.

Lights / Subject ID	1	2	3	4	5	6	7	8	9
L-3m	.092	-.113	-.144	-.092	.187	.024	.016	.137	-.049
L0m	.067	-.117	-.121	.023	.145	-.006	.065	.119	-.052
L+3m	.105	-.110	-.098	.019	.127	.044	.093	.182	-.057
Lights / Subject ID	10	11	12	13	14	15	16	17	18
L-3m	-.009	.131	-.257	.219	.171	-.090	.140	.022	-.046
L0m	.006	.157	-.098	.239	.186	.039	.169	-.019	-.039
L+3m	.064	.188	.043	.250	.219	.155	.153	-.020	-.008

Table 6.5 – Individual mean relative errors under different targets' distance.

Distances / Subject ID	1	2	3	4	5	6	7	8	9
6m	.002	-.133	-.130	-.003	.036	-.048	-.002	.118	-.034
7.5m	.078	-.137	-.144	-.004	.131	-.049	.056	.125	-.065
9m	.091	-.108	-.115	-.017	.143	.022	.072	.174	-.057
10.5m	.181	-.076	-.095	-.041	.302	.158	.106	.168	-.054
Distances / Subject ID	10	11	12	13	14	15	16	17	18
6m	.026	.113	-.059	.200	.160	.102	.105	-.075	-.046
7.5m	.010	.122	-.098	.185	.138	.047	.097	-.066	-.048
9m	.015	.144	-.155	.261	.210	.016	.172	.017	-.022
10.5m	.030	.256	-.104	.298	.260	-.027	.241	.102	-.009

subjects under different lights placements. The results show that the data of Subject 3 and Subject 13 distributed normally, while the data of Subject 16 under L-3m condition does not distributed normally: L-3m ( $W = 0.847$  ;  $p = 0.034$  ). A paired sample t-test was applied to data of Subject 3 and Subject 13. The result shows that there is no significant difference between L-3m and L0m for these two subjects. Subject 13 exhibit a significant higher relative error under L+3m condition compared to the L0m condition (Subject 13:  $t_{11} = 2.366$ ,  $p = 0.037$ ). In addition, A Mann-Whitney U was adopted to analyze relative errors of Subject 16. The result shows that L+3m exhibits significant higher relative errors than L0m ( $Z = 2.021$ ,  $p = 0.043$ ). This means that most of the subjects' (12 out of 18) individual data were coherent with the general result.

Moreover, we computed the mean relative error under different distances for each subject (see Table 6.5). Results show that 13 out of 18 subjects have the highest mean relative error under the distance of 10.5m. This means most of the subjects' data coherent with the general trends.

Similar to experiment 1, we also computed the individual mean completion time for each subject. Since only distances add a significant effect, to further analyze this effect, we

Table 6.6 – Individual mean completion time in distance judgments under different targets’ distances.

Distances / Subject ID	1	2	3	4	5	6	7	8	9
6m	6.162	5.864	7.287	9.205	8.295	8.125	11.711	3.999	9.144
7.5m	6.041	6.003	9.850	10.107	8.363	14.086	12.263	5.212	10.173
9m	7.172	7.269	11.585	7.186	11.133	14.807	13.424	5.212	13.605
10.5m	10.173	7.971	14.900	8.130	10.209	13.583	14.317	6.365	11.725
Distances / Subject ID	10	11	12	13	14	15	16	17	18
6m	13.283	10.545	5.801	5.346	8.803	8.925	3.507	3.567	5.382
7.5m	10.02	11.133	3.730	5.788	7.000	6.989	4.113	3.992	6.041
9m	10.514	12.934	3.935	6.802	10.015	9.296	4.679	6.344	7.649
10.5m	11.168	13.30	5.317	9.612	9.193	8.931	5.843	8.477	8.144

computed the mean completion time of each subject for each different targets’ distance (see Table 6.6). Results show that 14 out of 18 subjects have longer mean completion times for distances of 9m and 10.5m compared to 6m and 7.5m. This is also coherent with the global statistical analysis.

#### 6.1.4.4 Discussion

The present section is the first study to investigate the influence of the lighting misalignment on distance perception quantitatively in OST AR. As in the first experiment, we found a distance underestimation trend in AR OST-HMDs. However, high variance levels prevent us from confirming it statistically.

As expected, we found a significant influence of lighting coherence’s on distance perception in OST AR. The condition L+3m has a significantly higher distance judgment error than L0m. Although the mean relative error of L-3m is slightly lower than L0m, we did not find any significant difference between these two conditions. In addition, since L-3m leads to a better distance estimation accuracy than L+3m, this could imply that placing the virtual lights closer to the participants might be better than placing them farther away. Individual data were also observed to confirm these results, and we found that 12 out of the 18 subjects’ individual data were coherent with the general result.

Obviously, the direct consequence of changing virtual lights’ locations is the modification of the virtual shadows’ positions. In condition L+3m, the virtual lights were farther away from the participants and as a consequence virtual shadows were closer to them. Since our participants wore an OST-HMD, this created a perceptual conflict. Indeed, in OST AR, participants still see the real lights. Thus participants had to estimate distance

of the virtual cube based on the real lights' positions and the virtual shadows (that were closer to them than they should have been). This could lead participants to believe that objects were closer to them and thus to underestimate the egocentric distance to the virtual cube. In other words, the lights' location L+3m might have increased the degree of distance underestimation in OST AR.

In a similar, but opposite way, condition L-3m should have lessened underestimation of distance perception in OST AR. However, results showed a non-significant lower relative error in condition L-3m compared to L0m (the same relative error is significant between L-3m and L+3m). This means that we did not notice a cancelling, let alone a reduction of egocentric distance estimation in OST AR when moving the virtual lights closer to the participant than in reality.

As for the distances' influence on the relative errors, we found some significant results that distance perception at 10.5m is significantly worse than the condition of 6m and 7.5m. We can also observe a global trend that the mean relative error increases with the distance between the cube and participants. This trend is consistent with the results of the first experiment. Individual data for each subject (see Table 6.4) show that 13 out of 18 subjects have the highest mean relative error at 10.5m, which is coherent with the global trend.

Moreover, our results reveal that trials' completion time is significantly longer under farther targets' distances (9m and 10.5m) compared to nearer targets' distances (6m and 7.5m), while lighting placement has no significant effect on completion time. Participants' individual data were also consistent with these global trends.

### 6.1.5 Discussion

Our first hypothesis was, following some previously reported results, that people underestimate egocentric distances in OST AR. While we found a trend towards underestimation in both experiments, we cannot confirm it statistically due to the high variance levels of our data. This may be due to the task used in our experiments. Indeed, while our matching task (participants had to match the position of a virtual cube with that of a real target) is easier to evaluate and gives rise to very precise results, it may not be the best for egocentric distance estimation. In order to obtain more conclusive results, one should probably rely on other tasks such as verbal report, bisection tasks, blind walking, pointing or even throwing tasks. Nevertheless, it has been shown that tasks can have a strong effect on distance evaluation (Swan et al., 2015) and verbal report is known to



be less accurate. In order to confirm the phenomenon of underestimation in the present study, we computed individual data for each subject in both of the two experiment, and we found that the number of subjects who underestimate distances in AR accounts for the largest proportion in both experiments.

In addition, our research verified that virtual objects shadows improve depth judgment accuracy in OST AR (**H2**), and that different levels of realism for virtual shadows have different influences on distance perception in OST AR (**H3**). Regarding **H2** and **H3**, while our results confirmed that the presence of shadows improve the accuracy of distance perception in OST AR (Mann-Whitney U Test for **H2** showed that the shadows OFF condition has a significant higher relative error than all the other shadows ON conditions). We found an interesting result in that there is no significant difference between the three types of virtual shadows used, rejecting **H3**. Hence, the results show that different realism levels of virtual shadows have no significant influence on the distance perception in OST AR. This is interesting because it implies that in OST AR application developers could rely on “cheap” unrealistic shadows instead of realistic costly virtual shadows. This is especially important since most OST AR HMDs use standalone computers with limited CPUs/GPUs. The study also reveals that errors on perceived distances increase along with the distance between the cube and participants.

Regarding hypothesis **H4**, namely that lighting misalignment between the real and the virtual environments negatively impact depth perception in OST AR, we only found a significant result in condition L+3m where virtual lights were farther away (+3m) from the participants than real lights. In that case, distance perception was significantly underestimated by participants. A possible explanation to this phenomenon is that by moving away virtual lights, the virtual shadows come closer to the participant thus impacting distance estimation and leading to an underestimation. Surprisingly, the opposite effect did not happen in condition L-3m where virtual lights were closer than real lights. One could wonder why the same amount of modification in positions of virtual lights in two opposite directions lead to different effects on egocentric distance perception. A possible reason is that underestimation is a complex phenomenon caused by comprehensive factors in AR on general which have been proved in some previous research. Thus, moving the lights location closer cannot relieve the underestimation significantly. This result is intriguing since it could also mean that OST AR environments have a higher tolerance to lighting misalignment when the virtual lights are closer to the participants than when they are farther away from them.

In addition to the above hypotheses, a global trend was observed in both experiments that the mean relative error increases with the distance between the cube and participants. Results show that relative errors at 10.5m are significantly higher than those under other target's distances (6m, 7.5m and 9m) in experiment 1, while the distance of 10.5m corresponds to significant higher relative errors than the distance of 6m and 7.5m in experiment 2. Our results also reveal that completion time of trials increases along with the distance between the virtual cube and the participants, while the shadow type and lights placement have no significant effect on completion time.

### 6.1.6 Conclusion

This section focused on how virtual objects' shadows as well as coherence between virtual and real lighting influences distance perception in OST AR. Two experiments were designed with an OST HMD in order to elicit the impact of these factors. Participants had to match the position of a virtual object displayed in the OST-HMD with a real target. Distance judgment accuracy was recorded under the different shadows and lighting conditions.

The results showed that virtual objects' shadows improve distance judgment accuracy in OST AR and that different levels of lighting misalignment between real and virtual lights have different influences on distance perception in OST AR scenes. Surprisingly, no impact of the shape of virtual shadows on distance judgment accuracy was reported.

Thus, the use of shadows can be underlined as a strong depth cue in AR, with a correct position of the lighting sources being an important factor. However, the specific shape of the rendered shadows seems to be less important than its actual position.

## 6.2 Studying the influence of accommodation distance through the use of Retinal Projection Displays

The inaccurate accommodation distance provided by the OST HMDs has been reported to be the main factor of distance misperception in AR with such displays (Singh et al., 2018). However, new AR devices that do not use fixed optics to provide virtual images are emerging. Some of these devices rely on retinal laser projection and have a almost infinite depth of focus. As such, Retinal Projection Displays provide no accommodation cue. As a consequence, one may wonder whether this lack of accommodation cue

might thus be beneficial for distance perception compared to incorrect accommodation distance.

In this section, we present an experiment which compares distance estimation with real and virtual objects, using an OST HMD and a RPD. As such, we evaluate the influence of accommodation on the perception of distances and also aims to answer the question: Is the absence of accommodation cue better than a wrong one?

## 6.2.1 Methods

We conducted a user experiment to evaluate the egocentric distance perception in RPDs compared with OST HMDs. Participants had to perform a blind-reaching task; they had to align their finger with a target without seeing their hand. Two groups were involved in this experiment. The first group was wearing a RPD while the other group used an OST HMD. In each group, participants had to perform the blind reaching task with both virtual and real targets in order to compare their egocentric distance perception in AR with real perception. In the following, the apparatus, the protocol, as well as the experimental design are presented.

### 6.2.1.1 Participants

Twenty one right-handed participants took part in the study (18 males, 3 females, mean age: 24.0 (SD: 3.0), range: 21–34 years). The eye dominance was assessed using the method proposed by Durand and Gould (1910) in which participants were asked to look at a distant object through a hole made with their two hands. Five participants had a dominant left eye and 16 had a dominant right eye. Participants' IPD was measured using the technique detailed in Willemsen et al. (2008), ranged between 48 and 69 mm (mean: 59.5, SD: 5.76). Participants were students of the university or members of the laboratory and naive about the purpose of the experiment. They were rewarded 1,000 JPY for their participation. All participants had normal or corrected-to-normal vision, 8 were wearing glasses. Nineteen out of 21 had already used a VR HMD before and 13 of them had already used an OST AR HMD before.

This study was conducted in accordance with principles as stated in the declaration of Helsinki. The experiment was approved by the ethical committee of the Tokyo Institute of Technology.

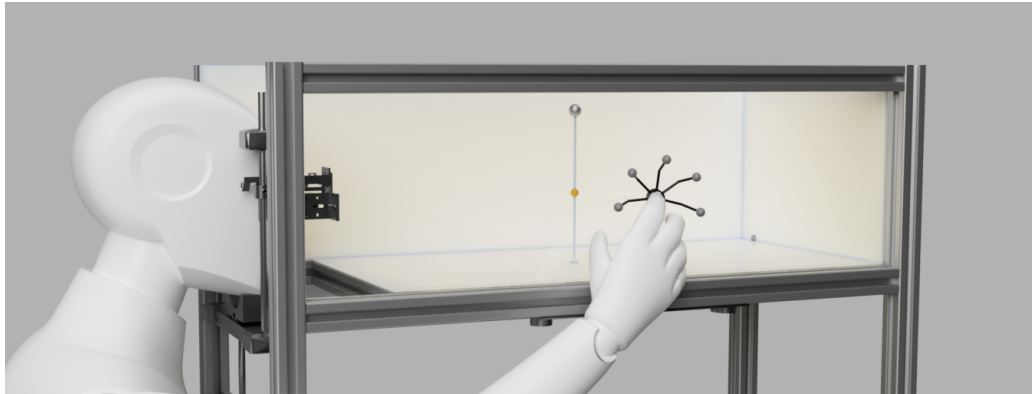


Figure 6.15 – Overview of the experimental setup. The participant had his head on a chinrest and was able to see virtual elements with an OST device or a RPD. He was asked to point a target (virtual or real) with his finger (wearing a tracked ring). The frame was covered by white panels to prevent the participant from seeing his hand. (Here, one panel is transparent for illustration purpose).

#### 6.2.1.2 Apparatus

**Overall structure and tracking system.** A structure was built (30 cm wide, 65 cm long and 18 cm high, see Figure 6.16) and covered by white cardboard panels to prevent the use of visual cues from the background. A moving part carrying a magnet was placed beneath the “floor” cardboard panel. Another magnet, fixed to the base of the white pole, allowed the experimenter to place the target at a specific distance (see Figure 6.17). The absolute position of the white pole was recorded using the reflective ball located at the top of the white target pole. The average reported placement error was 5 mm in depth which provides a sufficient precision for the placement since the following analysis is made using the veridical position. The participant was wearing a ring carrying a 5-sphere motion tracking constellation to measure the position of his finger tip (see Figure 6.16). The motion tracking system was a 6-camera OptiTrack system. Under the experimental conditions, the precision of the system once calibrated was always below 0.1 mm. The headset used to display the virtual images (either a RPD or an OST HMD) was fixed on a chin rest. The height of the chair and the height of the chin rest were adjusted at the beginning of the experiment to maintain the same eye level for every participant (i.e. for each participant, the sphere was always visible in the center of their field of view). See Figure 6.15 for a global view of the experimental setup.

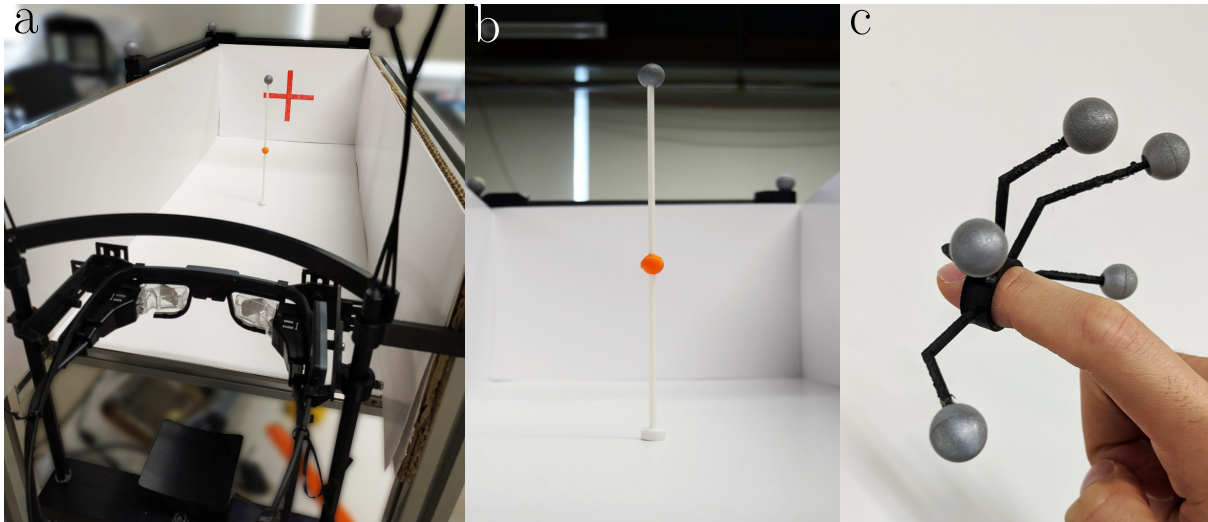


Figure 6.16 – (a) Overview of the experimental system, here with the RPD. The red cross is used only for the calibration and removed during the experiment itself. (b) The real pole with the orange sphere used as target with the reflective tracking marker on the top. (c) The 3D printed ring used to support 5 reflective tracking markers and provide the position of the fingertip.

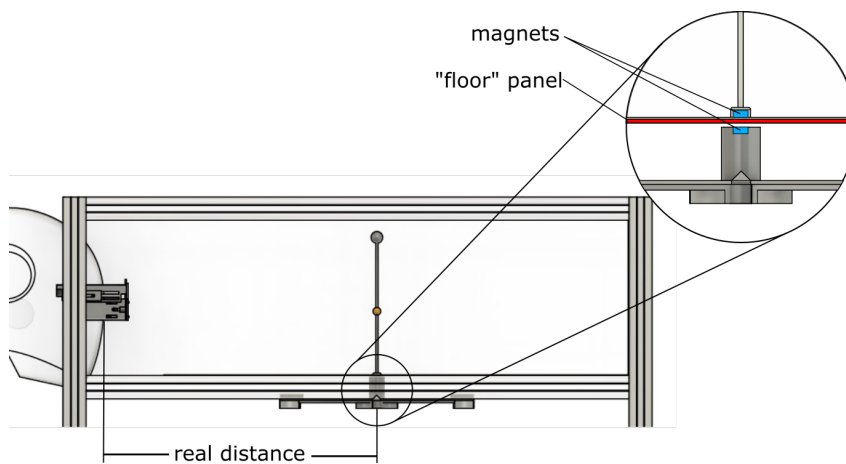


Figure 6.17 – Diagram of the moving element under the structure which allows the experimenter to move the target using two magnets.



Figure 6.18 – Close-up of the custom binocular headset with its two RPD modules mounted on an eye tracking device.

**Real and virtual target.** The real stimulus was an orange 3D-printed sphere with a 10 mm radius, placed on a 1.5 mm radius white pole. The center of the ball was at 85 mm height above the frame floor, so as to be at eye height (see Figure 6.16). The white pole continues above the 3D-printed orange sphere to support an OptiTrack motion tracking ball marker, placed at 180 mm high (see Figure 6.16). Given the characteristics of the experimental setup, the base of the pole and the motion tracking marker were out of the participant’s sight. The virtual stimulus was designed to reproduce the appearance of the real stimulus, including the pole. Colors, lightning and shading were carefully chosen to be as close as possible to the real target.

**Rendering displays.** One group of participants used the Epson Moverio BT-30C AR glasses<sup>5</sup>. The Moverio headset has a  $22.8^\circ \times 12.8^\circ$  FoV with a resolution of  $1280 \times 720$  pixels per eye.

The other group of participants was using a custom mounted binocular RPD (see Figure 6.18). This headset consisted in two QD Laser RETISSA Displays, one for each eye. Each display module had a resolution of  $1024 \times 600$  pixels and an approximate horizontal FoV of  $20^\circ$ . These two modules have been assembled directly on Tobii Pro Glasses 2<sup>6</sup> and the whole unit can be worn directly on the head. To the best of our knowledge, this custom HMD is the first reported prototype of a binocular AR headset using RPD technology and capable of eye tracking. However, like the OST headset, this headset was rigidly mounted on the structure to ensure optimal alignment. In addition, eye tracking

---

5. <https://epson.com/moverio-bt30c-usb-c-compatible-smart-glasses>

6. <https://www.tobii.com/>

data was not collected during this experiment. The positioning of both RPD modules ensured that the stimuli was always in the FoV for both eyes. Moreover, each module has to project the image onto the observer’s retina and was manually positioned and rotated along every axis to ensure the best viewing conditions. This adjustment was made at the beginning of the experiment (during the installation process) for each participant. Note that participants kept wearing the AR device (the OST or the RPD) at all time during the experiment. This means that whenever a real object was showed to them, it was seen through the AR device.

The AR scene was built using Unity3D v2019.1.8f1. The spheres and their support were displayed using the standard Unity shader, they were lit with a single frontal directional light and without texture. This provided a coherent ambient and specular lightning with no visible shadow (since the light was frontal).

**Calibration procedure.** The calibration process was divided into several steps. Before the experiment, the FoV of the AR device was assessed during an iterative calibration process on an optical bench. Since both AR displays are always the same for every participants, the FoV of each device remains the same for the whole study. The other parameters of the calibration (rotation and translation) are adjusted for each eye separately by the participant. This was achieved by aligning, in each eye display, a virtual cross with a real calibration cross drawn on the “back” panel of the setup (see Figure 6.16).

At the end of this phase, the calibration was checked by visually assessing the alignment between a real and a virtual object placed at the same location. The calibration cross was removed after the calibration procedure and the participant was able to see the target as depicted in Figure 6.16.b.

### 6.2.1.3 Task and Procedure

Participants started by reading and filling out a short form containing written instructions about the experiment and giving their consent. After some verbal explanations, they carried out the calibration procedure as described above. Then, participants performed one trial using a real target and one trial using a virtual target to get used to the experimental procedure. Finally, participants had to perform four blocks of 16 trials. Each block consisted in 8 trials with a real target and 8 trials with a virtual target. In each group of participants, half of them started with the virtual target and the other half started with the real target. Participants were allowed to have a break after the two first blocks.

Table 6.7 – Summary of Independent and Dependent Variables.

Independent Variables				
	<i>observers</i>	21	(random variable)	
<b>C1</b>	<i>distance</i>	8	pseudo-random, from 30 to 50 cm	
<b>C2</b>	<i>object type</i>	2	virtual or real object	(within subject)
<b>C3</b>	<i>device</i>	2	OST HMD or RPD	(between subject)
	<i>repetition</i>	4		
Dependent Variables				
<b>D1</b>	<i>reported distance</i>	Projected distance between the eye and the fingertip of the participant along the depth axis		

Each trial was performed as follows: the target was displayed in front of the participant at a certain depth. If the target was virtual, it was directly displayed by the AR device. If the target was real, the experimenter moved the part under the frame to place the magnet at the chosen distance and then placed the real target over it at the correct location thanks to the magnetic guidance. The participant was then asked, without seeing his hand, to place his fingertip on the side of the setup, so as to align it with the target. They were free to use either hand but all our participants were right-handed and as such used their right hand. The target remained visible during this step. When the participant considered the alignment as correct, he reported it to the experimenter and the depth answer of the participant was recorded. Finally, the virtual target disappeared or the real target was taken off. The experiment was not time-constrained and the participant was able to observe the target for as long as necessary. On average, one trial lasted 8.3 seconds.

The total average time per participant was 45 min, including instructions, calibration, experiment, breaks, post-questionnaire and debriefing. Participants wore the AR displays for ~30 min.

#### 6.2.1.4 Experimental Design

We used a mixed-model design with the following conditions:

- **C1**: The distance between the participant’s eye and the target. This distance was pseudo-randomly picked between 30 and 50 cm. For a given trial, the total range was divided into eight segments and one distance was randomly picked inside each segment.
- **C2**: The type of the object: real or virtual.



- **C3**: The device used: RPD or OST HMD. Each participant performed the experiment with only one device.

The distance between the participant and the target (**C1**) was chosen to be in the near distance field and easily reachable to ensure the precision of the reporting technique. The closest distance was limited by the specificity of the AR devices used. The size of the target and the dimensions of all the surrounding elements were chosen to ensure that the target was always fully visible by both eyes and was the only element which could be seen by the participant during the experiment.

In summary, participants were presented with 64 trials: 8 distances (**C1**)  $\times$  2 target types (**C2**)  $\times$  4 repetitions. For each block, the order of **C1** was randomized and the first type of object (**C2**) was counterbalanced among participants. The participants were divided in two groups according to the AR device they were using (**C3**). The only dependent variable was the reported distance **D1**. Eventually participants rated the difficulty of the task after achieving all trials.

#### 6.2.1.5 Research Hypotheses

The key element of this experimental study is the egocentric distance perception of virtual elements, compared to real element distance perception. Given the literature, consistent results assess that, while using an OST HMD, egocentric distances of virtual objects are usually overestimated when compared with real objects' perception (Singh et al., 2018; Swan et al., 2015).

In this experiment, two types of rendering are compared. Their specificities rely on the accommodation cue they are providing. OST HMDs provide a fixed but incorrect accommodation cue which is suspected to be partially responsible for the depth overestimation in AR. On the other hand, RPDs have no focal plane since the image is always sharp and then the accommodation cue is lacking. However, since the eye has to accommodate at a certain distance, the accommodation demand of the eye is expected to be driven by vergence, through the accommodation-vergence reflex. As such, the accuracy of distance estimation of virtual objects is expected to be better with RPDs than with OST HMDs.

However, since this accommodation cue is lacking with RPDs the overall depth perception performance may be reduced. Indeed, the focal plane of OST HMDs is incorrect but fixed, which usually drive the depth overestimation by a fixed amount (Swan et al., 2015). However, the lack of focal plane for RPDs could confuse the viewer, leading to a distance more difficult to evaluate. Distance estimation is then expected to be less precise

with RPDs, i.e. with a larger spreading of the results.

Considering this forewords, our main research hypotheses were the following:

- **H1**: Egocentric distance perception is overestimated for virtual targets compared with real targets when seen with an OST HMD.
- **H2**: Egocentric distance perception in AR is more accurate (i.e. has a smaller bias) with RPDs than with OST devices.
- **H3**: The precision of egocentric distance estimation is better (i.e. has a smaller variability) with OST devices compared to RPDs.

#### 6.2.1.6 Statistical Analysis

The reported distance **D1** was analysed using a multiple linear regression analysis. The model included the independent variables **C1**, **C2** and **C3** as fixed effect and the *participant* as a random factor.

Since **C1** is continuous, a multiple linear regression approach allows to evaluate the depth estimation as a continuous variable, along another continuous variable (the veridical distance) and to highlight the differences across **C2** and **C3** which are categorical. Contrary to an ANOVA, a Linear Mixed-Effects (LME) analysis is able to consider **C1** as a continuous variable and then provides slopes and intercepts for each condition. This approach provides more information about the evolution of depth perception along the distances considered (Pedhazur, 1975; Swan et al., 2015).

For each participant and condition a linear regression was also computed. This supplementary element provides additional information at an individual level and also allows to evaluate the precision for each participant. Usually, the Sum of Squared Estimate of Errors (SSE) is used to represents the part of the dispersion not explained by the model. However, the SSE is dependent to the number of data points in each group. Here, not all groups have the same number of participants and therefore the same number of data points. So, for virtual and real targets for each participant, the Mean Squared Error (MSE) of the residuals is used instead to evaluate the precision of the participants. This result was analysed using a two-way ANOVA with the *device* (**C3**) considered as a between-subject factor and the *object type* (**C2**) of the target as a within-subject factor. Finally, the results of the Likert-scale estimation of the difficulty of the task was analysed. A Wilcoxon-Mann-Whitney test was used since the results did not pass the requirement for a parametric test.

## 6.2.2 Results

Overall, distances were underestimated by 20.3% on average. This finding is coherent with results reported in the literature for depth estimation with the chosen answering protocol (Swan et al., 2015).

**Effects of object type and device.** When analyzing the results, the LME model showed a significant main effect for *object type* of the target (**C2**) on the reported distance ( $F_{(1,1314)} = 186.56, p < 0.0001$ ). On average, distances of virtual objects were overestimated by 9.3%. The overestimation of virtual objects at near-field distances is a well documented effect and one of the motivation of the present study. Here, this effect is in the same magnitude (though close to the upper bound) as those reported in previous studies. As expected, the real distance **C1** had a significant effect on the reported distance ( $F_{(1,1314)} = 1064.93, p < 0.0001$ ). The *device* itself (**C3**) was found not significant.

The interaction effect *device*  $\times$  *object type* was found significant ( $F_{(1,1314)} = 96.25, p < 0.0001$ ). This implies that the overestimation of the virtual target position, compared to the real case, is different according to the device used. Finally, the interaction effect *object type*  $\times$  *position* was also found significant ( $F_{(1,1314)} = 26.05, p < 0.0001$ ) which shows that the overestimation of distances of virtual objects varies according to the position of the object, i.e. the slope of the linear regression is different if the object is real or virtual. The other interaction effects were non-significant.

To summarize, when comparing virtual and real objects, virtual objects are perceived 5.9 cm farther at 30 cm and 2.7 cm farther at 50 cm when seen through OST devices, while this bias is of only 2.4 cm at 30 cm and -1.2 cm at 50 cm with a RPD device. For each group, the global regressions are presented in Figure 6.19 and the individual regressions in Figure 6.20.

**Precision and Consistency.** To evaluate the effect of the device on the precision of the observers reported distance, we consider the quadratic means of the residuals of the linear regression (see subsection 6.2.1.6). While the *object type* of the objects significantly affects the precision of the results ( $F_{(1,19)} = 25.08, p < 0.0001$ ), the device itself has no effect ( $F_{(1,19)} = 0.89, p = 0.45$ ). The interaction term was also not significant. This result shows that the distance evaluation of virtual objects is less precise than the evaluation of real objects. Since the depth cues provided by the HMDs are not as precise and coherent as the ones provided by the real world, the depth estimation is usually altered, which

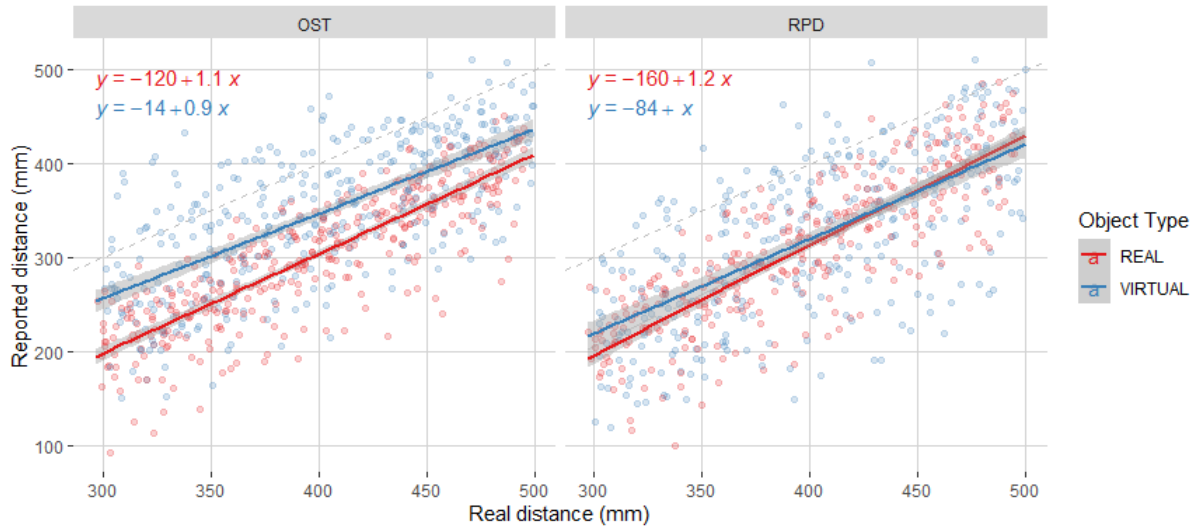


Figure 6.19 – Reported distance vs. real distance for each object type and device, with linear regression for each group.

leads to this decrease in precision. However, the type of the *device* does not affect the precision of the evaluation. This result shows that RPDs' specific way of rendering virtual objects of does not significantly affect the precision of the depth estimation. In particular, the lack of accommodation cues does not provide a more confounding depth estimation compared to usual OST renderings.

Moreover, participants were asked to evaluate the difficulty of the task, using a Likert-scale ranging from 1 to 7. There was no significant difference ( $W = 39.5, p = 0.28$ ) between the results of participants who performed the task on the RPD ( $M = 4.30, SD = 1.49$ ) and those who used the OST device ( $M = 3.64, SD = 1.36$ ). Then, while each technology has its specific rendering, no difference of task difficulty is reported by users when performing a depth estimation task.

To summarize, results showed no difference in distance estimation difficulty between RPDs and OST devices, whether in terms of user perception or objective results.

### 6.2.3 Discussion

At first, our results show a global underestimation of distances for every condition, when compared with the veridical distance. Our setup is very similar to that of Singh et al. (Singh et al., 2012) which also reported an underestimation of distance on a reaching task but with a better accuracy. Moreover, this result is coherent with the effort-based

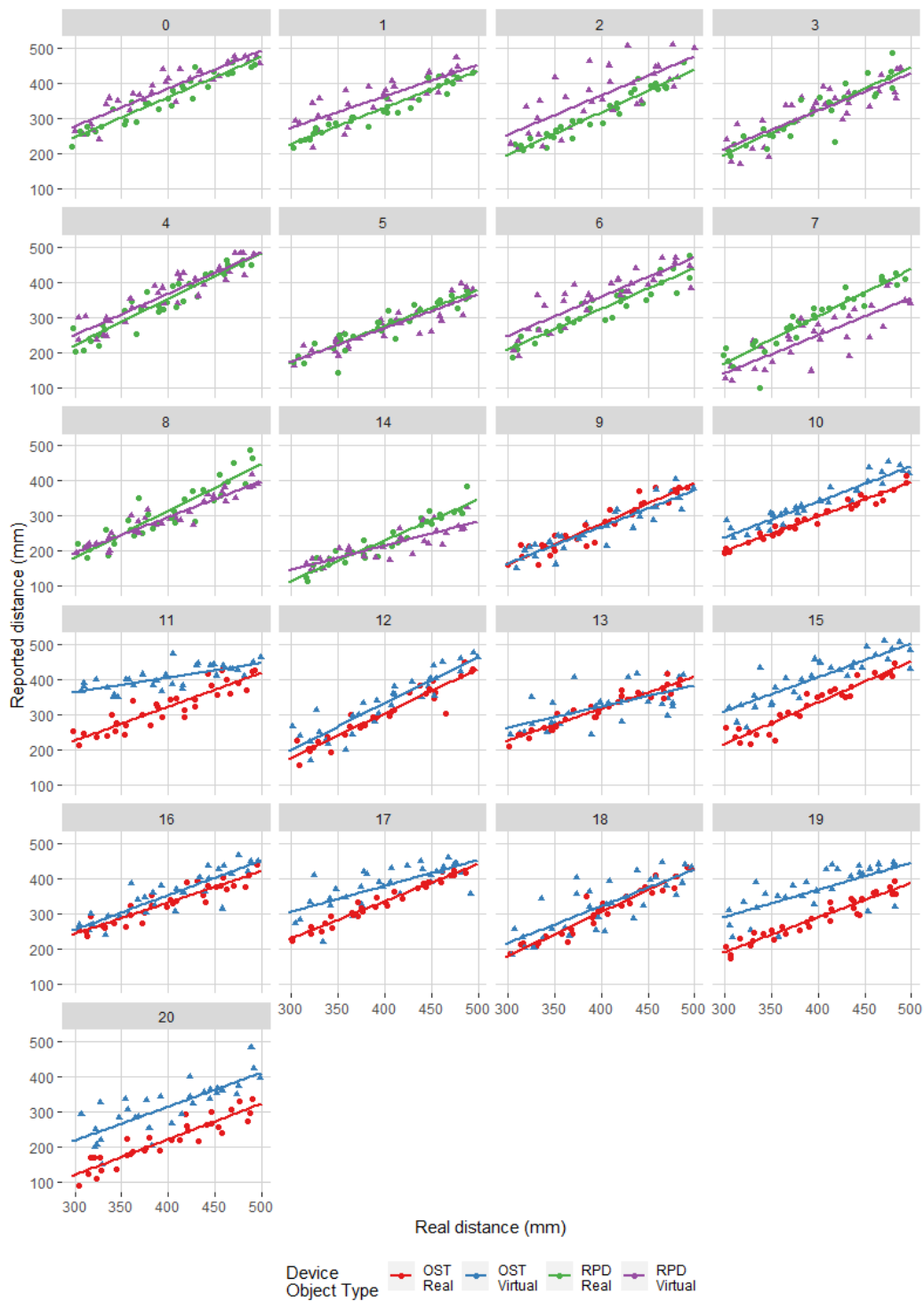


Figure 6.20 – Reported distance vs. real distance for each participant, with linear regression for both object type.

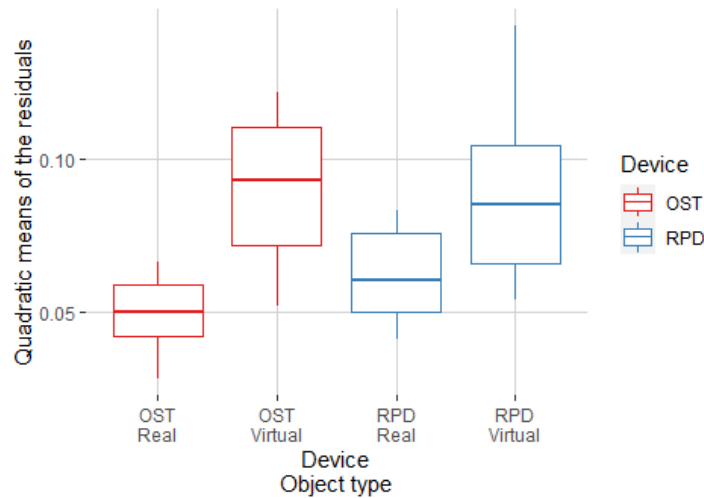


Figure 6.21 – Quadratic means of the residuals of the linear regression for each object type for each participant, aggregated by object type and device used.

theory of distance perception. Indeed, the resting position of the arm is closer to the body of the participants than to the field where the objects were displayed. As such, participants had to extend their arm to point at the object and, since no visual corrective feedback was provided, the effort needed to perform the task may induce the observed underestimation (Berkeley, 1709; Witt et al., 2004). Interestingly, we didn't find any difference between RPD and OST group regarding the real distance estimation. While the two groups were seeing the real objects through specific headsets, with different fields of view, bulkiness and in-built lenses, this did not led to a bias for the real object perception. Their answers remains underestimated by around 24% compared with the veridical distance, whichever was the device worn. This is a strong evidence that the overall underestimation observed in this study is conveyed by the nature of the reaching task and not by the device itself. In addition, the consistency of the answer among groups, for real object, is also an argument in favour of the use of this method for highlighting biases caused by virtual objects.

The results of our study also show a significant reduction of the overestimation bias of virtual images with RPDs, compared to OST devices. From an average 15% overestimation of virtual objects seen with the OST HMD compared to real estimation, this bias is reduced to only 3% with the RPD. This results are also supported by the individual data. Previous research underlined the importance of accommodation cue for distance estimation in OST displays (Singh et al., 2018). In our study, the focal distance of the Moverio BT-30C was 2.5 meters. As such, since the tested distances were under 50 cm, the overestimation

of virtual objects with the OST device can be explained by its bigger accommodation distance. It is still unclear if this effect is a direct consequence of the VAC or a specific bias which allow the accommodation cue to influence the distance estimation.

However, on the other hand, we did not observed this overestimation bias with RPDs. Since these devices does not provide any accommodation cue, the usual bias inherent in OST devices seams to be drastically reduced. When the observer were performing the estimation task, two scenarios can occur: (1) the accommodation cue can be perceived as missing or (2) the accommodation kept being driven by the vergence and as such perceived as always correct. While we cannot assess strongly for one option or the other given our results, participants never reported any discomfort regarding accommodation while wearing the RPD, during the experiment or as a comment in the following questionnaire. Thus, there is no obvious discomfort with accommodation when participants use RPDs, but it is unclear whether this is the main reason why they performed better for the estimation task.

Finally, the results of the precision and consistency of the answers, as well as the perceived difficulty of the task shed an interesting light on these results. Indeed, the VAC should not occur when using a RPD since the image is always sharp and then does not conflict with vergence. But, the VAC is known to induce fatigue, discomfort and overall making the distance estimation task more difficult (D. M. Hoffman et al., 2008). So, if the VAC was responsible for the overestimation reported for OST devices, we should report a decrease in precision, with more spreading of the data, for the OST group (compared to the RPD), which we did not. Those results suggest that the overestimation reported for OST devices is more likely to rely on a specific bias induced by accommodation, rather than being an effect of the Vergence-Accommodation Conflict (VAC).

On the other hand, we might have expected that the lack of a accommodation cue with RPDs would make the task more difficult for the observers. However, it turns out that participants not only gave more accurate answers, but also as precise ones. These results suggest that the absence of the accommodation depth cue is not specifically impacting in this case and that to have no accommodation cue at all is better than an incorrect accommodation cue.

## 6.2.4 Conclusion

In this section, we studied the use of an emerging technology for AR displays to tackle accommodation-related perception issues: Retinal Projection Display. With RPDs,

virtual images have no focal distance and the AR content is always in focus. We conducted the first reported experiment evaluating egocentric distance perception of observers using RPDs. We compared the precision and accuracy of depth estimation between real and virtual targets, displayed by either OST HMDs or RPDs.

Interestingly, our results showed that RPDs significantly improve depth perception in AR compared to OST HMDs. Indeed, the use of an OST device was found to overestimate the perceived distance by 16%, whereas the distance overestimation bias drops to 4% with RPDs. Besides, the task was reported with the same level of difficulty and no difference in precision.

The main difference between OST HMDs and RPDs is the accommodation depth cue they are able to provide. While OST HMDs have a fixed yet incorrect accommodation distance, RPDs provide an always sharp image and no accommodation cue. As such, our results suggest that no accommodation cue leads to a better distance estimation than an incorrect accommodation cue.





PART III

# Prospects and Discussion

---



Any interaction with an object (as here when grabbing a go stone) involves a sensory-motor loop and a constant dialogue between perception and action. In this chapter, we present an experimental protocol that exploits this sensorimotor loop in order to evaluate the impact of interaction techniques on distance perception in Augmented Reality.

# PROSPECTS – HOW TO COMPENSATE PERCEPTUAL BIASES? THE USE OF INTERACTION TECHNIQUES

---

## Contents

---

<b>7.1</b>	<b>Theoretical Background and Research Hypotheses . . . . .</b>	<b>180</b>
<b>7.2</b>	<b>Experimental Protocol Proposal . . . . .</b>	<b>181</b>
7.2.1	Apparatus . . . . .	181
7.2.2	Calibration . . . . .	182
7.2.3	Task and Procedure . . . . .	183
7.2.4	Experimental Design . . . . .	185
<b>7.3</b>	<b>Discussion . . . . .</b>	<b>186</b>

---

In this thesis, we focused our work on some specific factors in the “perception & recognition” steps of the perceptual process (see subsection 1.1.1 for more details). Thus, the next step of this process is “action”. Due to the interaction loop, action can have a major impact on the perceptual process. For some authors, perception can even be defined solely as a mean to anticipate action (J. J. Gibson, 1979).

On the other hand, interaction, and more specifically interaction techniques, remain an important issue in AR. Due to the mixed nature of augmented environments, interaction in AR combines issues of VR environments with some specific issues of AR.

In this section, we present an experimental protocol that aims at evaluating the effect of three different interaction techniques on distance perception in AR. This experiment was expected to be conducted in March/April 2020. Unfortunately, due to the current pandemic situation, it has not yet been conducted. The protocol will be detailed in detail but the results cannot be presented here.

## 7.1 Theoretical Background and Research Hypotheses

Interaction is often included in the perceptual process, as proposed by Goldstein and Brockmole (2017). However, the body and its interaction with the surroundings has always been closely related to the concept of perception. Berkeley (1709) defined the notion of distance and its perception by a person as “measured by the motion of [one’s] body”. Thus, since the earliest distance perception theories, the body and its action have a preeminent place in measuring distances.

While perception was first considered as a modular system with many independent sensors (Fodor, 1983; Marr, 1982), subsequent studies, behavioral (Grossenbacher & Lovelace, 2001; Mcgurk & Macdonald, 1976; Sekuler et al., 1997) but also brain imaging-based (Sadato et al., 1996; Uhl et al., 1991; Von Melchner et al., 2000), tend to show just the opposite. As reported by Shimojo and Shams (2001), “cross-modal interactions are the rule and not the exception in perception”. As such, body and its motion play a major role in the perceptual system.

Different experimental results showed the link between body and perception. Let us first focus on the body as a metric. Geuss et al. (2010) showed that the perceived size of an aperture is affected by the width of the participant’s shoulders. Other results showed that the height of an obstacle is affected by the size of the participant (Warren, 1984) or the size of the avatar (Banakou et al., 2013; Van der Hoort et al., 2011). Secondly, with the body motion and the effort needed to reach a distance. This so-called, “effort-based theory” assesses that distance judgments are influenced by the amount of energy observers anticipate to expend in order to reach the item to which distances are judged (Decety et al., 1989; Proffitt, 2006).

Moreover, the interaction technique chosen to report perceived distances has been reported to have a significant impact on the final answer. Indeed, reaching tasks lead to underestimated estimates compared to matching tasks (Swan et al., 2015). The authors suggested that this bias might rely on the length of the participant’s arm leading to this underestimation when no feedback is provided. To go further, some authors suggested that different response measures could come from fundamentally different representation of space. Thus, action and non-action based techniques are providing different results due to their difference in nature (Creem-Regehr & Proffitt, 1998; Philbeck & Loomis, 1997).

Finally, it is worth recalling the impact of knowledge on perception. Indeed, knowl-

edge is crucial in the process of building distance perception, and this is also true in its interaction with motion. Thus, experiments showed that experimented climbers (Warren, 1984) or tracers<sup>1</sup> (Taylor et al., 2011) perceived obstacles and heights as shorter than novices. Thus, the experience coming from motion also modifies perception.

One may therefore wonder whether it is possible to replicate this bias in AR to modify perception through selected interactions. Considering this foreword, three main research hypotheses emerge:

- **H1:** Interaction techniques involving different body movements induce a modification of distance perception.
- **H2:** This modification implies a recalibration of distance perception and leads to a decrease of the perceptual bias in AR.
- **H3:** Direct interaction, involving more body movements, is more likely to reduce this bias.

To tackle these hypotheses, we propose the following experimental protocol.

## 7.2 Experimental Protocol Proposal

In this section, we present an experimental protocol to assess how repeatedly interacting with the virtual/real content with an interaction technique changes the perception in AR. Participants will be asked to proceed to a distance estimation task in the personal space using real and then virtual targets. They will then perform an interaction task, chosen among three and before estimating again egocentric distances with of virtual targets. In the following, we detail the apparatus, the perception and interaction tasks, as well as the experimental design.

### 7.2.1 Apparatus

Participants seat on a chair in front of a white table. The table is covered by a white sheet of paper to prevent any effect of the material (texture, reflections, ...) and to fix markers for tracking purposes. Participants wear an OST HMD, namely the Microsoft Hololens, during all the experiment.

The interaction device used is a commercial clicker, integrated in a 3D-printed case. This custom device, designated as “stylus” in the following, has one button used by

---

1. people who practice *parkour*

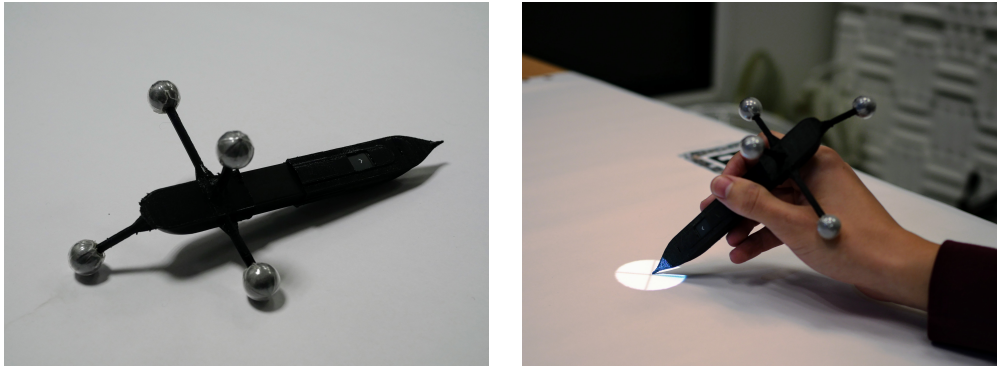


Figure 7.1 – Interaction device.

participants to validate their answer, a thin tip to precisely interact with the targets and a 4-sphere constellation of tracking markers. Its position is detected by a 6-camera ART tracking system <sup>2</sup>.

The real stimuli are projected by a commercial projector on the table, while the virtual ones are viewed on the table through the HMD. Both stimuli are white disks of the same size (3cm in diameter), crossed by a thin black cross marking the center. Their brightness is adjusted to be as close as possible. During the experiment, participants are instructed to point the center of the circle.

Overall, four systems are running during the experiment and are connected on a dedicated network: (1) the camera-projector system, with a dedicated C++ program running the calibration program explained thereafter; (2) the HMD, running a program created with Unity v2019.3.0f6 <sup>3</sup> (3) the tracking system, through the Dtrack software <sup>4</sup>, running on a Windows 7 PC; (4) the main experiment Python runner, which processes the experimental parameters and allows the experimenter to control the experiment.

## 7.2.2 Calibration

This calibration phase aims at determining the relation between the different coordinates systems. Five systems must be referenced in relation to each other: the headset, the physical workspace (the tabletop), the ART tracking system, the camera detecting the user's responses, and the projector displaying real objects.

Markers are placed on the table to ensure the calibration of the workspace. During

---

2. <https://ar-tracking.com/>

3. <https://unity.com/>

4. <https://ar-tracking.com/products/software/dtrack/>

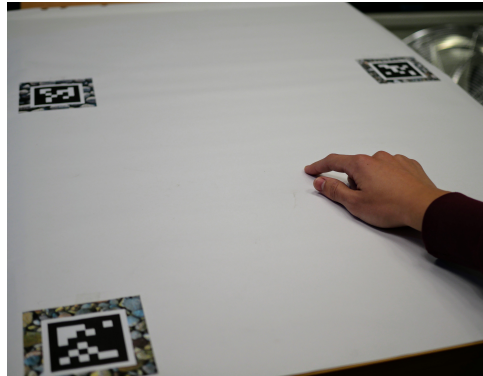


Figure 7.2 – Calibration markers and participant’s hand during perception task.

the calibration phase, the camera is calibrated using the markers and the Aruco computer vision library<sup>5</sup>. The HoloLens uses the Vuforia library<sup>6</sup> to detect the position of the workspace. The projector is calibrated by projecting markers, detected by the camera. Finally, participants are asked to point at the different markers’ corners to calibrate the stylus’ position. The entire calibration is checked by simultaneously displaying the real and virtual elements and making sure that they are superimposed.

### 7.2.3 Task and Procedure

During the experiment and after a presentation and calibration phase, the participant performs four different tasks:

1. a perception task using “real” stimuli (displayed by the projector);
2. a perception task using “virtual” stimuli (displayed in the HMD);
3. an interaction task (chosen among three different ones);
4. a perception task using “virtual” stimuli.

The “real” perception task will provide a baseline to evaluate the performance of the virtual estimates.

**Perception task.** This task is composed of several trials (see subsection 7.2.4). For each trial, the participant sees either a real (projected by a projector) or a virtual (displayed by the HMD) target presented for one second between 15 to 50 cm from the participant in the sagittal axis. Once the target disappears, the participant is asked to point the center

---

5. <https://www.uco.es/investiga/grupos/ava/node/26>

6. <https://developer.vuforia.com/>





Figure 7.3 – Interaction techniques: direct interaction (left), indirect pointing (middle), gaze pointing (right).

of the target with his finger. He can validate his answer by clicking on the stylus' button carried in the other hand. The position of the finger is captured using a camera and the pointed distance recorded.

**Interaction task.** This task is also divided in several trials (see subsection 7.2.4). For each trial, a target is presented to the participant. This target is displayed in the HMD and is placed at one of nine different positions in a  $3 \times 3$  square. This layout allows for covering the whole interaction space, ensures that each participant has the same interaction task and provides sufficient different positions to generate a random order without repetition. The central position is at 30 cm from the participant and the other possible target's position are distant from 10 cm one from another. Each of the nine positions is displayed in a random order.

For each target, the participant is asked to point its center using one of the three following interaction techniques:

- **Direct interaction:** Using the stylus' tip, the participant touches the center of the target.
- **Indirect pointing:** With the stylus acting like a laser pointer, the participant controls a virtual pointer on the table's surface. The participant validates his answer by clicking the button on the stylus.
- **Gaze pointing:** In this configuration, a virtual pointer on the table follows the participant's head movement. The stylus' button is still used to validate the answer.

The target is visible during the whole interaction task. As soon as the target is hit, the distance between the center of the target and the point selected by the participant is recorded, the target disappears and the next one is displayed.

Table 7.1 – Summary of Independent and Dependent Variables.

Independent Variables			
	<i>observers</i>	?	(random variable)
<b>C1</b>	<i>distance</i>	8	from 15 to 50 cm
<b>C2</b>	<i>interaction</i>	2	before or after (within subjects)
<b>C3</b>	<i>interaction type</i>	3	direct, indirect or gaze interaction (between subjects)
	<i>repetition</i>	6	
Dependent Variables			
<b>D1</b>	<i>reported distance</i>	Distance between the participant and his fingertip along the sagittal axis	
<b>D2</b>	<i>reported distance ratio</i>	Ratio between the reported answer for the virtual and the mean reported distance for the real stimuli	

### 7.2.4 Experimental Design

We propose a mixed-model design with the following conditions:

- **C1**: The distance between the participant and the target. Eight distances are considered, all in the personal space, ranging from 15 to 50 cm.
- **C2**: The time of the perception task: before of after the interaction task.
- **C3**: The interaction technique of the interaction task: direct interaction, indirect pointing or gaze pointing.

The distance between the participant and the target (**C1**) is chosen to be in the near distance field and easily reachable to ensure the precision of the reporting technique and to facilitate the interaction task.

In addition to the reported distance, the ratio between the reported distance in virtual compared to real (**D2**) is also included in the analysis to elicit the specific biases of virtual estimates.

In summary, participants are presented with 48 trials for a given perception task: 8 distances (**C1**)  $\times$  6 repetitions. Two perception tasks with virtual targets are conducted, one before and one after the interaction task (**C2**), which leads to 96 data points per participant. For each perception task, the order of **C1** is randomized. The participants are divided in three groups according to the interaction technique they use (**C3**). The two dependent variables are the reported distance for a virtual stimulus (**D1**) and the ratio between this distance and the mean distance reported with a real stimulus in the same condition (**D2**).

## 7.3 Discussion

Two main elements can be found by this experiment. First, the analysis of the distance estimates between the two perception tasks with virtual targets can show the effect of recalibration via the interaction task. Therefore, three tasks are proposed and can have a different impact on this recalibration. We are expected that the more the interaction task implies hand movement in the interaction field the better is the recalibration. However, interaction tasks can also induce specific bias. These biases can be amplified during the interaction and thus leading not to a recalibration of the space perception but to a more distorted perception. The first perception task using real targets will then show if the potential effect of the interaction is beneficial to the perception or will create an additional bias.

In addition to these main results tackling the research hypotheses, supplementary results can be drawn. On the one hand, our experimental design during the interaction task can provide data about the precision of the chosen interaction technique. Previous studies has shown that different interactions in AR can have different outcomes in term of precision and this experiment may be able to provide more insight on this topic. On the other hand, the evolution of precision during the interaction task itself could provide more elements the appearance of the potential bias or recalibration. Even if the final perception task might show a difference due to the interaction, the analysis of the precision during the interaction may elicit the learning process involved in this effect.

Finally, it has to be noted that the final perception task does not involve an evaluation of the perception with real targets.





What are the future challenges for perception in Mixed Reality? What are the contributions and openings of the different contributions of this thesis? How to use perception as a research tool for designing Mixed Reality applications? This chapter proposes to address these questions as well as a more global summary and openings of the work presented in this thesis.

# DISCUSSION

## Contents

<b>8.1</b>	<b>Space perception and Mixed Reality - Current state of knowledge and research challenges . . . . .</b>	<b>189</b>
<b>8.2</b>	<b>Contributions of this Thesis . . . . .</b>	<b>191</b>
<b>8.3</b>	<b>Limitations and Future Works . . . . .</b>	<b>193</b>
8.3.1	Experimental variations . . . . .	193
8.3.2	Perspectives on conducted experiments . . . . .	196
8.3.3	Long term perspectives . . . . .	199

## 8.1 Space perception and Mixed Reality - Current state of knowledge and research challenges

The principles underlying human perception have been debated for a long time in the scientific community. The initial theory of the internal representation of space in the form of a mind map (S. Kaplan, 1973) was later challenged by a topological approach of space (Kuipers, 1983), rejecting the idea of a metric representation in favor of a semantic representation of objects in relation to each other. However, these theories have also proved to be too restrictive. The most recent works agree on a mixed, more ecological conceptualization of perception. They focus on describing how different types of information, resulting from the interaction of the individual with his environment, can contribute to the construction of a mental representation (Tversky, 1993).

This approach is the one underlined by Goldstein and Brockmole (2017) in their model of perception and in particular of the perceptual process. This process can be divided into four stages: (1) stimuli (distal & proximal), (2) physiological response (receptors & neuronal processing), (3) perception & recognition, (4) action. This process can be seen as a loop, integrating external elements, processing them and then acting on them, but

it should also be noted that all these steps are interconnected and can influence each other. Finally, knowledge plays a major role in this process, influencing each step through learning and experience.

Perception is therefore a mental state, an internal representation, which can potentially provoke an action. However, the internal nature of this process is already a major challenge. On the one hand, action is not necessarily a faithful translation of this internal representation and there may be a difference between what is perceived and what is presented to someone. On the other hand, the action externalizing the perception can influence this same perception and thus bias its measurement.

The theory of distance perception is based on depth cues (Goldstein & Brockmole, 2017; Howard, 2002). These cues allow the human visual system to determine distances by combining them. There are three main categories of depth cues: (1) oculomotor cues, provided by the oculomotor system of the eyes, (2) pictorial cues, that rely on 2D information, and (3) motion cues, based on the movements of the objects or of the point of view. Thus, the weight of these depends on the distance between the object and the observer which defines three spaces, also characterized by the actions that can be carried out inside (Cutting, 2003): (1) the personal space, within 1.5 meters, within arm's reach, (2) the action space, up to 30 meters, where the observer can quickly interact with an object, (3) the vista space, beyond 30 meters, where interaction is difficult or even impossible.

The perception of distances in real environments is already a major challenge, in particular when trying to model it. Differences between real space and perceived space can be observed, and current models of perception are still struggling to transcribe these biases. Considering the case of environments containing both real and virtual elements, i.e. mixed environments, things become even more difficult. The virtual elements are displayed differently from the real ones, so they will cause different perception biases. But in addition, perception being a global process, these “simple” biases will also interfere with the perception of all other elements, even real ones.

In Mixed Reality (MR), the devices and rendering techniques are numerous. Head Mounted Displays (HMDs) are the most widely used devices. They can be non-see through for VR headsets or Optical See-Through (OST) or Video See-Through (VST) for AR headsets. However, other devices can be used, such as handheld devices, using a camera and a screen, or projection devices. Each of these types of devices has its own characteristics and therefore produces different effects on perception. Since virtual elements are simulated elements, techniques for creating virtual images and rendering choices can also

vary, which will also impact perception. Overall, finding general trends within all these potential effects is a major challenge.

Thus, to categorize perceptual biases in MR, their origins can be divided into four groups: (1) depth cues related biases, (2) intrinsic biases, (3) rendering induced biases and (4) environment and object specific biases. However, if this categorization allows a for systematic analysis of perception biases in MR, it also shows many configurations that have not yet been studied from a perceptual point of view. Moreover, while this classification allows for an individual analysis of the factors influencing perception in MR, much work remains to be done to be able to model MR perception in a systematic way, if indeed this can be done at all.

## 8.2 Contributions of this Thesis

In this thesis, we aimed at characterizing some perceptual biases that occur in Mixed Reality. In particular, we focused on the study of several factors inducing distance misperception in mixed environments. To do so, we have studied different factors, of various nature and origin. We investigated factors intrinsic to observers, factors related to display devices and rendering techniques, as well as variations in referential from egocentric to exocentric distance estimates.

In PART I, we have detailed the different elements of the state of the art related to perception and Mixed Reality.

In **Chapter 1** we presented an overview of the previous literature on perception and more especially on distance perception. First, we presented the process which transforms sensations into perception: the perceptual process. We also underlined the interconnections between the different stages of this process as well as the impact of knowledge. Then, we highlighted the challenges that the perceptual process has to face and the difficulty in measuring perception. Subsequently, we presented the specific case of distance perception, describing in detail the depth cues theory as well as more global theories on distance perception. Finally, we presented different reporting techniques that can be used to evaluate distance perception and the issues and possible biases due to these techniques.

In **Chapter 2** we detailed the theoretical background of Virtual Reality and Augmented Reality. Starting from the definitions and basic presentation of the mixed reality continuum, we highlighted the fundamental differences between actual reality and mixed reality. Then, we reviewed all the devices that allow to perform MR, highlight-



ing the different categories of devices and their characteristics in order to evaluate their impact on perception. Finally, we underlined two major challenges for MR: the vergence-accommodation conflict and the so-called “x-ray vision”.

In **Chapter 3** we presented an overview of the key studies in distance perception related to AR. For each element influencing the perception of distances, we detailed the impact of AR on this factor and thus the induced perceptual bias. For this, we have distinguished four types of biases: (1) depth cues related biases, (2) intrinsic biases, (3) rendering induced biases and (4) environment and object specific biases. Finally, we also presented the existing studies in distance perception for exocentric distances. This review highlights some major factors impacting the perception of distance in AR.

In PART II, we presented four series of experiments conducted to elicit several factors of perceptual biases in Mixed Reality.

In **Chapter 4** we presented a series of four user experiments aiming at highlighting an intrinsic perceptual bias in VR. This series of experiments showed that the spatial perception of observers is anisotropic in VR. The first experiment showed that virtual objects placed in front of the participant are perceived closer than objects placed on the side. The second experiment confirmed this effect and provided evidence that it increases with the viewing angle (the more the object is placed on the side and the farther it is perceived). Finally, two control experiments were conducted and showed that this effect is neither affected by the position of the object in the field of view nor by the amount of depth cues provided by the virtual environment.

In **Chapter 5** we studied the effect of AR on exocentric distance perception. We reported a user experiment aiming at evaluating distance perception between two objects lying in a frontoparallel plane in the medium-field perceptual space. Four conditions were tested in our study, involving two virtual objects, two real objects and the two symmetrical conditions with one real object and one virtual object. Results showed that two virtual objects were perceived closer one to another than two real objects, i.e. the exocentric distance was underestimated in AR compared to reality. Moreover, a non symmetrical behavior was found for the mixed condition.

In **Chapter 6** we reported two series of experiments, aimed at studying the impact of depth cues on distance perception in AR.

The first two experiments focused on the effect of shadows on depth perception. Both experiments showed an increase in distance perception performance when shadows were provided. The first experiment showed that the shape of the shadows has no significant

impact on distance perception. Simple circular shadows had the same results as more realistic shadows, with hard or soft edges. The second experiment showed that the position of lighting sources is an important factor for virtual shadows since a misalignment of virtual and real lights sources (and then inaccurate shadows position) can lead to a misperception of egocentric distances.

The last experiment reported in this chapter aimed at evaluating the effect of the fixed accommodation distance in Optical See-Through (OST) Head Mounted Displays (HMDs). In this experiment, distance perception of the participants was evaluated using a usual OST HMDs with a fixed focal distance and a Retinal Projection Display (RPD). RPDs do not have focal distance thanks to their laser projection. This study showed that while the participants overestimated distances in AR compared to reality with the classical OST HMD, their performance significantly increased when using the RPD. Besides, the task with the RPD was reported with the same level of difficulty and no difference in precision.

In PART III, we presented the perspectives and prospects related to perception in Mixed Reality.

In **Chapter 7** we presented a proposal of user experiment to evaluate the effect of interaction techniques on distance perception. First, we presented the motivations and the theoretical background on action-related perception theories, as well as the proposed research hypotheses of this study. Then, we presented the detailed protocol of this study and mentioned some of the expected results and potential prospects.

## 8.3 Limitations and Future Works

In this section, we first present the different experimental variations that could be made on each of the experiments detailed in this thesis in order to study more parameters. Then, we will present the issues and new questions related to perception raised specifically by each of our experiments. Finally, we present more global perspectives on the potential contribution of the study of perception to other major subjects in Mixed Reality.

### 8.3.1 Experimental variations

**Varying devices.** All our studies were conducted with one or two specific HMDs. Each device has its particular (characteristics, display technology, FoV, accommodation distance or the technical possibilities for rendering). These characteristics constrained our

experimental designs and setups, such as the distances of the virtual targets or the choices for rendering. Then, new experiments could be conducted using other AR devices: OST or VST HMDs, or screens, handheld or SAR devices.

**Different evaluation tasks.** The protocol used in each of our experiment was selected to ensure the evaluation of a specific element. For example, a Two-Alternative Forced Choice (2AFC) protocol was preferred to highlight potential biases, while an open-loop protocol was used to evaluate the extent of the distance estimation biases, without dwelling with the comparison and calibration effects of the closed-loop protocol. However, as stated in the first chapters of this thesis, the evaluation protocol can influence perception. For example, in many distance perception studies in AR, an open-loop action based answering protocol might be partly responsible for the overall underestimation of the perceived distances. Indeed, reaching tasks are generally used to report perceived distances, as they are not biased by any visual feedback, but suffer from proprioception-based biases since no correction is provided (Swan et al., 2015). Thus, new experiments using different protocols could be carried out, to compensate for the effects of one specific answering protocol. It is also possible to focus on some specific conditions using 2AFC reporting protocols as they are more accurate when assessing the existence of potential biases.

**Considering other distances.** As stated by the literature, space perception varies according to the distance field considered (Howard, 2002). Each of our experiment focused on one specific distance field, or even one distance, but space perception could also be studied for other distance fields. The weight of the depth cues varying according to distance, one could thus see an evolution of the effects studied previously in interaction with distance. More broadly, all variations in positions and orientations of the targets could be considered and studied. Moreover, we have proposed to move the targets, but we could also consider moving the observer and therefore his point of view, since it was shown in chapter 4 that this also has an effect on perception.

**Changing size, shape or rendering.** Moreover, the visual aspect of the considered objects could also influence the results. In our studies, we focused on very salient objects, with known sizes and simple shapes. Moreover, the rendering and shading parameters could be changed, in particular to observe the impact of the rendering quality on distance perception. Since the evaluation of distance in real or mixed environments has been reported to be dependent on the shape, size and rendering of the virtual object (Eggleston

et al., 1996; Phillips et al., 2009; Renner et al., 2013), these parameters should also be evaluated to extend the experiments presented in this thesis. Augmented objects – with mixed real and virtual parts – could even be used. In such configurations, the effect of other usual anchoring cues such as projected shadows or other auxiliary augmentations should also be evaluated.

**More ecological configurations allowing head movements.** In some of our experiments, the setup was completely fixed, from the targets to the head of the participant itself (fixed to the chin-rest). However, this configuration is not an ecologically valid condition. Usually, the observer is free to move his head and small movements of the head could help to evaluate distances. The observed objects could also be moving, providing more depth cues such as motion parallax. We choose to use simple environments to prevent for biases but a richer and more ecological environment could provide a specific blending of many more cues which could interact with each other. The evaluation of the specific interaction of different devices in such an ecological situation is a very exciting open question.

Beyond the distance, it is also possible to evaluate the **intrinsic characteristics** of objects, such as size, weight, etc. Most of these characteristics are related to distance perception but could also show specific biases in MR. It is therefore interesting to study them specifically, by studying their evolution according to their shadows, appearance, etc.

**Refer to real.** Finally, we conducted most of our experiments using real estimates as a baseline. This provided some comparable data, and could helped to isolate factors leading to egocentric distance errors in MR. However, we were not able to do this for all the experiments (especially the one evaluating the effect of shadows and the one on anisotropic perception), mainly for technical reasons (precisely placing real targets is usually harder to ensure compared to virtual ones). It would therefore be relevant to introduce a comparison with reality for all perception experiments in MR. This could allow both to have a reference to evaluate perception in virtual but also to bring out explanations and extensions to the general theories of perception, not specific to MR.

## 8.3.2 Perspectives on conducted experiments

### 8.3.2.1 An intrinsic bias: the visual space orientation?

In chapter 4, our setup for evaluating the anisotropy of space perception, involved a participant seated on a chair, which is very common to many consumer applications. Interestingly, it naturally induces a strong orientation of the VE. This specific direction is here enforced by the position of the participant, seated and holding a joystick in front of him. The entire body is constrained toward one specific direction. Some other VEs or user positions/orientations could therefore be considered in future work, without any strong *forward* direction, and involving other settings.

Moreover, the present experiment was designed to provide an evaluation of the anisotropy of distance perception in ecological conditions. It did not assess the global shape of the visual space, which is of course more complex. The task we chose is specific and cannot cover all the various numerical models proposed for the visual space. Our observed variation in distance perception according to the head direction already seems in favor of the non-Euclidean shape of the visual space. But future work could also extend our results, to confront them with the various models proposed regarding visual space perception. As an example, our experiments considered the depth perception of a virtual object always placed at ~3 meters. The evolution of the anisotropy according to the distance considered, the vertical position or other configurations in space seem other worthwhile topics to consider in order to better assess the shape of the visual space in Virtual Reality. Thus, more studies in real, virtual and even augmented environments are needed to provide a better characterization of this phenomenon and to identify the specific influence of VR or AR compared with reality.

### 8.3.2.2 Multiple objects imply multiple biases?

When studying exocentric distance perception in chapter 5, we choose to limit the layout of the objects to two objects placed in the frontoparallel plane. This layout could be extended by adding other objects to evaluate the impact of a virtual or a real object placed between the two original ones, or by moving the relative positions of the objects.

The objects' layout could also be modified by rotating the plane where the objects are placed. Indeed, we could not only consider extending our study on exocentric distance perception in the sagittal plane but also on all the possible angles between those two (frontoparallel and sagittal) layouts. This is motivated by previous evidence (see Loomis

et al., 1992) showing that physically equal distances in a sagittal plane on the ground are perceptually underestimated relative to those in a frontoparallel plane, even under full-cue conditions. We expect that placing virtual objects outside of the frontoparallel plane would lead to different exocentric distance estimations.

On the other hand, new targets could be organized in several layouts: two side-by-side or one behind the other, or even in a regular grid or some other random pattern. Those layouts would allow to evaluate the potential impact of references, occlusion, ordering or relative distances. While the literature provides an extensive corpus around egocentric distance perception in AR, the specificities of multi-targets distance evaluation and their interaction still require more evaluation. Indeed, when comparing two objects' depths, occlusions are a very strong cue. When using OST devices, this cue is not available for mixed conditions (with real and virtual objects) and the VAC would prevent the observer to adjust the accommodation in a natural way. With RPDs, the increment of vergence from one object to another could be associated with an increment in depth of the accommodation without inducing a conflict. Retinal projection could then be an even more accurate technology for depth perception when it comes to inter-objects estimations. As such, the effect of occluders or objects in other planes could be a very important question to tackle in AR, especially if one also considers the type of device used.

Finally, using a second object, as for matching tasks for example, could have a major impact on perception. Those tasks are different from the tasks used in our studies since they provide visual feedback during distance estimation. If the participant has to align the designated target with another object, say, a cursor, the nature of this cursor is fundamental. Indeed, if the cursor has the same type (real or virtual) as the target, then the potential bias would be the same and the alignment should be accurate. However, in the case where the target and the cursor have a different type, the depth cues provided by the two kinds of rendering would interfere, leading to potential unusual biases. Moreover, we already underlined the major influence of the accommodation cue in the specific case of HMDs. Here, the different nature of the target and the cursor might be even more impacting for distance perception since accommodation is a very strong cue for ordering (Mon-Williams & Tresilian, 2000).

Overall, presenting more than one object to the observer could lead to many different layouts and then have many different effects on distance perception. Many studies are still lacking to assess all the potential effects of these combinations of factors.

### 8.3.2.3 Shadows influence

It has to be noted that all MR technologies do not provide the same level of shadows' rendering. In particular, OST HMDs and projector-based AR devices could not remove light to display “dark” shadows. They have to make the global environment darker, or use local adjustment to create this effect of shadows. Some displays using a subtractive rendering technique exist but they are ready yet to do AR (Itoh et al., 2019). Displaying realistic shadows is therefore already a challenge in itself and much remains to be done to allow AR devices to display shadows properly.

In section 6.1, we provided a study of the impact of shadows on depth perception, focussing on the shape and position of the shadows. However, in order to further study this phenomenon, we envision different future leads focusing on the nature of the shadows. We could study the influence of shadows' rendering in OST AR, e.g. by increasing the brightness of the rest of the scene to make them more noticeable). We could also compare the impact of the use of “drop shadows”, i.e. shadows placed immediately below the objects, vs. realistic shadows on distance perception accuracy.

Results from such experiments could lead to the definition of guidelines for proper use of shadows when designing AR applications which may become more and more accessible in a near future.

### 8.3.2.4 Retinal Projection Displays

RPDs are lacking accommodation cues. However, the blur induced by the adjustment of the focal distance could be reproduced with dedicated rendering (Cholewiak et al., 2017; Dunn et al., 2018). Then, using blur to help depth estimation in RPD would be interesting to study. However, it would require very precise and fast eye tracking as well as an evaluation of the actual accommodation state of the eye. For the time being, those technologies are still subject to ongoing research but are improving quickly and their integration with RPDs could provide very accurate depth rendering in AR.

One last specificity of RPDs is the use of those displays for visually impaired people. Since the image is projected directly on the retina with a laser beam, the resulting image is always sharp, even for people suffering from eye accommodation disabilities such as myopia or hypermetropia. Then, virtual images rendered with RPDs are even sharper than real images. As such, while improving the quality of the image seen, it is still unclear if retinal projected rendering would significantly enhance distance perception for visually

impaired people or if this kind of rendering, drastically different from their usual viewing will disturb their perception of distance.

### **8.3.3 Long term perspectives**

I would like to insist on the plurality and the always different nature of the observed biases in MR: intrinsic biases, biases linked to technologies but also biases induced by rendering and visualization choices. Thus, it seems extremely complex, if not almost impossible, to characterize all the elements that can impact the perceptual process. However, I have to recall the purpose of perception: action. Indeed, all the biases highlighted above lead to changes in the use of MR technologies. An unsuitable focal length could lead to a change in the user's navigation strategy, a misperception of the shape of an object, impacting its handling, etc.

Therefore, I consider perception as an entry point that is at once original, relevant and promising for many issues already raised in MR. Thus, I want to highlight three research axes: Perception and Representation, Perception and Interaction, and Perception and Cooperation.

#### **8.3.3.1 Perception and Interaction**

Interaction is a major issue in MR and in particular in AR. The real and virtual nature of objects and environments require a constant adaptation of interaction techniques to adapt to these changes. The stakes are first technical: it is impossible for a virtual visualization system to move a real element and, conversely, a synchronous movement of a virtual object via a real object remains difficult to achieve. However, the stakes of interaction are also related to the users performing these interactions. I want to emphasize here the major importance of interaction in the perceptual process. Action is often considered as the finality of the perceptual process. Humans evaluate the size of an obstacle to overcome it, or the shape of an object to grasp it. However, the proximity between perception and action leads to an interdependence between these two processes. Thus, we have presented in our prospects a study aimed at modifying or even improving perception in Augmented Reality by using different interaction techniques. To go further, I propose to study more broadly the effects of rendering and of the environment (i.e. what is seen) on actions, interactions and affordances (i.e. what is done) in mixed environments. This could lead to the proposal of efficient interaction techniques more adapted to the uses and



users. These techniques could also benefit from the use of other sensory modalities, other than vision, to get closer to the multisensory nature of perception.

### **8.3.3.2 Perception and Representation**

Representation and display, whether of realistic elements or data, present major challenges for MR. How to accurately represent realistic objects with limited technology? How to take advantage of the limits of perception to propose ecologically valid representations? How to represent complex, potentially multidimensional and non-spatial digital data in a 3D virtual environment? These questions are already open research questions, but I wish to bring a different, perception-based interrogation to them. We have extensively presented the problems of perceptual biases in MR in this thesis. Within the framework of the study of representations (of data or realistic elements), it seems important to me to consider these representation techniques as levers to compensate perceptual biases. Ultimately, this approach would allow to improve data visualization and representation in general by relying on the capabilities and limits of the human perceptual system and to propose representations adapted to the context and to the observers. By taking advantage of the limits of perception in Virtual Reality, it would then be possible to build and represent virtual environments more adapted to human perception.

### **8.3.3.3 Perception and Cooperation**

Finally, I would like to consider the study of collaboration in mixed environments. This subject is already significantly studied within the scientific community and the new avenues opened up from the perspective of perception seem particularly exciting to me. I have already detailed some major elements and steps of the perceptual process, but broadening them in the context of multi-user augmented applications is a research question that is both original and stimulating. Can perceptual corrections and improvements observed for one user be transferred to other users cooperating in the same environment? Do several users in the same environment develop similar perceptual biases? Do interaction techniques used in mixed environments vary when several users are present? In a mixed environment cooperation involving body representation of participants, how do the issues of avatars and embodiment extend in the case of Augmented Reality? More generally, the cooperation of several agents in a mixed environment raises questions about the types and characteristics of cues and references used. In my opinion, these questions

present major challenges for MR, both in its theoretical approach and in its practical implementation.

In conclusion, I would like to present my work on the question of the study of perception in Mixed Reality not as a theoretical goal but as a powerful tool for questioning in a way that is both original and embodied the multiple problems posed by mixed environments.



# CONCLUSION

---

The study of human perception is a vast research field with many potential openings. As an inner process, a part of this process will always be hidden from observation and evaluation. A part of *magic* will always be involved in the transformation from sensation to perception. This will remain hidden from any observational study. Thus, it seems also impossible to find a global model of perception. Too many factors, contingent elements and an almost infinite number of interactions between these factors prevent the perceptual model from being completely defined. Indeed, we would have to be able to read or predict the mental states of the observers, their knowledge, etc. In my opinion, the search for such a model is not only pointless but also unhelpful. Much more remains to be done in order to use perception not as an end but as a tool. The study of specific biases allows to show the way to necessary technological advances. By using the characterization of perceptual biases as a metric, the technology could evolve to solve the most important problems related to perception and thus to the user.

All the studies in this thesis show that the current perceptual biases are always linked with technological deficiencies. There is still a long way to go to be able to provide mixed environments that are indistinguishable from real environments. However, this goal may one day be achieved, and researchers will finally be able to tackle the following question: “Is a virtual object that is indistinguishable from a real object still perceived differently?”.



PART IV

# Appendix

---



# RÉSUMÉ LONG EN FRANÇAIS

---

## Contexte

Les principes qui sous-tendent la perception humaine sont débattus depuis longtemps dans la communauté scientifique. La théorie initiale de la représentation interne de l'espace sous la forme d'une carte mentale (S. KAPLAN, 1973) a ensuite été remise en cause par une approche topologique de l'espace (KUIPERS, 1983), rejetant l'idée d'une représentation métrique au profit d'une représentation sémantique des objets les uns par rapport aux autres. Cependant, ces théories se sont également révélées trop restrictives. Les travaux les plus récents s'accordent sur une conceptualisation mixte et plus écologique de la perception. Ils s'attachent à décrire comment différents types d'informations, résultant de l'interaction de l'individu avec son environnement, peuvent contribuer à la construction d'une représentation mentale (TVERSKY, 1993).

Cette approche est celle soulignée par GOLDSTEIN et BROCKMOLE (2017) dans leur modèle de perception et en particulier du processus perceptif. Ce processus peut être divisé en quatre étapes : (1) les stimuli (distal & proximal), (2) la réponse physiologique (récepteurs & traitement neuronal), (3) la perception & la reconnaissance, (4) l'action. Ce processus peut être considéré comme une boucle, intégrant des éléments externes, les traitant puis agissant sur eux, mais il faut également noter que toutes ces étapes sont interconnectées et peuvent s'influencer mutuellement. Enfin, la connaissance joue un rôle majeur dans ce processus, en influençant chaque étape par l'apprentissage et l'expérience.

La perception est donc un état mental, une représentation interne, qui peut potentiellement provoquer une action. Cependant, la nature interne de ce processus constitue déjà un défi majeur. D'une part, l'action n'est pas nécessairement une traduction fidèle de cette représentation interne et il peut y avoir une différence entre ce qui est perçu et ce qui est présenté. D'autre part, l'action externalisant la perception peut influencer cette même perception et donc biaiser sa mesure.

La théorie de la perception des distances est basée sur les indices de profondeur



---

(GOLDSTEIN & BROCKMOLE, 2017; HOWARD, 2002). Ces indices permettent au système visuel humain de déterminer les distances en les combinant. Il existe trois grandes catégories d'indices de profondeur : (1) les indices oculomoteurs, fournis par le système oculomoteur des yeux, (2) les indices picturaux, qui reposent sur des informations en 2D, et (3) les indices de mouvement, basés sur les mouvements des objets ou du point de vue.

Ainsi, le poids de ces indices dépend de la distance entre l'objet et l'observateur, ce qui définit trois espaces, également caractérisés par les actions qui peuvent être effectuées à l'intérieur de (CUTTING, 2003) : (1) l'espace personnel, à moins de 1,5 mètre, à portée de main, (2) l'espace d'action, jusqu'à 30 mètres, où l'observateur peut rapidement interagir avec un objet, (3) l'espace visuel, au-delà de 30 mètres, où l'interaction est difficile, voire impossible.

La perception des distances dans des environnements réels est déjà un défi majeur, en particulier lorsqu'on essaie de la modéliser. Des différences entre l'espace réel et l'espace perçu peuvent être observées, et les modèles de perception actuels ont encore du mal à transcrire ces biais. Si l'on considère le cas des environnements contenant des éléments réels et virtuels, c'est-à-dire des environnements mixtes, les choses deviennent encore plus difficiles. Les éléments virtuels sont affichés différemment des éléments réels, ce qui entraîne des biais de perception différents. Mais en outre, la perception étant un processus global, ces biais "simples" vont également interférer avec la perception de tous les autres éléments, même réels.

En réalité mixte, les dispositifs et les techniques de rendu sont nombreux. Les visio-casques sont les dispositifs les plus utilisés. Ils peuvent être non transparents pour les casques de Réalité Virtuelle ou à vision directe (*Optical See-Through* (OST)) ou indirecte (*Video See-Through* (VST)) pour les casques de Réalité Augmentée. Toutefois, d'autres dispositifs peuvent être utilisés, tels que des appareils type smartphone ou tablette, utilisant une caméra et un écran, ainsi que des appareils de projection. Chacun de ces types d'appareils a ses propres caractéristiques et produit donc des effets différents sur la perception. Les éléments virtuels étant des éléments simulés, les techniques de création d'images virtuelles et les choix de rendu peuvent également varier, ce qui aura également un impact sur la perception. Dans l'ensemble, trouver des tendances générales parmi tous ces effets potentiels est un défi majeur.

Ainsi, pour classer les biais perceptuels en réalité mixte, leurs origines peuvent être divisées en quatre groupes : (1) les biais liés aux indices de profondeur, (2) les biais intrinsèques, (3) les biais induits par le rendu et (4) les biais spécifiques à l'environnement

---

et à l'objet. Cependant, si cette catégorisation permet une analyse systématique des biais de perception en réalité mixte, elle montre également de nombreuses configurations qui n'ont pas encore été étudiées d'un point de vue perceptif. De plus, si cette classification permet une analyse individuelle des facteurs influençant la perception en réalité mixte, il reste beaucoup de travail à faire pour pouvoir modéliser la perception en réalité mixte de manière systématique, si tant est que cela puisse être fait.

Dans cette thèse, nous avons cherché à caractériser les biais perceptuels qui peuvent se produire en réalité mixte. En particulier, nous nous sommes concentrés sur l'étude de plusieurs facteurs induisant une mauvaise perception des distances dans les environnements mixtes. Pour cela, nous avons étudié différents facteurs, de nature et d'origine diverses : des facteurs intrinsèques aux observateurs, des facteurs liés aux dispositifs d'affichage et aux techniques de rendu, ainsi qu'aux variations de référentiel (egocentrique/exocentrique).

## **Contributions**

### **Chapitre 1 – De la sensation à la perception ; la perception visuelle comme exemple, l'évaluation de la distance comme métrique**

Nous avons présenté un aperçu de la littérature précédente sur la perception et plus particulièrement sur la perception des distances. Nous avons tout d'abord présenté le processus qui transforme les sensations en perception : le processus perceptif. Nous avons également souligné les interconnexions entre les différentes étapes de ce processus ainsi que l'impact des connaissances sur celui-ci. Ensuite, nous avons souligné les défis auxquels le processus perceptuel doit faire face et la difficulté de mesurer la perception. Ensuite, nous avons présenté le cas spécifique de la perception à distance, en décrivant en détail la théorie des indices de profondeur et les théories plus globales sur la perception à distance. Enfin, nous avons présenté les différentes techniques de compte rendu qui peuvent être utilisées pour évaluer la perception à distance, ainsi que les problèmes et les biais possibles de ces techniques.

---

## **Chapitre 2 – Réalité mixte - Comment représenter un monde virtuel et le mélanger à la réalité**

Dans ce chapitre, nous avons détaillé le contexte théorique de la réalité virtuelle (RV) et de la réalité augmentée (RA). En partant des définitions et de la présentation de base du continuum de la réalité mixte, nous avons mis en évidence les différences fondamentales entre la réalité réelle et la réalité mixte. Ensuite, nous avons passé en revue tous les appareils permettant l'usage de la réalité mixte, en mettant en évidence les différentes catégories d'appareils et leurs caractéristiques afin d'évaluer leur impact sur la perception. Enfin, nous avons souligné deux défis majeurs pour la réalité mixte : le conflit vergence-accommodation et la vision dites "x-ray".

## **Chapitre 3 – Aperçu des principales études sur la perception à distance en réalité augmentée**

Ce chapitre vise à présenter un aperçu des principales études sur la perception des distances en réalité augmentée. Pour chaque élément influençant la perception des distances, nous avons détaillé l'impact de la RA sur ce facteur et donc le biais perceptif induit. Pour cela, nous avons distingué quatre types de biais : (1) les biais liés aux indices de profondeur, (2) les biais intrinsèques, (3) les biais de rendu induits et (4) les biais spécifiques à l'environnement et à l'objet. Enfin, nous avons également présenté les études existantes en matière de perception des distances pour les distances exocentriques. Cette revue met en évidence certains facteurs majeurs qui ont un impact sur la perception des distances en réalité augmentée.

## **Chapitre 4 – Biais intrinsèques : l'anisotropie de la perception des distances**

Dans ce chapitre, nous avons présenté une série de quatre expériences utilisateurs visant à mettre en évidence un biais perceptuel intrinsèque en réalité virtuelle. Cette série d'expériences a montré que la perception spatiale des observateurs en RV est anisotrope. La première expérience a montré que les objets virtuels placés devant le participant sont perçus plus près que les objets sur le côté. La deuxième expérience a confirmé cet effet et a fourni la preuve que l'effet augmente avec l'angle de vision (plus l'objet est placé sur le côté et plus il est perçu loin). Enfin, deux expériences de contrôle ont été menées et ont

---

montré que cet effet n'est pas affecté par la position de l'objet dans le champ de vision ni par la quantité de repères de profondeur fournis par l'environnement virtuel.

## **Chapitre 5 – Transfert des biais de perception : De l'égoцентриque à l'exocentrique ?**

Dans ce chapitre, nous avons étudié l'effet de la RA sur la perception exocentrique de la distance. Nous y avons présenté une expérience utilisateur visant à évaluer la perception des distances entre deux objets se trouvant face à l'observateur. Quatre conditions ont été testées dans notre étude, impliquant deux objets virtuels, deux objets réels et les deux conditions symétriques avec un objet réel et un objet virtuel. Les résultats ont montré que deux objets virtuels étaient perçus plus proches l'un de l'autre que deux objets réels, c'est-à-dire que la distance exocentrique était sous-estimée en réalité augmentée par rapport à la réalité. De plus, un comportement non symétrique a été trouvé pour la condition mixte.

## **Chapitre 6 – L'impact des indices de profondeur sur la perception égoцентриque de la distance en réalité augmentée**

Ce chapitre présente deux séries d'expériences, visant à étudier l'impact des indices de profondeur sur la perception des distances en RA. Les deux premières expériences étaient axées sur l'effet des ombres sur la perception de la profondeur. Les deux expériences ont montré une augmentation des performances de la perception des distances lorsque des ombres étaient présentes. La première expérience a montré que la forme des ombres n'a pas d'impact significatif sur la perception des distances. Les ombres circulaires simples ont donné les mêmes résultats que les ombres plus réalistes, qu'elles aient des bords nets ou flous. La deuxième expérience a montré que la position des sources de lumière est un facteur important pour les ombres virtuelles, car un mauvais alignement des sources de lumière virtuelles et réelles (et donc une position inexacte des ombres) peut conduire à une perception erronée des distances égoцентриques. La dernière expérience rapportée dans ce chapitre visait à évaluer l'effet de la distance d'accommodation fixe des visiocasques. Dans cette expérience, la perception des distances des participants a été évaluée en utilisant soit un visiocasque habituel avec une distance focale fixe, soit un dispositif de projection rétinienne (RPD). Les RPD n'ont pas de distance focale grâce à leur projection laser. Cette étude a montré que si les participants surestimaient les distances en RA par rapport à la réalité avec le visiocasque, leurs performances augmentaient de manière significative

---

lorsqu'ils utilisaient le système de projection rétinienne. En outre, la tâche avec le RPD a été rapportée avec le même niveau de difficulté et sans différence de précision.

## **Chapitre 7 – Perspectives - Comment compenser les biais de perception ? L'utilisation des techniques d'interaction**

Nous avons présenté une proposition d'expérience utilisateur pour évaluer l'effet des techniques d'interaction sur la perception des distances. Tout d'abord, nous avons présenté les motivations et le contexte théorique des théories de perception liées à l'action, ainsi que les hypothèses de recherche proposées dans le cadre de cette étude. Ensuite, nous avons présenté le protocole détaillé de cette étude et mentionné certains des résultats attendus et des perspectives potentielles.

## **Chapitre 8 - Discussion**

Ce dernier chapitre conclut le manuscrit par une discussion générale. Tout d'abord, il rappelle les principaux défis de la perception spatiale et de la réalité mixte. Deuxièmement, il résume les contributions de cette thèse. Enfin, il fournit une discussion générale des problèmes, des biais et des améliorations potentielles apportés par les différentes études et présente des perspectives potentielles.

## **Conclusion**

L'étude de la perception humaine est un vaste domaine de recherche qui offre de nombreuses possibilités. En tant que processus interne, une partie de ce processus sera toujours cachée à l'observation et à l'évaluation. Une part de la magie sera donc toujours impliquée dans la transformation de la sensation en perception. Ceci restera caché à toute étude observationnelle.

Ainsi, il semble également impossible de trouver un modèle global de perception. Trop de facteurs, d'éléments contingents et un nombre presque infini d'interactions entre ces facteurs empêchent de définir complètement le modèle perceptif. En effet, il faudrait pouvoir lire ou prédire les états mentaux des observateurs, leurs connaissances, etc. Il me semble que la recherche d'un tel modèle serait non seulement vaine mais aussi inutile. Il reste encore beaucoup à faire pour utiliser la perception non pas comme une fin mais comme un outil.

---

L'étude des biais spécifiques permet de montrer la voie vers les avancées technologiques nécessaires. En utilisant la caractérisation des biais perceptifs comme une mesure, la technologie pourrait évoluer pour résoudre les problèmes les plus importants liés à la perception et donc à l'utilisateur.

Toutes les études de cette thèse montrent que les biais perceptuels actuels sont toujours liés à des déficiences technologiques. Il reste encore beaucoup de chemin à parcourir pour pouvoir fournir des environnements mixtes qui ne se distinguent pas des environnements réels.

Cependant, cet objectif pourra peut-être un jour être atteint, et des chercheurs pourront enfin pour aborder cette question : “Un objet virtuel impossible à distinguer d'un objet réel serait-il malgré tout perçu différemment ?”.



# AUTHOR'S PUBLICATIONS

---

## Journals

- Gao, Y., Peillard, E., Normand, J. M., Moreau, G., Liu, Y., & Wang, Y. (2020). Influence of virtual objects' shadows and lighting coherence on distance perception in optical see-through augmented reality. *Journal of the Society for Information Display*, 28(2), 117-135.

## International conferences

- Peillard, E., Itoh, Y., Normand, J., Sanz, F. A., Moreau, G., Lécuyer, A., & Moreau, G. (2020). Can Retinal Projection Displays Improve Spatial Perception in Augmented Reality? In 2020 IEEE International Symposium on Mixed and Augmented Reality (ISMAR), (*in press*).
- Peillard, E., Argelaguet, F., Normand, J. M., Lécuyer, A., & Moreau, G. (2019). Studying Exocentric Distance Perception in Optical See-Through Augmented Reality. In 2019 IEEE International Symposium on Mixed and Augmented Reality (ISMAR) (pp. 115-122).
- Peillard, E., Thebaud, T., Normand, J.-M., Argelaguet, F., Moreau, G., & Lécuyer, A. (2019). Virtual Objects Look Farther on the Sides: The Anisotropy of Distance Perception in Virtual Reality. In 26th IEEE Virtual Reality Conference, VR 2019, Osaka, Japan. (**Best Paper Award**)





# BIBLIOGRAPHY

---

## A

- Adelstein, B. D., Lee, T. G., & Ellis, S. R., (2003), Head Tracking Latency in Virtual Environments: Psychophysics and a Model, *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, 4720, 2083–2087, <https://doi.org/10.1177/154193120304702001>
- Aksit, K., Lopes, W., Kim, J., Shirley, P., & Luebke, D., (2017), Near-eye varifocal augmented reality display using see-through screens, *ACM Transactions on Graphics*, 366, 1–13, <https://doi.org/10.1145/3130800.3130892>
- Al-Kalbani, M., Frutos-Pascual, M., & Williams, I., (2019), Virtual Object Grasping in Augmented Reality: Drop Shadows for Improved Interaction, *2019 11th International Conference on Virtual Worlds and Games for Serious Applications (VS-Games)*, 1–8, <https://doi.org/10.1109/VS-Games.2019.8864596>
- Allison, R. S., Gillam, B. J., & Vecellio, E., (2009), Binocular depth discrimination and estimation beyond interaction space, *Journal of Vision*, 91, 10–10, <https://doi.org/10.1167/9.1.10>
- Angell, R., (1974), The Geometry of Visibles, *Noûs*, 82, 87–117.
- Ashley, M. L., (1898), Concerning the Significance of Intensity of Light in Visual Estimates of Depth, *Psychological Review*, 56, 595–615, <https://doi.org/10.1037/h0068517>
- Avery, B., Sandor, C., & Thomas, B. H., (2009), Improving Spatial Perception for Augmented Reality X-Ray Vision, *2009 IEEE Virtual Reality Conference*, 79–82, <https://doi.org/10.1109/VR.2009.4811002>
- Aznar-Casanova, J. A., Matsushima, E. H., Da Silva, J. A., & Ribeiro-Filho, N. P., (2008), Can exocentric direction be dissociated from its exocentric distance in virtual environments?, *Perception & Psychophysics*, 703, 541–550, <https://doi.org/10.3758/PP.70.3.541>
- Azuma, R. T., (1997), A Survey of Augmented Reality, <https://doi.org/10.1162/pres.1997.6.4.355>

---

Azuma, R. T., Baillet, Y., Behringer, R., Feiner, S., Julier, S. J., & MacIntyre, B., (2001), Recent advances in augmented reality, *IEEE computer graphics and applications*, 216, 34–47.

## B

Baird, J. C., & Biersdorf, W. R., (1967), Quantitative functions for size and distance judgments, *Perception & Psychophysics*, 24, 161–166, <https://doi.org/10.3758/BF03210312>

Banakou, D., Groten, R., & Slater, M., (2013), Illusory ownership of a virtual child body causes overestimation of object sizes and implicit attitude changes, *11031*, <https://doi.org/10.1073/pnas.1306779110>

Bane, R., & Höllerer, T., (2004), Interactive tools for virtual X-ray vision in mobile augmented reality, *ISMAR 2004: Proceedings of the Third IEEE and ACM International Symposium on Mixed and Augmented Reality*, 231–239, <https://doi.org/10.1109/ISMAR.2004.36>

Bartz, D., Cunningham, D., Fischer, J., & Wallraven, C., (2008), The Role of Perception for Computer Graphics, *Eurographics state-of-the-art-reports*, 65–86, <https://doi.org/10.2312/egst.20081045>

Battro, A. M., Reggini, H. C., & Karts, C., (1978), Perspectives in open spaces: a geometrical application of the Thouless index, *Perception*, 75, 583–588, <https://doi.org/10.1068/p070583>

Battro, A. M., Di Pierro Netto, S., & Rozestraten, R. J., (1976), Riemannian geometries of variable curvature in visual space: visual alleys, horopters, and triangles in big open fields, *Perception*, 51, 9–23, <https://doi.org/10.1068/p050009>

Berkeley, G., (1709), *An essay toward a new theory of vision*, <https://doi.org/10.1037/11304-009>

Berning, M., Kleinert, D., Riedel, T., & Beigl, M., (2014), A study of depth perception in hand-held augmented reality using autostereoscopic displays, *ISMAR 2014 - IEEE International Symposium on Mixed and Augmented Reality - Science and Technology 2014, Proceedings*, 93–98, <https://doi.org/10.1109/ISMAR.2014.6948413>

Bian, Z., & Andersen, G. J., (2013), Aging and the perception of egocentric distance, *Psychology and Aging*, 283, 813–825, <https://doi.org/10.1037/a0030991>

- 
- Bian, Z., Braunstein, M. L., & Andersen, G. J., (2005), The ground dominance effect in the perception of 3-D layout, <https://doi.org/10.3758/BF03193534>
- Billinghurst, M., Clark, A., & Lee, G., (2015), A Survey of Augmented Reality, *Foundations and Trends® in Human-Computer Interaction*, 82-3, 73–272, <https://doi.org/10.1561/1100000049>
- Bimber, O., Emmerling, A., & Klemmer, T., (2005), Embedded entertainment with smart projectors, *Computer*, 381, 48–55, <https://doi.org/10.1109/MC.2005.17>
- Bimber, O., & Raskar, R., (2005), *Spatial augmented reality: Merging real and virtual worlds*, CRC press, <https://doi.org/10.1201/b10624>
- Bingham, G. P., Bradley, A., Bailey, M., & Vinner, R., (2001), Accommodation, occlusion, and disparity matching are used to guide reaching: A comparison of actual versus virtual environments., *Journal of Experimental Psychology: Human Perception and Performance*, 276, 1314–1334, <https://doi.org/10.1037//0096-1523.27.6.1314>
- Bingham, G. P., & Pagano, C. C., (1998), The necessity of a perception–action approach to definite distance perception: Monocular distance perception to guide reaching., *Journal of Experimental Psychology: Human Perception and Performance*, 241, 145.
- Biocca, F. A., & Lansing, E., (1998), Body : Adaptation to Visual Displacement in See-Through , Head-Mounted Displays, 73, 262–277.
- Blank, a. a., (1953), The Luneburg theory of binocular visual space., *Journal of the Optical Society of America*, 439, 717–727, <https://doi.org/10.1364/JOSA.43.000717>
- Blank, A. A., (1959), The Luneburg theory of binocular space perception, *Psychology: A study of a science*, 1, 395–426.
- Blank, A. A., (1978), Metric geometry in human binocular perception: Theory and fact, *Formal theories of visual perception*.
- Bodenheimer, B., Meng, J., Wu, H., Narasimham, G., Rump, B., McNamara, T. P., Carr, T. H., & Rieser, J. J., (2007), Distance Estimation in Virtual and Real Environments Using Bisection, *Proceedings of the 4th Symposium on Applied Perception in Graphics and Visualization*, 35–40, <https://doi.org/10.1145/1272582.1272589>
- Bowman, D., Kruijff, E., Jr., J. J. L., & Poupyrev, I., (2004), *3D User Interfaces: Theory and Practice*.
- Bruder, G., Argelaguet, F., Olivier, A.-H., & Lecuyer, A., (2015), Distance estimation in large immersive projection systems, revisited, <https://doi.org/10.1109/VR.2015.7223320>

- 
- Bruder, G., Argelaguet, F., Olivier, A.-H., & Lecuyer, A., (2016), CAVE Size Matters: Effects of Screen Distance and Parallax on Distance Estimation in Large Immersive Display Setups, *Presence*, 25 1, 1–16, [https://doi.org/10.1162/PRES{\\\_}\\\_a{\\\_}\\\_00241](https://doi.org/10.1162/PRES{\_}\_a{\_}\_00241)
- Buchmann, V., Nilsen, T., & Billinghamurst, M., (2005), Interaction With Partially Transparent Hands And Objects, *Sixth Australasian User Interface Conference*, 40, 17–20.
- Buck, L. E., Young, M. K., & Bodenheimer, B., (2018), A Comparison of Distance Estimation in HMD-Based Virtual Environments with Different HMD-Based Conditions, *ACM Transactions on Applied Perception*, 15 3, 1–15, <https://doi.org/10.1145/3196885>
- Burr, D. C., & Ross, J., (2002), Direct evidence that "speedlines" influence motion mechanisms., *The Journal of neuroscience : the official journal of the Society for Neuroscience*, 22 19, 8661–4, [https://doi.org/22/19/8661\[pil\]](https://doi.org/22/19/8661[pil])

## C

- Carnegie, K., & Rhee, T., (2015), Reducing Visual Discomfort with HMDs Using Dynamic Depth of Field, *IEEE Computer Graphics and Applications*, 35 5, 34–41, <https://doi.org/10.1109/MCG.2015.98>
- Carr, H. A., (1935), *An introduction to space perception.*, Longmans, Green.
- Casasanto, D., (2009), Embodiment of Abstract Concepts: Good and Bad in Right- and Left-Handers, *Journal of Experimental Psychology: General*, 138 3, 351–367, <https://doi.org/10.1037/a0015854>
- Chatterjea, R. G., Saha, A. K., & Biswas, D., (1974), Effect of variation in background, illumination and distance in depth perception, *Psychological Studies*, 19 2, 138–141.
- Cholewiak, S. A., Love, G. D., Srinivasan, P. P., Ng, R., & Banks, M. S., (2017), ChromaBlur: Rendering chromatic eye aberration improves accommodation and realism, *ACM Transactions on Graphics*, 36 6, 1–12, <https://doi.org/10.1145/3130800.3130815>
- Cidota, M. A., Clifford, R. M., Lukosch, S. G., & Billinghamurst, M., (2017), Using Visual Effects to Facilitate Depth Perception for Spatial Tasks in Virtual and Augmented Reality, *Adjunct Proceedings of the 2016 IEEE International Symposium on Mixed*

- 
- and Augmented Reality, ISMAR-Adjunct 2016*, 172–177, <https://doi.org/10.1109/ISMAR-Adjunct.2016.0070>
- Clemente, M., Rodríguez, A., Rey, B., & Alcañiz, M., (2014), Assessment of the influence of navigation control and screen size on the sense of presence in virtual reality using EEG, *Expert Systems with Applications*, 414 PART 2, 1584–1592, <https://doi.org/10.1016/j.eswa.2013.08.055>
- Collett, T. S., (1985), Extrapolating and interpolating surfaces in depth., *Proceedings of the Royal Society of London. Series B, Biological sciences*, 224 1234, 43–56, <https://doi.org/10.1098/RSPB.1985.0020>
- Cortes, G., Marchand, E., Brincin, G., & Lécuyer, A., (2018), MoSART: Mobile Spatial Augmented Reality for 3D Interaction With Tangible Objects, *Frontiers in Robotics and AI*, 5 AUG, 93, <https://doi.org/10.3389/frobt.2018.00093>
- Coules, J., (1955), Effect of photometric brightness on judgments of distance, *Journal of Experimental Psychology*, 50 1, 19–25, <https://doi.org/10.1037/h0044343>
- Creem-Regehr, S. H., & Proffitt, D. R., (1998), Two memories for geographical slant: Separation and interdependence of action and awareness, *Psychonomic Bulletin and Review*, 5 1, 22–36, <https://doi.org/10.3758/BF03209455>
- Creem-Regehr, S. H., Stefanucci, J. K., Thompson, W. B., Nash, N., & McCardell, M., (2015), Egocentric distance perception in the Oculus Rift (DK2). *SAP 2015: ACM SIGGRAPH Symposium on Applied Perception* (pp. 47–50), Association for Computing Machinery, Inc, <https://doi.org/10.1145/2804408.2804422>
- Creem-Regehr, S. H., Willemsen, P., Gooch, A. A., & Thompson, W. B., (2005), The Influences of Restricted Viewing Conditions on Egocentric Perception: Implications for Real and Virtual Environments, *Perception*, 34 2, 191–204.
- Cruz-Neira, C., Sandin, D. J., & DeFanti, T. A., (1993), Surround-screen projection-based virtual reality: The design and implementation of the CAVE, *Proceedings of the 20th Annual Conference on Computer Graphics and Interactive Techniques, SIGGRAPH 1993*, 135–142, <https://doi.org/10.1145/166117.166134>
- Cutting, J. E., (2003), Reconceiving perceptual space. *Looking into pictures: An interdisciplinary approach to pictorial space*. (pp. 215–238), MIT Press.

---

## D

- Da Silva, J. A., (1985), Scales for perceived egocentric distance in a large open field: comparison of three psychophysical methods., *The American journal of psychology*, 98 1, 119–144, <https://doi.org/10.2307/1422771>
- Daniels, N., (1974), *Thomas Reid's Inquiry: The Geometry of Visibles and the Case for Realism*, New York: B. Franklin.
- Davidson, R. J., (1992), Anterior cerebral asymmetry and the nature of emotion, *Brain and Cognition*, 20 1, 125–151, [https://doi.org/10.1016/0278-2626\(92\)90065-T](https://doi.org/10.1016/0278-2626(92)90065-T)
- Decety, J., Jeannerod, M., & Prablanc, C., (1989), The timing of mentally represented actions, *Behavioural Brain Research*, 34 1-2, 35–42, [https://doi.org/10.1016/S0166-4328\(89\)80088-9](https://doi.org/10.1016/S0166-4328(89)80088-9)
- Denion, E., Dugué, A.-E., Coffin-Pichonnet, S., Augy, S., & Mouriaux, F., (2014), Eye motion increases temporal visual field extent, *Acta Ophthalmologica*, 92 3, e200–e206, <https://doi.org/10.1111/aos.12106>
- Dewald, D. S., Evans, A. T., Welch, N., Gross, A., & Hill, G., (2016), 8-1: Invited Paper : The Avegant Glyph: Optical Design Considerations and Approach to Near-eye Display, *SID Symposium Digest of Technical Papers*, 47 1, 69–71, <https://doi.org/10.1002/sdtp.10609>
- Dey, A., Cunningham, A., & Sandor, C., (2010), Evaluating depth perception of photorealistic mixed reality visualizations for occluded objects in outdoor environments, *3DUI 2010 - IEEE Symposium on 3D User Interfaces 2010, Proceedings*, 127–128, <https://doi.org/10.1109/3DUI.2010.5444706>
- Dey, A., Jarvis, G., Sandor, C., & Reitmayr, G., (2012), Tablet versus phone: Depth perception in handheld augmented reality, *ISMAR 2012 - 11th IEEE International Symposium on Mixed and Augmented Reality 2012, Science and Technology Papers*, 187–196, <https://doi.org/10.1109/ISMAR.2012.6402556>
- Dey, A., & Sandor, C., (2014), Lessons learned: Evaluating visualizations for occluded objects in handheld augmented reality, *International Journal of Human Computer Studies*, 72 10-11, 704–716, <https://doi.org/10.1016/j.ijhcs.2014.04.001>
- Diaz, C., Walker, M., Szafir, D. A., & Szafir, D., (2017), Designing for Depth Perceptions in Augmented Reality, *2017 IEEE International Symposium on Mixed and Augmented Reality*, <https://doi.org/10.1109/ISMAR.2017.28>

- 
- Díaz, J., Ropinski, T., Navazo, I., Gobbetti, E., & Vázquez, P. P., (2017), An experimental study on the effects of shading in 3D perception of volumetric models, *Visual Computer*, 33 1, 47–61, <https://doi.org/10.1007/s00371-015-1151-6>
- Do, T. D., LaViola, J. J., & McMahan, R. P., (2020), The Effects of Object Shape, Fidelity, Color, and Luminance on Depth Perception in Handheld Mobile Augmented Reality, *2020 IEEE International Symposium on Mixed and Augmented Reality*, (in press).
- Dunn, D., Dong, Q., Fuchs, H., & Chakravarthula, P., (2018), Mitigating vergence-accommodation conflict for near-eye displays via deformable beamsplitters, *In* W. Osten, H. Stolle, & B. C. Kress (Eds.), *Digital Optics for Immersive Displays* (p. 104), <https://doi.org/10.1117/12.2314664>
- Dunn, D., Tippets, C., Torell, K., Kellnhofer, P., Akşit, K., Didyk, P., Myszkowski, K., Luebke, D., & Fuchs, H., (2017), Wide Field of View Varifocal Near-Eye Display Using See-Through Deformable Membrane Mirrors, *IEEE Transactions on Visualization and Computer Graphics*, 23 4, 1275–1284, <https://doi.org/10.1109/TVCG.2017.2657058>
- Durand, A. C., & Gould, G. M., (1910), A method of determining ocular dominance, *Journal of the American Medical Association*, 55 5, 369–370, <https://doi.org/10.1001/jama.1910.04330050007004>

## E

- Eckmeier, D., Kern, R., Egelhaaf, M., & Bischof, H.-J., (2013), Encoding of naturalistic optic flow by motion sensitive neurons of nucleus rotundus in the zebra finch (*Taeniopygia guttata*), *Frontiers in Integrative Neuroscience*, 7 SEP, 68, <https://doi.org/10.3389/fnint.2013.00068>
- Edgar, G. K., Pope, J. C., & Craig, I. R., (1993), Visual accommodation problems with head-up and helmet-mounted displays, *In* C. T. Bartlett & M. D. Cowan (Eds.), *International Society for Optics; Photonics*, <https://doi.org/10.1117/12.164705>
- Eggleston, R. G., Janson, W. P., & Aldrich, K. A., (1996), Virtual reality system effects on size-distance judgments in a virtual environment, *Proceedings - Virtual Reality Annual International Symposium*, 139–146, <https://doi.org/10.1109/vrais.1996.490521>



- 
- Elliott, D., (1987), The influence of walking speed and prior practice on locomotor distance estimation, *Journal of Motor Behavior*, 194, 476–485, <https://doi.org/10.1080/00222895.1987.10735425>
- Ellis, S. R., & Menges, B. M., (1997), Judgments of the distance to nearby virtual objects: interaction of viewing conditions and accommodative demand., *Presence (Cambridge, Mass.)*, 64, 452–460, <https://doi.org/10.1162/pres.1997.6.4.452>
- Ellis, S. R., & Bucher, U. J., (1994), Distance Perception of Stereoscopically Presented Virtual Objects Optically Superimposed on Physical Objects by a Head-Mounted See-Through Display, *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, 3819, 1300–1304, <https://doi.org/10.1177/154193129403801911>
- Ellis, S. R., Bucher, U. J., & Menges, B. M., (1995), The relationship of binocular convergence and errors in judged distance to virtual objects, *IFAC Proceedings Volumes*, 2815, 253–257, [https://doi.org/10.1016/S1474-6670\(17\)45241-4](https://doi.org/10.1016/S1474-6670(17)45241-4)
- Ellis, S. R., & Menges, B. M., (1998), Localization of virtual objects in the near visual field., *Human factors*, 403, 415–431, <https://doi.org/10.1518/001872098779591278>
- Epstein, R. A., & Baker, C. I., (2019), Scene Perception in the Human Brain, *Annual Review of Vision Science*, 51, 373–397, <https://doi.org/10.1146/annurev-vision-091718-014809>

## F

- Farne, M., (1977), Brightness as an indicator to distance: relative brightness per se or contrast with the background?, *Perception*, 63, 287–293, <https://doi.org/10.1068/p060287>
- Fasse, E. D., Hogan, N., Kay, B. A., & Mussa-Ivaldi, F. A., (2000), Haptic interaction with virtual objects: Spatial perception and motor control, *Biological Cybernetics*, 821, 69–83, <https://doi.org/10.1007/PL00007962>
- Fechner, G. T., Howes, D. H., & Boring, E. G., (1966), *Elements of psychophysics* (Vol. 1), Holt, Rinehart; Winston New York.
- Feiner, S. K., Webster, A. C., & Krueger III, T. E., (1995), Architectural Anatomy, *Presence - Teleoperators and Virtual Environments*, 43, 318–325.
- Fernandes, A. S., & Feiner, S. K., (2016), Combating VR sickness through subtle dynamic field-of-view modification, *2016 IEEE Symposium on 3D User Interfaces, 3DUI 2016 - Proceedings*, 201–210, <https://doi.org/10.1109/3DUI.2016.7460053>

- 
- Fincham, E. F., & Walton, J., (1957), The reciprocal actions of accommodation and convergence., *The Journal of physiology*, *137*3, 488–508, <https://doi.org/10.1113/jphysiol.1957.sp005829>
- Fischer, M. H., (2000), Do Irrelevant Depth Cues Affect the Comprehension of Bar Graphs?, *Applied Cognitive Psychology*, *14* 2, 151–162, [https://doi.org/10.1002/\(SICI\)1099-0720\(200003/04\)14:2<151::AID-ACP629>3.0.CO;2-Z](https://doi.org/10.1002/(SICI)1099-0720(200003/04)14:2<151::AID-ACP629>3.0.CO;2-Z)
- Fodor, J. A., (1983), *The Modularity of Mind*, <https://doi.org/10.7551/mitpress/4737.001.0001>
- Foley, J. M., (1968), Depth, size and distance in stereoscopic vision, *Perception & Psychophysics*, *3* 4, 265–274, <https://doi.org/10.3758/BF03212742>
- Foley, J. M., (1977), Effect of Distance Information and Range on Two Indices of Visually Perceived Distance, *Perception*, *6* 4, 449–460, <https://doi.org/10.1068/p060449>
- Foley, J. M., Ribeiro-Filho, N. P., & Da Silva, J. A., (2004), Visual perception of extent and the geometry of visual space, *Vision Research*, *44* 2, 147–156, <https://doi.org/10.1016/j.visres.2003.09.004>
- Fontana, G., (1420), *Bellicorum instrumentorum liber: cum figuris et fictitiis literis conscriptus*.
- Francis, G., & Kim, H., (2001), Perceived motion in orientational afterimages: Direction and speed, *Vision Research*, *41* 2, 161–172, [https://doi.org/10.1016/S0042-6989\(00\)00242-X](https://doi.org/10.1016/S0042-6989(00)00242-X)
- Franklin, N., & Tversky, B., (1990), Searching Imagined Environments, *Journal of Experimental Psychology: General*, *119* 1, 63–76, <https://doi.org/10.1037/0096-3445.119.1.63>
- Fuchs, P., & Moreau, G., (2003), Résultats de recherche Résultats Web Le traité de la réalité virtuelle - Volume 1, *Les Presses de l'Ecole des Mines de Paris*, *1*, 131.
- Fuhrmann, A., Hesina, G., Faure, F., & Gervautz, M., (1999), Occlusion in collaborative augmented environments, *Computers and Graphics (Pergamon)*, *23* 6, 809–819, [https://doi.org/10.1016/S0097-8493\(99\)00107-7](https://doi.org/10.1016/S0097-8493(99)00107-7)
- Furmanski, C., Azuma, R. T., & Daily, M., (2002), Augmented-reality visualizations guided by cognition: Perceptual heuristics for combining visible and obscured information, *Proceedings - International Symposium on Mixed and Augmented Reality, ISMAR 2002*, 215–224, <https://doi.org/10.1109/ISMAR.2002.1115091>

---

## G

- Gagnon, H. C., Na, D., Heiner, K., Stefanucci, J., Creem-Regehr, S. H., & Bodenheimer, B., (2020), The Role of Viewing Distance and Feedback on Affordance Judgments in Augmented Reality, *Proceedings - 2020 IEEE Conference on Virtual Reality and 3D User Interfaces, VR 2020*, 922–929, <https://doi.org/10.1109/VR46266.2020.00112>
- Galanter, E., & Galanter, P., (1973), Range estimates of distant visual stimuli, *Perception & Psychophysics*, *14* 2, 301–306, <https://doi.org/10.3758/BF03212395>
- Gauss, C. F., (1828), *Disquisitiones generales circa superficies curvas* (Vol. 1), Typis Dieterichianis.
- Geuss, M. N., McCardell, M. J., & Stefanucci, J. K., (2016), Fear Similarly Alters Perceptual Estimates of and Actions over Gaps (R. Canal-Bruland, Ed.), *PLOS ONE*, *11* 7, e0158610, <https://doi.org/10.1371/journal.pone.0158610>
- Geuss, M. N., Stefanucci, J. K., Creem-Regehr, S. H., & Thompson, W. B., (2012), Effect of viewing plane on perceived distances in real and virtual environments., *Journal of Experimental Psychology: Human Perception and Performance*, *38* 5, 1242–1253, <https://doi.org/10.1037/a0027524>
- Geuss, M. N., Stefanucci, J. K., Creem-Regehr, S. H., & Thompson, W. B., (2010), Can I Pass?: Using Affordances to Measure Perceived Size in Virtual Environments, *Proceedings of the 7th Symposium on Applied Perception in Graphics and Visualization - APGV '10*, 61–64, <https://doi.org/10.1145/1836248.1836259>
- Gibson, E. J., & Bergman, R., (1954), The effect of training on absolute estimation of distance over the ground., *Journal of Experimental Psychology*, *48* 6, 473–482, <https://doi.org/10.1037/h0055007>
- Gibson, E. J., Bergman, R., & Purdy, J., (1955), The effect of prior training with a scale of distance on absolute and relative judgments of distance over ground., *Journal of Experimental Psychology*, *50* 2, 97–105, <https://doi.org/10.1037/h0048518>
- Gibson, J. J., (1950), *The perception of the visual world.*, Houghton Mifflin.
- Gibson, J. J., (1979), The Theory of Affordances, In Boston: Houghton Mifflin (Ed.), *The Ecological Approach to Visual Perception* (pp. 127–137).
- Gilinsky, A. S., (1951), Perceived size and distance in visual space, *Psychological Review*, *58* 6, 460–482, <https://doi.org/10.1037/h0061505>

- 
- Glasser, A., Campbell, M. C. W. et al., (1998), Presbyopia and the optical changes in the human crystalline lens with age, *Vision research*, 382, 209–230.
- Gogel, W. C., (1969), The sensing of retinal size, *Vision Research*, 99, 1079–1094, [https://doi.org/10.1016/0042-6989\(69\)90049-2](https://doi.org/10.1016/0042-6989(69)90049-2)
- Gogel, W. C., (1976), An indirect method of measuring perceived distance from familiar size, *Perception & Psychophysics*, 206, 419–429, <https://doi.org/10.3758/BF03208276>
- Gogel, W. C., (1981), The role of suggested size in distance responses, *Perception & Psychophysics*, 302, 149–155, <https://doi.org/10.3758/BF03204473>
- Gogel, W. C., & Tietz, J. D., (1973), Absolute motion parallax and the specific distance tendency, *Perception & Psychophysics*, 132, 284–292, <https://doi.org/10.3758/BF03214141>
- Goldstein, B. E., (2005), Pictorial perception and art, *Blackwell Handbook of Sensation and Perception*, 344–378.
- Goldstein, B. E., & Brockmole, J. R., (2017), Introduction to Perception. *Sensation and Perception* (Tenth Edit, pp. 3–19).
- Gooding, L., Miller, M. E., Moore, J., & Kim, S.-H., (1991), Effect of viewing distance and disparity on perceived depth, In J. O. Merritt & S. S. Fisher (Eds.), *Stereoscopic Displays and Applications II* (pp. 259–266), SPIE, <https://doi.org/10.1117/12.46314>
- Grechkin, T. Y., Nguyen, T. D., Plumert, J. M., Cremer, J. F., & Kearney, J. K., (2010), How does presentation method and measurement protocol affect distance estimation in real and virtual environments?, *ACM Transactions on Applied Perception*, 74, 1–18, <https://doi.org/10.1145/1823738.1823744>
- Grossenbacher, P. G., & Lovelace, C. T., (2001), Mechanisms of synesthesia: Cognitive and physiological constraints, [https://doi.org/10.1016/S1364-6613\(00\)01571-0](https://doi.org/10.1016/S1364-6613(00)01571-0)

## H

- Hardy, L. H., Rand, G., & Rittler, M. C., (1947), The ishihara test as a means of detecting and analyzing defective color vision, *Journal of General Psychology*, 361, 79–106, <https://doi.org/10.1080/00221309.1947.9918108>
- Heinrich, F., Bornemann, K., Lawonn, K., & Hansen, C., (2019), Depth Perception in Projective Augmented Reality: An Evaluation of Advanced Visualization Techniques,

---

25th ACM Symposium on Virtual Reality Software and Technology on - VRST '19, 1–11, <https://doi.org/10.1145/3359996.3364245>

- Hoffman, D. M., Girshick, A. R., Akeley, K., & Banks, M. S., (2008), Vergence–accommodation conflicts hinder visual performance and cause visual fatigue, *Journal of Vision*, 83, 33, <https://doi.org/10.1167/8.3.33>
- Hoffman, W. C., (1966), The Lie algebra of visual perception, *Journal of Mathematical Psychology*, 31, 65–98, [https://doi.org/10.1016/0022-2496\(66\)90005-8](https://doi.org/10.1016/0022-2496(66)90005-8)
- Hoffman, W. C., (1980), Subjective geometry and geometric psychology, *Mathematical Modelling*, 14, 349–367, [https://doi.org/10.1016/0270-0255\(80\)90045-7](https://doi.org/10.1016/0270-0255(80)90045-7)
- Holliman, N. S., Dodgson, N. A., Favolora, G. E., & Pockett, L., (2011), Three-dimensional displays: A review and applications analysis, *IEEE Transactions on Broadcasting*, 572 PART 2, 362–371, <https://doi.org/10.1109/TBC.2011.2130930>
- Howard, I. P., (2002), *Seeing in depth, Vol. 1: Basic mechanisms.*, University of Toronto Press.
- Hu, X., & Hua, H., (2014), High-resolution optical see-through multi-focal-plane head-mounted display using freeform optics, *Optics Express*, 2211, 13896, <https://doi.org/10.1364/oe.22.013896>
- Hua, H., & Javidi, B., (2014), A 3D integral imaging optical see-through head-mounted display, *Optics Express*, 2211, 13484, <https://doi.org/10.1364/oe.22.013484>
- Huang, T. H., Yang, Y. H., Liao, H. I., Yeh, S. L., & Chen, H. H., (2012), Directing visual attention by subliminal cues, *Proceedings - International Conference on Image Processing, ICIP*, 1081–1084, <https://doi.org/10.1109/ICIP.2012.6467051>
- Hubel, D. H., & Wiesel, T. N., (1965), Receptive fields and functional architecture in two nonstriate visual areas (18 and 19) of the cat., *Journal of Neurophysiology*, 282, 229–289.
- Hubona, G. S., Wheeler, P. N., Shirah, G. W., & Brandt, M., (1999), The Relative Contributions of Stereo, Lighting, and Background Scenes in Promoting 3D Depth Visualization, *ACM Transactions on Computer-Human Interaction*, 63, 214–242, <https://doi.org/10.1145/329693.329695>

## I

- Ijsselstein, W., De Ridder, H., Freeman, J., Avons, S. E., & Bouwhuis, D., (2001), Effects of stereoscopic presentation, image motion, and screen size on subjective and

- 
- objective corroborative measures of presence, *Presence: Teleoperators and Virtual Environments*, 103, 298–311, <https://doi.org/10.1162/105474601300343621>
- Indow, T., (1967), Two Interpretations of Binocular Visual Space: Hyperbolic and Euclidean., *Annals of the Japan Association for the Philosophy of Science*, 51–64.
- Indow, T., (1979), Alleys in visual space\* 1, *Journal of Mathematical Psychology*, 193, 221–258.
- Indow, T., (1991), A Critical Review of Luneburg 's Model With Regard to Global Structure of Visual Space, 983, 430–453.
- Interrante, V., Ries, B., & Anderson, L., (2006), Distance Perception in Immersive Virtual Environments, Revisited, *IEEE Virtual Reality Conference (VR 2006)*, JANUARY 2006, 3–10, <https://doi.org/10.1109/VR.2006.52>
- Itoh, Y., Langlotz, T., Iwai, D., Kiyokawa, K., & Amano, T., (2019), Light Attenuation Display: Subtractive See-Through Near-Eye Display via Spatial Color Filtering, *IEEE Transactions on Visualization and Computer Graphics*, 255, 1951–1960, <https://doi.org/10.1109/TVCG.2019.2899229>

## J

- Jaeae-Aro, K.-M., & Kjelldahl, L., (1997), Effects of image resolution on depth perception in stereo and nonstereo images, In S. S. Fisher, J. O. Merritt, & M. T. Bolas (Eds.), *Stereoscopic Displays and Virtual Reality Systems IV* (p. 319), SPIE, <https://doi.org/10.1117/12.274474>
- Jerome, C. J., & Witmer, B. G., (2005), The Perception and Estimation of Egocentric Distance in Real and Augmented Reality Environments, *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, 2249–2252, <https://doi.org/10.1037/e512752010-001>
- Jiang, Y., & Mark, L. S., (1994), The effect of gap depth on the perception of whether a gap is crossable, *Perception & Psychophysics*, 566, 691–700, <https://doi.org/10.3758/BF03208362>
- Johnson, P. V., Parnell, J. A. Q., Kim, J., Saunter, C. D., Love, G. D., & Banks, M. S., (2016), Dynamic lens and monovision 3D displays to improve viewer comfort, *Opt. Express*, 2411, 11808–11827, <https://doi.org/10.1364/OE.24.011808>
- Jones, J. A., Swan, J. E., Singh, G., & Ellis, S. R., (2011), Peripheral visual information and its effect on distance judgments in virtual and augmented environments, *Pro-*

---

*ceedings of the ACM SIGGRAPH Symposium on Applied Perception in Graphics and Visualization - APGV '11*, 29, <https://doi.org/10.1145/2077451.2077457>

Jones, J. A., Swan, J. E., Singh, G., Kolstad, E. E., & Ellis, S. R., (2008), The effects of virtual reality, augmented reality, and motion parallax on egocentric depth perception, *APGV 2008 - Proceedings of the Symposium on Applied Perception in Graphics and Visualization*, 9–14, <https://doi.org/10.1145/1394281.1394283>

Julesz, B., (1971), *Foundations of cyclopean perception.*, U. Chicago Press.

## K

Kalkofen, D., Mendez, E., & Schmalstieg, D., (2007), Interactive focus and context visualization for augmented reality, *2007 6th IEEE and ACM International Symposium on Mixed and Augmented Reality, ISMAR*, <https://doi.org/10.1109/ISMAR.2007.4538846>

Kalkofen, D., Mendez, E., & Schmalstieg, D., (2009), Comprehensible visualization for augmented reality, *IEEE Transactions on Visualization and Computer Graphics*, 15(2), 193–204, <https://doi.org/10.1109/TVCG.2008.96>

Kalkofen, D., Veas, E., Zollmann, S., Steinberger, M., & Schmalstieg, D., (2013), Adaptive ghosted views for Augmented Reality, *2013 IEEE International Symposium on Mixed and Augmented Reality, ISMAR 2013*, 1–9, <https://doi.org/10.1109/ISMAR.2013.6671758>

Kaplan, G. A., (1969), Kinetic disruption of optical texture: The perception of depth at an edge, *Perception & Psychophysics*, 6(4), 193–198, <https://doi.org/10.3758/BF03207015>

Kaplan, S., (1973), Cognitive maps in perception and thought, *Image and environment: Cognitive mapping and spatial behavior*, 63–78.

Kappers, A. M. L., & Koenderink, J. J., (1999), Haptic Perception of Spatial Relations, *Perception*, 28(6), 781–795, <https://doi.org/10.1068/p2930>

Kelly, J. W., Cherep, L. A., & Siegel, Z. D., (2017), Perceived Space in the HTC Vive, *ACM Transactions on Applied Perception*, 15(1), 1–16, <https://doi.org/10.1145/3106155>

Kelly, J. W., Hammel, W., Sjolund, L. A., & Siegel, Z. D., (2015), Frontal extents in virtual environments are not immune to underperception, *Attention, Perception, and Psychophysics*, 77(6), 1848–1853, <https://doi.org/10.3758/s13414-015-0948-8>

- 
- Kenyon, R. V., Sandin, D., Smith, R. C., Pawlicki, R., & Defanti, T., (2007), Size-Constancy in the CAVE, *Presence: Teleoperators and Virtual Environments*, 162, 172–187, <https://doi.org/10.1162/pres.16.2.172>
- Kim, H. J., Lee, S. K., Piao, M. L., Kim, N., & Park, J. H., (2015), Three-dimensional holographic head mounted display using holographic optical element, *2015 IEEE International Conference on Consumer Electronics (ICCE)*, 132–133, <https://doi.org/10.1109/ICCE.2015.7066351>
- Kim, H., & Francis, G., (1998), A computational and perceptual account of motion lines, *Perception*, 277, 785–797, <https://doi.org/10.1068/p270785>
- Kitaoka, A., (1998), Apparent contraction of edge angles, *Perception*, 2710, 1209–1219, <https://doi.org/10.1068/p271209>
- Knapp, J. M., & Loomis, J. M., (2004), Limited Field of View of Head-Mounted Displays Is Not the Cause of Distance Underestimation in Virtual Environments, *Presence: Teleoperators and Virtual Environments*, 135, 572–577, <https://doi.org/10.1162/1054746042545238>
- Koenderink, J. J., Van Doorn, A. J., Kappers, A. M. L., & Lappin, J. S., (2002), Large-scale visual frontoparallels under full-cue conditions, *Perception*, 31 12, 1467–1475, <https://doi.org/10.1068/p3295>
- Koenderink, J. J., van Doorn, A. J., Kappers, A. M., Doumen, M. J., & Todd, J. T., (2008), Exocentric pointing in depth, *Vision Research*, 48 5, 716–723, <https://doi.org/10.1016/j.visres.2007.12.002>
- Koenderink, J. J., van Doorn, A. J., & Lappin, J. S., (2000), Direct measurement of the curvature of visual space, *Perception*, 29 1, 69–79, <https://doi.org/10.1068/p2921>
- Konrad, R., Cooper, E. A., & Wetzstein, G., (2016), Novel Optical Configurations for Virtual Reality, *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems - CHI '16*, 1211–1220, <https://doi.org/10.1145/2858036.2858140>
- Koretz, J. F., & Handelman, G. H., (1988), How the Human Eye Focuses, *Scientific American*, 259 1, 92–99.
- Kruijff, E., Swan, J. E., & Feiner, S., (2010), Perceptual issues in augmented reality revisited, *2010 IEEE International Symposium on Mixed and Augmented Reality*, 3–12, <https://doi.org/10.1109/ISMAR.2010.5643530>



- 
- Kudoh, N., (2005), Dissociation between Visual Perception of Allocentric Distance and Visually Directed Walking of its Extent, *Perception*, *34* 11, 1399–1416, <https://doi.org/10.1068/p5444>
- Kuhl, S. A., Thompson, W. B., & Creem-Regehr, S. H., (2006), Minification influences spatial judgments in virtual environments, *APGV '06: Proceedings of the 3rd symposium on Applied perception in graphics and visualization*, *1212*, 15, <https://doi.org/10.1145/1140491.1140494>
- Kuipers, B., (1983), The cognitive map: Could it have been any other way? *Spatial orientation* (pp. 345–359), Springer.
- Kurt, W., (2007), Virtual reality: an interview with Jaron Lanier, *Serials Review*, *33* 3, 190–195, <https://doi.org/10.1016/j.serrev.2007.05.009>
- Kytö, M., Mäkinen, A., Häkkinen, J., & Oittinen, P., (2013), Improving relative depth judgments in augmented reality with auxiliary augmentations, *ACM Transactions on Applied Perception*, *10* 1, 1–21, <https://doi.org/10.1145/2422105.2422111>
- Kytö, M., Mäkinen, A., Tossavainen, T., & Oittinen, P., (2014), Stereoscopic depth perception in video see-through augmented reality within action space, *Journal of Electronic Imaging*, *23* 1, 011006, <https://doi.org/10.1117/1.JEI.23.1.011006>

## L

- Lampton, D. R., McDonald, D. P., Singer, M., & Bliss, J. P., (1995), Distance Estimation in Virtual Environments, *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, *39* 20, 1268–1272, <https://doi.org/10.1177/154193129503902006>
- Landy, M. S., Maloney, L. T., Johnston, E. B., & Young, M., (1995), Measurement and modeling of depth cue combination: in defense of weak fusion, *Vision Research*, *35* 3, 389–412, [https://doi.org/10.1016/0042-6989\(94\)00176-M](https://doi.org/10.1016/0042-6989(94)00176-M)
- Landy, M. S., Maloney, L. T., & Young, M. J., (1991), Psychophysical estimation of the human depth combination rule, In P. S. Schenker (Ed.), *Sensor Fusion III: 3D Perception and Recognition* (pp. 247–254), SPIE, <https://doi.org/10.1117/12.25260>
- Langbehn, E., Raupp, T., Bruder, G., Steinicke, F., Bolte, B., & Lappe, M., (2016), Visual blur in immersive virtual environments, *Proceedings of the 22nd ACM Conference*

- 
- on *Virtual Reality Software and Technology - VRST '16*, 241–250, <https://doi.org/10.1145/2993369.2993379>
- Lanier, J., (2017), *Dawn of the new everything: A journey through virtual reality*, Random House.
- Lappin, J. S., Shelton, A. L., & Rieser, J. J., (2006), Environmental context influences visually perceived distance, *Perception & Psychophysics*, *68*4, 571–581, <https://doi.org/10.3758/BF03208759>
- LaViola Jr, J. J., Kruijff, E., McMahan, R. P., Bowman, D., & Poupyrev, I. P., (2017), *3D user interfaces: theory and practice*, Addison-Wesley Professional.
- Lee, C., Bonebrake, S., Höllerer, T., & Bowman, D. A., (2010), The role of latency in the validity of AR simulation, *Proceedings - IEEE Virtual Reality*, 11–18, <https://doi.org/10.1109/VR.2010.5444820>
- Levin, C. A., & Haber, R. N., (1993), Visual angle as a determinant of perceived inter-object distance, *Perception & Psychophysics*, *54*2, 250–259, <https://doi.org/10.3758/BF03211761>
- Li, B., Zhang, R., & Kuhl, S., (2014), Minication affects action-based distance judgments in oculus rift HMDs, *Proceedings of the ACM Symposium on Applied Perception - SAP '14*, <https://doi.org/10.1145/2628257.2628273>
- Lin, C. J., & Woldegiorgis, B. H., (2015), Interaction and visual performance in stereoscopic displays: A review, *Journal of the Society for Information Display*, *23*7, 319–332, <https://doi.org/10.1002/jsid.378>
- Lin, C. J., Woldegiorgis, B. H., Caesaron, D., & Cheng, L. Y., (2015), Distance estimation with mixed real and virtual targets in stereoscopic displays, *Displays*, *36*, 41–48, <https://doi.org/10.1016/j.displa.2014.11.006>
- Lin, J., Cheng, D., Yao, C., & Wang, Y., (2017), Retinal projection head-mounted display, *Frontiers of Optoelectronics*, *10*1, 1–8, <https://doi.org/10.1007/s12200-016-0662-8>
- Linn, M. C., & Petersen, A. C., (1986), A meta-analysis of gender differences in spatial ability: Implications for mathematics and science achievement, *The psychology of gender: Advances through meta-analysis*, 67–101.
- Lisle, L., Tanous, K., Kim, H., Gabbard, J. L., & Bowman, D. A., (2018), Effect of volumetric displays on depth perception in augmented reality, *Proceedings - 10th International ACM Conference on Automotive User Interfaces and Interactive Vehicular Applications, AutomotiveUI 2018*, 155–163, <https://doi.org/10.1145/3239060.3239083>

- 
- Liu, J. M., Narasimham, G., Stefanucci, J. K., Creem-Regehr, S. H., & Bodenheimer, B., (2020), Distance Perception in Modern Mobile Augmented Reality, *2020 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW)*, 196–200, <https://doi.org/10.1109/VRW50115.2020.00039>
- Liu, S., Cheng, D., & Hua, H., (2008), An optical see-through head mounted display with addressable focal planes, *2008 7th IEEE/ACM International Symposium on Mixed and Augmented Reality*, 33–42, <https://doi.org/10.1109/ISMAR.2008.4637321>
- Livingston, M. A., Ai, Z., & Decker, J. W., (2009), A user study towards understanding stereo perception in head-worn augmented reality displays, *2009 8th IEEE International Symposium on Mixed and Augmented Reality*, 53–56, <https://doi.org/10.1109/ISMAR.2009.5336496>
- Livingston, M. A., Ai, Z., Swan, J. E., & Smallman, H. S., (2009), Indoor vs. Outdoor Depth Perception for Mobile Augmented Reality, *2009 IEEE Virtual Reality Conference*, 55–62, <https://doi.org/10.1109/VR.2009.4810999>
- Livingston, M. A., Swan, J. E., Gabbard, J. L., Hollerer, T. H., Hix, D., Julier, S. J., Baillot, Y., & Brown, D., (2003), Resolving multiple occluded layers in augmented reality, *Proceedings - 2nd IEEE and ACM International Symposium on Mixed and Augmented Reality, ISMAR 2003*, 56–65, <https://doi.org/10.1109/ISMAR.2003.1240688>
- Livingston, M. A., Zambaka, C., Swan, J. E., & Smallman, H., (2005), Objective measures for the effectiveness of augmented reality, *IEEE Proceedings. VR 2005. Virtual Reality, 2005.*, 2005, 287–288, <https://doi.org/10.1109/VR.2005.1492798>
- Lombard, M., Reich, R., Grabe, M., Bracken, C., & Ditton, T., (2000), Presence and television., *Human Communication Research*, 26 1, 75–98, <https://doi.org/10.1111/j.1468-2958.2000.tb00750.x>
- Long, M., Burnett, G., Hardy, R., & Allen, H., (2015), Depth discrimination between augmented reality and real-world targets for vehicle head-up displays, *Proceedings of the 7th International Conference on Automotive User Interfaces and Interactive Vehicular Applications - AutomotiveUI '15*, 72–79, <https://doi.org/10.1145/2799250.2799292>
- Loomis, J. M., (2014), Three theories for reconciling the linearity of egocentric distance perception with distortion of shape on the ground plane, *Psychology and Neuroscience*, 7 3, 245–251, <https://doi.org/10.3922/j.psns.2014.034>

- 
- Loomis, J. M., Da Silva, J. A., Fujita, N., & Fukusima, S. S., (1992), Visual Space Perception and Visually Directed Action, *Journal of Experimental Psychology: Human Perception and Performance*, 184, 906–921, <https://doi.org/10.1037/0096-1523.18.4.906>
- Loomis, J. M., & Knapp, J. M. J., (2003), Visual perception of egocentric distance in real and virtual environments. *Virtual and adaptive environments: Applications, implications, and human performance issues* (pp. 21–46), CRC Press, <https://doi.org/10.1201/9781410608888>
- Loomis, J. M., & Philbeck, J. W., (2008), Measuring spatial perception with spatial updating and action., *Carnegie Symposium on Cognition, 2006, Pittsburgh, PA, US*.
- Loomis, J. M., Philbeck, J. W., & Zahorik, P., (2002), Dissociation between location and shape in visual space., *Journal of experimental psychology. Human perception and performance*, 285, 1202–12.
- Looser, J., Grasset, R., & Billinghurst, M., (2007), A 3D flexible and tangible magic lens in augmented reality, *2007 6th IEEE and ACM International Symposium on Mixed and Augmented Reality, ISMAR*, 1–4, <https://doi.org/10.1109/ISMAR.2007.4538825>
- Luneburg, R. K., (1950), The Metric of Binocular Visual Space, *Journal of the Optical Society of America*, 40, 627, <https://doi.org/10.1364/JOSA.40.000627>
- Luneburg, R. K., (1947), *Mathematical analysis of binocular vision.*, Princeton University Press.
- Luo, X., Kenyon, R., Kamper, D., Sandin, D., & DeFanti, T., (2007), The effects of scene complexity, stereovision, and motion parallax on size constancy in a virtual environment, *Proceedings - IEEE Virtual Reality*, 59–66, <https://doi.org/10.1109/VR.2007.352464>

## M

- Machover, C., & Tice, S., (1994), Virtual Reality, *IEEE Computer Graphics and Application*, January, 15–16.
- Maimone, A., Georgiou, A., & Kollin, J. S., (2017), Holographic near-eye displays for virtual and augmented reality, *ACM Transactions on Graphics*, 364, 1–16, <https://doi.org/10.1145/3072959.3073624>

- 
- Mamassian, P., Knill, D. C., & Kersten, D., (1998), The perception of cast shadows, [https://doi.org/10.1016/S1364-6613\(98\)01204-2](https://doi.org/10.1016/S1364-6613(98)01204-2)
- Marr, D., (1982), *Vision. A Computational Investigation Into the Human Representation and Processing of Visual Information*, W. H. Freeman; Company, <https://doi.org/10.1086/413352>
- Martens, T. G., & Ogle, K. N., (1959), Observations on accommodative convergence; especially its nonlinear relationships., *American journal of ophthalmology*, *47* 1 Part 2, 455–62.
- McCandless, J. W., Ellis, S. R., & Adelstein, B. D., (2000), Localization of a Time-Delayed, Monocular Virtual Object Superimposed on a Real Environment, *Presence: Teleoperators and Virtual Environments*, *9* 1, 15–24, <https://doi.org/10.1162/105474600566583>
- Mcgurk, H., & Macdonald, J., (1976), Hearing lips and seeing voices, *Nature*, *264* 5588, 746–748, <https://doi.org/10.1038/264746a0>
- Medeiros, D., Sousa, M., Mendes, D., Raposo, A., & Jorge, J., (2016), Perceiving Depth: Optical Versus Video See-through, *Proceedings of the 22nd ACM Conference on Virtual Reality Software and Technology - VRST '16, 02-04-Nove*, 237–240, <https://doi.org/10.1145/2993369.2993388>
- Mendez, E., Kalkofen, D., & Schmalstieg, D., (2006), Interactive context-driven visualization tools for augmented reality, *Proceedings - ISMAR 2006: Fifth IEEE and ACM International Symposium on Mixed and Augmented Reality*, 209–218, <https://doi.org/10.1109/ISMAR.2006.297816>
- Mendez, E., & Schmalstieg, D., (2009), Importance masks for revealing occluded objects in augmented reality, *Proceedings of the 16th ACM Symposium on Virtual Reality Software and Technology - VRST '09*, <https://doi.org/10.1145/1643928.1643988>
- Mershon, D. H., Kennedy, M., & Falacara, G., (1977), On the Use of ‘Calibration Equations’ in Perception Research, *Perception*, *6* 3, 299–311, <https://doi.org/10.1068/p060299>
- Messing, R., & Durgin, F. H., (2005), Distance Perception and the Visual Horizon in Head-Mounted Displays, *ACM Transactions on Applied Perception (TAP)*, *2* 3, 234–250, <https://doi.org/10.1145/1077399.1077403>
- Mikkola, M., Boev, A., & Gotchev, A., (2010), Relative importance of depth cues on portable autostereoscopic display, *MoViD '10*, *6* 3, <https://doi.org/10.1145/1878022.1878038>

- 
- Milgram, P., & Kishino, F., (1994), A taxonomy of mixed reality visual displays, *IEICE Transactions on Information Systems*, *E77-D12*, 1–15.
- Milgram, P., Takemura, H., Ustimi, A., & Kishino, F., (1994), Augmented Reality: A class of display on the reality-virtuality continuum, *Telemanipulator and Telepresence Technologies*, *2351*, 282–292.
- Milner, A. D., Goodale, M. A., & Vingrys, A. J., (1995), The visual brain in action (Vol. 27), *Oxford Psychology Series*. Oxford University Press, Oxford.
- Moehring, M., Gloystein, A., & Doerner, R., (2009), Issues with virtual space perception within reaching distance: Mitigating adverse effects on applications using hmds in the automotive industry, *Proceedings - IEEE Virtual Reality*, 223–226, <https://doi.org/10.1109/VR.2009.4811027>
- Mohler, B. J., Bülthoff, H. H., Thompson, W. B., & Creem-Regehr, S. H., (2008), A full-body avatar improves egocentric distance judgments in an immersive virtual environment, *APGV '08: Proceedings of the 5th symposium on Applied perception in graphics and visualization*, August, 194, <https://doi.org/http://doi.acm.org/10.1145/1394281.1394323>
- Mohler, B. J., Creem-Regehr, S. H., & Thompson, W. B., (2006), The influence of feedback on egocentric distance judgments in real and virtual environments, *Proceedings of the 3rd symposium on Applied perception in graphics and visualization - APGV '06*, 9, <https://doi.org/10.1145/1140491.1140493>
- Mohler, B. J., Creem-Regehr, S. H., Thompson, W. B., & Bülthoff, H. H., (2010), The Effect of Viewing a Self-Avatar on Distance Judgments in an HMD-Based Virtual Environment, *Presence: Teleoperators and Virtual Environments*, *193*, 230–242, <https://doi.org/10.1162/pres.19.3.230>
- Mon-Williams, M., & Tresilian, J. R., (2000), Ordinal depth information from accommodation?, *Ergonomics*, *433*, 391–404, <https://doi.org/10.1080/001401300184486>

## N

- Nabiyouni, M., Scerbo, S., Bowman, D. A., & Höllerer, T., (2017), Relative Effects of Real-world and Virtual-World Latency on an Augmented Reality Training Task: An AR Simulation Experiment, *Frontiers in ICT*, *3 JAN*, 30, <https://doi.org/10.3389/fict.2016.00034>

- 
- Naceri, A., & Chellali, R., (2012), The effect of isolated disparity on depth perception in real and virtual environments, *Proceedings - IEEE Virtual Reality, March 2012*, 107–108, <https://doi.org/10.1109/VR.2012.6180905>
- Naceri, A., Chellali, R., & Hoinville, T., (2011), Depth Perception Within Peripersonal Space Using Head-Mounted Display, *Presence: Teleoperators and Virtual Environments, 203*, 254–272, [https://doi.org/10.1162/PRES{\\\_}a{\\\_}00048](https://doi.org/10.1162/PRES{\_}a{\_}00048)
- Nakayama, K., (1996), Binocular visual surface perception., *Proceedings of the National Academy of Sciences of the United States of America, 932*, 634–9, <https://doi.org/10.1073/PNAS.93.2.634>
- Nawrot, M., & Stockert, C., (2010), Motion parallax in movies: Background motion, eye movement signals, and depth, *Journal of Vision, 58*, 644–644, <https://doi.org/10.1167/5.8.644>
- Nestmann, S., Karnath, H.-O., Bühlhoff, H. H., & Nikolas de Winkel, K., (2020), Changes in the perception of upright body orientation with age (R. Baurès, Ed.), *PLOS ONE, 155*, e0233160, <https://doi.org/10.1371/journal.pone.0233160>
- Niehorster, D. C., Li, L., & Lappe, M., (2017), The accuracy and precision of position and orientation tracking in the HTC vive virtual reality system for scientific research, *i-Perception, 83*, 1–23, <https://doi.org/10.1177/2041669517708205>
- Niketeghad, S., & Pouratian, N., (2019), Brain Machine Interfaces for Vision Restoration: The Current State of Cortical Visual Prosthetics, <https://doi.org/10.1007/s13311-018-0660-1>
- Nilsson, N. C., Serafin, S., & Nordahl, R., (2014), Establishing the range of perceptually natural visual walking speeds for virtual walking-in-place locomotion, *IEEE Transactions on Visualization and Computer Graphics, 204*, 569–578, <https://doi.org/10.1109/TVCG.2014.21>
- Norman, J. F., Todd, J. T., Perotti, V. J., & Tittle, J. S., (1996), The Visual Perception of Three-Dimensional Length, *Journal of Experimental Psychology: Human Perception and Performance, 221*, 173–186, <https://doi.org/10.1037/0096-1523.22.1.173>

## O

- Ogle, K. N., (1953), Precision and validity of stereoscopic depth perception from double images., *Journal of the Optical Society of America, 4310*, 907–913, <https://doi.org/10.1364/josa.43.000906>

---

## P

- Peck, T. C., Sockol, L. E., & Hancock, S. M., (2020), Mind the Gap: The Underrepresentation of Female Participants and Authors in Virtual Reality Research, *IEEE Transactions on Visualization and Computer Graphics*, 26 5, 1945–1954, <https://doi.org/10.1109/TVCG.2020.2973498>
- Pedhazur, E. J., (1975), *Multiple Regression in Behavioral Research*. (Vol. 70), <https://doi.org/10.2307/2285468>
- Peer, A., & Ponto, K., (2016), Perceptual space warping: Preliminary exploration, *2016 IEEE Virtual Reality (VR)*, 261–262, <https://doi.org/10.1109/VR.2016.7504753>
- Peer, A., & Ponto, K., (2017), Evaluating perceived distance measures in room-scale spaces using consumer-grade head mounted displays, *2017 IEEE Symposium on 3D User Interfaces, 3DUI 2017 - Proceedings*, 83–86, <https://doi.org/10.1109/3DUI.2017.7893321>
- Philbeck, J. W., & Loomis, J. M., (1997), Comparison of two indicators of perceived egocentric distance under full-cue and reduced-cue conditions., *Journal of Experimental Psychology: Human Perception and Performance*, 23 1, 72–85, <https://doi.org/10.1037/0096-1523.23.1.72>
- Phillips, L., Ries, B., Interrante, V., Kaeding, M., & Anderson, L., (2009), Distance perception in NPR immersive virtual environments, revisited, *APGV '09: Proceedings of the 6th Symposium on Applied Perception in Graphics and Visualization*, 11, <https://doi.org/10.1145/1620993.1620996>
- Ping, J., Thomas, B. H., Baumeister, J., Guo, J., Weng, D., & Liu, Y., (2020), Effects of shading model and opacity on depth perception in optical see-through augmented reality, *Journal of the Society for Information Display*, jsid.947, <https://doi.org/10.1002/jsid.947>
- Plato, (400), *The Republic: Book VII*.
- Plopski, A., Moser, K. R., Kiyokawa, K., Swan, J. E., & Takemura, H., (2016), Spatial consistency perception in optical and video see-through head-mounted augmentations, *Proceedings - IEEE Virtual Reality, 2016-July*, 265–266, <https://doi.org/10.1109/VR.2016.7504755>
- Plumert, J. M., Kearney, J. K., Cremer, J. F., & Recker, K., (2005), Distance perception in real and virtual environments, *ACM Transactions on Applied Perception*, 2 3, 216–233, <https://doi.org/10.1145/1077399.1077402>



- 
- Pointon, G., Thompson, C., Creem-Regehr, S. H., Stefanucci, J., Joshi, M., Paris, R., & Bodenheimer, B., (2018), Judging action capabilities in augmented reality, *Proceedings of the 15th ACM Symposium on Applied Perception - SAP '18*, 1–8, <https://doi.org/10.1145/3225153.3225168>
- Prablanc, C., Echallier, J. F., Komilis, E., & Jeannerod, M., (1979), Optimal response of eye and hand motor systems in pointing at a visual target, *Biological cybernetics*, *35*, 113–124.
- Predebon, J., (1991), Spatial judgments of exocentric extents in an open-field situation: Familiar versus unfamiliar size, *Perception & Psychophysics*, *50*, 361–366, <https://doi.org/10.3758/BF03212228>
- Proffitt, D. R., (2006), Embodied Perception and the Economy of Action, *Perspectives on Psychological Science*, *1*, 110–122, <https://doi.org/10.1111/j.1745-6916.2006.00008.x>
- Purdy, J., & Gibson, E. J., (1955), Distance judgment by the method of fractionation, *Journal of Experimental Psychology*, *50*, 374–380, <https://doi.org/10.1037/h0043157>

## R

- Ramachandran, V. S., & Cavanagh, P., (1985), Subjective contours capture stereopsis, *Nature*, *317*, 527–530, <https://doi.org/10.1038/317527a0>
- Reichelt, S., Häussler, R., Fütterer, G., & Leister, N., (2010), Depth cues in human visual perception and their realization in 3D displays, In B. Javidi, J.-Y. Son, J. T. Thomas, & D. D. Desjardins (Eds.), *Three-Dimensional Imaging, Visualization, and Display 2010 and Display Technologies and Applications for Defense, Security, and Avionics IV* (76900B), International Society for Optics; Photonics, <https://doi.org/10.1117/12.850094>
- Reid, T., (1764), Inquiry Into the Human Mind, *Philosophical works*.
- Renner, R. S., Velichkovsky, B. M., & Helmert, J. R., (2013), The perception of egocentric distances in virtual environments - A review, *ACM Computing Surveys*, *46*, 1–40, <https://doi.org/10.1145/2543581.2543590>
- Richardson, A. R., & Waller, D., (2005), The effect of feedback training on distance estimation in virtual environments, *Applied Cognitive Psychology*, *19*, 1089–1108, <https://doi.org/10.1002/acp.1140>

- 
- Rieser, J. J., Ashmead, D. H., Talor, C. R., & Youngquist, G. A., (1990), Visual perception and the guidance of locomotion without vision to previously seen targets., *Perception*, *19* 5, 675–689, <https://doi.org/10.1068/p190675>
- Rieser, J. J., Pick, H. L., Ashmead, D. H., & Garing, A. E., (1995), Calibration of human locomotion and models of perceptual-motor organization., *Journal of Experimental Psychology: Human Perception and Performance*, *21* 3, 480–497, <https://doi.org/10.1037/0096-1523.21.3.480>
- Ritter, M., (1979), Perception of depth: Processing of simple positional disparity as a function of viewing distance, *Perception & Psychophysics*, *25* 3, 209–214, <https://doi.org/10.3758/BF03202988>
- Rogers, S., (1995), Perceiving pictorial space, *Perception of space and motion*, 119–163.
- Rolland, J. P., Biocca, F., Barlow, T., & Kancherla, a., (1995), Quantification of adaptation to virtual-eye location in see-thru head-mounted displays, *Proceedings Virtual Reality Annual International Symposium '95*, 56–66, <https://doi.org/10.1109/VRAIS.1995.512480>
- Rolland, J. P., & Cakmakci, O., (2009), Head-Worn Displays: The Future Through New Eyes, *Optics and Photonics News*, *20* 4, 20, <https://doi.org/10.1364/opn.20.4.000020>
- Rolland, J. P., Gibson, W., & Ariely, D., (1995), Towards Quantifying Depth and Size Perception in Virtual Environments, *Presence: Teleoperators and Virtual Environments*, *4* 1, 24–49, <https://doi.org/10.1162/pres.1995.4.1.24>
- Rolland, J. P., Krueger, M. W., & Goon, A., (2000), Multifocal planes head-mounted displays, *Applied Optics*, *39* 19, 3209, <https://doi.org/10.1364/ao.39.003209>
- Rolland, J. P., Meyer, C., Arthur, K., & Rinalducci, E., (2002), Method of Adjustments versus Method of Constant Stimuli in the Quantification of Accuracy and Precision of Rendered Depth in Head-Mounted Displays, *Presence: Teleoperators and Virtual Environments*, *11* 6, 610–625, <https://doi.org/10.1162/105474602321050730>
- Rosales, C. S., Pointon, G., Adams, H., Stefanucci, J., Creem-Regehr, S. H., Thompson, W. B., & Bodenheimer, B., (2019), Distance Judgments to On- and Off-Ground Objects in Augmented Reality, *2019 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*, 237–243, <https://doi.org/10.1109/VR.2019.8798095>
- Ross, J., (2004), The perceived direction and speed of global motion in Glass pattern sequences, *Vision Research*, *44* 5, 441–448, <https://doi.org/10.1016/j.visres.2003.10.002>

- 
- Rousset, T., Bourdin, C., Goulon, C., Monnoyer, J., & Vercher, J. L., (2015), Does virtual reality affect visual perception of egocentric distance?, *2015 IEEE Virtual Reality (VR)*, 277–278, <https://doi.org/10.1109/VR.2015.7223403>
- Ryu, J., Hashimoto, N., & Sato, M., (2005), Influence of resolution degradation on distance estimation in virtual space displaying static and dynamic image, *Proceedings - 2005 International Conference on Cyberworlds, CW 2005, 2005*, 43–50, <https://doi.org/10.1109/CW.2005.54>

## S

- Sadato, N., Pascual-Leone, A., Grafman, J., Ibañez, V., Deiber, M. P., Dold, G., & Hallett, M., (1996), Activation of the primary visual cortex by Braille reading in blind subjects, *Nature*, *380 6574*, 526–528, <https://doi.org/10.1038/380526a0>
- Sahm, C. S., Creem-Regehr, S. H., Thompson, W. B., & Willemsen, P., (2005), Throwing versus walking as indicators of distance perception in similar real and virtual environments, *ACM Transactions on Applied Perception*, *21*, 35–45, <https://doi.org/10.1145/1048687.1048690>
- Sandor, C., Cunningham, A., Dey, A., & Mattila, V.-V., (2010), An Augmented Reality X-Ray system based on visual saliency, *2010 IEEE International Symposium on Mixed and Augmented Reality*, 27–36, <https://doi.org/10.1109/ISMAR.2010.5643547>
- Santos, M. E. C., de Souza Almeida, I., Yamamoto, G., Taketomi, T., Sandor, C., & Kato, H., (2016), Exploring legibility of augmented reality X-ray, *Multimedia Tools and Applications*, *75 16*, 9563–9585, <https://doi.org/10.1007/s11042-015-2954-1>
- Schmalstieg, D., & Hollerer, T., (2016), *Augmented reality: Principles and practice*, <https://doi.org/10.1109/VR.2017.7892358>
- Schmidt, R. A., & Wrisberg, C. A., (2008), *Motor learning and performance: a situation-based learning approach*.
- Schmidt, S., Bruder, G., & Steinicke, F., (2017), Moving Towards Consistent Depth Perception in Stereoscopic Projection-based Augmented Reality., *Icat-Egve*, 161–168.
- Schmidt, S., Bruder, G., & Steinicke, F., (2016), Illusion of depth in spatial augmented reality, *2016 IEEE VR 2016 Workshop on Perceptual and Cognitive Issues in AR, PERCAR 2016*, <https://doi.org/10.1109/PERCAR.2016.7562417>

- 
- Sedgwick, H. A., (1991), The effects of viewpoint on the virtual space of pictures, *Pictorial Communication In Real And Virtual Environments*, 460.
- Sei, I., Yuto, K., Shinnosuke, M., Asako, K., & Fumihisa, S., (2020), Shadow induction on optical see-through head-mounted displays, *Computers & Graphics*, 3219, <https://doi.org/10.1016/j.cag.2020.07.003>
- Sekuler, R., Sekuler, A. B., & Lau, R., (1997), Sound alters visual motion perception [6], <https://doi.org/10.1038/385308a0>
- Shimojo, S., & Shams, L., (2001), Sensory modalities are not separate modalities: Plasticity and interactions, [https://doi.org/10.1016/S0959-4388\(00\)00241-5](https://doi.org/10.1016/S0959-4388(00)00241-5)
- Singh, G., Ellis, S. R., & Swan, J. E., (2018), The Effect of Focal Distance, Age, and Brightness on Near-Field Augmented Reality Depth Matching, *IEEE Transactions on Visualization and Computer Graphics*, 26 2, 1385–139, <https://doi.org/10.1109/TVCG.2018.2869729>
- Singh, G., Swan, J. E., Jones, J. A., & Ellis, S. R., (2011), Depth judgment tasks and environments in near-field augmented reality, *Proceedings - IEEE Virtual Reality*, 2, 241–242, <https://doi.org/10.1109/VR.2011.5759488>
- Singh, G., Swan, J. E., Jones, J. A., & Ellis, S. R., (2012), Depth judgments by reaching and matching in near-field augmented reality, *2012 IEEE Virtual Reality Workshops (VRW)*, 165–166, <https://doi.org/10.1109/VR.2012.6180933>
- Slater, M., (2009), Place illusion and plausibility can lead to realistic behaviour in immersive virtual environments, *Philosophical Transactions of the Royal Society B: Biological Sciences*, 364 1535, 3549–3557, <https://doi.org/10.1098/rstb.2009.0138>
- Smith, W. M., (1953), A Methodological Study of Size-Distance Perception, *Journal of Psychology: Interdisciplinary and Applied*, 35 1, 143–153, <https://doi.org/10.1080/00223980.1953.9712847>
- Srinivasan, M. V., Zhang, S. W., Lehrer, M., & Collett, T. S., (1996), Honeybee navigation en route to the goal: Visual flight control and odometry, *Journal of Experimental Biology*, 199 1, 237–244.
- State, A., Chen, D., Tector, C., Brandt, A., Hong Chen, Ohbuchi, R., Bajura, M., & Fuchs, H., (1994), Observing a volume rendered fetus within a pregnant patient, *Proceedings Visualization '94*, 6, 364–368, <https://doi.org/10.1109/VISUAL.1994.346295>
- Stefanucci, J. K., Creem-Regehr, S. H., Thompson, W. B., Lessard, D. A., & Geuss, M. N., (2015), Evaluating the Accuracy of Size Perception on Screen-Based Displays :

- 
- Displayed Objects Appear Smaller Than Real Objects, *Journal of Experimental Psychology: Applied*, 21 3, 215–223.
- Steinicke, F., Bruder, G., & Kuhl, S., (2011), Realistic perspective projections for virtual objects and environments, *ACM Transactions on Graphics*, 30 5, 1–10, <https://doi.org/10.1145/2019627.2019631>
- Sugano, N., Kato, H., & Tachibana, K., (2003), The effects of shadow representation of virtual objects in augmented reality, *Proceedings - 2nd IEEE and ACM International Symposium on Mixed and Augmented Reality, ISMAR 2003*, 76–83, <https://doi.org/10.1109/ISMAR.2003.1240690>
- Sutherland, I., (1965), The Ultimate Display, *Proceedings of IFIPS Congress 1965, New York*, 506–508.
- Swan, J. E., Jones, J. A., Kolstad, E., Livingston, M. A., & Smallman, H. S., (2007), Ego-centric depth judgments in optical, see-through augmented reality, *IEEE Transactions on Visualization and Computer Graphics*, 13 3, 429–442, <https://doi.org/10.1109/TVCG.2007.1035>
- Swan, J. E., Kuparinen, L., Rapson, S., & Sandor, C., (2017), Visually Perceived Distance Judgments: Tablet-Based Augmented Reality Versus the Real World, *International Journal of Human-Computer Interaction*, 33 7, 576–591, <https://doi.org/10.1080/10447318.2016.1265783>
- Swan, J. E., Livingston, M. A., Smallman, H. S., Brown, D., Baillot, Y., Gabbard, J. L., & Hix, D., (2006), A Perceptual Matching Technique for Depth Judgments in Optical, See-Through Augmented Reality, *IEEE Virtual Reality Conference (VR 2006)*, <https://doi.org/10.1109/VR.2006.13>
- Swan, J. E., Singh, G., & Ellis, S. R., (2015), Matching and Reaching Depth Judgments with Real and Augmented Reality Targets, *IEEE Transactions on Visualization and Computer Graphics*, 21 11, 1289–1298, <https://doi.org/10.1109/TVCG.2015.2459895>
- Swenson, H. A., (1932), The relative influence of accommodation and convergence in the judgment of distance, *Journal of General Psychology*, 7 2, 360–380, <https://doi.org/10.1080/00221309.1932.9918473>

---

## T

- Takeichi, H., Watanabe, T., & Shimojo, S., (1992), Illusory Occluding Contours and Surface Formation by Depth Propagation, *Perception*, *21* 2, 177–184, <https://doi.org/10.1068/p210177>
- Taylor, J. E. T., Witt, J. K., & Sugovic, M., (2011), When walls are no longer barriers: Perception of wall height in parkour, *Perception*, *40* 6, <https://doi.org/10.1068/p6855>
- Teghtsoonian, M., & Teghtsoonian, R., (1970), Scaling apparent distance in natural indoor settings, *Psychonomic Science*, *20* 6, 281–283, <https://doi.org/10.3758/BF03335669>
- Teghtsoonian, R., & Teghtsoonian, M., (1970a), Scaling apparent distance in a natural outdoor setting, *Psychonomic Science*, *21* 4, 215–216, <https://doi.org/10.3758/BF03332449>
- Teghtsoonian, R., & Teghtsoonian, M., (1970b), The Effects of Size and Distance on Magnitude Estimations of Apparent Size, *The American Journal of Psychology*, *83* 4, 601, <https://doi.org/10.2307/1420693>
- Teghtsoonian, R., & Teghtsoonian, M., (1978), Range and regression effects in magnitude scaling, *Perception & Psychophysics*, *24* 4, 305–314, <https://doi.org/10.3758/BF03204247>
- Thompson, W. B., Fleming, R., Creem-Regehr, S. H., & Stefanucci, J. K., (2011), Visual Perception from a Computer Graphics Perspective. *Visual Perception from a Computer Graphics Perspective* (p. 540), CRC Press, <https://doi.org/10.1201/b10927>
- Thompson, W. B., Willemsen, P., Gooch, A. A., Creem-Regehr, S. H., Loomis, J. M., & Beall, A. C., (2004), Does the Quality of the Computer Graphics Matter when Judging Distances in Visually Immersive Environments?, *Presence: Teleoperators and Virtual Environments*, *13* 5, 560–571, <https://doi.org/10.1162/1054746042545292>
- Thomson, J. A., (1983), Is continuous visual monitoring necessary in visually guided locomotion?, *Journal of Experimental Psychology: Human Perception and Performance*, *9* 3, 427–443, <https://doi.org/10.1037/0096-1523.9.3.427>
- Trepkowski, C., Eibich, D., Maiero, J., Marquardt, A., Kruijff, E., & Feiner, S., (2019), The effect of narrow field of view and information density on visual search performance in augmented reality, *26th IEEE Conference on Virtual Reality and 3D User Interfaces, VR 2019 - Proceedings*, <https://doi.org/10.1109/VR.2019.8798312>

---

Tversky, B., (1993), Cognitive maps, cognitive collages, and spatial mental models, *European conference on spatial information theory*, 14–24.

## U

Uhl, F., Franzen, P., Lindinger, G., Lang, W., & Deecke, L., (1991), On the functionality of the visually deprived occipital cortex in early blind persons, *Neuroscience Letters*, *124* 2, 256–259, [https://doi.org/10.1016/0304-3940\(91\)90107-5](https://doi.org/10.1016/0304-3940(91)90107-5)

## V

Van der Hoort, B., Guterstam, A., & Ehrsson, H. H., (2011), Being barbie: The size of one's own body determines the perceived size of the world, *PLoS ONE*, *6* 5, <https://doi.org/10.1371/journal.pone.0020195>

Viega, J., Conway, M. J., Williams, G., & Pausch, R., (1996), 3D magic lenses, *Proceedings of the 9th annual ACM symposium on User interface software and technology - UIST '96*, 51–58, <https://doi.org/10.1145/237091.237098>

Vishton, J., & Cutting, P., (1995), Perceiving Layout and Knowing Distances: The Integration, Relative Potency, and Contextual Use of Different Information about Depth, *Perception of Space and Motion*, *22* 5, 69–117, <https://doi.org/http://dx.doi.org/10.1016/B978-012240530-3/50005-5>

Vishwanath, D., Girshick, A. R., & Banks, M. S., (2005), Why pictures look right when viewed from the wrong place, *Nature Neuroscience*, *8* 10, 1401–1410, <https://doi.org/10.1038/nn1553>

Von Melchner, L., Pallas, S. L., & Sur, M., (2000), Visual behaviour mediated by retinal projections directed to the auditory pathway, *Nature*, *404* 6780, 871–876, <https://doi.org/10.1038/35009102>

## W

Wagner, M., (1985), The metric of visual space, *Perception & Psychophysics*, *38* 6, 483–495, <https://doi.org/10.3758/BF03207058>

- 
- Walker, J., Zhang, R., & Kuhl, S. A., (2012), Minification and gap affordances in head-mounted displays, *Proceedings of the ACM Symposium on Applied Perception - SAP '12*, 124, <https://doi.org/10.1145/2338676.2338706>
- Wallace, G. K., (1959), Visual Scanning in the Desert Locust *Schistocerca Gregaria* Forskål, *Journal of Experimental Biology*, 363, 512–525.
- Waller, D., (1999), Factors affecting the perception of interobject distances in virtual environments, *Presence: Teleoperators and Virtual Environments*, 86, 657–670, <https://doi.org/10.1162/105474699566549>
- Waller, D., & Richardson, A. R., (2008), Correcting Distance Estimates by Interacting With Immersive Virtual Environments: Effects of Task and Available Sensory Information, *Journal of Experimental Psychology: Applied*, 141, 61–72, <https://doi.org/10.1037/1076-898X.14.1.61>
- Ware, C., (2008), Visual Thinking: For Design, *Ergonomics*, 531, 138–139, <https://doi.org/10.1080/00140130903458285>
- Warren, W. H., (1984), Perceiving affordances: Visual guidance of stair climbing, *Journal of Experimental Psychology: Human Perception and Performance*, 105, 683–703, <https://doi.org/10.1037/0096-1523.10.5.683>
- Wartenberg, C., & Wiborg, P., (2003), Precision of Exocentric Distance Judgments in Desktop and Cube Presentation, *Presence: Teleoperators and Virtual Environments*, 122, 196–206, <https://doi.org/10.1162/105474603321640941>
- Webb, R. H., Hughes, G. W., & Pomerantzeff, O., (1980), Flying spot TV ophthalmoscope, *Applied Optics*, 1917, 2991, <https://doi.org/10.1364/ao.19.002991>
- Willemsen, P., Colton, M. B., Creem-Regehr, S. H., & Thompson, W. B., (2004), The effects of head-mounted display mechanics on distance judgments in virtual environments, *Proceedings of the 1st Symposium on Applied perception in graphics and visualization - APGV '04*, 35, <https://doi.org/10.1145/1012551.1012558>
- Willemsen, P., Colton, M. B., Creem-Regehr, S. H., & Thompson, W. B., (2009), The effects of head-mounted display mechanical properties and field of view on distance judgments in virtual environments, *ACM Transactions on Applied Perception*, 62, 1–14, <https://doi.org/10.1145/1498700.1498702>
- Willemsen, P., Gooch, A. A., Thompson, W. B., & Creem-Regehr, S. H., (2008), Effects of stereo viewing conditions on distance perception in virtual environments, *Presence: Teleoperators and Virtual Environments*, 171, 91–101, <https://doi.org/10.1162/pres.17.1.91>



- 
- Witther, J., & Hollerer, T., (2005), Pictorial Depth Cues for Outdoor Augmented Reality, *Ninth IEEE International Symposium on Wearable Computers (ISWC'05)*, 92–99, <https://doi.org/10.1109/ISWC.2005.41>
- Witmer, B. G., & Kline, P. B., (1998), Judging Perceived and Traversed Distance in Virtual Environments, *Presence*, *7*(2), 144–167, <https://doi.org/10.1162/105474698565640>
- Witmer, B. G., & Sadowski, W. J., (1998), Nonvisually Guided Locomotion to a Previously Viewed Target in Real and Virtual Environments, *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *40*(3), 478–488, <https://doi.org/10.1518/001872098779591340>
- Witt, J. K., Proffitt, D. R., & Epstein, W., (2004), Perceiving distance: a role of effort and intent., *Perception*, *33*(5), 577–90, <https://doi.org/10.1068/p5090>
- Woldegiorgis, B. H., & Lin, C. J., (2017), The accuracy of distance perception in the frontal plane of projection-based stereoscopic environments, *Journal of the Society for Information Display*, *25*(12), 701–711, <https://doi.org/10.1002/jsid.618>
- Wu, H., Adams, H., Pointon, G., Stefanucci, J., Creem-Regehr, S. H., & Bodenheimer, B., (2019), Danger from the deep: A gap affordance study in augmented reality, *26th IEEE Conference on Virtual Reality and 3D User Interfaces, VR 2019 - Proceedings*, 1775–1779, <https://doi.org/10.1109/VR.2019.8797965>
- Wul, B., Ool, T. L., & He, Z. J., (2004), Perceiving distance accurately by a directional process of integrating ground information, *Nature*, *428*(6978), 73–77, <https://doi.org/10.1038/nature02350>

## Y

- Yonas, A., Goldsmith, L. T., & Hallstrom, J. L., (1978), Development of sensitivity to information provided by cast shadows in pictures, *Perception*, *7*(3), 333–341, <https://doi.org/10.1068/p070333>

## Z

- Ziemer, C. J., Plumert, J. M., Cremer, J. F., & Kearney, J. K., (2009), Estimating distance in real and virtual environments: Does order make a difference?, *Attention, perception & psychophysics*, *71*(5), 1095–106, <https://doi.org/10.3758/APP.71.5.1096>

---

Zollmann, S., Kalkofen, D., Mendez, E., & Reitmayr, G., (2010), Image-based ghostings for single layer occlusions in augmented reality, *9th IEEE International Symposium on Mixed and Augmented Reality 2010: Science and Technology, ISMAR 2010 - Proceedings*, 19–26, <https://doi.org/10.1109/ISMAR.2010.5643546>





---

**Titre :** Vers une caractérisation des biais perceptifs en réalité mixte : une étude de facteurs altérant la perception des distances

**Mot clés :** réalité augmentée, réalité virtuelle, perception

**Résumé :** Réalité Virtuelle, Réalité Augmentée, Réalité Mixte, ces mots comme les applications qui les accompagnent entrent peu à peu dans l'usage commun. Cependant, la réalité proposée par ces technologies n'est pas identique à notre réalité ordinaire. Le présent ouvrage se propose de mettre en évidence certains biais perceptifs en réalité mixte. Dans un premier temps nous étudierons un biais perceptif lié à l'observateur : l'anisotropie de la perception egocentrique des distances en réalité virtuelle. Dans un second temps nous examinerons la perception exocentrique des distances en Réalité Augmentée (RA). En effet la sous-estimation des distances egocentriques est un phénomène souvent observé et il est donc intéressant d'étudier son potentiel transfert à la perception exocentrique. Puis nous étudierons plus

en avant d'autres biais potentiels en RA en s'attachant en particulier à évaluer l'impact des indices de profondeur sur la perception des distances. En particulier, nous analyserons dans ce chapitre l'effet de deux indices de profondeurs en RA : l'impact de la position et de la forme des ombres sur la perception des distances, puis l'influence de l'accommodation sur la perception des distances en utilisant une technologie d'affichage spécifique : les dispositifs de projection rétinienne. Enfin nous discuterons le potentiel impact des techniques d'interaction sur la perception des distances et proposerons un protocole permettant d'évaluer l'effet de certaines interactions sur la perception des distances en RA, afin peut-être de parvenir à rapprocher celle-ci de la perception réelle.

---

**Title:** Toward a Characterization of Perceptual Biases in Mixed Reality: A Study of Factors Inducing Distance Misperception

**Keywords:** Mixed Reality, Augmented Reality, Virtual Reality, perception

**Abstract:** Virtual Reality, Augmented Reality, Mixed Reality, these words as well as their applications are gradually becoming common usage. However, the reality proposed by these technologies is not identical to our regular reality. This work aims to highlight some perceptual biases in Mixed Reality. First we study a perceptual bias linked to the observer: the anisotropy of the egocentric distances perception in virtual reality. In a second part, we study the exocentric perception of distances in Augmented Reality (AR). Indeed the underestimation of egocentric distances is a phenomenon frequently observed and it is therefore interesting to consider its potential transfer to

the exocentric perception. Then we further study other potential biases in AR by focusing in particular on evaluating the impact of depth cues on the perception of distances. In particular, we investigate in this chapter the effect of two depth cues in AR: the impact of the position and shape of shadows on distance perception, and then the influence of accommodation on distance perception using a specific display technology: retinal projection devices. Finally, we discuss the potential impact of interaction techniques on distance perception and propose a protocol to evaluate the effect of certain interactions on distance perception in AR, in order to perhaps bring it closer to real perception.