Virtual Reality Usability Design

David Gerhard
Wil J. Norton

CRC Press
Taylor & Francis Group

Virtual Reality Usability Design

The development of effective and usable software for spatial computing platforms like virtual reality (VR) requires an understanding of how these devices create new possibilities (and new perils) when it comes to interactions between humans and computers. *Virtual Reality Usability Design* provides readers with an understanding of the techniques and technologies required to design engaging and effective VR applications.

The book covers both the mechanics of how human senses and the mind experience immersive virtual environments, as well as how to leverage these mechanics to create human-focused virtual experiences. Deeply rooted in principles of human perception and computational interaction, the current and future limitations of these replacements are also considered.

Full of real-world examples, this book is an indispensable guide for any practising VR developer interested in making efficient and effective interfaces. Meanwhile, explorations of concrete theory in its practical application will be useful for VR students and researchers alike.

Taylor & Francis
Taylor & Francis Group
http://taylorandfrancis.com

Virtual Reality Usability Design

David Gerhard
Wil J. Norton

CRC Press
Taylor & Francis Group
Boca Raton London New York

CRC Press is an imprint of the
Taylor & Francis Group, an **informa** business

First edition published 2023
by CRC Press
6000 Broken Sound Parkway NW, Suite 300, Boca Raton, FL 33487-2742

and by CRC Press
4 Park Square, Milton Park, Abingdon, Oxon, OX14 4RN

CRC Press is an imprint of Taylor & Francis Group, LLC

Library of Congress Cataloging-in-Publication Data

Names: Gerhard, David, author.
Title: Virtual reality usability design / David Gerhard, Wil Norton.
Description: Boca Raton : CRC Press, 2023. | Includes bibliographical references and index. | Summary: "The development of effective and usable software for spatial computing platforms like Virtual Reality requires an understanding of how these devices create new possibilities (and new perils) when it comes to interactions between humans and computers. Virtual Reality Usability Design provides readers with an understanding of the techniques and technologies required to design engaging and effective VR applications. The book covers both the mechanics of how human senses and the mind experience immersive virtual environments, as well as how to leverage these mechanics to create human-focused virtual experiences"-- Provided by publisher.
Identifiers: LCCN 2022033995 (print) | LCCN 2022033996 (ebook) | ISBN 9781032198699 (hardback) | ISBN 9781032191324 (paperback) | ISBN 9781003261230 (ebook)
Subjects: LCSH: User interfaces (Computer systems) | Haptic devices. | Virtual reality. | System design.
Classification: LCC QA76.9.U83 G45 2023 (print) | LCC QA76.9.U83 (ebook) | DDC 005.4/37--dc23/eng/20220830
LC record available at https://lccn.loc.gov/2022033995
LC ebook record available at https://lccn.loc.gov/2022033996

ISBN: 978-1-032-19869-9 (hbk)
ISBN: 978-1-032-19132-4 (pbk)
ISBN: 978-1-003-26123-0 (ebk)

DOI: 10.1201/9781003261230

Typeset in Latin Roman
by KnowledgeWorks Global Ltd.

Publisher's note: This book has been prepared from camera-ready copy provided by the authors.

To Diane and Arthur

Contents

Section I Understanding Virtual Reality and Users

Chapter 1 ▪ What Makes Virtual Reality Remarkable? 3

1.1 DEFINING VIRTUAL REALITY 3
1.2 VIRTUAL ENVIRONMENTS 6
1.3 THE ORIGINS OF VR 8
1.4 A SHORT HISTORY OF HUMAN COMPUTER INTERACTION 14
1.5 WHAT VR CAN DO FOR HUMAN–COMPUTER INTERACTION 17
1.6 WHERE INTERACTION DESIGN FITS IN 22

Chapter 2 ▪ Making the Virtual Seem Real 25

2.1 THE FEELING OF BEING THERE 25
2.2 IMMERSION 27
2.3 PRESENCE 43
2.4 COMPONENTS OF PRESENCE 45
2.5 OTHER SENSATIONS OF REALITY WITHIN VR 52
2.6 MEASURING PRESENCE 53
2.7 SUMMARY 56

Chapter 3 ▪ Sensation and Perception 59

3.1 PROVIDING THE PERCEPTION OF A VIRTUAL WORLD 59
3.2 THE PERCEPTUAL PROCESS 60
3.3 A SENSORY VIEW OF THE HUMAN BODY 66
3.4 QUANTIFYING STIMULUS 69
3.5 PERCEPTIVE TASKS 70
3.6 SUMMARY 77

CHAPTER 4 ▪ Supporting Primary Senses 79

4.1 VISUAL SENSORY INPUTS 79
4.2 RESOLUTION 106
4.3 AUDITORY SENSORY INPUTS 110
4.4 THE PHYSICS OF SOUND 115
4.5 SOUND AND GAMES 116
4.6 SUMMARY 118

CHAPTER 5 ▪ Supporting Peripheral Senses 119

5.1 THE GUSTATORY SYSTEM—TASTE 119
5.2 THE OLFACTORY SYSTEM—SMELL 126
5.3 SOMATOSENSORY SYSTEM 132
5.4 THE PROPRIOCEPTIVE SYSTEM 141
5.5 THE VESTIBULAR SYSTEM 142
5.6 OTHER INTEROCEPTIVE SENSES 143
5.7 SUMMARY 144

CHAPTER 6 ▪ Perceiving Space and Scale 145

6.1 INTERPRETATION OF SPACE AND DEPTH CUES 146
6.2 SPATIAL AUDIO REPRODUCTION 157
6.3 BIOLOGICAL MAPPING OF SPACE 162
6.4 NON-EUCLIDEAN GEOMETRY 164

CHAPTER 7 ▪ Further Psychological Effects of Inhabiting a Virtual Environment 169

7.1 EFFECTS OF INTERACTIVITY 169
7.2 EMBODIMENT ILLUSION 178
7.3 COMPONENTS OF EMBODIMENT ILLUSION 178
7.4 EMOTION AND EMPATHY 186
7.5 SUMMARY 191

Section II Designing Virtual Interactions

Chapter 8 ▪ Experience Usability 195

8.1 INTRODUCTION 195
8.2 USABILITY THEORY 198
8.3 HUMAN FACTORS IN USABILITY 203
8.4 MULTIUSER CONSIDERATIONS 211
8.5 DESIGNING FOR ACCESSIBILITY 216
8.6 SUMMARY 234

Chapter 9 ▪ Fictions of Physics 237

9.1 THE RULES OF A SIMULATION 237
9.2 ADVANTAGES OF A CONSISTENT SET OF PHYSICS 238
9.3 WORLD INCONSISTENCY BOUNDARIES 240
9.4 IMPACTS OF WORLD INCONSISTENCY BOUNDARIES 243
9.5 RESOLVING WORLD INCONSISTENCY ISSUES 245
9.6 CREATING BOUNDARIES FOR PLAYER SAFETY 249
9.7 INTERACTION PATTERNS FOR BOUNDARY HANDLING 251
9.8 PHYSICS-BASED INTERACTIONS 257
9.9 SUMMARY 259

Chapter 10 ▪ Locomotion and Navigation 261

10.1 INTRODUCTION 261
10.2 PURPOSES OF LOCOMOTION 264
10.3 ARTIFICIAL LOCOMOTION STRATEGIES 271
10.4 PHYSICAL LOCOMOTION STRATEGIES 277
10.5 QUANTIFYING LOCOMOTIVE EFFICACY 284
10.6 PHYSIOLOGICAL EFFECTS OF MOVING IN VR 291
10.7 SUMMARY 297

Chapter 11 ▪ Activities and Interactions 299

11.1 AFFORDANCE 299
11.2 FAMILIARITY 310
11.3 MAPPING 321

11.4 COMMON INTERACTIONS IN VR 334
11.5 SUMMARY 343

CHAPTER 12 ▪ Information Display 345

12.1 INTRODUCTION 345
12.2 CAUTIONARY TALES 349
12.3 PRINCIPLES OF INFORMATION DISPLAY 350
12.4 INFORMATION DISPLAY IN VR 362
12.5 SUMMARY 363

CHAPTER 13 ▪ Translating Traditional Interfaces for VR 365

13.1 A SUMMARY OF TRADITIONAL UI 365
13.2 TRADITIONAL INTERFACES IN VR 373
13.3 GESTURES 383
13.4 SUMMARY 386

Bibliography 387

Index 393

I

Understanding Virtual Reality and Users

What Makes Virtual Reality Remarkable?

1.1 DEFINING VIRTUAL REALITY

Virtual reality (VR) is a medium that seeks to replicate the many sensations we humans experience when we interact with the physical world. Because it seeks to replicate reality, VR is defined by its resemblance to the physical world—while other computer interfaces may use audio output and visual display to represent digital information to the user, VR uses these technologies to "trick" the user into feeling as if the digital world is somehow real. The "virtual" part of *virtual reality* refers to that the world displayed on a VR device is of our own creation, as opposed to the physical world. However, by this logic, any fictional world is *virtual*, be it expressed through literature, visual arts, television, or video games. For a display to be considered virtual reality, the world displayed must be interactive, convincing, and similar to the physical world in terms of form, not just content.

There's a wide diversity of definitions for the term virtual reality, but most people have an intuitive sense of what counts as virtual reality and what does not. Different virtual reality systems support the simulation of different senses at different levels of accuracy, and the applications (or **experiences**) supported across these systems vary widely in the accuracy of their own simulation. Regardless, these systems are still all considered to belong to the category of VR—therefore any definition of the term must be abstract enough to allow for this variety.

Any definition for virtual reality requires us to first take a step back and define **reality**. A common sense definition of reality might refer to the world around us; the space and time that we live in—the place and experiences that we call "the real world." In a more general sense, however, reality refers to that set of things we can independently and objectively agree upon, and specifically the set of things which is not somehow "fake." Reality is often thus defined in a negative sense, and we seek to exclude the false from the real by experimenting to determine the objective and the independent. When discussing virtual reality, however, a much more succinct definition of reality is available—"reality" is the physical world in our immediate vicinity that we can experience with our senses; the virtual version of reality, then, exists

DOI: 10.1201/9781003261230-1

when we block out the sensations from the physical world and replace them with sensations from some virtual world, such that we might believe these new sensations, even if only subconsciously. Although a variety of media exist which attempt to block out and replace the physical world (big screen TVs and loud radios are examples), virtual reality equipment distinguishes itself from other media by attempting to produce a complete illusion of another physical place, from the perspective of and tied to the perceptions of a single user. Two people cannot share the same experience of reality, since we are looking at the world with different eyes from different perspectives. Similarly, no two people can experience the same virtual reality at the same time, since the world is presented to each user through their own apparatus, with their own personal illusions.

Notice the use of the word illusion in this last sentence. The objects and worlds you interact with in a virtual reality system aren't part of the physical world—they're a simulation, despite how much they may replicate their physical counterparts. This is the meaning of the word **virtual** in the context of virtual reality—the simulation seems like a reality but isn't. The word "virtual" originally comes from the Latin "virtus" meaning "excellence" or "potency." Other words with a similar origin are "virtuous" and "vertical," both suggesting a sense of aiming towards perfection. Over time, the meaning of the word "virtual" progressed from "representing the best example of an effect" to "capable of producing an effect" and then "capable of imitating an effect" and finally to the connotation it has today: "being something in essence or effect, but not in reality." Another sense of the word virtual is "almost a particular thing or quality." For example, we say that something is "virtually non-existent" if it's almost gone. When used in this sense, "virtual" can be replaced by "almost" with little to no loss of meaning. In the case of virtual reality, the simulation presented is "almost reality" to the mind—for a simulation to be considered virtual reality, the mind has to interpret it as if it were a physical reality. If our mind interprets a simulation as if it were a physical space we could interact with, then for all practical purposes it's almost reality—virtual reality.

So, back to the definition—for an electronic system to count as virtual reality, it has to be capable of creating the illusion for the user that they are in a different physical place. No matter how high definition your TV is, it will never feel like you're anywhere but your living room. Virtual reality systems are different—in order to make it look like you're in a physical space, a VR system must be able to block out and replace what you can see in any direction. Often, this is done with a head-mounted display that tracks the position of your head to detect the direction of your gaze and replaces your field of view with a render of the virtual world from your point of view. Not only must this display change the rendered image when you move your head and look around, it must update this view so rapidly that you don't notice the change at all. Indeed, as we will discuss later in the book, if the display does not refresh fast enough, it can cause a form of disorientation that is similar to seasickness.

1.1.1 Vision and other senses

Although visual replacement is usually what people think about when they consider a VR headset, sight isn't the only way that people receive information from their environment—what we feel, hear, taste, and smell all tell us the information about the physical space we occupy. In fact, humans have many more "senses" than the five just mentioned that provide our brain with information about our surroundings as well as our own bodies—the internal tension of our muscles, our sense of balance, and the system that allows us to feel acceleration and motion, just to name a few. Providing any non-visual sense with information that adds to the illusion of a physical place could also be considered virtual reality. For much longer than we've had optical displays in virtual reality headsets, we've been able to generate "virtual" audio that convincingly replicates the experience of listening to sound in a physical space—if you closed your eyes, you'd believe you were on a busy city street or listening to a performance in a concert hall[1]. Is this, then, a virtual reality device? By our definition, it is—if you close your eyes, the equipment gives you the illusion that you are somewhere else. However, since vision is such a major component of our sensation of the world around us (for those of us with typical vision), and since stereophonic headphones have been around for decades, we don't typically imagine being pulled into a virtual world when we put on our headphones for the train ride home. When people talk about a virtual reality system, at minimum we are usually discussing a system that gives the visual illusion that the user is in another place. This is the lowest bar for something to be considered a virtual reality system in the public discourse—of course, a system that can simulate additional senses is even better, and in fact most VR systems also provide audio simulation as well.

1.1.2 Interactions

The emergence of VR as a medium presents new, unique ways of interacting with computers. Being able to digitally generate and display content that a user's mind may interpret as a three-dimensional space introduces new ways to create and improve the way we interact with the digital world, for training, entertainment, employment, productivity, and even for therapeutic experiences. As our computing technology becomes more and more advanced, VR presents a way to merge our interactions with computers with our familiar interactions with the world around us. Even so, the existence of VR as a medium is not a guarantee that such software will be more useful, entertaining, or easier to interact with. Applications running on virtual reality systems are just like any other computer application—their actual usefulness depends on the experience, knowledge, skills and choices of the developers. This book attempts to explain the basic skills and knowledge needed to effectively develop applications involving virtual environments. Much as someone developing a physical product must practice industrial design to create a useful and elegant object to meet

[1] Although stereo sound by itself is engaging, it is often insufficient to convince a listener that they are somewhere else. In Chapter 4 we discuss audio technologies like HRTF and Ambisonics that are very convincing.

the user's needs, someone developing a virtual reality application must also have a solid knowledge of interaction design.

This book assumes you already have some knowledge regarding basic programming and computer graphics principles or that you have supplemental texts for these matters. Although it is not necessary to have any prior computer science knowledge to read this book, an elementary understanding of these fields is a requirement for anyone looking to develop virtual reality applications. This first chapter serves as a general introduction to the scope of virtual reality and attempts to explore some of the unique strengths and challenges of this medium. We start the chapter with a general discussion of terminology for different types of virtual environments and discuss some of the historical developments and uses of virtual reality systems. We talk about the specific problems virtual reality is well suited to address in human computer interaction (HCI), discuss some of the current applications of the technology, and try to imagine a few future use cases for virtual environments. Finally, we finish the chapter with a short discussion on why interaction design is an essential part of the development of any virtual reality application.

1.2 VIRTUAL ENVIRONMENTS

If virtual reality equipment creates the illusion that a digital world occupies physical space, the digital world displayed is referred to as a **virtual environment**. A virtual reality system obscures the physical environment, using displays or other equipment to block sensations from the physical world, while simultaneously providing new inputs to the occluded senses.

It is possible to provide simulated sensory input without first blocking out physical sensory input. **Mixed reality** (MR) equipment allows a user to experience elements from a virtual environment while allowing them to still receive stimuli from the physical environment, although in this case the virtual environment and the physical environment must be aligned. If VR is when the environment displayed to a user is purely virtual and physical reality is when the environment is entirely physical, then MR includes everything in between.

Overlapping with the real
The Milgram Kishino continuum, described below, applies primarily when considering different levels of augmented reality, where the virtual world is well aligned with the physical world the user is in. When the virtual world is not aligned with the physical world, each sense must be occluded as much as possible, and the goal at that point is to remove or isolate distractions from the physical world, rather than considering a continuum between them.

The Milgram Kishino Virtuality Continuum describes the spectrum of mixed reality, shown in Figure 1.1. Any device capable of generating a virtual environment could be described as being somewhere along this continuum.

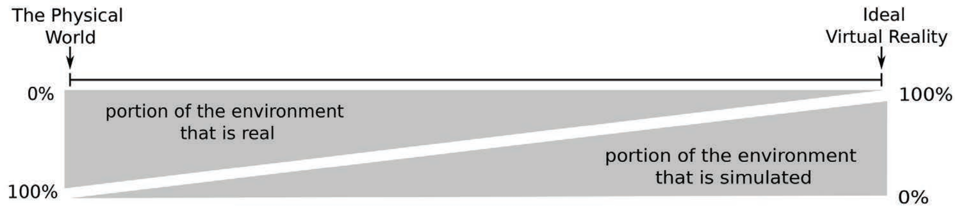

Figure 1.1: The Milgrim Kishino Virtuality Continuum.

On this continuum, "ideal" virtual reality is at the extreme right end—meaning that no stimuli from the outside world reach the user. Almost no practical VR system exists that fits this definition—even in a high-quality virtual reality headset, the user might still be able to hear sounds from the real world or see between the headset and their nose. A device that would qualify as pure virtual reality on the spectrum is at the very least more than a few years past today's technology and may never be practically achievable. Despite the way the continuum defines VR, we'll use the term "virtual reality" to refer to systems that don't quite meet the bar set by Milgram and Kishino—if the intention of a system is to replace outside inputs with simulated ones, we'll call it a virtual reality system (even if it doesn't quite achieve this goal).

On the extreme left hand of the continuum is the physical world — the unmediated experience of the physical environment around us. Everything in between this and ideal virtual reality can be considered to be a form of MR.

MR is a broad category. Although some devices are marketed generally as "Mixed Reality," it is also common to hear other, more specific terms being used. Although the Milgram Kishino continuum is well recognized in academia, there is some ambiguity about where exactly certain classifications belong on the continuum. Some common descriptions for various parts of the continuum are as follows:

- **Extended Reality** (XR) – A catch-all term that refers to any application that includes some form of digital world. This category includes both VR and MR.
- **Augmented Reality** (AR) – Often used in marketing as a synonym for MR. In academia, AR often refers to applications where digital elements do not seem to be placed in the "real world," but are simply overlaid.
- **Augmented Virtuality** (AV) – Used to describe applications for "VR devices" where physical objects are used to replicate difficult-to-simulate parts of the virtual environment. For example, an augmented virtuality tennis application may have VR visuals, but use a physical tennis racket-shaped controller to enhance the realism.

Figure 1.2 shows these various classifications labelled on the Milgram Kishino continuum.

Although this book primarily refers to designing interactions for virtual reality, many of the same design considerations that apply to the applications we discuss are equally valuable when developing MR or AR software. These mediums all relate

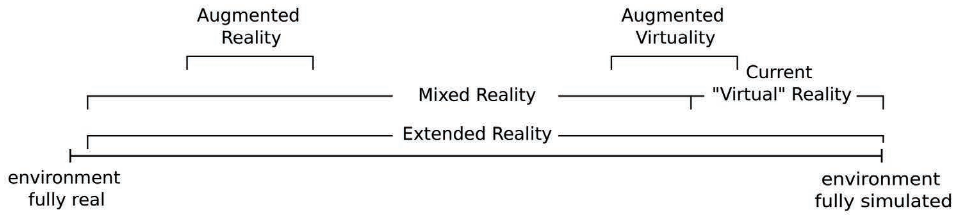

Figure 1.2: Portions of the Milgram Kishino Continuum and how they relate to common spatial computing categorizations.

in that to some degree; they involve or display a world that has similar sensory properties as physical three-dimensional space and as such may be referred to under the blanket term of **spatial computing** technologies. Since all of these mediums are defined by the use of a virtual environment, many of the chapters related to general interaction and interface design for virtual environments are applicable for technologies at any point on the continuum. Further, while most of our case studies focus on virtual reality applications, we have included a few examples showing how the same principles can be applied to mixed reality cases as well. The appendix further suggests resources that include more information on MR design outside of the scope of this book.

The Milgram Kishino continuum is a good basis for high level of comparisons between systems, but there are lots of ambiguities in how it compares systems. What if one system only simulates audio, but another only presents a visual display—which one simulates a higher portion of the environment? As there is no easy way to define if the audio or visual sense represents "more" of the real environment, it might make more sense to consider a specific system on several different continua, separated by the sense being simulated. A MR system might overlay virtual graphics over the real world (falling somewhere near the middle of the continuum for vision), but might not provide virtual sound at all (and would therefore be at the extreme left of the continuum for audio). Even this definition of the reality continuum is not necessarily useful for determining how much of the world has been replaced with simulated stimuli—as the middle 4 percent of the human field of view contains 90 percent of the total resolution of the eye, a continuum that says a headset with a display encompassing only this portion of vision is only 4 percent simulated is not an accurate measure of how much of an environment we have replaced—and neither is a continuum that places this display at the 90 percent simulated mark. A large portion of Chapter 2 and lesser portions of subsequent chapters in the first half of this book are dedicated to describing more objective measures of where a system truly fits between virtual and physical.

1.3 THE ORIGINS OF VR

One of the first recorded depictions of virtual reality in a form that we'd recognize today was in a 1949 novel. *Pygmalion's Spectacles* by Stanley Grauman Weinbaum is regarded as the first instance of fiction describing what we would now call "virtual reality." The story is about a man possessing a pair of eyeglasses capable of seeing into a false world as convincing as the physical one. In Weinbaum's novel, the titular spectacles were of supernatural origin, but it was not long before the fictional narrative exploring the concept of virtual reality took on a technological nature. Science fiction stories began depicting similar realities generated by electronic devices.

In almost every fictional depiction of virtual reality devices, the common ground between the systems described is that they are able to support simulations that look and act just like the physical world. Ever since science fiction planted this image in the minds of the general public, the work of countless researchers and engineers has brought us virtual reality devices that come closer and closer to achieving this goal. Eyesight could be considered our "dominant" sense—if the information we see disagrees with what we hear, feel, or smell, we tend to trust our eyes. It makes sense then that many of the earliest virtual reality advances were focused on providing input to the eyes that better replicated the way we see three-dimensional space. People had been able to create paintings to depict the visual world since prehistoric times, but even when we developed the technology required to perfectly capture a scene in time (photography), a flat photograph still wouldn't fool anyone into thinking they were actually looking through a window. There are two major reasons why no static image on a flat plane can seem fully three dimensional:

1. **Perspective** – In the physical world, even the slightest rotation or translation of our head allows us to see new sides of objects and causes the sides we do see to distort in very specific ways, according to the laws of perspective. In order for a scene to seem three dimensional, we would have to be able to track the user's head movements, as well as update the image to match the changing position of the user's head.

2. **Stereospsis** – The pupils of our eyes are separated by a small distance, varying between 3 cm and 5 cm from person to person. This causes our eyes to each receive a slightly different perspective of a scene. One flat image from a single perspective isn't enough to emulate stereospsis.

1.3.1 Prehistory: 1800s

Out of these problems, stereospsis was the easiest to address. The stereoscope, a device that was capable of displaying a separate image to each eye, had been invented by Sir Charles Wheatstone, in 1832. Stereoscopes saw further improvements throughout the 19th century—by 1850, over 250,000 stereoscopes had been sold. Although they had began to decline in popularity by the 1870s, stereoscopes continued to get better at depicting stereoscopic scenes—with the advent of the camera, a likeness of the physical world could be depicted by taking two separate images, with one taken from a position 5 cm to the right of the other. A stereoscopic image generated by this method would look quite similar to standing and looking at the subject of the photograph in real life—as long as the user didn't move their head. Stereoscopes survived the 19th century to live on today—not only is stereoscopy a key feature in modern

VR headsets, but modern stereoscopes, like the Viewmaster, still sell well despite remaining relatively unchanged since Sir Charles' original device. Both an early and a modern stereoscope can be seen in Figure 1.3.

Figure 1.3: Left: an early 19th century stereoscope (*Auckland Museum CC BY*). Right: a modern Viewmaster stereoscope, circa 1970 (*Jamiecat CC BY 2.0*).

The second major hurdle in making an image seem three dimensional, tracking head movement and changing perspective, took longer to solve. Early pioneers created devices that tried to emulate the physical world without the use of head tracking. For these devices to be convincing, the user had to remain stationary and their view had to be locked in place. Although the visual sense had yet to be perfected, some inventors still tried to add the support for the simulation of more senses in their "virtual worlds." One of the most ambitious simulators from the pre-head tracking era was Morton Heilig's *Sensorama* (shown in Figure 1.4). In addition to coloured stereoscopic film, a *Sensorama* booth included a seat that would move, fans to simulate wind, and even scents to match the film. Of course, the user still wasn't able to move their head, or else the lack of a change in perspective would make it apparent they weren't in a physical place.

1.3.2 Early Prototypes: 1960s

It wasn't until 1968 that head tracking and perspective tracking were achieved in a virtual reality system. In that year, Ivan Sutherland and three students at MIT's Lincoln Laboratory developed the *Sword of Damocles*, the direct ancestor of today's VR **head-mounted displays** (HMDs). Ivan Sutherland had been involved in research related to virtual reality for several years prior to the invention of the *Sword of Damocles* and had come up with much of the early theory of VR, including his definition of the Ultimate Display (discussed more in Chapter 2)—research that laid the groundwork for the *Sword of Damocles*. The headset was tracked via a combination of mechanical linkages and ultrasonic sensors to determine the head rotation and position of the user. Cathode ray tubes near the user's eyes projected stereoscopic vector graphics and would update appropriately when the user's head moved (the

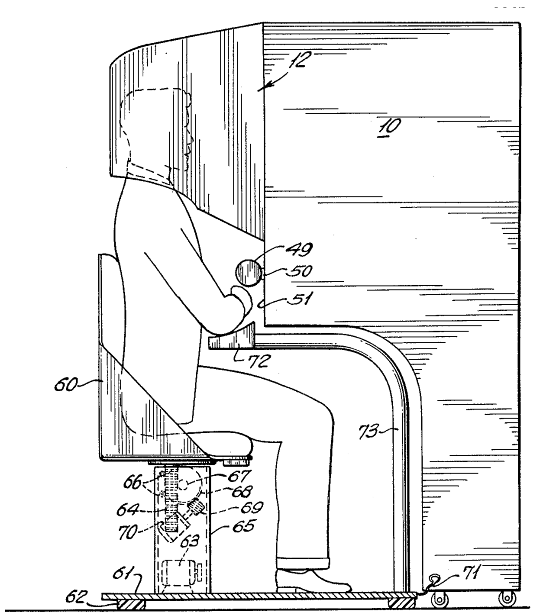

Figure 1.4: The patent drawing for Sensorama (1962).

user was still able to see the room at the same time as the vector graphics, so the *Sword of Damocles* could be considered a mixed reality headset). Due to its weight, the headset had to be suspended by a mechanical arm from the ceiling of the laboratory. With the exception of motion controls, the *Sword of Damocles* was capable of all of the same functions as the first wide-selling consumer VR headsets of the 2010s were—without, of course, the accompanying tracking accuracy or resolution.

Ivan Sutherland – The Father of Computer Graphics

Ivan Sutherland may be an important figure in the history of virtual reality development, but his work in VR only represents a small fraction of the overall impact he had on the field of computer science as a whole.

Sutherland's research for his PhD involved his development of *Sketchpad*, an interface which allowed users to make engineering drawings on a computer with a stylus. What made this really impressive was that *Sketchpad* was made in *1963*—before consumer text-based computers even existed! Sketchpad was not only the first computer-aided drafting (CAD) program to exist, but also the first program to use a graphical user interface (GUI). Interacting with *Sketchpad* had more in common with an iPad (2010) than the computers of the day, which for the most part still used punch cards as input.

Sutherland was given the A.M. Turing Award for his work developing *Sketchpad.* After completing his PhD, Sutherland would continue to be responsible for many more important breakthroughs across various fields of computer science.

VR would continue to be developed after the *Sword of Damocles*, primarily within universities and government institutions. Due to the prohibitive expense of equipment at the time, most applications of the technology were exclusive to research, but other uses would begin to emerge. One of the biggest applications for virtual reality in the 1970s was in training applications, particularly in the military. Flight simulators further developed virtual reality technology, and the military also started developing AR displays for jet pilot helmets. As a result of all the interest in virtual reality, headsets continued to get lighter and became capable of supporting better graphics and higher resolutions.

Aside from the continued development of head-mounted displays, researchers continued to innovate new ways to experience virtual environments. Research into haptic technology, allowing a user to move around their hand and feel force feedback, began as early as 1967—with the GROPEHaptic project. Although the GROPEHaptic also had to be mounted to the ceiling for tracking, later haptic technologies followed a similar path as HMDs, freeing themselves from the ceiling by becoming lighter and cheaper. By the 1980s, Scott Fisher and other members of the Virtual Environment Workstation Project at NASA's Ames Research Center had done a work on developing a space station maintenance simulator that included tracked haptic gloves and an HMD. In fact, by around this time, both HMDs and tracked gloves had matured to the point where they were being sold on the commercial market.

1.3.3 The first consumer generation: 1990s

The first wave of consumer VR equipment was marked by the founding of VPL research in 1984. Standing for "Virtual Programming Languages," VPL was founded by Jaron Lanier and would eventually develop and sell the DataGlove (a tracked glove), the EyePhone (an HMD, not to be confused with another important innovation in consumer electronics), and the DataSuit (a full body tracking garment). The EyePhone 1 cost around $9,400 and was only capable of refreshing the display at a rate of around six or seven frames per second. Although the device did see sales, they were primarily in industry—product developers and architects found the device useful for viewing designs before prototypes were made. VPL continued to operate until they shut down in 1990.

However, the 1990s saw a multitude of new VR companies established. By 1995, even Nintendo had developed a VR console, the Virtual Boy. Despite the relatively low prices of virtual reality equipment during the 90s wave, the limitations of these devices prevented them from ever truly catching on with the public. VR consoles had low frame rates, small fields of view, and in the case of the virtual boy, only supported graphics in two colours—black and red. Nintendo discontinued the Virtual Boy by 1996, after a disappointing 770,000 worldwide sales. Meanwhile, fiction kept on depicting VR systems that were capable of supporting simulations that looked just like the real world, in novels like Neal Stephenson's *Snow Crash* and William Gibson's *Neuromancer*. When you consider the dissonance between the VR systems of fiction and the actual implementations of the time, it's no wonder consumers of 90s VR were disappointed.

While consumer VR went silent, VR researchers and enthusiasts continued to develop the technology. Advancements in graphical quality and resolution in other fields allowed VR system prototypes to solve many of the issues present in early consumer models. The advent of mobile computing lead to smaller and smaller devices that were capable of better and better performance. The rise of the internet lead to wide-scale discussion and dissemination of virtual reality research—virtual reality flourished, outside of the public view.

1.3.4 The Second Consumer Generation: 2010s

In 2012, Palmer Luckey launched the crowdfunding campaign for the Oculus Rift, a headset he had been developing on his own for several years. By the time it was released to the public in 2016, the Oculus Rift (Figure 1.5) had a resolution of 1080×1220 pixel resolution, which was impressive for the time, and a frame rate of 90Hz—additionally, it originally cost $600, inexpensive compared to similar devices of the past (for reference, the VPL EyePhone cost upwards of $250,000). The combination of leading-edge hardware and a consumer-friendly price point, as well as the key feature of head-tracked field of view, was enough to capture the imagination of the public. Over 2.5 million units of the Oculus Rift were sold by 2017. Other companies began producing similar VR headsets, like the HTC Vive and Playstation VR, and these devices also saw similar commercial success—as of 2019, Playstation VR had sold 7 million units worldwide.

Figure 1.5: Palmer Luckey wearing the Oculus Rift (2016).

Since then, VR technology has continued to be developed. The second wave of consumer VR devices generated more interest in research and development of VR devices, and affordable headsets allowed a much wider audience to have access to the technology.

Prior to the second wave, the majority of VR research and development was focused on improving the hardware of VR systems to a point where the technology would be useful. In the early stages of VR, a system would be made to forward the state of VR, in which case the software would usually be simple tech demos or be custom made for a very specific purpose (such as military training). In the current state, consumer VR equipment is capable of running any manner of programs—although many of the existing programs for second wave consoles could be categorized as entertainment, VR has the potential to be a great medium for solving lots of different problems. In the next section, we explore why VR is such an exciting medium for human–computer interaction, by framing it as one of the many stages of how people have interacted with the digital world.

1.4 A SHORT HISTORY OF HUMAN COMPUTER INTERACTION

Both the methods through which humans have interacted with computers and the tasks people have used computer for have changed significantly since computers were first invented. The earliest practical computers, electromechanical devices such as Herman Hollerith's census tabulators (1884), used punch cards for input. The holes in punch cards would be used to encode data, which would then be tabulated by the machine and displayed on dials to an operator. Although the Hollerith machines represented a large increase in efficiency over the census tabulating methods of the time, they were still cumbersome to use. Data first had to be translated to the format of the punch cards using a separate punching device and a translation table. After the cards were fed into the machine, the output on the dials would have to be read and added by hand to a running total. Not only did data have to be manually processed by humans in order for the machine to understand it, but the output from the machine had to be further processed in order to create the total! People interacted with a Hollerith machine in a machine-centred fashion—it was up to the user to translate inputs to the machine's language and to translate outputs back to a human usable format.

Computers continued to evolve in terms of processing power, speed, and range of function over time, and incremental progress was made in human computer interaction as well. However, it wasn't until 1946 that an important leap in providing input to computers was made—the keyboard. Typewriters had existed as a commercial product since 1867, so by 1946 people had grown accustomed to using them for writing text. When John Presper Eckert and John Mauchly were developing a computer that required text input at the University of Pennsylvania, they decided to use a teletype machine (seen in Figure 1.6) to punch their cards. The end result, the *Electronic Numerical Integrator and Computer*, allowed operators to skip the step of translating instructions into punched holes—instructions could be entered in digits and English. By this time, computers had also gained the ability to print out their output, using primitive displays or ticker tape. Now, humans were able to "speak to" computers in a way that was more familiar—written language.

Figure 1.6: A teletype unit, similar to the one that would have been used with the ENIAC (*Eric Fischer CCBY2.0*).

Communicating with computers via written language was the most common method of interaction for a long time. Command line interfaces (CLIs) were the sole method of interaction in the most successful personal computers of the early 1970s. Although these command line interfaces were easy to use for a large amount of tasks involving a computer, they did have their shortfalls. While text input was useful for things like word processing and programming, using it for drawing graphics or moving files could be slow and taxing. Further, by the 1970s, computers were no longer exclusively used by computer scientists and researchers. While computer scientists were content with learning verbose technical commands to communicate with their computers, these same commands were frustrating for laypeople. The solution was to introduce a new method that people and computers could use to communicate: the GUI.

The first GUI was developed by Ivan Sutherland for his doctoral dissertation, *Sketchpad*, which allowed users to draw and see shapes on a screen using a stylus. A second important invention followed the GUI—the mouse and pointer. The Xerox Alto (1974), the first personal computer to support a GUI-based operating system, was also the first personal computer to come with a mouse. Software developers designing for GUIs now had the freedom to use **interaction metaphors**, the process of making one task a metaphor for another, in order to make it easier for newcomers to understand how to operate computers. The interaction metaphor you are likely most familiar with is the **desktop** metaphor, developed for the Alto at Xerox PARC and made widely available in the Apple Macintosh (shown in Figure 1.7). Prior to computers, people would commonly keep physical files in manila folders and were used to rearranging their papers on the top of their physical desk—the desktop. When Jef Raskin was designing the interface of the Apple Macintosh, he used the metaphor of the traditional physical file system to make it easier for users to grasp how directories and digital files worked. On the Macintosh, files would be represented by little icons that looked like pages of paper and directories would instead become “folders.” If no programs were open, the user would see the desktop—a space where “folders” and “files” could be dragged around and rearranged with the mouse like their physical counterparts. By changing the way users interacted with the computer to be more like how they interacted with the world, Jef Raskin made it easier for new users to figure out how to navigate the file system—after all, they had been doing it in the physical world their whole lives.

Figure 1.7: The Apple Macintosh (1984), one of the first personal computers to use the desktop interaction metaphor. *Marcin Wichary CCBY2.0.*

The development of HCI reveals a pattern—people have difficulty formatting their inputs and reading outputs in the way the computer requires it to be done, so they change the interface to the computer to better align with human needs. This brings us to VR—which, in essence, is just another interface for computers. Like the keyboard, the GUI, and the mouse before it, VR represents a leap in bringing the way we input and receive information from computers closer to the way we interact with the physical world. Just as these prior inventions made computers more accessible for the public by making the operation of computers similar to the way we conduct other tasks in the physical world (typing, organizing files, and pointing), VR does this as well—by making the way we interact with computers match the way we move and occupy physical space.

1.5 WHAT VR CAN DO FOR HUMAN–COMPUTER INTERACTION

In an human–computer interaction context, "Virtual Reality" refers to a set of outputs and inputs for computers that approximate the sensations of interacting in a physical reality. As VR is a categorical term, discussing the HCI impacts of virtual reality is different than discussing the impact of the invention of the mouse. Although the computer mouse exists in several different brands and forms, all mice appear very similar in function—all are moved on a surface to move a cursor, have (at least) two buttons, and allow for scrolling. A mouse may have more features than this (for example, more buttons along the side to bind hotkeys too), but those are often unrelated to its function as a mouse. In contrast, when you talk about a "virtual reality system," the exact features supported by the hardware are ambiguous. It's probably safe to assume that a VR system supports a head-mounted display, but not all head-mounted displays support the same visual simulation. A **degree of freedom**, when used in regard to a VR headset, refers to an axis on which the headset display updates with movement. The x, y, and z axis represent the three translational degrees of freedom, while rotation around each of those axes (pitch, roll, and yaw) represent the three rotational degrees of freedom—all of these degrees are shown in Figure 1.8. A 6-DOF display updates accordingly when the headset is moved in any of these ways. While some headsets are able to simulate accurate changes in perspective for both head rotation and translation (displays with six degrees of freedom or **6-DOF** tracking), many VR devices exist that only update properly for rotational movement of the head (**3-DOF** tracking).

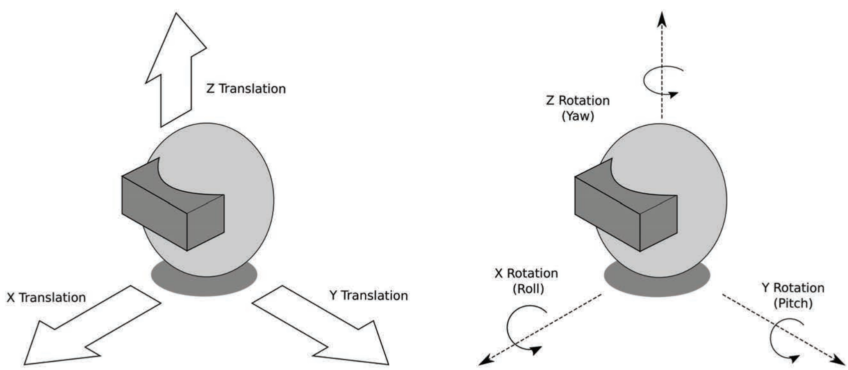

Figure 1.8: The six possible degrees of freedom a headset can move in.

Outside of the visual realm, a VR system may support motion controls with tracked hands, or it may not. A VR system could include audio or haptic feedback, both, or neither.

Each of these features could also vary greatly in quality—it's impossible for us to assume the features that any one virtual reality system could have. In order to broadly discuss how VR might impact HCI, we'll examine VR from mostly a theoretical perspective—all that we'll assume for now is that VR systems utilize some technology to simulate a virtual environment for at least one of the user's senses. When we do mention features of theoretical VR systems, we'll discuss the power of these specific features—VR systems that do not support these features will likely not have the same benefits for interaction. However, the features that we explicitly mention in this section are ones that typically exist in current consumer VR systems, and we envision that such features will only continue to improve over time. If VR ever becomes like the mouse, where every system shares a common set of features, we can say with near certainty that the features mentioned below will be included.

We've broken the ways VR could impact HCI into two major categories. The first category includes the HCI implications of using VR *to make interacting with computers more like interacting with the physical world.* Just as the keyboard made text input similar to writing on a typewriter, and the desktop made arranging file systems similar to organizing papers, we'll detail the ways in which VR will make interacting with computers even more similar to interacting with the physical world (as well as the benefits of this). Although we discuss these aforementioned implications in the context of VR, many of these HCI impacts have been seen before in mice, GUIs, and other products that mimic familiar interactions. However, the same is not true for the second category of VR's HCI effects—these benefits are unique to VR and are unrelated to mimicking real-world interactions. Instead, we care about the raw capabilities of the computer and what it can do to us—*how we can benefit from "fake" versions of the physical world.* Together, these categories encompass many of the reasons why VR has the potential to be such a beneficial technology.

1.5.1 Making Interacting with Computers More Like Interacting with the Physical World

VR is often lauded as a fix to many problems with existing computer interaction,and we list some of the potential benefits here. It should be noted, however, that not all opportunities offered by VR systems are net positives, and as with any new technology, the changes that come from new ways of doing things often come with trade-offs. Each of these benefits will be discussed in more detail later in this text, at which point we will also highlight and discuss the potential trade-offs inherent in these advances.

Benefit 1: VR Applications Could Be Easier to Learn. Skimming through our short history of human computer interaction makes it apparent that computers used to be a lot harder to use—punch cards took a lot of training to learn how to use, and command line interfaces were intimidating to new computer users. We think of computers now as quite easy to use, but many people still have difficulty interacting with computers. You may recall learning to type—as natural as typing may seem to an experienced user, you might remember how difficult and slow it felt to first try typing while growing up. The keyboard was easier to learn than punch cards as it mimicked the operation of a typewriter, but a lot of this ease relied on the user already knowing how to type. However, VR is even easier to learn than modern desktop computing.

Since a virtual environment resembles the physical world, VR applications allow us to interact with the computer in the same way we interact with a physical world—a skill we know very well. Unlike a keyboard, which was initially easy only for those who had the prerequisite experience (using a typewriter), VR uses a prerequisite metaphor everyone is familiar with (the physical world). In a VR system that supports hand tracking, a user can reach out and press buttons or keys, pull levers, or perform other tasks in the same way that they would in the physical world. Of course, in order for it to be easy for someone to learn how to interact with a VR application, we have to make the actions match physical equivalents they've already learned to do. If you've never pressed a button in real life, we can't assume pressing one in VR would be intuitive—but most people do have experience with buttons. If we utilize physical metaphors that are common to most users, there would be no need to learn how to operate a VR interface—the user would already know from their prerequisite experience. Just as keyboards made computers accessible to a larger audience and GUIs made computers accessible to an even larger audience beyond that, VR has the potential to make it so that even more people can benefit from computing, by lowering the knowledge barrier to entry.

Benefit 2: VR could reduce the ergonomic stress of interacting with computers. When personal computers were invented, few people imagined how much we'd use them. Computers are helpful for completing tasks more efficiently, but weren't initially designed with ergonomics in mind. Extended use of computer keyboards has been shown to cause pain in the wrists, arms, and neck; repetitive use of a mouse can cause carpal tunnel syndrome; looking at a screen for hours a day can cause eye strain and sleep problems; and even sitting at a desk for extended periods can cause heath issues as well. Sitting for extended periods has become more and more common as computer use has increased, and with it, the health problems associated with sitting have increased as well. Sitting for long periods has been linked to increased risks of cancer, heart disease, and back problems. These problems aren't due to computers, per se—they're due to the way computers have been designed to be used. Computers only evolved to the form they are in today due to manufacturing constraints, performance requirements, and the occasional HCI concern addressed after the fact. It doesn't help that most people aren't terribly concerned about the ergonomics of the computer they're purchasing—they're more concerned about what the computer can do. Similarly, people weren't trying to think of a more ergonomically friendly method of computer interaction when they came up with VR—the ergonomic benefits of VR were an unintended side effect. Regardless, VR-based interaction presents a solution to many of the ergonomic problems caused by current computer interaction systems. Users of VR interact with the computer in the same way they interact with the physical world—by standing, reaching, grabbing, and poking, using their entire body. Many of the ergonomic problems attributed to modern computing come from the fact that using a computer requires us to perform movements that evolution left us ill suited for. Our spine shapes evolved to support us while standing, while computers require us to spend our time sitting. Our wrists evolved to perform a wide range of tasks, not to perform the same motion thousands of times a day in order to click a mouse. VR allows us to interact with computers through actions that better match the ones we evolved to perform—simply because the virtual world that a VR user interacts with is similar to the physical world we evolved in. VR is not a universal solution to computer-based ergonomic problems—some of the problems caused by modern computers, like repetitive actions and starting at artificial light for long periods of time, are still present in VR. Further, there are some ergonomic problems that are unique to VR. For example, wearing a heavy headset has been shown to cause neck discomfort and strain (although lighter headsets have helped to reduce the impact of this). We'll discuss more ergonomic considerations for designers of VR applications later in the book. However, by moving to an interaction model closer to the way we interact with the physical world, the adoption of VR as a computing interface has the potential to eliminate a large portion of the ergonomic issues caused by modern computing.

Benefit 3: VR provides new ways of interacting with content. The introduction of the GUI and mouse was particularly impactful in computing because these new ways of interacting with content on the computer allowed for easier ways to perform tasks. Creating an image for a manufacturing blueprint would have been extremely tedious on a command line interface (CLI), but on a GUI, drawing such an object would have been a lot easier. CLIs weren't usually used as an alternative to hand-drawn engineering drawings, but once GUIs were invented, computer-aided drafting surpassed hand-drawn drafts as the dominant method of blueprint production. Developing new ways to interact with the computer made it possible to perform tasks on a computer that weren't possible before, in addition to making existing tasks easier. Similarly, VR creates even more ways to interact with computers—allowing us to perform some existing tasks more efficiently, in addition to making the range of tasks we can use computers for even larger. Interaction with digital 3D objects is possible on desktop computers, but isn't ideal—it's often difficult to determine the orientation of an object in space or to visualize it as a 3D object without the help of (often cumbersome) 3D view port controls. In VR, we can see these objects in three dimensions, not just a 2D projection—modelling or evaluating 3D CAD models could be much less cumbersome than it is on a 2D screen. The inclusion of allowing movement in the digital space allows for new possibilities—instead of watching a video to learn dance steps, a user could follow outlines for foot placement within a virtual environment, making learning techniques easier. As the amount of ways we can interact with a computer increases, the range and effectiveness of the tasks we can perform with it increase as well—different inputs are better suited for different tasks.

Benefit 4: VR is more suitable for displaying certain content than 2D screens. We've just discussed how it's easier to create 3D content within a 3D environment, but 3D interfaces are also more suitable for displaying 3D items. VR is well suited for displaying 3D content, especially where it may be vital or otherwise difficult to understand how a three-dimensional object fits into a larger object or world. For example, it may be easier to grasp how a 3D part fits into a mechanical assembly if you are able to see it in a 3D environment—there's no ambiguity over where the part attaches or exists within relation to the the other parts, in contrast to when the 3D environment is projected onto a 2D screen. If an image of a 3D object is projected onto a 2D screen, shading and shadow can provide some insight about the depth and relation of the object to 3D space, but examples like common optical illusions show that the depth of a 3D object on a 2D plane can be ambiguous. One such ambiguous shape is shown in Figure 1.9.

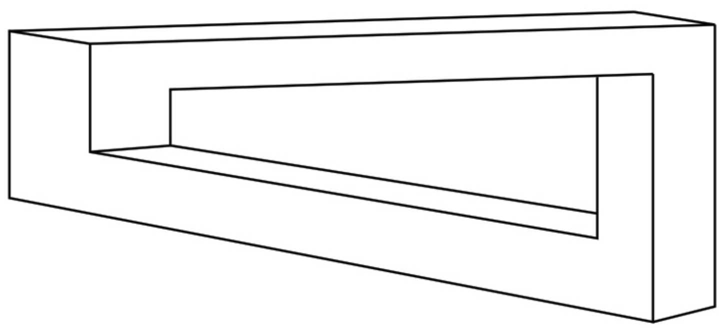

Figure 1.9: A shape that presents depth ambiguities when viewed in a 2D projection.

VR may still allow for some spatial ambiguity, but the addition of a third visual dimension reduces the number of possible visual ambiguities. Further, just as we can display a projection of 3D content on a 2D screen, we can display a projection of 4D (or even higher dimensional) geometry within a virtual environment. Of course, we'd be losing some information in showing the 4D object on a display that supports fewer dimensions than the shape has, but with a 4D shape, there's a larger degree of dimensional separation (and therefore ambiguity) when displayed on a 2D screen rather than a 3D interface.

Benefit 5: VR maps physical motion to real-world motion. In desktop computing, the actions we perform with the mouse or keyboard often map only indirectly to the motions on screen. For example, the plane in which the screen resides is usually (but not always) orthogonal to the plane in which we move the mouse—we look at the screen and we move the mouse on our desk. Whenever we're using a mouse, we implicitly have to translate the movement on one plane to the other—taking a little bit of cognitive load. With motion controls, motion is one to one—pointing occurs on the same plane as the motion. This removes the cognitive load associated with this translation and even makes pointing in VR a little bit faster (under certain conditions discussed in later chapters).

1.5.2 Making Worlds That Feel Physical By Using Computers

The previous five benefits of VR mentioned above can be viewed as the most recent steps in the history of human–computer interaction that we briefly discussed—making the form computers better match the needs of humans. These previous benefits are all due to how VR changes the way a computer is communicated with—because of either the way we provide a VR system with input or interpret its output. These benefits look at the VR output/input system as a replacement for other computer output/input systems. However, the output displayed on a VR device can be though of from a different perspective—instead of using a VR system to replace computer interactions, we can use it to replace interactions in the physical world.

Not everything is practical or possible to do in the physical world. Flight simulators, an early use case of VR, are a good example of this: although it's possible for a rookie to learn to fly using real planes, this would be an expensive (and dangerous) approach. Using VR, a novice can learn the controls for a plane in an environment that replicates much of the stimuli present while physically flying, but without the risk. VR is particularly useful for training due to the sense of **presence** it supports—users of a VR device are likely to instinctively interpret their surroundings as a physical environment (even if they know it isn't real). Because of presence, people are more likely to be mentally or emotionally affected by content if it's displayed in VR compared to other mediums. Training someone in a VR simulation causes both more accurate reactions in training compared to other non-physical mediums, and similarly, VR training carries over better to an actual scenario than other non-physical mediums. Due to the realistic physical interpretations VR invokes in a user's mind, it is useful for applications like exposure therapy. Curing someone's fear of heights by taking them to the edge of an actual roof may be dangerous, but a virtual roof can give similar benefits without the risk. The fear need not even be physical—studies have shown that many people become nervous while speaking to virtual crowds. Finally, the (relative) lack of expense of VR equipment makes many things possible for entertainment or leisure that would otherwise be too dangerous or costly for many. Simulations of hang gliding, visiting exotic locations, and scuba diving are all free assuming you have a VR system and do a good job approximating many of the sensations experienced in the real activity.

1.6 WHERE INTERACTION DESIGN FITS IN

Hopefully the previous few sections have given you a general understanding of how VR fits in to the history of human computer interaction and where VR may take HCI in the future. However, we've spoken a lot about the hardware advances of VR, and some specific use cases of such technology, but relatively little about VR software itself.

Any medium is only as good as the content available for it—a lack of desirable media for a platform is part of why Laserdisc lost out to VHS, why the Philips CDI flopped, and why a lot of people weren't overenthusiastic when modern VR emerged—the content available was lacking. No matter how much potential VR hardware has, it is squandered unless the software running on the platform is effective and was made with an understanding of the limitations of the user and the hardware.

Regardless of how well implemented an application's backend is, the interface is a bottleneck for the usefulness of any program. The user only sees the interface, and if the interface is designed in a way that makes it impossible to use certain functions of the software, the end result is the same as if those functions didn't exist. VR is a unique case when it comes to interface design—a VR experience isn't like a typical web or desktop interface, but it's not quite physical either. In order to be effective, a VR interface must successfully blend cues from the physical and digital worlds to accomplish a task. In order to design a spatial interface, we can borrow from traditional UI design as well as from industrial design. We can take design cues from older, physical design disciplines, such as architecture, or from newer fields, like web design. In VR, an interface may be visible or the program might be making decisions using the users' input (say, motion) without them even knowing. In this sense, a VR application can have an interface, or a traditional interface may be completely absent—although in either case, the user is still **interacting** with the experience. Over the first half of the book, we'll continue to discuss how the user interprets VR experiences, with a few case studies on how to use this knowledge to design effective VR applications. The second part of the book takes this knowledge and applies it to many different types of interfaces—both those that have been traditionally developed for computer applications or physical products and those that are only possible through the use of VR. By the end of the book, you'll have the tools you need to understand how good interaction design can make VR experiences enjoyable, effective, and easy to use.

CHAPTER 2

Making the Virtual Seem Real

2.1 THE FEELING OF BEING THERE

At a mall in an unnamed town in the Midwest, an interested but cautious consumer approaches a kiosk with trepidation. A sign beckons them to "Try the exciting immersive new technology of *Virtual Reality!*". They reflect on the virtual reality (VR) trends of the 1990s, remembering that the technology wasn't really that great, and although computer technology has improved significantly since then, they can't imagine how strapping a smartphone-sized screen to their head could provoke either excitement or immersion. After all, they have a phone. They have played games on their phone. It can't be that much better.

So they surreptitiously circle the kiosk for 20 minutes, hoping to watch someone else try it; eventually, a new shopper wandering the mall is drawn into the flashy advertising. Here, now, our sceptic will see how people really react to this new VR. The new volunteer dons the headset, immediately takes two steps back, looks up, and says "... woah..." They then proceed to waive their arms around wildly, step forward with caution, and whirl around in fear. They are behaving as if this new VR really is exciting—as if it really is immersive. Our sceptic decides to give it a try.

They approach the kiosk and don the headset, and to their amazement, they are teleported across reality to a new world filled with monsters; they grab a weapon leaning against the wall beside them, and they dive into the fray.

Many people trying VR for the first time will remark on how "real" the virtual world feels. This realism is one of the major characteristics of VR and a key component in many of its emerging use cases, which are already far beyond the confines of gaming and entertainment. The success of VR-based training, treatment, and teaching largely depends on the whether or not the experience is "real enough" to convince the brain to modify itself as a result of the stimuli it is experiencing.

2.1.1 Realism in Digital Media

Picture for yourself the most realistic and most believable digital environment you can imagine. What would it be like? For example, imagine a digital forest rendered

DOI: 10.1201/9781003261230-2

with such fidelity that you can watch the wind gently rustle the leaves of each tree. If you close your eyes, you can hear each bird and insect as if you were actually in the woods. You can feel the breeze on your arm as you hold your hand to block out the bright sun shining through the canopy, and you can smell the freshness of the damp earth beneath your feet. To replicate this virtual environment "perfectly," the software simulating this forest would require detailed data and physics calculations that can mirror the mechanics of the real world, down to the movement of individual leaves on a tree; although it is probably not necessary to simulate the movement of sap within the tree.

If this scene were displayed on a television, however, no one would be tricked into thinking it were real. It might be a recording of a real scene, but even at such a high level of representational detail, even if the TV could somehow display smells and weather effects, it would be impossible to mistake the simulation for an actual forest, because most of your sensory processes would still tell you that you are sitting on a couch in your living room, "watching" a forest, rather than walking through one. For this simulation to be "realistic," distractions and sensations from the real world must be removed, as much as possible, so that other sensations can be replicated in their place. This is the twofold role of virtual reality: exclude the real world so it can be replaced by the virtual world.

A virtual reality hardware system, such as a VR headset, is able to support more realistic content than a TV screen can because it is capable of blocking sensations from the real world and providing more realistic simulated data to more senses. However, even though VR systems have the potential to more realistically represent a virtual world, it is possible to imagine a traditional software experience that might be more immersive than one in VR. A VR game that poorly represents audio, makes mistakes on depth rendering, or has poor performance may feel less "real" than our perfectly simulated forest projected on a flat screen. In fact, depending on the content and presentation, a well-written book can draw the reader into a more complete feeling of "immersion" than a poorly developed software experience. Figure 2.1 presents this simple idea visually: the **realism** of a simulation can be compared as a simple sum of the realism of each aspect of that simulation. We will develop more concrete and informative analysis than the abstract measure of "realism" and build more comprehensive comparisons between simulations; but for now this is a reasonable heuristic: a simulation can be very good at one thing, but unless it is also good enough at the other aspects of simulation, it will not be immersive.

At this point, you may be able to infer that, from the perspective of human sensation, the "realism" of a given simulation is related to both the number of senses the system simulates and how realistically each of these sensations is simulated. A perfectly realistic simulation would require every human sense to be simulated with perfect accuracy. However, accurately simulating the sensations of a given world can be a difficult task in practice, and in reality there are tricks and shortcuts that can be used to create a believable experience without having to simulate every sense.

The first half of this book covers the theory required for an understanding of how to accurately simulate human senses. The current chapter covers the basics of the psychological theory behind the "reality" of virtual reality—why a good VR

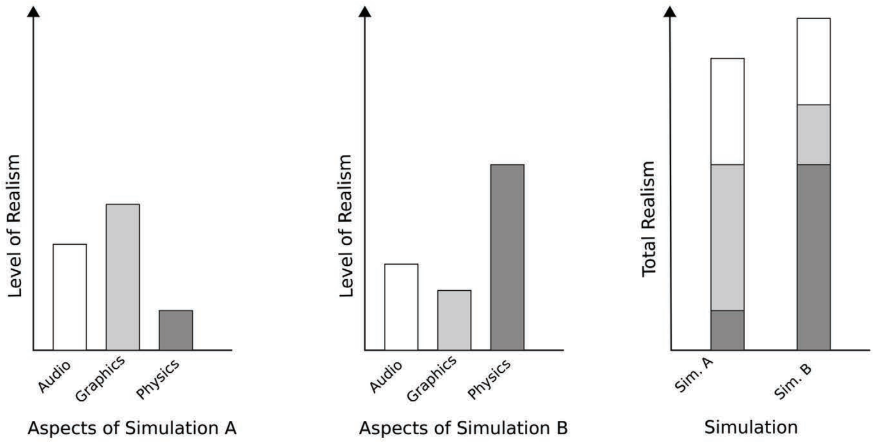

Figure 2.1: Comparing the realism of two different simulations by adding the levels of realism for each aspect of the simulations.

experience can make a user feel as if they are genuinely in a different location, a feeling called **presence**. We begin with a discussion of the concept of presence in a more precise manner. Following that, we explore two psychological effects that combine to contribute to the feeling of presence—namely, the place and plausibility illusions. Finally, we end the chapter with a discussion of how to measure these illusions in VR experiences and provide some tools and examples to help developers construct the sensation of presence within VR projects.

2.2 IMMERSION

VR experiences are capable of inducing a stronger sensation of "being there" than many other artificial media. Although VR users may instinctively try to reach out and touch a virtual table, TV viewers rarely make the mistake of reaching out to grab something shown on the television. This feeling of being in a real environment requires a medium that is capable of displaying a set of sensations that the brain could mistake for a physical space.

Immersion occurs when a sensory input is replaced with a reasonable virtual simulation of that sense—the sensory unit of the body responsible for interpreting this now-virtualized input has been immersed. Looking at a small black and white TV far away has low immersion; looking at a head-tracked VR display has much higher immersion.

What do you mean by the word "Immersion"?
The interpretation of the terms "presence" and "immersion" can be inconsistent across the literature, when talking about VR. Some researchers use the terms

as we have defined them in this book, while others may reverse their definitions or use them to refer to different scopes entirely. For example, people often talk of being immersed in a good book or a TV show, and by this they mean they are fully invested in the narrative. People also refer to video games as immersive if they can imagine themselves in the scenario being generated—to avoid confusion, we will use the term “narrative engagement” to describe these situations. However, throughout this book, when we use the term, we are referring to the ability of displays to replicate the range and field of real-life sensory input. If you find yourself encountering these terms elsewhere, don’t assume that they are being used the way they are in this book, but seek out the definitions presented by the authors.

Mel Slater has been researching presence in VR since 1992—the date of his first conference proceedings on the experience of presence in virtual reality. Since then, his work has focused on investigating the sensation of presence in virtual reality, along with its applications. Mel Slater’s work on defining the aspects that make up presence has resulted in one of the most widely accepted definitions for these terms. His distinction of the various phenomena responsible for presence is the basis of the definitions put forward in this book.

While a television screen can show a flat image, VR displays support stereoscopy, which is the ability to present a slightly different image to each eye. Since the fields of view of our two eyes overlap, the brain can interpret the slight differences between them as a proxy for the distance from the object being observed. A flat screen does not provide this depth information, but other signifiers like occlusion, shading, size, and brightness can lead to a passable illusion of depth in two dimensions. A stereoscopic display results in the brain receiving an image that is more closely aligned to the expectations of reality, and therefore more convincing and engaging. A 3D TV or movie theatre provides such an experience, but only to a small subset of the user’s complete field of view. If the screen could be enlarged, then the experience would be better, but the screen and the scene it displays are also fixed in space and unresponsive to the user, meaning that if the user looks away, they won’t see the screen anymore—instead, they might be met with a quizzical expression from their seat neighbour. Additionally, if the user moves their head from side to side, they might expect the perspective to change, but since the TV does not have information about the location or direction of the user’s view, it cannot change what it is showing. If, however, the view itself can be made to move with the user such that their vision is replicated stereoscopically regardless of where they look, then the experience is even more convincing and engaging and closer to the way visual stimuli are processed in the real world, than images on a TV ever were. This is why current VR hardware is more like a pair of goggles than a TV screen—it can move with the user.

In order to understand how to replace a human sense input with appropriate information from a virtual world, we must first have an understanding of the processes by which humans use senses to perceive the world. A detailed discussion of the various sensation categories available to humans will be presented in later chapters, but a generalized abstract model of human sensation will be sufficient for our discussion of immersion. The process of human perception, whether from the real world or a virtual reality simulation, encompasses several layers of processing from the point where the human receives the sensation to the point where their mind interprets it. Figure 2.2 presents an abstract model of this process in a flowchart.

First, **sensation** is the physical act of a user's sensory organs receiving information from a display[1]. All of the information we have about the world comes from sensations received from our surroundings. Sensation consists of the reception, processing, and interpretation of some form of outside energy. For example, sensory cells in the eye—rods and cones—send signals to the nervous system when they are stimulated by associated frequencies of electromagnetic energy. Our ears respond to mechanical energy created by pressure differences in the air and convert those pressure waves to nerve impulses that are interpreted as sounds. Any differences we experience between sensations of the same type—for example, seeing one object as blue and another as red—are due to differences in the characteristics of the energy being received.

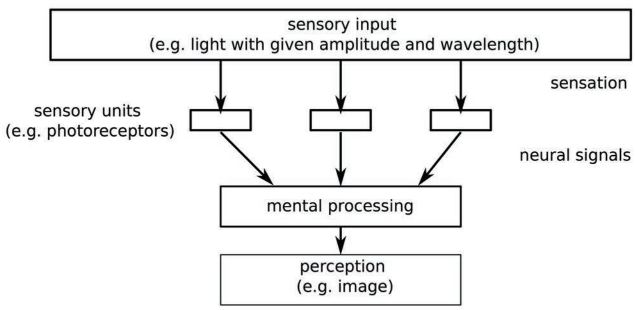

Figure 2.2: A flowchart depicting an abstract model for perception.

2.2.1 Displays

In the physical world, sensory stimulus comes from objects around us. Waves of energy created, reflected, or otherwise changed by objects in a person's physical environment result in the sensation of the world. For a virtual environment to be sensed by the user, it must first be rendered into physical energy corresponding to the sense being replaced. We use the term **display** to refer to any device that converts digital information into stimuli that can be sensed by the user. Displays exist for many types of stimuli—for example, a monitor is a display that outputs visual stimuli, while speakers generate auditory stimuli. Under this definition, a computer-controlled

[1]Since vision is such a predominant sense, much of the language of sensation is inherited from visual terminologies, like the word "display." A sensory display is any device used to replicate the stimulus to trigger a sense. Headphones can be considered an auditory display, in this sense of the word.

heater could be considered a display—it uses digital values to control the heat energy in a given area. The act of translating a digital representation to a physical stimuli by a display is referred to as **transduction**.

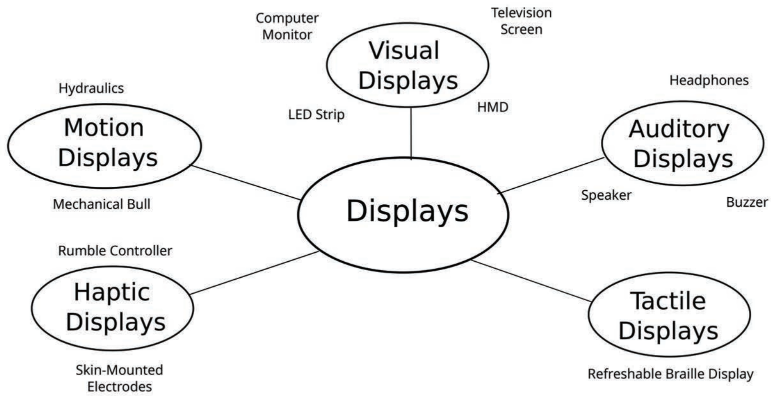

Figure 2.3: Different types of displays.

2.2.2 Sensory Inputs and Sensory Units

We define **sensory input** as the physical energy that the human body receives from the outside world (either from the real world or from a display) and that subsequently gives rise to sensation. A sensory input usually consists of a wave of a particular type of energy with a particular set of characteristics (amplitude, frequency, and phase) that correlates with or is a proxy for some real-world event. The term "input" considers this energy from the perspective of the user—energy that is considered a sensory input is the input to the sensation/perception process. A signal intended to stimulate a sensory input is the *output* of a display.

Sensory inputs are received and registered by **sensory units**, individual biological sensors that trigger a neurological signal when activated by a specific type of energy with specific characteristics. Sensory units are usually the size of individual cells, and we have a lot of them—there are over 120 million individual sensory units just in the eyes. The sense of sight, then, refers to a complicated system of different types of sensory units sensing many types of sensory inputs, which is integrated into a coherent concept by the brain.

Sensory Abilities and Immersion

Differing sensory abilities can also influence the immersion felt by the user of a simulation. Consider a user without the use of vision. One could argue that the blind user does not have the same visual immersion as a seeing player would, as they cannot see the world being simulated. However, it could be said that their visual experience of the VR game is more realistic than the seeing player—the seeing player can detect inaccuracies between the sensory