ABSTRACT

ŞENER, ELIF. A Comparison of Memory Performances for Expository Scientific Prose and Diagram in Flat vs. Spatially Distributed Layouts in Virtual Reality. (Under the direction of Dr. Matthew O. Peterson.)

Unlike desktop computers, reading in virtual reality allows the reader and the information to cohabit in the three-dimensional space. However, virtual reality typography rarely exploits the new affordances three-dimensionality offers. This study addressed the need to understand the contributions of spatially distributed layouts on expository text memory. An experiment with 40 college students was conducted to compare recall and recognition performances in two conditions; a flat layout without any spatial attributions to test baseline memory ability and a condition using a virtual reality-based, spatial layout. It was hypothesized that spatial layout assist in retrieving information from scientific expository text, which tends to have a difficult word or syntax structure and high conceptual density. Multiple regression analysis showed that using spatially distributed layouts to retrieve information did not provide an increase in both recall and recognition of the participants. A thematic analysis of interviews suggested that customizing the text layout through interactions would be helpful for information retrieval.

A Comparison of Memory Performances for Expository Scientific Prose and Diagram in Flat vs. Spatially Distributed Layouts in Virtual Reality

by
Elif Şener

A dissertation submitted to the Graduate Faculty of
North Carolina State University
in partial fulfillment of the
requirements for the Degree of
Doctor of Philosophy

Design

Raleigh, North Carolina
2023

APPROVED BY:

Dr. Deborah Littlejohn	Dr. Karen Chen
Dr. Lam Pham	Dr. Matthew O. Peterson Chair of Advisory Committee

BIOGRAPHY

Elif Şener is a Fulbright research scholar, instructor, and Ph.D. candidate at North Carolina State University, Department of Graphic and Experience Design. After earning her bachelor's degree in Visual Communication Design at the University of Economics, she started her master's program in the same field at Sabanci University, Istanbul. Simultaneously, she taught the history of modern art and visual literacy in her master's program. In 2018 she was awarded the Fulbright Scholarship, and since 2020, she has been an instructor of record in Graphic and Experience Design at NC State University.

In her practice and research, Elif has been interested in challenging the idea of typography as the servant of the content. For that, she gradually moved from traditional production to emerging technology. Starting with creating two-dimensional, static forms, her work became interactive and generative. Her current research interest is in investigating the influence of experience–oriented and technology–driven paradigms and practices on our interactions with complex information systems. Elif is passionate about contributing to the democratization of XR technologies by adopting a user-centered approach.

ACKNOWLEDGEMENTS

First, I would like to thank my advisor, Dr. Matthew Peterson, and committee members, Dr. Lam Pham, Dr. Karen Chen, and Dr. Deborah Littlejohn, for providing guidance and feedback throughout this process. You were extremely generous with your time and supervision. Likewise, many thanks to Tsailu Liu for being such an amazing educator to look up to and a great mentor.

Thanks to all my students, who always inspired me to do better. Many thanks to those who spared time to help me with my dissertation: Anna Lawrence, Joel Weiss, Tucker Baumgartner, Bhavana Veeravalli, and Anabel Russo. My sincere thanks to the Fulbright family, who never wavered in their support. I am honored to be a part of it.

My dear colleagues, Meichun Liu, Payod Panda, Grace Wonaphotimuke, Raunak Mahtani, Meichun Liu, Derek Bangle, and Renae Mantooth, thank you for all your emotional and intellectual support and encouragement. I would also like to thank my dear friend Ömer Mert Candan for teaching me how to talk to computers. It was a pleasure to work with all of you.

I want to thank my amazing cousins İpek Karaoğlu and Sevil Mert Batu and my caring brother Hamit Şener. Thank you for your absolute support and love. Many thanks to my wonderful mother, Memnune Şener, and father, Feriz Şener. I would never imagine achieving anything in life without you. The incredible grandma İsmet Karaoğlu, special thanks to you for always looking after me my whole life. Finally, the brightest woman I knew, my dear grandma, Şükriye Şener, I know you have been watching and protecting me.

Special thanks to Masoome Haaghani. You have been incredibly good and patient with me. I can not imagine seeing the finish line without you. Mario Hernandez, you have never let me down over the years. You are one of the best people Raleigh has brought to me. Sebastian Marcin Feculak, you have a unique power to take me out of my darkness. I am deeply indebted to you for that.

I know I am not the best at showing my appreciation to you. Nathan Sanders, you have been all that anyone could be in every way. Thank you.

Last but not least, I want to thank me for believing in me, for doing all this hard work, for having no days off, for never quitting, for being a giver, and trying to give more than I receive, for trying to do more right than wrong, and for just being me at all times.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

DEFINITION OF TERMS

Character: The smallest unit of written language with semantic values, referring to an abstract meaning or shape (Smura et al., 1989).

Glyph: A recognizable abstract graphic symbol unrelated to any particular design (Smura et al., 1989).

Immersion: Immersion refers to the degree to the VR system's sensory modalities which match that of the real-world (Slater, 1999). The more the VR system provides fidelity similar to its real-world counterparts, the more immersive it becomes.

Immersive Environments: Immersive digital environments are similar to virtual reality but without the implication that actual "reality" is being simulated. An immersive digital environment can be a model of reality. Still, it can also be a complete fantasy user interface or abstraction, as long as the user is immersed within it (Handa et al., 2012).

Typography: Merriam-Webster dictionary defines typography as "the style, arrangement, or appearance of typeset matter." While this definition is partially valid, it omits important aspects of the discipline. This dissertation uses the definition of typography from Baines & Haslam (2005), which defines it as the mechanical notation and arrangement of language. Mechanical refers to the deliberate steps taken towards a defined aim through automatization that creates reproducible units -be it print or digital; notation points to the symbolic code, i.e., the Latin alphabet; and language is defined as the way those symbols come together to signal complex meanings to those who are knowledgeable (Baines & Haslam, 2005).

Legibility: Legibility includes how individual characters can be distinguished. It describes how a typeface's glyphs can be correctly identified as characters and words. Although legibility influences readability, the two terms should not be confused (Dyson & Kipping, 1997).

Readability: It refers to a person's comfort level when interacting with text created in legible type (Dyson, 2014a). The person's unique characteristics, the text itself, and the typography, both the typeface and how it is used, affect the readability of the text.

virtual reality: A computer-generated virtual environment that allows users to travel through virtual settings that might be shown on a head-mounted display or a vast projection screen (Slater, 2003). The primary difference between VR environments and other digital media is immersion.

CHAPTER

1

INTRODUCTION

1.1 Focus and Scope

By a considerable margin, reading is the most common way to learn during college education (Verhoeven & Perfetti, 2008). While a printed format was the fixed medium for reading for a long time, it gradually changed with the emergence of personal computers and the development of versatile formats (Jury, 2009). One of the recent technologies to become more popular over the past decade is virtual reality (VR). Unlike other electronic media, VR allows users to enter a three-dimensional virtual world and interact with it. It has the unique power of invoking spatial memory, which is often quite good for compensating for other unstable types of memory (Bowman et al., 1998; Yang et al., 2020). However, compared to interactive and image-based representations, the impact of 3D typography on human memory has received less investigation, especially for lengthy reading materials with cognitively demanding concepts. Therefore, this research investigates how setting non-spatial and expository information in a three-dimensional landscape affects human memory functions, specifically recognition and recall.

Since reading is a major source of information acquisition (Fang, 2006a), typography, which is concerned with structuring and arranging visual language, is a crucial element of reading. When reading takes place in VR, several affordances come with the technology, shaping the way we consume typography. First, in spatial typography, elements become spatial and temporal. Second, they can be sculptural and 3D (Miller, 1996). While the latter mainly affects how individual letterforms are designed (i.e., extruded letters), the prior direction describes a navigational affordance for the text.

Human spatial navigation can be considered egocentric or allocentric depending on the navigator's viewpoint (Sanders et al., 2008). The egocentric frame of reference is the ability to perceive spatial relationships with respect to one's body by maintaining a dynamic spatial orientation (Brandt et al., 1973). By contrast, allocentric navigation offers a "map-like" navigation that relies on fundamental geometrical relationships between the objects, where the navigator's relationship with the elements in terms of distance, angle, and paths is fixed (Brandt et al., 1973). Outside of controlled experiment settings, navigation requires using allocentric and egocentric reference frames (Grech et al., 2018).

Literature suggests a difference in the level of automation between the two navigation frames. It appears that the tasks that require allocentric processing are less automatic than those in the egocentric frame of reference (Andrade & Meudell, 1993; Köhler et al., 2001). This is due to the attentional demands of allocentric reference, which requires a detachment from egocentric perspectives (Ruggiero et al., 2009; Huttenlocher & Presson, 1973). Detachment from self makes the navigation task more difficult, especially when the observer does not move physically but assumes a hypothetical movement (Huttenlocher & Presson, 1973).

Traditionally, in preparing complex textual systems, designers manipulate various properties of typographic stimuli to construct an easily recognized and navigated information landscape. Well-established and functional typographic conventions usually aim to provide the fastest access to information when reading a lengthy text (Haslam, 2006; Warde, 1955), following a top-left to bottom-right approach. That way, readers can confidently wander their gaze across the pages and navigate the information. Throughout this interaction, the spatial relationships of elements are kept allocentric, whether the reader's position is fixed or flexible. Conversely, a spatially distributed layout in VR can challenge these conventions by placing the reader in the middle of a complex information landscape and moving along the spatially distributed text layout. This environment can be discovered by a complementary navigational system in which the two navigation techniques (egocentric and allocentric) alternate, thus evoking more automated retention of the information (Skupin & Fabrikant, 2003; Yang et al., 2020). Therefore, unlike screen reading, one can process VR typography architecturally, creating a unique self-directed path through the text.

Designers also use several graphics organizers to contribute to the ease with which information is integrated into a mental model for traditional text. Graphic organizers are used depending on the text type and structure (Hartley & Trueman, 1983). Psychology models of text traditionally focus on two significant types of text; expository and narrative. Expository texts are characterized by difficult word or syntax structure, high conceptual density, and need for prior knowledge, suggesting a challenge to recall the content (Clinton et al., 2020). Within expository texts, five structural patterns reveal; description, collection, comparison/ contrast, causation, and problem/solution (Meyer, 1975). Text structures refer to how information in the text is organized. Research shows that while some readers are immediately aware of the structural pattern in expository text, some benefit from the physical organization of the content components (Richgels et al., 1987). Especially for scientific expository text, diagrams are such visual aids that help (1) deconstruction the bits of information from text, (2) label these chunks based on criteria presented by the text, (3) connect these labeled pieces by relevance and finally (4) create an organized visual space which supports the understanding of text (Horn, 1998). As a result, the content represented by the diagram can be incorporated into a mental model easier than in prose format. Because diagrams are spatial, their organization differs from prose's spatial arrangement.

It is clear from the literature that space supports human cognitive abilities such as recall, reasoning, and thinking (Andrews et al., 2010; Hegarty & Just, 1993; Bowman & McMahan, 2007), and we often use it to retain less complicated knowledge. Recalling a piece of information by its place on a notebook or spatially arranging the documents on a computer desktop according to their relevance are examples from our everyday lives (Polys & Bowman, 2004; Robertson et al., 1993; Ni et al., 2006) which are mainly using allocentric frame of reference. VR affords a highly controlled environment to discover the use of space in relation to one's body and its effect on the memory of expository text. In the current landscape, many successful VR applications seek a realistic experience that depends on high-fidelity sensory stimuli, i.e., high-level immersion, to produce a sense of presence (Bowman & McMahan, 2007). However, the application space's critical features depend on the task's nature. For example, for a military training application where soldiers use a virtual city filled with virtual enemies and friendly troops (Ahir et al., 2020), VR can provide a level of realism that is cost-efficient compared to real-world exercises. On the contrary, learning from a lengthy expository text and navigating through the information space requires being purposefully abstract (Bowman & McMahan, 2007; Bowman et al., 2003) since the level of realism does not enable the reader to perform the task better. Therefore, discovering the value of typographic representation in VR requires a novel approach to its form.

Memory is an information processing system where data is encoded, stored, and retrieved (Anderson et al., 2013). Recognition and recall are the two categories of memory retrieval that require different mental processes (Baddeley & Hitch, 1974). Word recognition in the context of reading refers to a linguistic process where written and spoken words are converted into linguistic representations through word recognition (Anderson et al., 2013). The mental process of retrieving knowledge from the past is called recall, and it has three types: free, serial, and cued-recall. A cued recall occurs when a person is given cues during the testing phase to aid in retrieving memories (Anderson et al., 2013).

The memory of reading first involves a visuoperceptual distinction of letterforms (Huttenlocher & Presson, 1973). A word is made up of letters that are arranged in a specific order (Kintsch & Bates, 1977). As a result, the reader must be able to distinguish between letters based on their positions. However, when multiple textual elements such as paragraphs, sections, and pages are included in the text, aspects of visual experience separate from spatial information (Kintsch, 1982; Kintsch & Van Dijk, 1978). For example, reading involves visual and spatial components, but processing a piece of information by its place in a notebook differs from how letterforms are processed using their serial positions (Anderson et al., 2013; Newcombe & Frick, 2010).

The spatial representation of text relates to the dynamic production of mental models. Wilson & Rutherford, (1989, p.6) give a comprehensive definition of the term: "A mental model is a representation formed by a user of a system and/or task, based on previous experience as well as current observation, which provides most (if not all) of their subsequent system understanding and consequently dictates the level of task performance." For desktop reading, such as an article in PDF format, mental models are built for the content of the text rather than the spatial representation (Butcher, 2006). Because many traditional desktop typesettings follow a familiar structure that matches readers' previous observations and experiences, a new mental model is rarely built. Moreover, the lack of a prior mental model for spatial reading may reduce readers' capability of processing content easily (Brewer, 1988) and make it hard to incorporate into a mental model. However, mobility and spatial ability are likely to act through physiological and behavioral mechanisms developed throughout human evolution (Cashdan & Gaulin, 2016) therefore increasing the likelihood of intuitively discovering text as an architectural system.

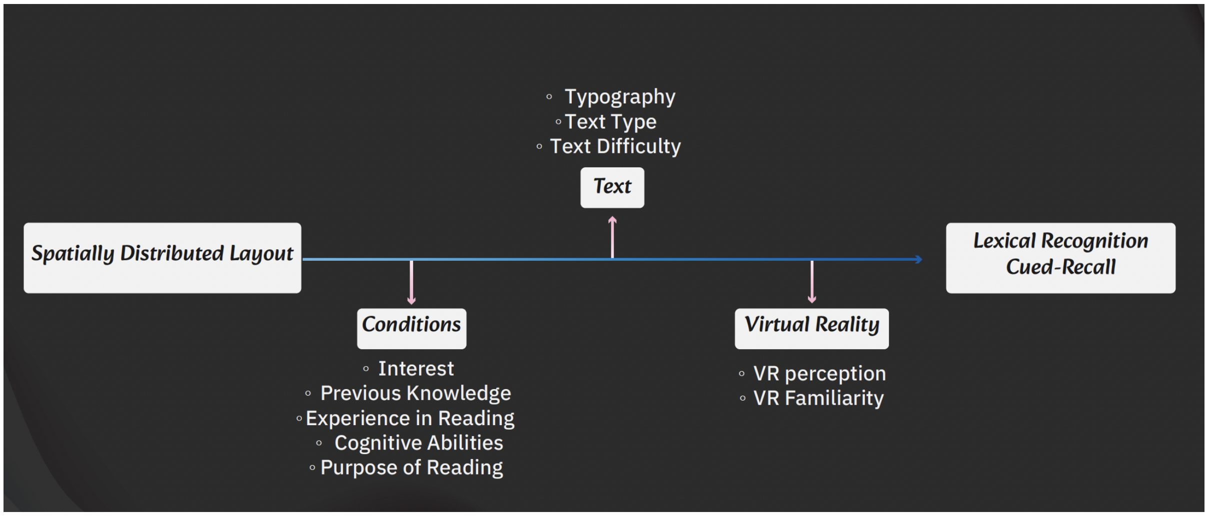
Today, many application spaces use typography but rarely exploit the aforementioned affordances three-dimensionality offers for spatial information presentation (Raja et al., 2004; Polys & Bowman, 2004). This may be due to spatial typography lacking tradition accumulated by the practitioners, its relatively short history, and a lack of theoretical base on which investigations in VR reading are built. As a result, VR typography has to draw on adjacent literature created by generations of knowledge and technique, which is focused on two dimensions. Nevertheless, unless there is empirical evidence of the usefulness of VR reading, the development of such history for VR text will be slower. Evidence of this slow pace is in current VR applications, which use text as an integral part of their system, but still apes screens, barely considering prose text beyond labels, annotations, and user-interface elements even after almost four decades after releasing the first commercially available VR headset. An important contributor to this delay is the slow adaptation of VR headsets for everyday use due to their high prices and the negative physical effects that induce eye strain and nausea (Nichols & Patel, 2002). Nevertheless, according to the International Data Corporation (IDC)'s research (Rydning et al., 2018), the VR market is projected to grow exponentially, being at the center of digital transformation in the near future. For this future to be realized, designers must synthesize, investigate, and iterate to make VR a comfortable reading tool (Small et al., 1994; Small, 1999).

1.2 Conceptual Framework

Figure 1.1 illustrates the conceptual framework representing the expected relationship between spatially distributed layouts and the two memory performances. In the context of this research, the relationship between text memory and spatial typography can be affected by several factors with several qualities within those factors; (1) characteristics of the reader, (2) text-related, and (3) VR-related factors.

Four factors can affect the memory of text concerning conditions. First, people who are

Figure 1.1: Conceptual Framework



Note: Conceptual framework representing the factors affecting the relationship between spatial layout and cued-recall and lexical recognition performances. Notice that there is no sequential relationship between the *Conditions, Text, and Virtual Reality*. They rather interact with each other as elements of spatially distributed layouts.

interested in a subject show greater attention, persistence, and performance in that domain (Hidi, 2001). Therefore, learning outcomes improve if individuals experience a mere passing curiosity or have an enduring fascination with a topic. For example, Fastrich et al. (2018) presented participants with trivia questions and asked them to provide their best guess for the answer and their interest in learning the correct one. One week later, participants were given a memory test on the questions. Results revealed a positive relationship between interest and long-term memory.

The second factor concerning conditions is previous knowledge. Prior knowledge may facilitate the processing of incoming information by providing a structure into which the new information can be integrated (Shing & Brod, 2016). Although it sometimes leads to an elaborated memory trace, prior knowledge can also hinder knowledge acquisition, particularly when the to-be-learned information is inconsistent with the presuppositions of the learner (Brod, 2021). Therefore, considering prior knowledge and knowing how it affects memory processes is essential for optimizing memory.

Experience in reading is the third factor that affects text memory. Studies demonstrate that poor readers read significantly less than their average-reading peers, emphasizing the impact of practice on the reading task itself (Perfetti, 1985). Moreover, skilled and less skilled readers' working memory performance cause deficits for less skilled readers across verbal and nonverbal conditions, suggesting a central processing deficiency (Swanson et al., 1989). The research also indicates that reading competency is best understood as dependent upon general language comprehension skill (Swanson & Berninger, 1995). In other words, the reader's proficiency in the language, such as vocabulary, syntax, and semantics, is important in predicting reading memory performance (Christoffels et al., 2006). Therefore, reading in a foreign or second language depends on reading ability and exposure to that second language.

The fourth factor is cognitive abilities, which refer to skills needed to acquire knowledge, manipulate information, and reasoning (Bjork & Bjork, 1996). Throughout the reading process, attention, visual processing, auditory processing, working, short-term and long-term memory, and sensory integration plays an important role, and these operations must be automatic for successful reading (Bjork & Bjork, 1996).

And finally, the purpose of reading affects how readers interpret texts differently based on why they are reading (Linderholm & van den Broek, 2002). In recent investigations of college students' use of reading purpose to guide text processing, students used different cognitive processes and strategies when reading for study versus when reading for entertainment (Linderholm, 2006). In particular, individuals read slowly and focus on cognitive processes and tactics, including inferences, paraphrasing, and rehearsing of text content when reading for study.

Text-related factors are shown in three categories; typography, text type, and text difficulty. Theories like the Interactive Model of Reading (Rumelhart, 1977) or Schema Theory of reading (Adams & Collins, 1977) illustrate that reading includes a micro-level perceptual task that involves recognizing patterns and organizing the information to pursue meaning and ideas. Several research indicates an effect of typographic parameters such as layout on text performance (Lonsdale et al., 2006; Dyson, 2004), different line lengths on comprehension (Dyson & Haselgrove, 2001) or tighter than standard character spacing on reading speed (Chung, 2002). This puts typography in a rightful place to consider as an important factor affecting reading processes.

The second factor concerning text is *the text type*. Psychology models of text traditionally focus on two significant types of text; expository and narrative. Expository texts are characterized by difficult word or syntax structure, high conceptual density, and need for prior knowledge, suggesting a challenge to recall the content (Clinton et al., 2020). Science books are examples of expository texts that aim to inform and communicate explicit ideas about a specific topic. Scientific texts often introduce a general theme and then elaborate on it in length, resembling a pyramid approach from general to specific elaborations (Meyer, 1975). Many college-level textbooks include prose text to explain general concepts and use diagrammatic representations to depict subsequent concepts visually (Hegarty & Just, 1993). A diagram refers to schematic illustrations in which the relationships between the textual elements are expressed graphically (Hegarty et al., 1991). In a diagrammatic representation, information is organized by location, which dictates specific inferences, and the relations signaled in diagrams depend on the convention of interpretation (Hegarty et al., 1991). For example, students have to learn how to interpret the layout of a diagram to understand the relationships conveyed through spatial metaphors, unlike prose which is linear in the reading method.

Text difficulty can be analyzed using quantitative measures such as syntactic simplicity, word concreteness, referential cohesion, and deep cohesion (Solnyshkina et al., 2017). The degree of challenge a text poses to the reader decreases the likelihood of remembering it. There are several online tools to measure text difficulty based on various metrics. Coh-matrix (Graesser et al., 2004), for example, is a computer tool that analyzes texts on over 200 measures of cohesion, language, and readability. Flesch and Kincaid is another one that reveals the required education, or grade level, to be able to understand a given text (Flesch, 2007).

Finally, two VR-related factors can affect the relationship between spatial layout and memory functions. They are VR familiarity and perception. VR studies involving performance tasks may sometimes lead to distractions for being engrossed by the VR environment, triggering a wow factor (Chaudhry, 2021). With a typographic composition, this factor may emerge because it is not usually how people read. However, it may also create discomfort due to less-developed graphics and hardware than desktop computers, leading to a negative experience. In this case, it is essential to discuss and prepare the readers for what they will receive. Moreover, the wow factor can relate to previous VR experience and familiarity. Those with previous experience in VR tend to have more control, while those receiving their first VR experience acknowledge that they have at least moderate control over the environment, which may cause frustration, and in return, it can affect memory (Jerald, 2015).

1.3 Research Aim, Objectives, and Questions

The primary **aim** of this dissertation is to investigate the effect of spatially distributed layouts for expository text on human memory functions, recognition, and recall. It is **hypothesized** that when the self and the information co-inhabit the same space, our spatial memory can be activated and employ a new way of encoding semantic information more intuitive than desktop reading. A secondary aim is to compare the results for two different text structures: diagram and prose because diagrams are thought to have a different effect on memory performances than sentential representations because of the differences in their spatial and structural connections and content (Larkin & Simon, 1987). Therefore, research question 1, "How do the spatially distributed layouts in VR impact lexical recognition and cued-recall of sentence location performances of prose scientific text?" aims to identify the contribution of spatially distributed layouts in recognition and recall test scores and the research question 2, "How does the impact of spatially distributed layouts in VR change the lexical recognition and cued-recall of sentence location performances based on prose vs. diagram structures of the scientific text?" aims to compare and contrast the effect of spatially distributed layouts for prose and diagram.

A between-subject experiment with 40 college students investigated whether a spatially distributed text influences cued-recall of sentence location and lexical recognition performance compared to flat/traditional layouts.

1.4 Significance

First, this research will help address the current shortage of research in this area and provide real-world value to students and learners in improving the memory of text and enhancing learning. This will be achieved by creating a controlled experiment procedure and a design sketch for VR reading to ask fundamental questions about the value of having text in space.

Secondly, this research will contribute to the body of knowledge on VR typography's unique and undiscovered challenges. Creating a compelling vision of how VR can redefine the reading paradigm is a new activity for practitioners and researchers. Therefore, some of the challenges regarding reading are expected to be discovered through this research. For example, because VR hardware is still evolving, the effort required to process the text increases concerns about VR sickness and nausea, especially for prolonged reading activities. Then, there could be others who are unknown to most system designers and researchers, which will be discovered through this research. Nevertheless, at this point in time, there is an immense opportunity for innovation and provoking a paradigm shift in how we experience reading. Instead of waiting for these challenges to disappear through a technology shift or the imposition of rigid rules inherited from designers and legacy technologies, this research reinforces a vision of the future of VR-mediated typography, which is necessary to instantiate ideas in the designers and researchers who will witness that future, while positioning users in the center of this discovery process.

And finally, this research will provide an experimental methodology that combines the latest typographic practice with science which is generated under controlled and defined settings to reveal the potential of VR reading. Given the short history of VR typography and empirical research on the cognitive value that comes with it, such methodology might be significant for future studies.

CHAPTER

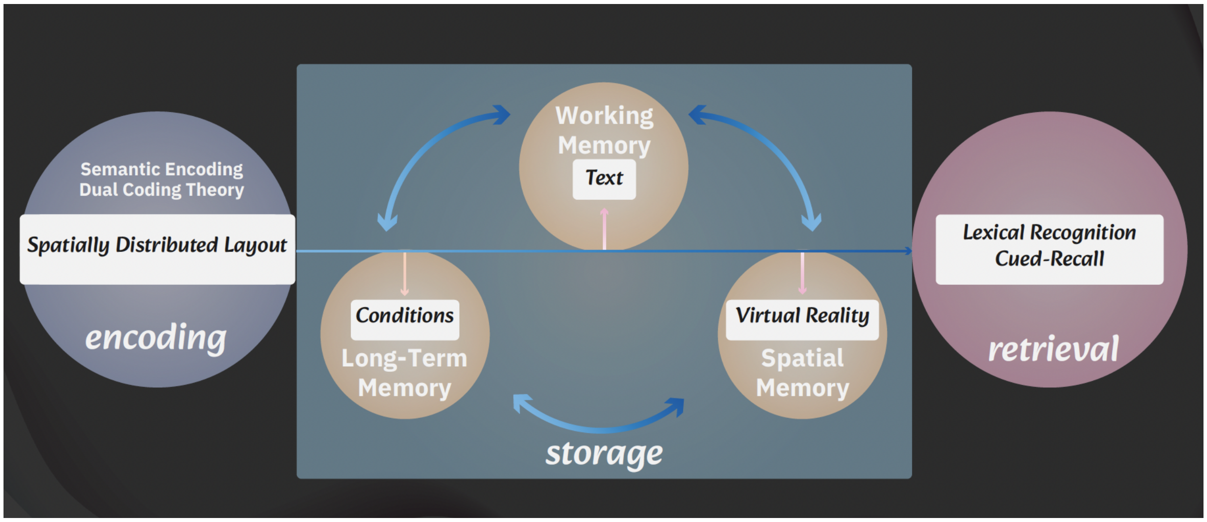
2

LITERATURE REVIEW

Figure 2.1 reveals the scope of this section, which is produced after a foundational review of existing theories. It serves as a roadmap for developing the arguments in this research. The main focus of the literature review is on understanding the role of different types (long-term, spatial, semantic, and working memory) and operations (encoding, storage, and retrieval) of memory, which may be activated or used through spatial reading.

Literature suggests that words' perceptual processing does not occur until after the meaning is registered in the semantic memory (Jones et al., 2015). Once the participants see the spatial layout, they are expected to register the meaning and typographic features in different cognitive subsystems; spatial and verbal (Clark, 1987). Dual Coding Theory suggests that integrating spatial and verbal information is in and of itself a form of deeper processing (Clark, 1987). Working memory is temporarily activated once the perceptual and verbal information is registered, and it works as a rehearsal buffer which can be described as a storage system that can hold information (Baddeley & Hitch, 1974). The spatial layout is expected to be registered at the visuospatial sketch pad, creating a higher level of memory that is positively correlated with the memory of text. Together with that information, familiar

Figure 2.1: Theoretical Framework



*Note.*Theories that served for developing the arguments this research use and their relationship with variables of interests and moderators.

features with what has been stored in long-term memory, such as previous knowledge, the purpose of reading, topic interest, and experience, create temporary representations in working memory (Shing & Brod, 2016; Brod, 2021; Perfetti, 1985; Swanson et al., 1989; Bjork & Bjork, 1996). Because this is a new way of reading, the novel mental model is formed until a cohesive, repeatable action sequence for the reading is possible (Norman & Bobrow, 1979; Piaget & Cook, 1952). Due to the use of spatial cognition, this process is expected to be fast and easy (Yang et al., 2020; Grech et al., 2018). And finally, this whole process is expected to be affected by the perception and familiarity of the participants with such a spatial representation of reading. All these factors are expected to change the duration and information storage capacity, and for retrieval to succeed, all three stages need to be intact (Baddeley & Hitch, 1974; Lockhart, 2002; Eysenck, 1978).

2.1 Factors Influencing Memory

2.1.1 Depth of Processing and Meaningfulness of Information

A general inference one might make about text memory usually involves studying and practicing the material (Anderson et al., 2013). For example, Rundus (1971) asks participants to recite a list of items aloud and shows that the more participants rehearse the words, the more likely they are to remember them. Craik & Lockhart (1972) were critical of this view and claimed that memory depends on the rehearsal and the depth of information processing. His theory called "Depth of Processing" was proven by several studies. Depth of processing defines semantic processing, which happens when we encode the meaning of a word, as a strong encoding method over structural (e.g., typography of the word) and phonemic (e.g., rhyming words together) processing. For example, Kapur et al. (1994) showed that participants who were asked to remember if the words contained a particular letter (shallow processing) remembered 25% fewer words than those who were asked to remember whether the word described living things (deep processing). However, the theory was criticized for only describing the processes rather than explaining the mechanisms underlying the theory, leading to better memories (Eysenck, 1978). In recent years, research brought more empirical evidence to this criticism, and it appears that deeper coding produces longer retention due to elaborative processes.

2.1.2 Elaborative Processes

Elaborative processing enriches the memory representation of an item by relating it to the pre-existing network of semantic associations (Anderson et al., 2013). Frase & Schwartz (1975) presents evidence of elaborative processing with text material by comparing two groups. The experimental group receives questions about the stimulus text and is prompted to find answers to these questions, and the control group studies the text as they normally would. A memory test consisting of the relevant questions with the previously read questions and other irrelevant ones reveals that 76% of the experiment group answered the relevant questions correctly. In contrast, the control group answered only 52%. For the irrelevant questions, the experiment group still performed slightly better (64%) than the control group (57%). Thus, participants did much better on the topics for which they were given questions beforehand.

Mental imagery is another effective method for forming meaningful elaborations. *Method of loci* is an ancient technique first described by the Greek poet Simonides, using spatial mental models (Wang & Thomas, 2000). According to the story, Simonides delivered a poem at a banquet that praised the gods, Castor and Pollux. The roof fell when he left the banquet, killing all the attendees. As the only witness alive, Simonides was called to identify mangled bodies of people. In doing so, Simonides used where each person sat during the banquet. This was documented as a useful memory technique to remember information by organizing it into locations. To use the method of loci, one first imagines a familiar route with fixed locations, e.g., a path from your childhood home to the grocery store. Then as one mentally walks through her path she associates the objects with the fixed locations. A grocery list composed of items like cat food, milk, bananas, and bread can be associated with this loci, for instance, by imagining a package of cat food fed to the neighbor's cat and envelopes lying in a puddle of milk in front of the post office on the way. The method is highly flexible, and the same loci can be used several times for different sets of items (Christie & Just, 1976). Two important principles underlying this method's effectiveness relate to the imposition of an organization to a formerly unorganized list, but also the connections between the loci and the items forcing us to process the material meaningfully, elaborately, and by the use of imagery (Anderson et al., 2013).

The effectiveness of this method for scientific text was tested through VR. In their research, Yang et al. (2020) compare participants' recall and recognition performances for scientific text between a control condition without an advanced memory strategy as a baseline measure, a second condition using an image-based variant of the method of loci, and finally, an experiment condition using a VR-based variant. They use a coffee shop scene in which the pieces of text are attached in different locations. The results show that the VR-based technique moderately increased recall accuracy and recognition precision. In this case, the loci were unfamiliar to the participants. The experiment group had to study the coffee shop beforehand, which accounts for the cognitive difficulties of setting personally intimate loci.

2.1.3 Comprehension

Besides the depth of processing and the meaningfulness, other factors determine the levels of memory. According to Craik & Lockhart (1972), memory is inevitable and an imperfect by-product of text comprehension. How the information is comprehended has implications for how it is remembered and what was understood in the first place (Harris et al., 2013).

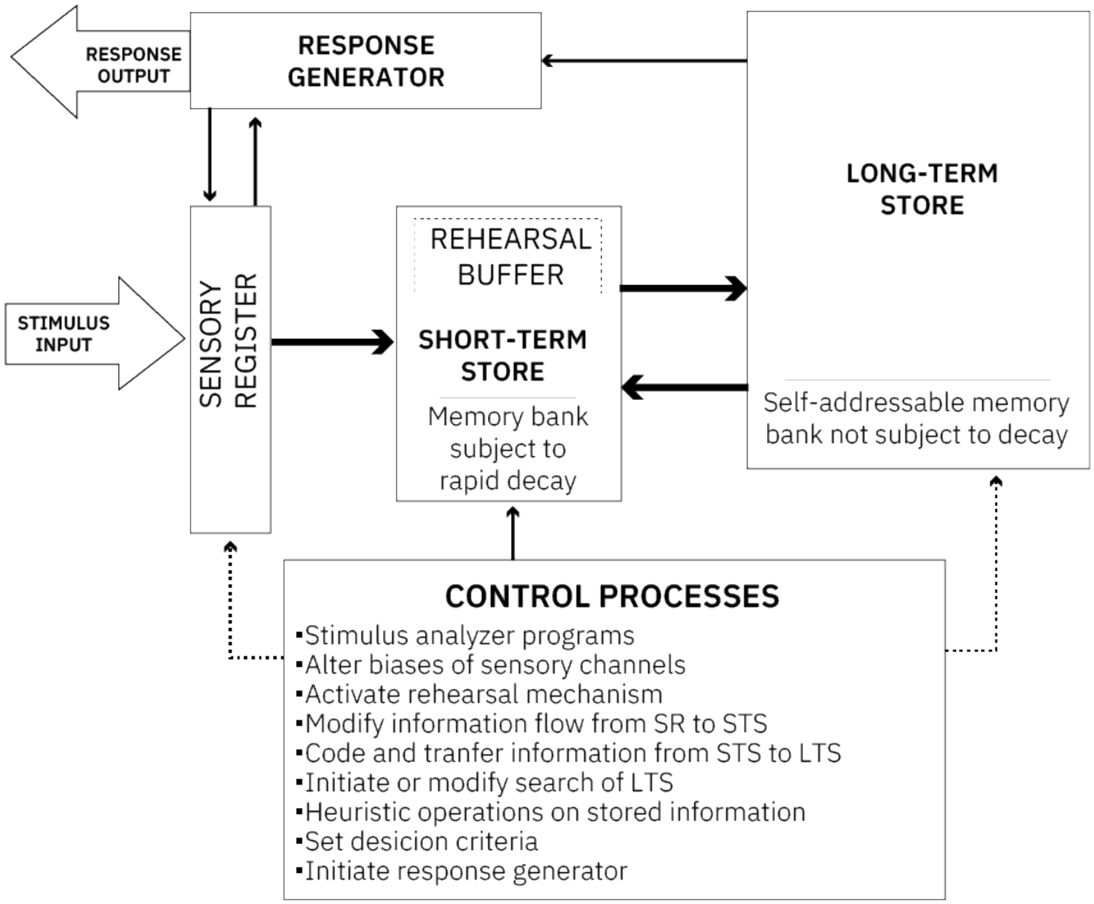
Comprehension is first affected by the factors related to the reader's characteristics such as topic interest, previous knowledge of the topic, motivation, experience in reading, and/or cognitive abilities (including aging) (Kintsch, 1994; Guthrie & Kirsch, 1987), as well as the purpose of reading (Roehling et al., 2017). Secondly, text characteristics, including lexical density, syntactic complexity, and semantic abstractness (Sadeghi, 2007) play an important role in comprehension. Lexical density refers to the statistical measures of the lexical richness of texts (Daller et al., 2003). Syntactic complexity refers to the level of sophistication and challenge of a reading selection or other type of text (Frantz et al., 2015), and semantic abstraction defines the mechanism for representing things in a simplified manner in the text (Nagels et al., 2013). Moreover, text type, topic, genre, and the writer's style have also been recognized as factors affecting reading comprehension (Alderson, 2000). Finally, the effect of context refers to situational and environmental elements in which the text is read (Sadeghi, 2007). They can vary from environmental distractions, the time during which reading occurs, and social elements (Irwin, 2021; Sadeghi, 2007).

2.1.4 Long and Short-Term Memory Interactions

As can be seen, comprehension involves complex processes and is affected by many factors. During this process, the text is understood as new perceptual inputs connected to past knowledge or experience from our memory (Harris et al., 2013). Traditionally, information was thought to arrive in sensory memory; after being processed in short-term memory, it is then registered in long-term memory (Atkinson & Shiffrin, 1968). Recent views pose that encoding information with familiar features with what has been stored in long-term memory creates temporary representations in working memory (Cowan, 2014). Indeed, some researchers think working memory is a fragment of long-term memory temporarily activated as the information is registered (Shiffrin & Schneider, 1984; Cowan, 2014). In this case, a stimulus arrives at sensory memory and is matched with the long-term memory representation, and together they form a mental representation in the working memory. At the same time, a new stimulus can be connected to a different stimulus in the long-term memory while rehearsing. It then forms new associations that can be stored as new knowledge (Cowan, 2014). Applied to text memory, this means that text memory is created during the reading process and by a constant interaction with our previous knowledge and experiences stored in long-term memory instigated by the stimulus's comprehension.

According to the Atkinson & Shiffrin (1968) *Modal Model of Memory*, long-term memory is the final stage of sensory registration and short-term storage. A rehearsal buffer is described as a storage system that can hold information and is characterized by its capacity, duration, code, and the reasons why the information is remembered (Atkinson & Shiffrin, 1968). Figure 2.2 shows the processes of the model: (1) the sensory information store holds onto the information about the scene (e.g., color, form, or angles) for a very short period of time, (2) the short-term memory store retains the information if there is no rehearsal for only about 10-20 seconds, and finally (3) long-term memory storage contains information which goes back in time as far as years, months, and days (Atkinson & Shiffrin, 1968; Tulving, 1974).

Figure 2.2: Modal Model of Memory



*Note.*Modal Model of Memory focuses on how rehearsal helps information move from short-term to long-term memory. (Atkinson & Shiffrin, 1968)

The effect of sensory and perceptual attributions of stimuli on long-term memory was thought of as relatively transient (Sachs, 1974) and was challenged by different studies (Tardif & Craik, 1989; Kolers, 1973, 1976, 1979). In his study Kolers (1976), had his subjects read 160 pages of text in which words were printed upside-down. A year later, subjects reread 49 pages of the same text with an addition of 49 new pages taken from the same source. Subjects reread the old pages 5% faster than the new pages, which (according to Kolers) was due to the retention of pattern-analyzing skills specific to the reading passage. His reasoning for these results was two-fold. First, he claimed participants' pattern-analyzing skills were highly specific to the particular passages read. It can not be explained simply by retaining the newly developed reading skills for the upside-down typography. Secondly, he refers to his earlier experiment (Kolers, 1976), which revealed that semantic content was a less useful aid for reading transformed typography than analyzing typographical forms as objects. Tardif & Craik (1989) later ran a similar experiment to understand the association between the newly acquired transformed typography reading skills and retention of meaning. The results showed that pattern-analyzing operations or semantic factors could not account for memory solely; rather, both perceptual and conceptual factors play a role in long-term storage (Tardif & Craik, 1989). They concluded that the human information processing system is flexible with the degree of transformation from one occasion to another, depending on how encoding takes place and the readers' goals (Tardif & Craik, 1989).

Therefore, long-term memory does not only function as a place for information to be stored, but it also becomes a place where the ways in which responses to previous encounters are saved.

2.2 Reading Processes

The number of theories explaining what happens during reading processes is excessive, and an overview of each seems impossible. Although not exhaustive, the theories detailed here offer valuable insights into the complex nature of reading. Moreover, they all refer to a traditional mode of reading which usually refers to print media and mainly investigates the phenomenon from a cognitive perspective rather than a psycholinguistic or typographic perspective.

2.2.1 Bottom-Up Processing

Early theorists like Gibson (1969) focused on perceptual factors and noted that readers construct meaning sequentially, using bottom-up processing. This theory suggests that the reader's construction of the higher-level meaning starts from the smallest text units (letters, sounds, etc.), and comprehension occurs through a meaningful assembly of these smaller units of information (Kong et al., 2018).

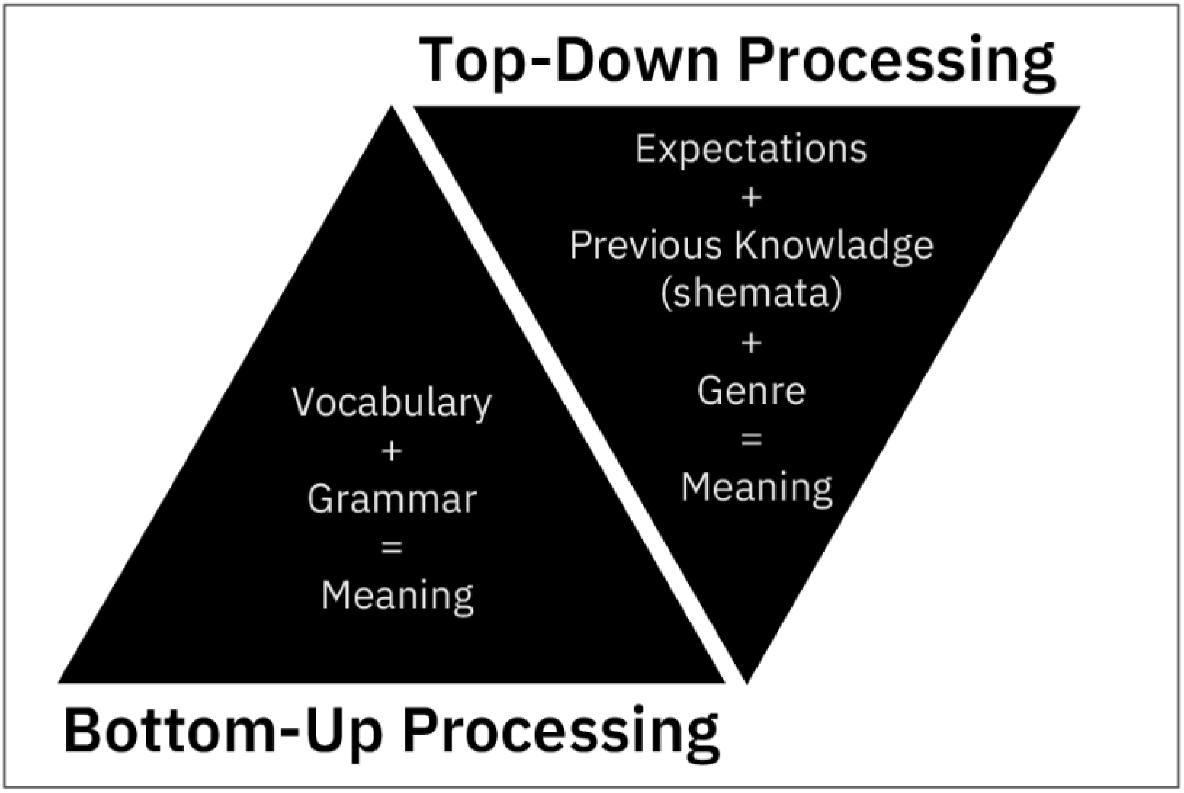
Gough et al. (1972) was one of the early proponents of this theory who presented a model of what happens in the first second of reading when letters are reflected on the retina. The model consisted of three main stages: (1) the iconic representation of the written words, (2) recognition of the letters, and (3) the positioning of word meanings in the memory through their phonological recording, where they turn into sentences eventually. According to this model, all stages occur serially where the higher processes do not interfere with lower processes(Gough et al., 1972).

Other researchers criticized the model for its claims on the serial processing of words and recognition of each word through the phonological form. Investigations on the word processing order (Reicher, 1969) and phonological recording hypothesis (Gough, 1984) argued otherwise, that the parallel processing was possible and the recognition of the words was unmediated by the phonological form.

2.2.2 Top-Down Processing

In contrast to bottom-up processing, this theory proposes that the readers use their prior knowledge, experience, and expectations to guide their processing (Goodman, 1976). Also known as the concept-driven model, this approach emphasizes a method of hypothesis verification based on the reader's prior knowledge and the writer's constructs or ways of organizing the knowledge (Goodman, 1976). The reader then builds transactions with the text, which results in a transformation of schemata in the pursuit of meaning. The process does not necessarily require a precise perception of all the elements in a word but rather a selection process that involves finding the most effective cues.

Figure 2.3: Bottom-up and Top-down Processes of Reading



Note. The bottom-up approach emphasizes the development of fundamental abilities and the matching of sounds with the letters, syllables, and words that comprise the text. The top-down approach, relies on the reader's background information to understand written material.

This theory was criticized for emphasizing contextual cues and identifying reading as a "psycholinguistic guessing game" (Goodman, 1976, p. 9). A body of research supports that readers use contextual cues only when orthographic and phonemic cues are minimal (Liu, 2010) and good readers still rely on graphical information (Stanovich, 1980).

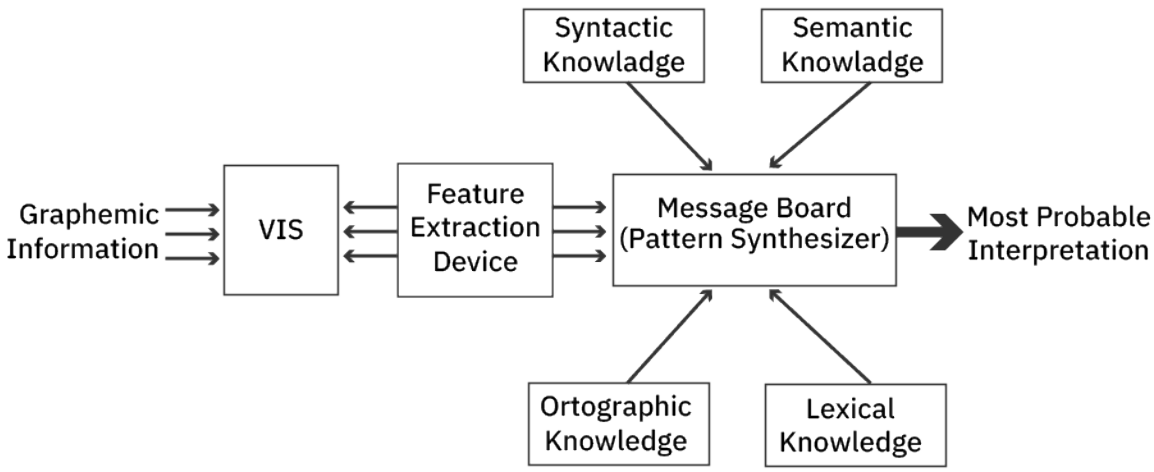
Figure 2.3 compares the bottom-up approach, which concentrates on developing fundamental abilities, such as matching sounds with letters, syllables, and words in the text composition, and the top-down approach focuses on the reader's prior knowledge and emphasizes the importance of schemata, or prior experiences and background knowledge, in understanding the text.

2.2.3 Interactive Model of Reading

Proposed by Rumelhart (1977), the interactive model aims to combine bottom-up and top-down processing, recognizing that both approaches are important for effective reading. Processing in an interactive model of reading starts with a set of expectations about what information is likely through visual input. Such expectations or prior hypotheses are formed based on the reader's knowledge of the words, letters, phrases, etc., and larger parts of discourse, including non-linguistic aspects. As the visual information from the page meets the eye, the expectations consistent with the input strengthen, and those inconsistent weaken (Rumelhart, 1977; Rumelhart & McClelland, 2017). In other words, if the interactive processing uses the information acquired from the text to activate the relevant parts of the reader's preexisting knowledge structure, it produces the highest level of comprehension (Kintsch et al., 1999; Kintsch & Walter Kintsch, 1998).

Figure 2.4 illustrates the stages in the Interactive Model of reading. First, the graphemic information enters the system and is registered by the visual information store (VIS). A feature extraction device extracts the necessary features from the VIS. The sensory or non-sensory information, such as the orthographic structure of the language, including a probability of various string characters about these features, is then served to the pattern synthesizer. The pattern synthesizer then uses all this information to create the most probable interpretation of the graphemic input. Thus, the reading process is positioned as the by-product of applying all knowledge sources.

Figure 2.4: Interactive Model of Reading



Note. The interactive reading model is a reading theory that acknowledges the simultaneous interplay of bottom-up and top-down processes during reading. (Rumelhart, 1977).

2.2.4 Schema Theory of Reading

Except for the bottom-up approach, which was criticized for overlooking the reader's prior knowledge, previous models emphasize background knowledge as an important factor in the reading process. The schema theory formalizes the role of background knowledge in the reading process(Shen, 2008). Piaget (2003) defines *schema* as "a cohesive, repeatable action sequence possessing component actions that are tightly interconnected and governed by a core meaning." A simple example of a schema can be ordering food at a restaurant where a pattern of behavior developed from an initial structure that directs the steps to be taken appropriately: looking at the menu, ordering the food, eating, and paying the bill (Schank, 1975).

Schemata are powerful in their influence. For instance, Brewer & Treyens (1981) asked subjects to wait in an office for 30 seconds, and the researchers asked the subjects what they saw in the office when they were removed. Most subjects reported seeing objects like books, filing cabinets, and coffee cups, although they did not exist in the environment. In this case, the presumption is that a schema for "office" includes such objects (Hampson & Morris, 1996).

An essential idea in the schema-theoretic account of reading is that it involves an integrated activity of schemata at both high and lower levels of analysis (Adams & Collins, 1977). A lower-level schema for reading starts from recognizing the letters, words, and, finally, sentences. Concurrently, higher-level schemata generate more meaningful and comprehensive levels of presentation (Anderson & Pearson, 1984). At the letter level, for example, schemata for "L" can be as concrete and specific as a vertical line connected to a smaller horizontal line from the bottom right. On the other extreme, they can be abstract and general. Tentative problem-solving schemata by Rumelhart (1977) states: "E causes P to want G." *Cause* and *want* in this schema are themselves schemata like the letters are (Anderson & Pearson, 1984). As the lower level schemata are activated, they are bound to activate the schemata at different levels (Anderson & Pearson, 1988).

The fact that lower and higher-level schemata work simultaneously puts an emphasis on pattern-analyzing mechanisms, both on the single-letter level and on whole-word patterns. As we read more, our schemata become more and more complex. Words become paragraphs, paragraphs become columns, columns become pages, and pages become books. This nested approach allows lower-level (such as perceptual elements) and higher-level schemata (like problem-solving schemata) to unite with meaning. It provides a structure to mentally visualize the inter-relationships between the two levels of processing (Anderson et al., 2013).

As one of the oldest forms of documentation, reading books involves schemata that are widely familiar to most of us on a perceptual level. Preexisting information in our heads about what composes a paragraph (introducing a single theme indicated by a new line, indentation, or numbering) or a title (a description of page content) makes it easier to form lower-level schemata almost instantaneously as we store this procedural information in our minds. The schemata of reading from a book are somewhere in our memory. In other words, schemata depend on past experiences and are accessed to guide the process of understanding current situations (McVee et al., 2005; Piaget & Cook, 1952; An, 2013). When text becomes spatial, reading schema is challenged due to the lack of experience, even for a skilled reader.

2.2.5 Mental Models and Reading

A recurring argument against schema theory focuses on its insufficiency to account for specific aspects of human knowledge and their meaning (Brewer, 1988). By Piaget & Cook (1952)'s definition, schemata serve as the means to assimilate information into preexisting cognitive structures. The essential difference between schemata and mental models is that schemata are permanent in contrast to mental models. For accomplishing a task or solving a problem, mental models serve as tools of accommodation that involve reconstruction of the cognitive structure, which is specifically useful for understanding new situations or an obscure problem (Al-Diban, 2012). In other words, schema uses prior data structures in memory, while mental models utilize such information dynamically, forming new mental representations (Mishra & Brewer, 2003; Wilson & Rutherford, 1989). However, if the new information is assimilated into the activated schema, constructing a mental model will be unnecessary (Al-Diban, 2012).

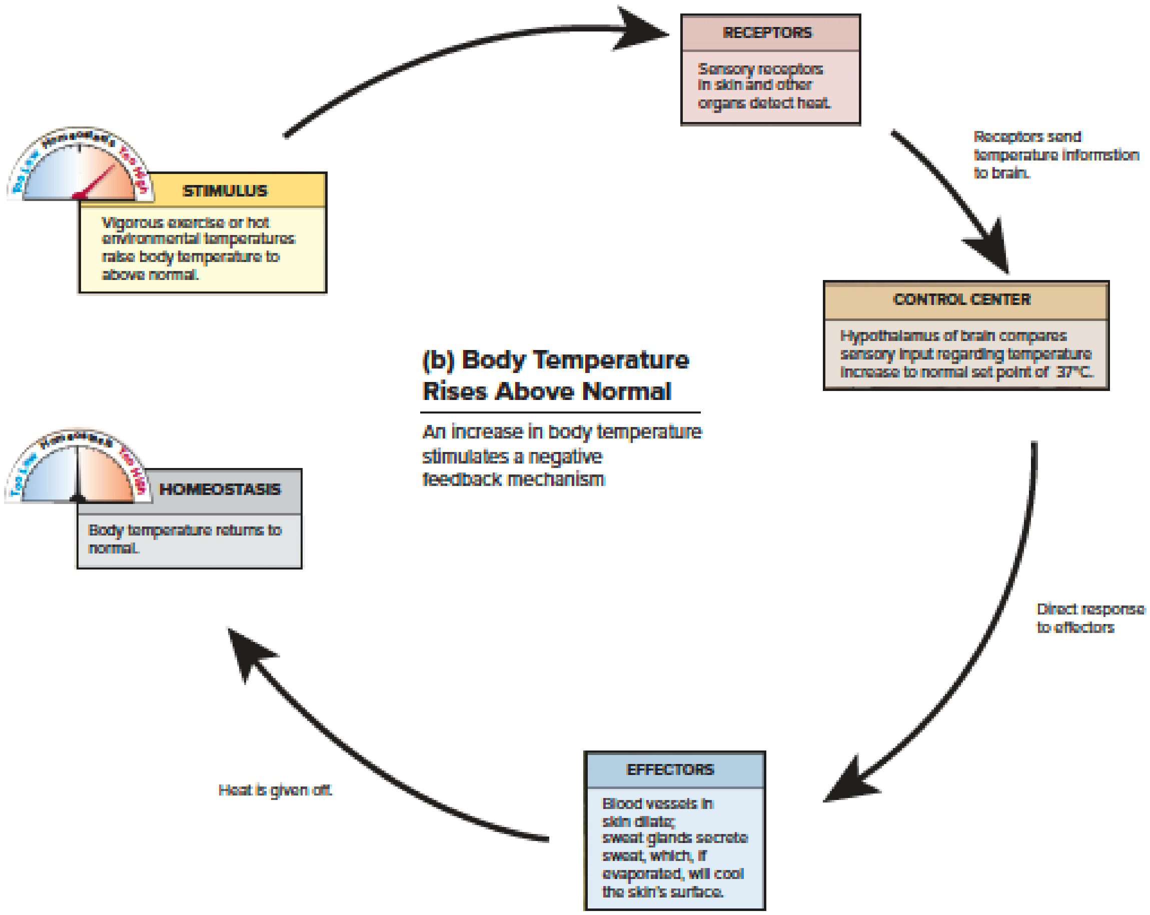
Mental models aid people in making inferences and predictions about the structure and internal relationships of any system and its behaviors. Some of these predictions include the system's behavior in future instances (Borgman, 1986). According to Norman (1981), interactions with the system develop mental models based on the clues gathered during this time. If there is no reference on which to base the mental model, a new model is constructed and is likely to be flawed. In other words, people trained with a conceptual model should use the system better than those who do not.

2.2.6 Mental Models and Typography

Although writing systems vary across cultures and times, conceptual knowledge about the typographic structures is not hard to learn to recognize (Bringhurst, 1996). The principles that unite all the texts are based on the visual structures which are unique to mankind. Typography is central in providing visual cues for users' navigation by implementing consistent visual components such as headings, subheadings, and page numbers (Haslam, 2006; Lupton & Miller, 1999). Similarly, typographic hierarchy instantly signals the most important to least important information by distinguishing between major sections and subsections (Lorch, 1989). Through these strategies, designers help readers form a more accurate mental model for the navigation and structural understanding of the text. However, it is important to note that these strategies change according to different text structures (Bringhurst, 1996).

Consider Figure 2.5, which shows the components of the homeostatic control mechanism. Unlike in prose text, here, the organization of elements and their relationships introduce the reader to a dynamic and particular representation that helps them to create a mental model which helps them infer meaning in the reading process. In this case, a cause-and-effect relationship exists between each element of the homeostatic control mechanism. Each element is created by the movement of the eye jumping from one panel of text to another. The spatial distribution of the components in that example (1) logically organizes the causes, (2) parents different elements of the system to one another, and (3) signals qualitative relationships between the entities. This mental model can be constructed numerous times using different spatial configurations, such as a tree diagram for the same content.

Figure 2.5: Textual Diagram



Note. An example of a textual diagram from McKinley et al. (2013). Diagrams are schematic drawings that graphically depict logical relationships, such as tables, networks, or embedded boxes.

Norman & Bobrow (1979) suggest a semi-formal model for memory retrieval based on the level of variability in the specification (i.e., the *description* of a prototype). According to this theory, people understand descriptions in terms of previously known *prototypes*, and retrieval starts with a description sought from memory. They suggest that what remains are those schemata that are efficiently formed into compounds or stored in memory. To that end, they argue that text-driven diagrams activate schemata in memory. Still, specific mental models are built dynamically based on the relationships between the elements of diagrams (i.e., descriptions). Thus, how mental models are built for different text structures depends on the person's memory capacity and variability for such specifications. As a result, reading embodies a number of aspects of memory (Estes, 1977; Just & Carpenter, 1980).

2.3 Text Memory

At the beginning of the 1800s, the first important accomplishment of experimental psychology in understanding the reading process revealed that there are more complex processes one undertakes through the interaction of perception and memory (Estes, 1977).

According to Craik & Lockhart (1972), encoding text involves several levels of processing, such as graphemic, phonemic, lexical, syntactic, and semantic analyses, with each leaving traces in the memory. A comprehensive understanding of text memory must include a description of the possible various traits of each of these levels. The primary interest of this section is the graphemic level of text memory. Craik & Lockhart (1972) argues that encoding surface forms of graphemes (also called glyphs) is an example of shallow processing, and their longevity is less than other kinds of levels. Depth of processing is defined as the "meaningfulness extracted from the stimulus." Graphemes are expected to be relatively transient in their effect on memory.

The following literature criticized the notion that sentences are only encoded and recognized based on their semantic features through a series of studies. Kolers (1973), showed subjects' 14 sentences that appeared twice in normal typography, 14 that appeared twice in inverted typography, and showed that sentences in inverted format remembered better. Kolers & Ostry (1974) revealed that information about typography was retrieved on a recognition test after 32 days of initial reading, arguing that memory for sentences can occur regarding the pattern-analyzing activities used in their encoding. Roediger & Blaxton (1987) discovered that consistent visual presentation conditions increased memory performance, i.e., memory was better when items were studied and tested in the same font style than when they were changed between study and test.

Several researchers also have examined the effects of different types of signaling devices, including titles, headings, and number signals, on reading and memory processes. Brandt et al. (1973) demonstrated that readers attend more closely to statements they perceive as title-relevant and use them as context cues. Hartley & Trueman (1983), looked at the position of the headings and their contribution to the text recall and showed that headings aided recall of the text, although the position of the headings (marginal or embedded) had no effect.

However, the precision of the encoding process is vague in most of these studies (Eysenck, 1978). For example, it can be argued that because inverted typography takes more time to read, it may require greater cognitive effort and lead to information retention for longer periods. More recent research tested this argument using Sans Forgetica, a perceptually difficult-to-process typeface that claimed to affect students' memory positively. The research defeated this claim and revealed that it did not elevate text recall compared to an unmodified pure reading condition (Geller et al., 2020; Taylor et al., 2020). A practical elaboration of the encoding process should occur to understand the mechanisms behind these claims and their reliability.

2.3.1 Encoding Typography

Three common operations of human memory include encoding, storage, and retrieval (Palladino et al., 2001). These are sequential events. First, the information is gathered, then stored for a while, and finally, retrieved from memory (Palladino et al., 2001). The information is stored in three ways: first in the sensory stage, then in short-term memory, and ultimately a few portions of the information are transferred to long-term memory. Typography specifically affects how we encode visual information at the sensory stage (Dyson & Kipping, 1997). Weinstein & Mayer (1983) define encoding as transforming basic perceptual experience into internal connections that allow thinking about the information. Two important characteristics of encoding are *selectivity*(Kahneman, 1970) and *elaboration* (Bransford et al., 2014).

Selectivity is strongly related to attention. Because our attention is limited (Treisman, 1986), organisms selectively attend to certain stimuli or their aspects more than others (Kahneman, 1970; Baddeley, 1982). By choosing which stimuli to attend to, organisms control their behavior. Treisman (1969) suggests a taxonomy for selected operations of attention. According to this taxonomy, the four variants of selective attention are inputs, targets, attributes of the objects, and responses. Given many particular stimuli, subjects may be asked to perform a task by making certain discrimination at different stages. For example, if one has to find capitalized letters from a text, selection will be directed by size discrimination. At the same time, responses will be controlled by discrimination of letter shape (Kahneman, 1970).

Elaboration refers to the effect of previous knowledge and experience in obtaining new information (Stein et al., 1984). This allows for the use of prior knowledge to facilitate retention. However, we don't always encode what "really" happened, but what we *interpret* to have happened (Loftus, 1992). The processes that lead to false interpretations may stem from self-referential encoding or a reduction in the brain's perceptual capacity due to sleep deprivation, multitasking, etc. (Sperling, 1963).

Encoding in the reading process starts with extracting the physical features of text (Just & Carpenter, 1980). McKoon & Ratcliff (1992) in their *minimalist hypothesis* of text processing identifies two sources that produce the representations for text to be encoded: explicit statements from the text and general knowledge. Any piece of information which is not explicitly stated in the text is described as *inference*, which they label as *automatic* or *strategic*. Automatic inferences are encoded without special goals, and strategic inferences are encoded when the readers adopt special goals. Secondly, they distinguish between the *local* and *global* inferences. Local inferences involve information close to the text, while global inferences involve widely separated ones. Therefore, local inferences live in working memory because they are quickly and easily available. In contrast, global inferences are not automatically encoded due to being widely dispersed on the global scale (McKoon & Ratcliff, 1992).

The visual representation of text usually facilitates the interaction between global and local inferences. To achieve intuitive navigation and hierarchy of information, typographers use type in different placement, sizes, and tonal boldness (Clair & Busic-Snyder, 2012). They establish an order of importance and a sense of navigation within the page. For example, running headers and footers usually are the quietest elements of a page. They are frequently placed on the margins and use light sans serif, italic, or small caps letters (Clair & Busic-Snyder, 2012), signaling a sensible sequence of the content. When readers read the body text, they discreetly build a sense of orientation and navigation for the global structure of the book. In contrast, bold letters are used to call attention to certain words by increasing the point size or changing the typeface to something different than the body text (e.g., bold fonts, decorative fonts, etc.). They highlight critical details which have central importance within the text.

Waller (1979) defines devices that aid readers in accessing and reading the text in many ways as *access structures*. Such structures determine where a reader's attention is directed and indicate what information is most emphasized. In other words, access structures imply small-to-large links between typographic elements and emphasize central ideas by deemphasizing tangential ideas Meyer & Rice (1982). Meyer & Rice (1982) suggests that knowing hierarchical text structures within the text can help readers select the most important information through encoding. Moreover, they allow for the transfer and construction of ideas in working memory (Meyer & Rice, 1982). In fact, research suggests that the reader updates the information supplied by the text as they read (Glenberg et al., 1987a; Bower et al., 1979). However, while local inferences can be handled on a sentence level within working memory, readers are reminded of earlier stages purposely when the text gets lengthy (Waller, 1987). This is a consequence of the limited processing capacity of working memory (Baddeley, 1982). Treisman (1986) suggests that meaningful wholes precede the parts and properties while looking at a visual field. In part, the evidence for such operation lies in the physiological processes of sensory data in different areas of specialization in the human brain. For example, while one area might concern itself with the orientation of lines and edges, others can be concerned with color or direction of movement. Therefore, deeper and simultaneous encoding of each piece of information can not be achieved due to the limited processing capacity of humans (Meyer & Rice, 1992).

2.3.2 Semantic Encoding

Tulving and Donaldson (1972) proposed a model distinguishing episodic memory from semantic memory. Semantic memory embodies the organized knowledge of individuals, whereas episodic memory refers to the memory of personal events. Semantic memory stores words and verbal symbols and allow for their manipulation. Their meanings and referents, relations among them, and ways to manipulate symbols, concepts, and relations are processed in semantic memory (Donaldson & Tulving, 1972).

Initially, Tulving's theory stated that the semantic system does not record the perceptible properties of the input but rather registers the meanings of the signals. Later expansions proposed that perceptual information is not encoded in episodic memory until an interpretation of its meaning is processed in semantic memory (Donaldson & Tulving, 1972). In other words, encoding perceptual information does not occur until the semantic memory interprets it. Therefore, the output of the semantic memory becomes the input for the episodic system, and it is then stored in parallel (Donaldson & Tulving, 1972).

The effect of semantic processing on memory has been studied widely (Hyde & Jenkins, 1969) and found to be superior for semantic tasks over non-semantic tasks (Delogu et al., 2009; Stein et al., 1984). Previously mentioned Craik & Lockhart (1972)'s levels of processing model accounts for this superiority claiming if the information is processed deeper, memory trace will last longer. For example, describing the typography of a word is considered a superficial way of analysis, whereas having the subjects rate the meaning of the words on a scale of pleasantness is considered a deeper semantic level analysis (Bellezza et al., 1977).

According to Clark (1987) *Dual Coding Theory* (DCT, which combines verbal and nonverbal stimuli, is one method for rich or elaborate processing. For instance, Frase & Kammann (1974) shows that sorting animal names according to a dual criterion resulted in better recall than sorting them according to a single feature. DCT states that the human mind processes imagery and verbal information differently along different channels, resulting in distinct representations of information processed in each channel. These representations' mental codes are used to organize incoming information for later use. Therefore, imagery and verbal codes can be derived from memory when recalling information. Moreover, later research revealed that perceptually disparate spatial and verbal information can also form a cohesive mental representation. For instance, Brunyé et al. (2008) give participants toy assembly sequences in three different formats picture, text, and multimedia format which includes a spatially drawn model of the toy along with the text. Results show that multimedia representation of the sequences provides learning benefits as a function of their inherent redundancy and their presentations, allowing for an integration of picture and text elements into a cohesive mental model. They conclude that "integrating spatial and verbal information is in and of itself a form of deeper processing."

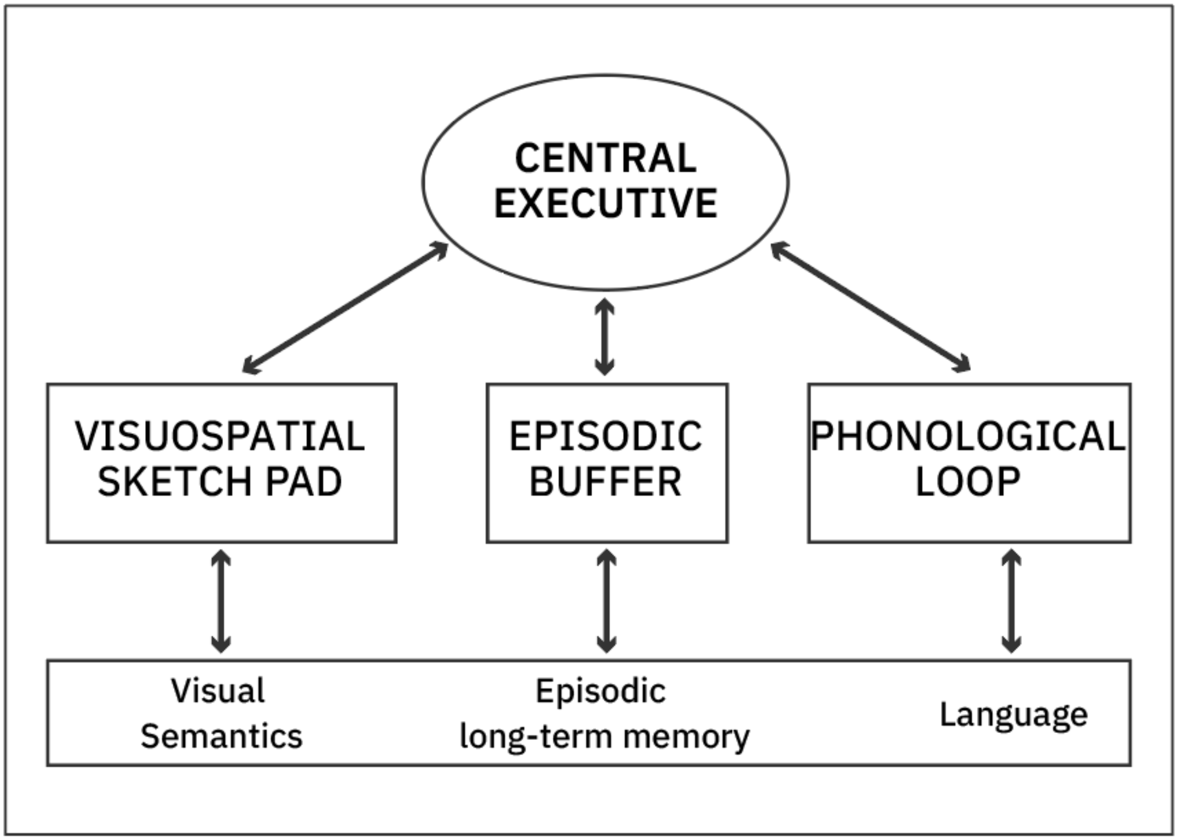
Several theories propose that sentence retention differs from word retention due to the additional support sentences get from long-term representations involved in sentence comprehension (Jefferies et al., 2004). Therefore, accessing relevant and already learned knowledge in the memory dramatically improves reading comprehension and memory of text (Stein et al., 1984).

In addition, syntactic and semantic factors function akin to *chunking* by which the predictability of words in a sentence is increased (Jefferies et al., 2004). Chunking refers to taking individual pieces of information and grouping them into larger units.Bartlett (1995)'s experiments revealed that memory storage for prose is not verbatim but constructive, where the information is more abstractly encoded and is related more to the general knowledge and less to the specific details with the passage of time.

2.3.3 Working Memory

Working memory is composed of four components (a) the central executive, (b) the phonological loop, (c)the episodic buffer, and (d) the visuospatial sketchpad (VSSP), whose coordination is accountable for temporary manipulation and storage of information (Baddeley & Hitch, 1974; Baddeley & Wilson, 2002).

Figure 2.6: Model for Human Working Memory



Note. Baddeley (1998)'s model for human working memory and its components. The bottom rectangle displays areas representing systems that are capable of accumulating long-term knowledge (visual semantics, episodic long-term memory, language); the upper two areas are fluid systems (attention and storage).

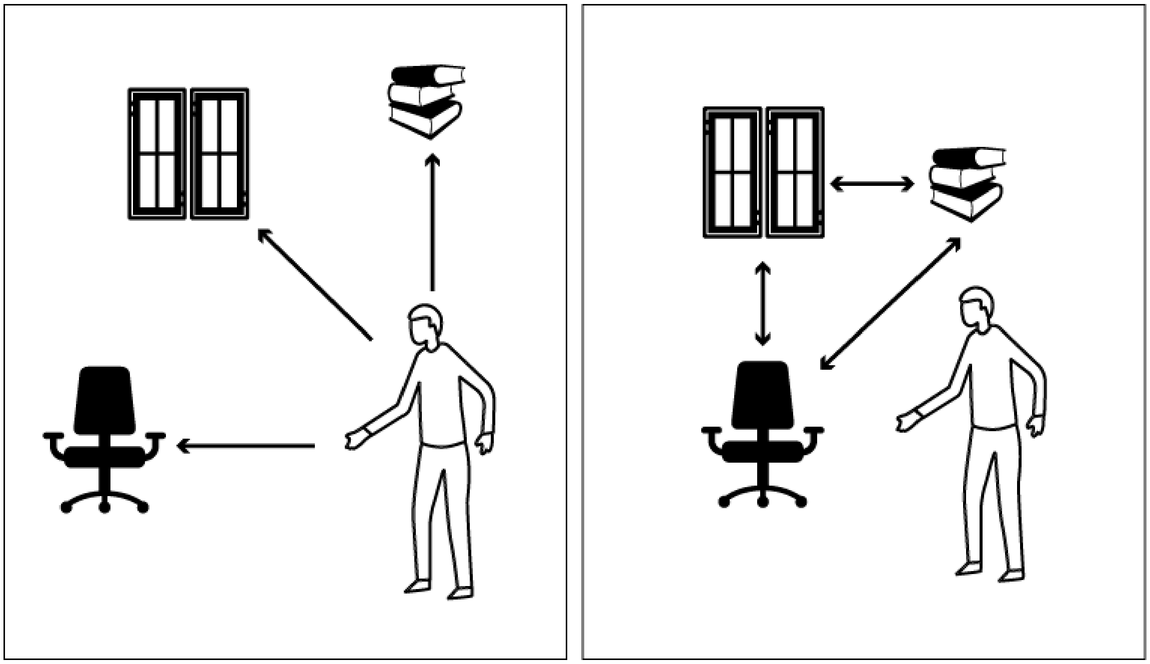
Figure 2.6 depicts the revised working memory model. The central executive comprises two subordinate systems; the phonological loop and the visuospatial sketchpad (Baddeley, 2012). Speech-based information goes to the phonological store and the articulatory control process linked to speech production. It also stores verbal information from the phonological store. The visuospatial sketchpad integrates spatial, visual, and kinesthetic information. It helps track where we are in relation to other objects as we move through our environment (Baddeley, 2012). The episodic buffer is a limited capacity system in which various codes are used to turn information from a range of sources into a single structure or episode (Baddeley & Wilson, 2002; Baddeley, 2003). Finally, VSSP displays and manipulates visual and spatial information in long-term memory (Baddeley, 2012, 2003).

For VSSP, a distinction between spatial and object codes is drawn by (Smith et al., 1995), connecting the two distinct processing paths for encoding where and what information. Only a few studies suggest the involvement of VSSP in text processing, most of which comes from data on individual differences. For example, a higher level of visuospatial memory was found to be correlated with comprehension of spatial text (Gyselinck et al., 2000) and those individuals present a pattern of performance related to imagery used during the text processing (Denis et al., 2002). Moreover, Kruley et al. (1994) revealed that text accompanied by imagery interfered with memorizing the spatial information about some dots on a grid more than text without pictures. They concluded that visual and spatial tasks compete in the comprehension process due to the limited capacity of VSSP.

2.3.4 Spatial Memory

Spatial memory is described as the cognitive processes of perceiving, encoding and retrieving landmarks in the environment for an organism to navigate and survive in the world (Grech et al., 2018). There are two types of reference frames used in spatial cognition, egocentric and allocentric (Ruggiero et al., 2009) (Figure 2.7). The egocentric frame of reference places the organism at the center of the organization of surrounding space (Grech et al., 2018). This is also identified as an orientation-dependent reference frame (Press & Issues, 1998). The allocentric frame of reference, also called orientation-independent (Press & Issues, 1998), specifies spatial information independently of the organism's position (Grech et al., 2018).

Figure 2.7: Egocentric and Allocentric Reference Frames



Note. Spatial reference frames. Egocentric (self-centered) and allocentric (world-centered) encoding of spatial environment.

The main difference between the two frames of reference is the cues that help organisms access the required location. While the egocentric frame of reference is biased by the relation between the organism's location, the allocentric frame of reference depends on the external objects chosen as the anchor points. Hasher & Zacks (1979) suggests that spatial encoding and retrieval are fundamental to the survival of living organisms, and both operate nearly automatically. However, the literature suggests a difference in the level of automation between the two. It appears that the tasks which require allocentric processing are less automatic than the egocentric processing Andrade & Meudell (1993); Köhler et al. (2001). This is due to the attentional demands of allocentric reference, which requires a detachment from egocentric perspectives (Ruggiero et al., 2009).

2.3.5 Spatial Memory in Human-Computer Interaction

In the field of graphic design and HCI, techniques to improve the understanding of abstract information through the use of different frames of reference are specifically measured through search task performances by comparing 2D and 3D displays (Salzman et al., 1999). Search tasks consist of two sub-tasks: target location (time/effort to find the target) and orienting the icons (creating a sense of direction to find the target) (McCormick & Wickens, 1995). Research suggests that the allocentric view can improve the understanding of global information, whereas the egocentric view can help improve the understanding of local information (McCormick & Wickens, 1995; Barfield et al., 1993). For the target location, the salient objects, amongst other elements, will determine how easy it is for a search task to be completed (Anderson & Pearson, 1984). Empirical evidence to support this claim comes from abstract data represented in 3D space such as scientific visualizations (Merwin et al., 1994), studies of altitude and depth (Yeh & Silverstein, 1990), analyzing multi-dimensional data (Liu et al., 2020).

One study which looks at letterforms in relation to frames of reference compares the use of a stationary monitor and hand-based input device in VR, then asked subjects to perform a search task to find certain camouflaged letterforms within a room (Pausch et al., 1997). Time and accuracy results show that VR users were 41% quicker in deciding if the target was not present, and concluded that this was due to having a mental model for the space enhanced by the ego-centric frame of reference.

The type of information presented, the spatial ability of users, the design of the environment, as well as the way users interact with their environment play an important role in determining which frame of reference maximizes search task performance (McCormick & Wickens, 1995). Moreover, sometimes it can prove beneficial to alternate between frames of reference to facilitate understanding of information by the user (Salzman et al., 1999).

Human perceptual systems try to provide a unanimous representation of our environment by integrating inputs from different modalities, which seldom are perceived in isolation (Delogu et al., 2009). For example, several studies show that spatial encoding processes occur simultaneously with the semantic encoding of the text. When asked to identify substantive information about the text and the sentence location in the passage, individuals accurately detected both(Christie & Just, 1976; Giulia Cataldo & Oakhill, 2000). In each of these examples, spatial attributes of the text play a discriminative function and a mnemonic aid for substantive recall. It is also consistent with (Bower, 1970)'s finding of the classical mnemotechnics in which spatial recall is defined as an effective retrieval system.

2.4 Expository Text Structures

Textual information comes in a variety of forms. However, the two broadest genres are called narrative and expository texts. Expository texts, which tend to have a high density of information and cognitively demanding concepts (Fang, 2005, 2006b) differ from narrative texts like stories in how they present and organize content. This differentiation led many researchers to believe that narrative and expository texts differ in their potential for information retention.

Science books are examples of expository texts whose primary intent is to inform and communicate explicit ideas about a specific topic (Graesser et al., 2011). They often introduce a general theme and then elaborate on it in length, resembling a pyramid approach from general to specific elaborations (Goodman, 1976). Many college-level textbooks include prose text to explain general concepts and use diagrammatic representations to depict subsequent concepts visually (Wise et al., 1995). The distinction between prose and diagrams is often drawn from the fact that reading prose is sequential, whereas diagrams are ordered by their location on a plane (Larkin & Simon, 1987). Therefore, the representation of the two supports different operators in recognizing patterns, making inferences, and controlling strategies (Larkin & Simon, 1987; Jian & Wu, 2015). Consequently, they are treated differently in their processing.

Chandler & Sweller (1991) in *Cognitive Load Theory* suggests that separating text and diagrams when referring to mutual information is ineffective because mentally integrating the information before comprehension can increase cognitive load, reducing retention. Chandler & Sweller (1991) looks at conventional and integrated instructions on the internal wiring of electrical light circuits. Based on the results, they hypothesized that in areas where the mental integration of text and diagrams was unnecessary, integrated instructions were no longer adequate (Chandler & Sweller, 1991). Therefore, subjects dedicated attention and mental sources to meaningful information.

2.4.1 Comparison of Diagrams and Prose

A diagram refers to the schematic illustrations in which the relationships between the textual elements are expressed graphically (Scott & Power, 2002). In 2000, The Report of the National Reading Panel published a scientific review of research on comprehension (Panel et al., 2000). The body of literature they surveyed starts from 1970 and groups the studies based on the instruction types they used. Amongst 16 categories of instructions (e.g., teacher preparation, listening actively, story structure, etc.), the use of graphic and semantic organizers was 1 of 7 categories that proved to be useful in improving recall and comprehension of the text for Social Studies and Science content areas (Panel et al., 2000).

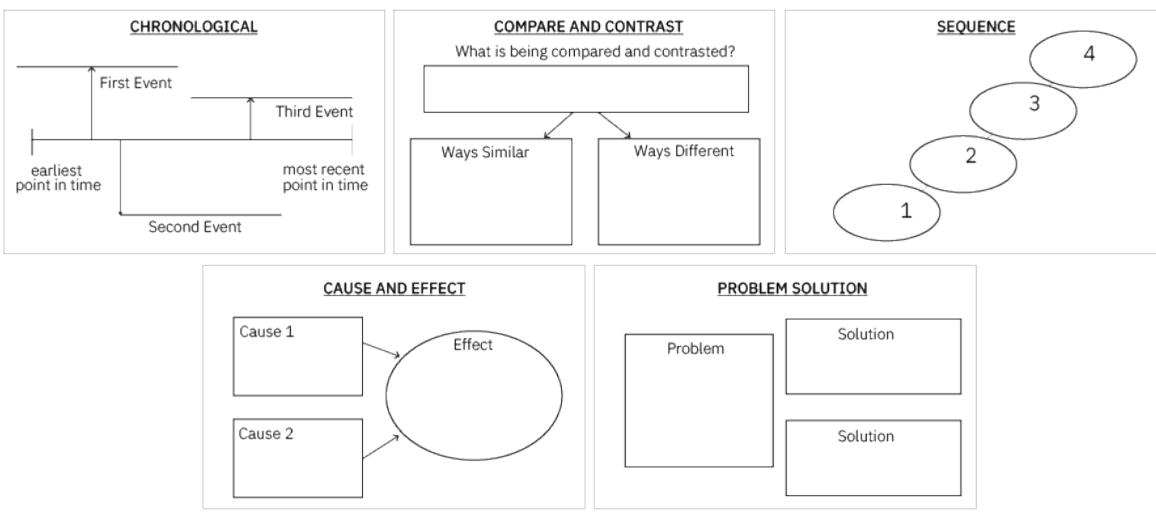
Those studies found that groups instructed on using graphic organizers comprehended and recalled the text better than groups without the instruction (Berkowitz, 1986; Baumann, 1984; Alvermann & Boothby, 1983). Four kinds of primary steps are taken in the creation of diagrams: (1) deconstructing the bits of information from the text, (2) labeling these chunks based on criteria presented by the text, (3) connecting these labeled pieces together by relevance and finally (4) creating an organized visual space which supports the understanding of text (Horn, 1998).

Many graphic design and psychology experts agree that how graphic devices work has yet to be discovered (Guri-Rozenblit, 1989; Waller, 1987). In creating these diagrams, intuition and only a small portion of isolated rules are applied (Alesandrini, 1984; Nunberg, 1990; Levie & Lentz, 1982) and most often, text and diagrams are seen as opposites (Waller, 1982).

Hegarty et al. (1991) defines three broad categories of diagrams: *iconic*, *charts and graphs*, and *schematic*. Iconic diagrams are usually concrete and refer to the spatial relations within referent objects (e.g., a heart diagram showing each individual part). Charts and graphs depict a set of quantitative facts or records. Finally, schematic diagrams illustrate abstract concepts and rely on their components and organization. In other words, the entities are not physical representations of components but have a meaning by convention. An example is electrical conductance and its symbols which were developed in the early 1900s. These symbols were organized to show functional electrical connectivity features instead of physical features (Hegarty et al., 1991). Venn diagrams, flow charts, and linguistic tree diagrams are examples of schematic diagrams (Hegarty et al., 1991). Chandler & Sweller (1991) use iconic diagrams for all six experiments they run. Bits of information that are not understandable unless integrated into a pictorial drawing or illustration are common in science education. However, some concepts in science can still be intelligible by the logical relationship between the text labels without referencing a concrete diagram Hegarty et al. (1991).

In a diagrammatic representation, information is organized by its location (oftentimes present in the same field of view), which dictates specific inferences (Larkin & Simon, 1987). Hegarty et al. (1991) suggests that the relations signaled in schematic diagrams depend on the convention of interpretation. For example, students have to learn how to interpret the layout of a diagram to understand the relationships conveyed through spatial metaphors. Figure 2.8 shows a set of diagrams for different kinds of text structures depicting the organizational patterns found in prose text.

Figure 2.8: Text Structures



*Note.*Text Structures Review Sheet. Students fill out appropriate graphic organizers based on text structures.

Sentential representations (or prose) differ from diagrammatic representations in their sequential nature, corresponding to the simple formal language (Larkin & Simon, 1987). Diagrammatic representations maintain information about the spatial and structural connections between textual components, while sentential representation does not. Therefore, they use different information-processing systems (Larkin & Simon, 1987), and what influences their effectiveness differs from prose. Literature suggests that representational context (Mayer, 2004), a spatial grouping of information (Larkin & Simon, 1987), the degree of correspondence between the text and the accompanying diagram (Mayer, 2004), prior knowledge, spatial ability, and user characteristics (Mayer & Gallini, 1990) all play a role in a diagram's efficacy. When in harmony with text, students' skills, and background, diagrams increase deep comprehension of the represented information (Butcher, 2006).

2.5 Overview of Typography in virtual reality Environments

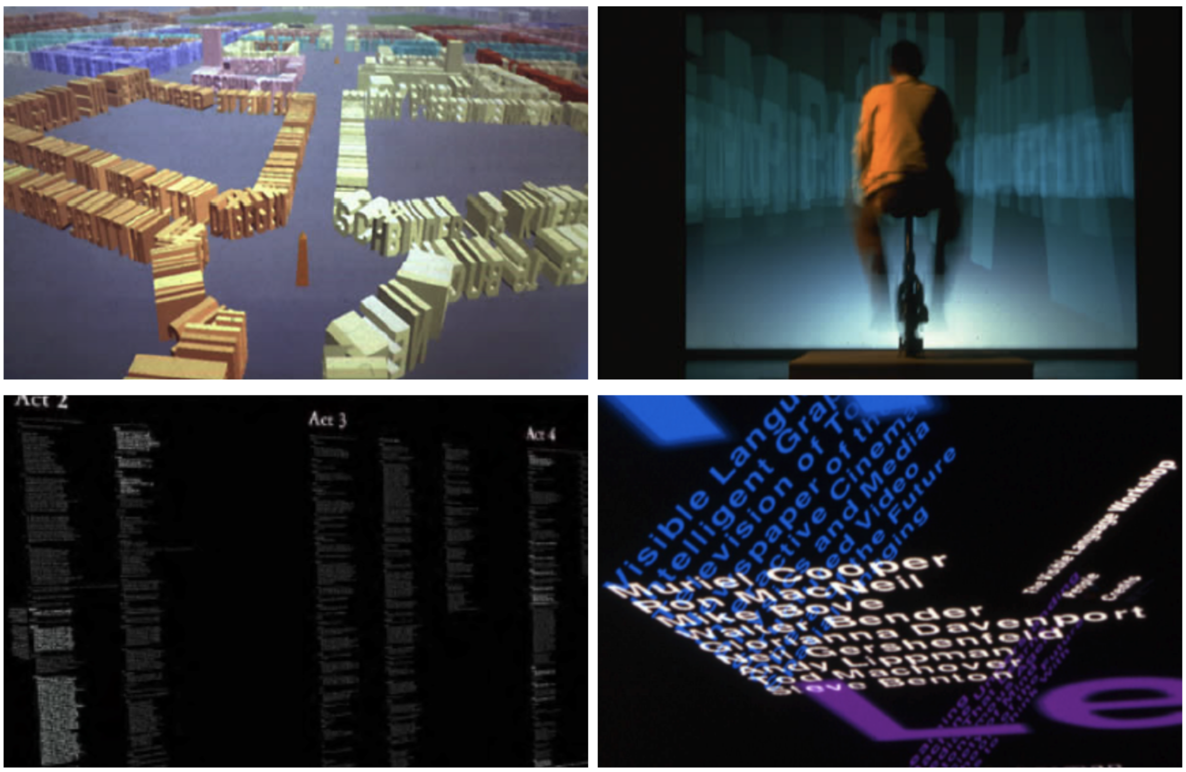
Typography development has historically focused on printed media (Small, 1999; Ishizaki & Mitchell, 1996). When this fixed presentation medium was gradually changed to electronic media, the potential of text was changed as well (Hillner, 2009).

An attempt to systematically discover immersive typography at MIT's Visible Language Workshop (VLW) between 1985 and 1994 was initiated by Muriel Cooper. VLW was mostly concerned with how text should be treated in 3D spaces (Cooper, 1989). Designers in the VLW tried to uncover the properties of typography when it moves in relation to readers' movement and raised issues such as legibility, orientation, and distortion of type due to changing viewpoints. (Allport et al., 1995; Wong, 1996; Small et al., 1994; Ishizaki & Mitchell, 1996). Figure 2.9 displays snapshots from some of the projects made by the members of the VLW group.

- **Legible City:** The legible city is an installation created by Jeffrey Shaw in 1989. In the installation, visitors ride a stationary bicycle by looking at a screen that takes them through a city consisting of 3D letters. As the visitor chooses the path for her/his journey, letters create words and sentences. The layout corresponds to several existing city plans and replaces their architecture with letterforms. Traveling through the cities is a reading process, but the meaning is randomly built by the user's choice.

- **Virtual Shakespeare:** David Small uses the plays of Shakespeare to create a meaningful information space. He uses different typographic treatments for certain sections, such as color, different typefaces, sizes, and various orientations in space. Users can see all the information simultaneously and modify their distance from the text. All textual elements are flat, but they flow in the space.

Figure 2.9: Work from Visible Language Workshop



*Note.*Top two images are from Legible City by Jefferey Shaw (1989), the bottom left Virtual Shakespeare project by David Small (1999) shows the entire content of A Midsummer Night's Dream by Shakespeare, and the Information Landscapes project by VLW (1994).

- **Information Landscapes:** Muriel Cooper and her students presented the Information Landscapes project at the TED5 conference in California in 1994. The project was a continuous flight through a changing 3D space. It explored the use of dynamic typography and 3D graphics to display hierarchical, financial, and geographic data. The Information Landscapes aimed to apply two-dimensional typographic techniques to design 3D information graphics (Small et al., 1994).

Jaron Lanier coined the term virtual reality in 1986 (Lanier, 1992). Researchers at the VLW neither used virtual reality nor adopted head-mouthed displays even though the first fully developed head-mouthed display *Virtuality* had been invented by 1990 (Heim, 1995). Consequently, the projects produced at the VLW neglect the affordances of VR, even though they investigate 3D space and its issues.

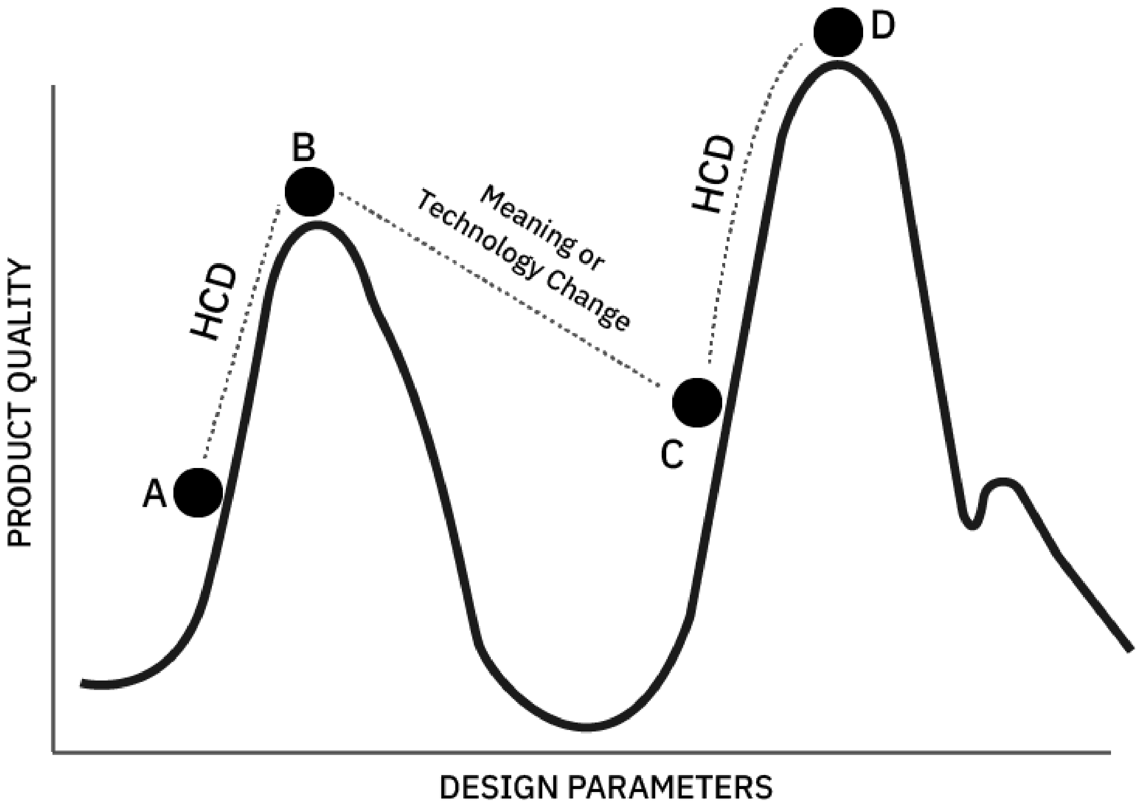
2.5.1 Design Research and virtual reality

Since it was first invented in Mesopotamia, the text has been a man-made, abstract system of symbols (Lévy, 1998). One of the main forms in which we consume text is the book, whose paradigm has been challenged with the arrival of screens (Small et al., 1994). Technological advances such as e-readers have used some of the advantages of computer screens to customize and create interactive textual environments. However, despite these technological advances, typography is still read as an image from a flat surface, and dimensional treatments for text can be considered radical given the current ways we consume it. Especially for emerging technologies like VR, in which the design space is largely uncharted, the imposition of old desktop paradigms for text representations on users precludes the possibility of discovering new and more effective user interfaces (UIs) which have not been previously imagined or considered.

The Hill-Climbing Paradigm by Norman & Verganti (2014) provides a framework for the growth of product quality with respect to various design parameters. In Figure 2.10, point A represents a product that undergoes incremental changes by using human-centered design (HCD) and design research (DR), which eventually arrives at point B, where the maximum quality of the design space is realized. The next bigger hill represents a higher potential for this product, which requires a more radical innovation caused by a meaning or technology shift. When this shift happens, the product arrives at point C which is inferior to point B. This product is then expected to undergo HCD and DR processes to reach its maximum potential point D.

Norman fails in his search to find a product that refutes this paradigm. In other words, radical technological innovations often ignore users' or even society's needs, and therefore they are investigated without prior design research or formal analysis of needs.

Figure 2.10: The Hill Climbing Paradigm



*Note.*The Hill-Climbing Paradigm applied to incremental and radical innovation (Norman & Verganti, 2014). A specific product might begin at " A. " The product experiences a succession of incremental changes through Human-Centered Design and Design Research, eventually bringing it to its highest quality of the design space, point "B." Radical innovation is necessary to go to a different hill with a higher potential, and this happens either through technological advancement or a shift in meaning, leading to point "C" on a bigger hill.

For many years, VR has been seen as a radical technology. Historically, we have not been closer to realizing its vision. According to the McKinsey survey (Aiello et al. 2022), the average US customer expects to spend four hours in VR in the next five years. This meaning shift results from advancement in hardware and software and deliberate attention paid by big companies like Meta.

However, VR is in the process of arriving at point C, where it needs to go through a series of incremental innovations through HCD and DR to maximize its potential. Typography in VR is neither affected by the *radical innovations* in technology nor used HCD to deeply understand what it can or can not do in VR space. Design space is largely uncharted and is still waiting to realize or improve ideas that have not been previously imagined or considered. This holds true, especially for information-rich environments where virtual objects represent abstract concepts rather than real-world objects humans have evolved to manipulate in the physical environment.

Norman & Verganti (2014) adds to Hill-Climbing Paradigm by saying, "...to make matters more complex, when the product is at point “C,” there is no way of knowing if indeed there is a superior level (“D”) or if this is an inferior spot in the design space." To understand if point D is possible for VR text, we must ask ourselves what is gained by putting large bodies of text in a 3D virtual space.

2.6 Affordances of virtual reality

2.6.1 Immersion and Presence

The primary characteristics distinguishing VR from other means of displaying information are *immersion* and *presence*. Immersion refers to the degree to which the VR system's sensory modalities match that of the real world (Slater, 1999). The more the VR system provides fidelity similar to its real-world counterparts, the more immersive it is. Immersion is the objective measure of the virtual environment, and it depends on the display technologies and rendering software with all of its sensory modalities (Bowman & McMahan, 2007). Presence refers to the subjective and psychological reaction to the VR system (Slater, 1999). Presence is context-dependent and relates to the subjective feeling of "being there" (Bowman & McMahan, 2007). While a body of literature looks at the psychological response to a VR system (i.e., presence), immersion allows for a practical approach to measure and control its levels (Bowman & McMahan, 2007). Components of immersion can vary from auditory to haptic displays (Bowman & McMahan, 2007; Bowman et al., 2003). Bowman & McMahan (2007) list the components of visual immersion as the following (included but not limited to):

- Field of view (FOV)—the size of the visual field (in degrees of visual angle) that can be viewed instantaneously
- Field of regard (FOR)—the total size of the visual field (in degrees of visual angle) surrounding the user
- Display size
- Display resolution
- Stereoscopy—the display of different images to each eye to provide an additional depth cue
- Head-based rendering—the display of images based on the physical position and orientation of the user's head (produced by head tracking)
- Realism of lighting
- Frame rate
- Refresh rate

Slater & Wilbur (1997) suggest a theoretical framework for measuring the level of immersion, including all other dimensions. According to the model, there are five aspects of VR; inclusiveness, extensiveness, surrounding, vividness, and matching that influence but do not solely determine the user's perceptual experience. Figure 2.11 presents each aspect that determines the level of immersion. *Inclusiveness* is the level of exclusion from physical reality. *Extensiveness* denotes the variety of sensory modalities supported. *Surrounding* describes how much of the VR environment is panoramic as opposed to being restricted to a small area. *Vividness* represents the fidelity, resolution, and diversity within a given modality (for instance, the visual and color resolution). Richness, information content, resolution, and display quality are all factors in vividness. Finally, *Matching* describes the match between the information generated on the displays and the proprioceptive feedback about the participant's body motions.

Figure 2.11: Framework for Immersive Virtual Environments

	Aspect of immersion				
Level of immersion	*Inclusive*	*Extensive*	*Surrounding*	*Vivid*	*Matching*
Low	Numerous signals indicating the presence of device(s) in the physical world (e.g., use of a joystick or mouse to control the VE, direct instruction from an experimenter during the task)	Only accommodates 1 sensory modality (e.g., auditory, visual, motor/proprioceptive); stimuli are not spatially oriented	Computer monitor presentation with limited field of view	Low fidelity and visual/color resolution; display may replicate features of the simulated environment, but not in a detailed or specific manner	No motion capture; visual experience does not match proprioceptive feedback
Moderate	Some signals indicating the presence of device(s) in the physical world (e.g., noise from a computer fan, weight and movement restriction from wearing a safety harness)	Accommodates 1–2 sensory modalities (e.g., auditory, visual, motor/proprioceptive); stimuli may or may not be spatially oriented	Large-screen projection with extended field of view	Moderate fidelity and visual/color resolution; display replicates some features of the simulated environment, but some detail may be missing	Body segment motion capture (e.g., head, hand); visual experience somewhat altered to match proprioceptive feedback based on head or body segment movement
High	Limited signals indicating the presence of device(s) in the physical world (e.g., the weight of an HMD or an eye-tracking device)	Accommodates >2 sensory modalities (e.g., auditory, visual, motor/proprioceptive); stimuli are spatially oriented	Head-mounted device or surround projection	High fidelity and visual/color resolution; display closely replicates multiple features of the simulated environment in great detail (e.g., correctly placed, dynamic shadows)	Full-body motion capture; visual experience altered to closely match proprioceptive feedback based on whole body movement

Note. Slater & Wilbur (1997)'s Immersion Framework which suggests that the degree of immersion may be objectively measured by the qualities of the technology. It has dimensions including how well a display system can give a participant the impression of inclusiveness, extensiveness, surroundings, vividness, and the illusion of a match between the real world.

Especially for cognitively demanding tasks, high immersion in VR does not always work as expected. Research shows a prominent disadvantage of VR is its *overwhelming* nature, which leads to cognitive sources being divided between the learning material and extraneous features of the environment (Baceviciute et al., 2020; Webster, 2016; Mikropoulos & Natsis, 2011). Even when VR spaces are moderately immersive, different sensory modalities, like auditory input, reduce knowledge retention, yet they increase immersion (Baceviciute et al., 2020). Similarly, the additional task of remembering spatial cues hinders cognitive tasks like recalling for those unfamiliar with the VR environment presented to them (Yang et al., 2020). Moreover, limited bodily movement and lack of self-paced learning were found to be disadvantageous (Parong & Mayer, 2018; Baceviciute et al., 2020) despite the fact they would decrease the immersion. Participants still report increased bodily movement as an expectation for the VR environment (Baceviciute et al., 2020; Parong & Mayer, 2018). Lack of self-paced learning is an important effect related to the active/passive viewing of the content, which affects memory and favors the active viewing (Hine & Tasaki, 2019). Active viewing is described as the change in the VR environment based on the participant's input (Hine & Tasaki, 2019). In Webster (2014)'s study, PowerPoint reduces the processing load of the students due to high and easy control of the material, which stems from its familiarity.

Furthermore, Webster (2014)'s study where a *Corrosion Prevention and Control* (CPC) training is implemented using VR shows a discrepancy between the usability study and pre-and post-test results. 88% of the soldiers believed VR training to be effective, despite 15% of the sample for the VR group showing a decrease in the post-test. This is explained as a result of the *wow factor* becoming a disadvantage which sometimes overshadows the true purpose of the environment (Webster, 2014). In addition, uncomfortable equipment and hardware issues (jittery scenes or low-quality rendering, etc.) were recurring points made by the participants in the study. Even though this can affect learning, arguably, researchers are trying to design for a near future where these concerns will be mitigated by the advancements in technology.

Learning in VR is not without its advantages, however. It is a potent tool for creating an environment representing a phenomenon that can not be seen with the human eye (Parong & Mayer, 2018) or helping users build environments quickly and cost-effectively while still providing engagement (Yang et al., 2020; Webster, 2014). From this perspective, science as a domain of interest benefits from VR because many scientific concepts are impossible to see with the human eye, and they carry spatially relevant information (Parong & Mayer, 2018). Moreover, VR environments allow for multiple representations of the same concept (Yang et al., 2020) and can give immediate feedback to increase situational interest and motivation (Parong & Mayer, 2018), all of which can lead to the creation of an external space where the internal processes become evident or concrete and in turn, they can reduce the mental demand in manipulating the concepts.

Regarding the design of such environments, it is important to know what is gained cognitively from having more realistic and interactive representations of information before marking VR as superior to screen (Scaife & Rogers, 1996). It is a complex task for systems designers. It requires reconciling cognitive theories, instructional design, and available technologies. Also, given the relatively short history of research which looks at the design parameters in a VR environment, it is hard to come up with recommendations that will fit it all or report a distinctive superiority of one media for VR content. Such prescriptions can not work across contexts and user groups since they vary immensely. Instead, adaptable and customized VR systems can prove more useful in reducing the cognitive demands required to perform a task as an alternative to more traditional ways of displaying content. For example, reading scientific text from VR can be a supplemental tool that matches the knowledge in the head (from the traditional reading) to the knowledge in the three-dimensional (3D) space. This environment can be highly diverse in size, form, and the type of relationships established, along with multiple ways of interaction. When such internal representations are solidified using space and different senses, they are more likely to simulate a possible future due to the power of sensory information provided by it. Thus, the way in which the virtual environment is designed can aid users in understanding the critical features of the applications and performing a task effectively by being purposefully abstract (Bowman & McMahan, 2007; Bowman et al., 2003). Scientific visualization (Dasilva et al., 2001), design prototyping (Tutt & Harty, 2013), business intelligence (Simpson et al., 2016), and metaphor visualizations (Knight & Munro, 1999) are just a few application areas that use the immersion Vr provides.

2.6.2 Embodiment and Ego-Centric Reference Frame

The embodiment has frequently been associated with a sense of self-location and body ownership (Newport et al., 2010; Lopez et al., 2008). However, it also has multiple definitions based on various contexts resulting from various applications and multidisciplinary use (Kilteni et al., 2012). Therefore, its conceptualization depends on the defined term's perspective. For example, from a psychological point of view, it can be described as mechanisms of the brain that underlie the representation of one's body (Berlucchi & Aglioti, 1997), or in robotics, it is seen as the degree of virtual or physical representation of the body in relation to virtual agents with all its capacity to sense and act in virtual worlds.

In the literature, the embodiment is associated with the term *sense of location*. Sense of location relates to the experience of being inside a body and not specifically being inside a world, and the latter is related to presence. Kilteni et al. (2012) draws a distinction between self-body and self-environment and explains, "an example of self-location could be the feeling that one's self is located inside the biological body or an avatar's body; whereas the analogous feeling for presence would be the feeling of one's self-being located in a physical or virtual room, even if this does not require a body representation in the form of an avatar."

Despite the fact that self-location and presence address different spatial issues, they can be seen as complimentary ideas that together make up a person's spatial representation (Kilteni et al., 2012). As described before, human navigation is inherently egocentric, and self-location is largely determined by that default visuospatial perspective.

In two-dimensional environments, the relationships between typographic elements dictate navigational and structural understanding of the text, independent of the reader's relative position, which is linear and external. Unlike screen reading, we process VR text architecturally, creating a unique self-directed relationship with the text. When we wander through prose, we intuitively understand the layout by drawing a direct analogy from real-world systems.

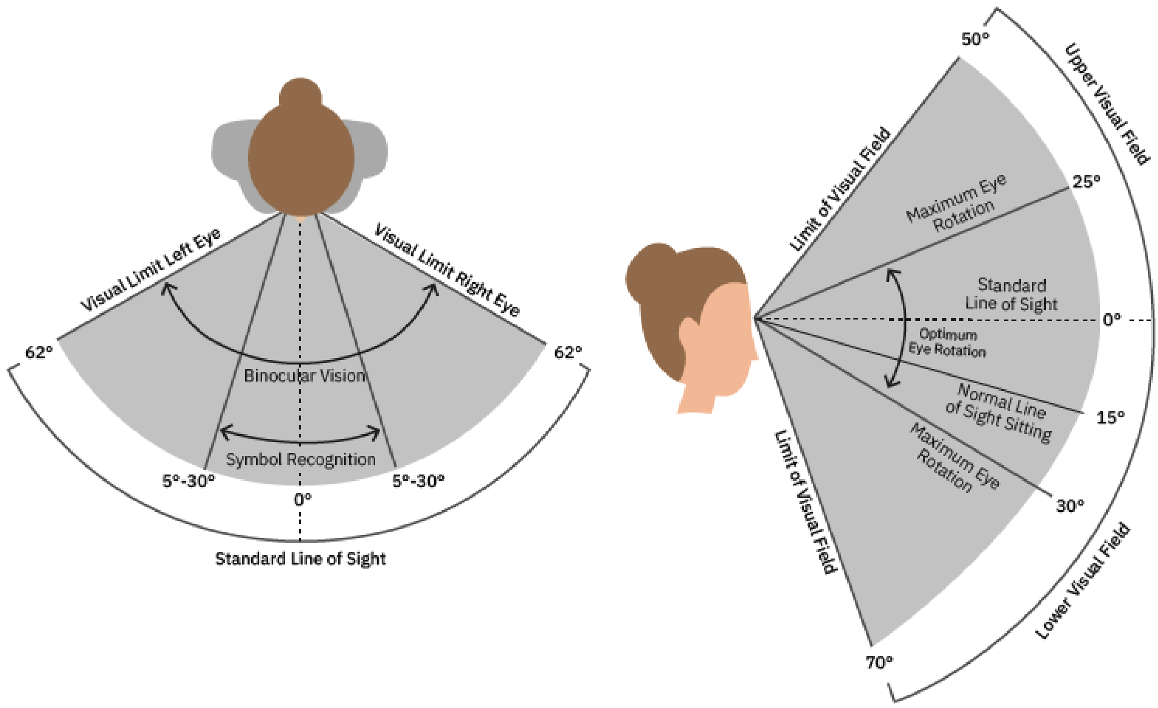
> "You are, in a sense, in an architectural construct, but you don't have the constraints to believe a physical building. So, you can use the abstract conceptual issues and the physical cues people are accustomed to." (Cooper, 1989)

Embodiment and ego-centric frame of reference can particularly be beneficial if the reader is positioned as an active viewer in relation to the elements of the text. The most accurate way of testing for egocentric strategies is to perform a test in a highly controlled environment, which ensures the removal of visual cues that could be used for allocentric strategies (Grech et al., 2018). However, empirical measurement of such effect exceeds the scope of this proposal.

2.6.3 Field of View

Field of view (FOV) is the angular measure of what can be seen at a particular time (Jerald, 2015). FOV plays a critical role in the presence, and as the FOV increases, the feeling of presence does too (Abrash, 2021).

Figure 2.12: Vertical and Horizontal Field of Views in Human Visual Systems



*Note.*The size of the visible environment at any particular time is known as the field of vision (FOV). (Ball et al., 1988)

As can be seen in Figure 2.12, the human eye has approximately 120° by 120° horizontal and vertical FOV. With the rotation of the eye to either side, an additional 50° is gained on each side. When it comes to the vertical direction, the human eye can not see nearly as much as in the horizontal direction because the eyes are not vertically additive, and our torso and forehead get in the way (Ball et al., 1988). Therefore we only see 50° above and 70° below with a total of 120° (Ball et al., 1988).

Although we use central vision to see details, peripheral vision is also important to function in the real world (Jerald, 2015). This area in total is also described as *useful field of view* (UFOV), where the information can be captured without eye or head movements (within one eye fixation) (Ball et al., 1988). UFOV measure varies across situations and individuals (Ball et al., 1988). Reading primarily occurs in that area using *Foveal vision* where 100% acuity is possible. Fovea extends 2° across the fixation point. Beyond this area, *parafoveal vision* extends 10°, and the rest is the peripheral region (Rayner & Bertera, 1979). Although we can not attend the parafoveal and peripheral regions with the same acuity, we use them to obtain information about word length and shapes and the beginning and end of words to construct coherent sentences with this available information. Therefore, the way we read and look at images is different.

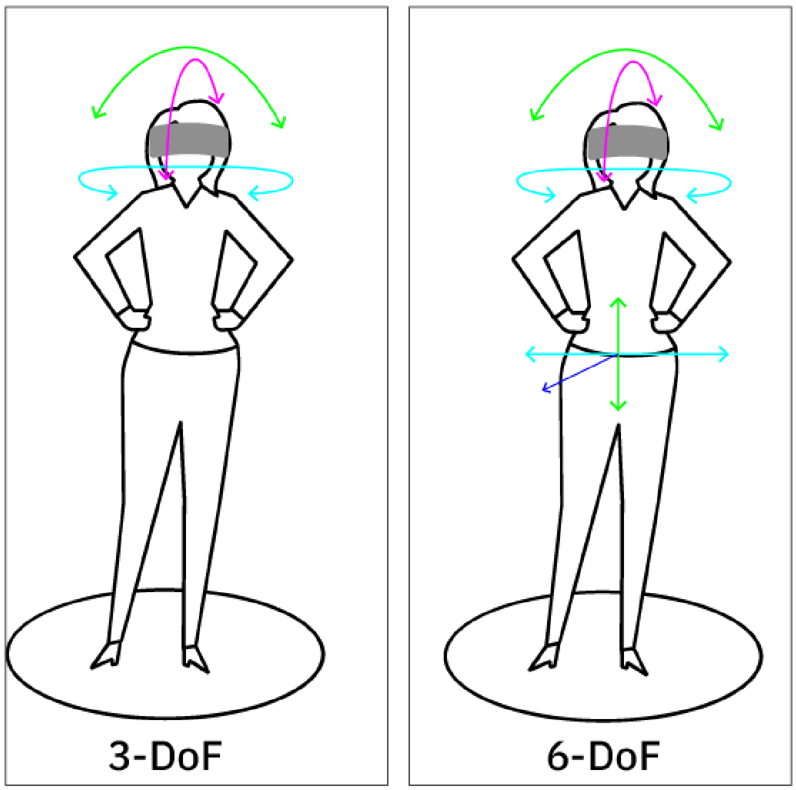
Figure 2.13: Field of View Measures for Different virtual reality Headsets

Headset Name	Horizontal FOV	Vertical FOV
HTC Vive Cosmos / Elite	84 (72 - 95)	87 (83 - 88)
HTC Vive Pro	82 (74 - 90)	91 (87 - 95)
Valve Index	108 (107 - 110)	105 (95 - 108) *
Oculus Quest	95 (88 - 103)	92 (86 - 100)
Oculus Rift S	86 (77 - 89)	85 (67 - 98) **
Oculus Rift CV1	87 (83 - 90)	84 (83 - 93)
Pimax Artisan	128 (125 - 130)	95 (94 - 95)
Pimax 8K Series ***	158 (155 - 160) ****	105 (102 - 109)
Samsung Oddysey+	103 (103 - 104)	107 (106 - 108)
Star VR One	174 (171 - 177)	114 (106 - 122)

Note. The world's first crowdsourced database of VR headsets is created by *iNFINITE Production*. There are 127 valid measurements of 24 different headsets in the database. From iNFINITE Production, 2020, (`https://www.infinite.cz/blog/VR-Field-of-View-measured-explained`).

iNFINITE is a production and development studio focusing on immersive technologies. They released a crowd-sourced database using a tool that measures FOVs for different headsets (Figure 2.13). Their current measure for horizontal and vertical FOV includes different ranges. The variance in reported degrees of FOV is explained by several factors. Reaching the maximum degrees of FOV depends on the face gasket design and head-strap fit, affecting the optimal distance eyes need from the lenses, which changes per individual. However, for VR applications, 80 degrees of FOV is considered enough to achieve *a state of presence* (Slater, 2003). For a desktop computer with a 1280x1024 resolution, horizontal FOV is 40° whereas vertical FOV is 30° (Pausch et al., 1997). But these values can be set at different degrees across media; however, the information display changes.

Figure 2.14: Degrees of Freedom in VR



Note. The difference between three degrees of freedom (rotational movement) and six degrees of freedom (rotational and translational movement) with a VR headset.

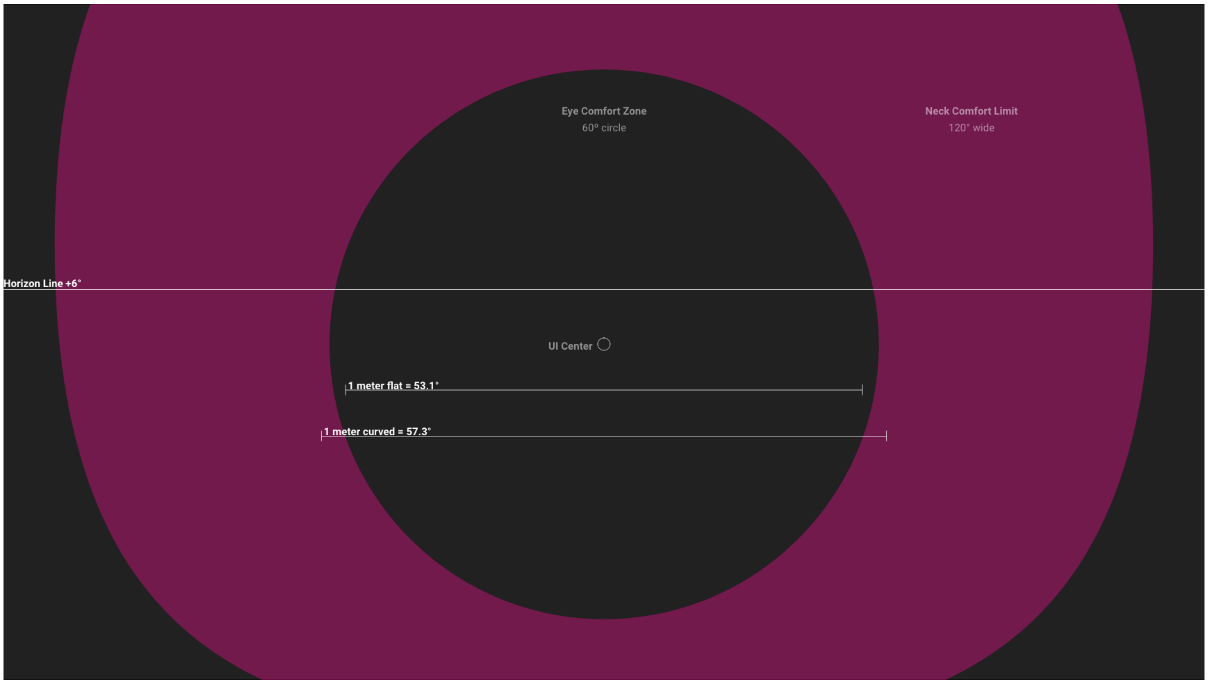
In 2017 Google UX designer Chris McKenzie reported a set of guidelines for text readability in VR[1]. They found that 1.3° were comfortable reading with 3 degrees of freedom

[1] `https://www.youtube.com/watch?v=ES9jArHRFHQ`

(DoF). DoF refers to the extent to which one can move in 3D space. There are six possible DoFs in VR that describe the possible movement ranges of an object (Rivas Méndez et al., 2018) (Figure 2.14). 3-DoF headsets allow for looking to the left or right, moving the head up and down, and turning to the left or right (e.i., sitting or stationary position). 3-DOF headsets do not determine transitional movement, which defines the user's movement (Rivas Méndez et al., 2018). Many early headsets like Oculus Go and Google Cardboard have 3-DoF. 6-DoF headsets can provide translational movement that captures whether the user moved their head or the body forward, backward, horizontally, vertically, or up and down (Rivas Méndez et al., 2018). This allows users to explore locations and a more immersive experience (Rivas Méndez et al., 2018). Current headsets such as Oculus Quest and HTC Vive support translational movement detection.

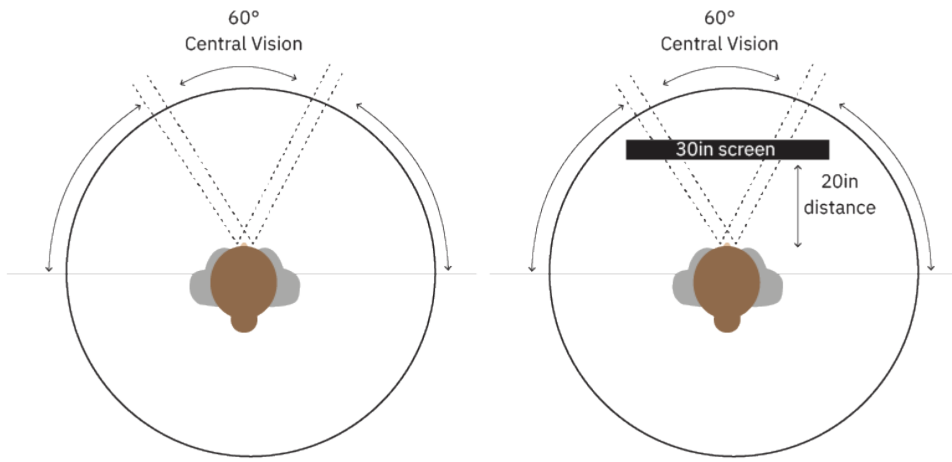
McKenzie also mentioned some ergonomics of VR reading; (1) assuming 3 DoF eyes can move 30-35° in four major directions, (2) with neck movement overall rotation increase 120°, and lastly (3) visual area where people can read comfortably is at about 6° below horizon line (Figure 2.15). For a 30-inch computer screen with an average distance of 1.6 feet between the user, the focal point will narrow, but FOV will still be kept at 60° (Figure 2.16).

Figure 2.15: Google Daydream's Comfort Zones for Designing Screen-based UIs



Note. Comfort zones based on Google Daydream. The guide suggests these degrees be somewhat applicable to all head-mounted displays. The focal point will get smaller for a 30-inch computer screen with an average user distance of 1.6 feet, but the field of view will remain at 60 degrees. From *Google VR References*, n.d.(`https://developers.google.com/vr/design/sticker-sheet`).

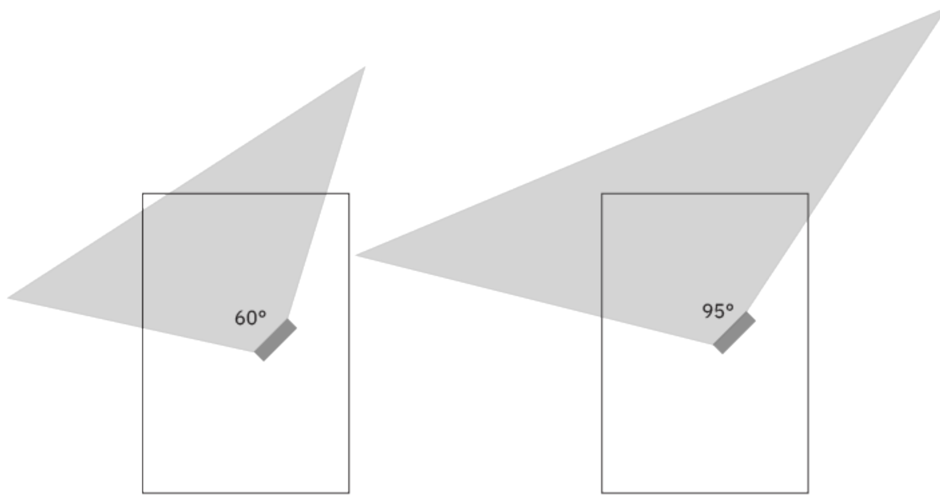
Figure 2.16: The Change in Horizontal Field of View Based on the Distance of Screen



Note. Horizontal field of view when a 30-inch screen is placed 20 inches distance within the same field.

A small field of view leads to a small peripheral vision which affects getting strong motion cues and delivering information about direction, velocity, and orientation during movement (LaViola Jr et al., 2017). Although not the only factor, a low FOV can also cause motion sickness (Jerald, 2015).

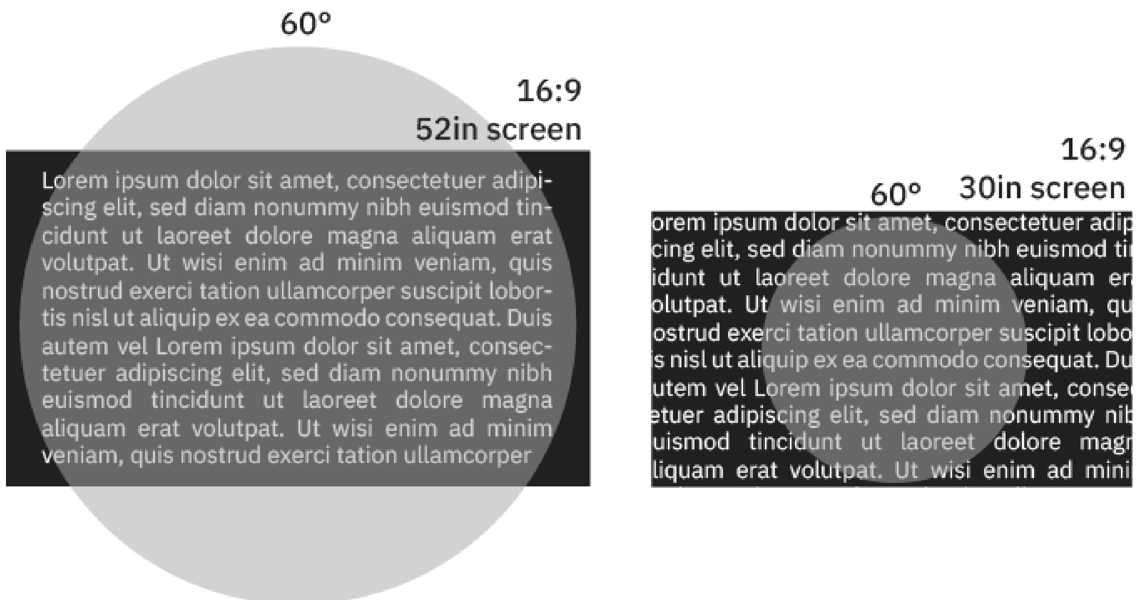
Figure 2.17: Effect of Increased Field with a Fixed Camera



*Note.*Increased angle seeing the scene. More information fits on the same screen size; hence scene gets smaller.

Computers use 90-95° FOV, especially for games. The designer or developer can change FOV through a lens change, allowing for a bigger frame within the same scene without changing the camera position (Figure 2.17).

Figure 2.18: Difference in Field of Views for 52 vs. 30-inch Screens



*Note.*To fit the same amount of information in this scenario, the scene must get smaller.

In this case, the scene will get smaller to fit the same amount of information. When it comes to reading, even though the FOV is a variable that can be set at different angles, it gets harder to recognize the letters in the UFOV due to the shrinkage of the cone of vision. An illustration of this change is shown in Figure 2.18. Therefore, 2D screens can have the same FOV as the VR headsets but the information displayed will be illegible; due to shrinkage in the content, and this value varies.

CHAPTER

3

METHODS

This chapter gives an overview of the experiment and design justification. It follows with a detailed description of the experimental procedure and materials.

3.1 Research Questions and Objectives

This research aims to find the differences in memory performance based on two layout conditions; flat and spatial in VR. Memory tasks include lexical sentence recognition and cued-recall of sentence location. Lexical recognition refers to the ability to identify a familiar stimulus that has been encountered previously amongst distractors (Donaldson, 1992). Cued-recall refers to the retrieval of stimulus information with the help of cues (Muir & Hawes, 2013). In both cases, stimuli were sentences extracted from the text participants read. The stimulus material used in the experiment was two explanatory, college-level scientific passages extracted from the "Anatomy and Physiology; An Integrative Approach" textbook by McKinley et al. (2019) about "Homeostasis" and "Labor in Humans."

The two research questions were:

RQ1. How do the spatially distributed layouts in VR impact lexical recognition and cued-recall of sentence location performances of prose scientific text?

The passages included diagrams that differ from sentential representations in their spatial and structural connections and the content (Larkin & Simon, 1987). They are hypothesized to show a different impact on memory performances due to their inherent spatial arrangements. Therefore the secondary research question was:

RQ2. How does the impact of spatially distributed layouts in VR change the lexical recognition and cued-recall of sentence location performances based on prose vs. diagram structures of the scientific text?

3.2 Design Space and Challenges

A systematic way to approach the design space for text in VR is to examine what constitutes on-screen text. This approach will be limited to scientific prose with textual diagrams, which will be used in this research.

Figure 3.1 organizes the typographic variables used in the book *Homeostasis: Keeping Internal Conditions Stable* (McKinley et al., 2013) which was used to extract the labor and homeostasis passages. The table first lists the typographic elements under three main groups from macro to micro scale; page, paragraph, and glyphs. Typographic elements are listed based on these categories and are nested within them. Functions define how each typographic element in the three main groups function for the organization and the readability of the text. The last column indicates the portion of these aspects that are relevant to the VR space.

Several aspects of type which exist on the 2D plane become obsolete when put in 3D space with constantly changing viewpoints. For example, margins (gutter and page), running footers (page number and chapter name), recto (right page), and vecto (left page) are not defined in space unless it mimics a high-fidelity book. A fundamental question is how to define the design space for type, given the lack of a page. The decisions regarding the size and position of the text area, margin and gutter, and type sizes are all informed by the dimensions of a page. This complex design question needs to be investigated independently with rigorous research and formal analysis of needs.

Figure 3.1: An Analysis of Typographic Elements for "Homeostasis: Keeping Internal Conditions Stable

Macro to Micro Levels

	Typographic Elements	Function	VR Space	
Page	Margins → *Gutter*, *Page*	Defines the text area.	Only space between paragraphs remains.	VR 360 Layout (due to increased Field of View)
	Running Footers → *Page Number*, *Chapter No/Name* Recto and Vecto	Establishes the sequence of reading. Fascilitate search within the document.	Page in VR becomes what reading pattern dictates.	
Paragraph	Image Caption/Figure Numbers	Refer and explain to the diagrams.	Remain the same.	
	Title Subtitle	Help reader percieve overall structure and aids it's recall, search and retrieval.		
	Leading → *12pt*	Maintain space and jump smoothly to the next line without repeating.		
	Space before/after → *10pt*	Seperates paragraphs.		
	Column → *Double*, *Single*	Sperates text content vertically (into one or more pieces).		
	Indentation → *18pt*	Indicates a paragraph break or offsets a selection of text or tabular matter.		
	Tabular Material	Listing items with particular order or not.		
	Font → *Bold*, *Italic*	Create contrast to indicate a change in information category.	Changed to use affordences of space.	
Glyphs	Typeface → *STIX Math Jax Regular*, *Proxima Nova*	Creates unity throughout the document. Can carry different connotations by style.	Size range will remain the same, but numbers will change based on scale of the environment	VR 360 Layout (due to increased Field of View)
	Kerning → *10pt* Tracking → *10pt*	Controls the density of type.		
	Type Size → *10pt, 12pt, 17pt*	Allows for baselines of letters to accurately align across the page.		
	Type Color → *Blue*, *Red*, *Black*, *Grey*	Can be used decoratively or in aiding the reader's ability to follow text.	Remain the same.	
	Type Case → *Upper Case*, *Lower Case*	Differentiate types of information.		
	Special Characters → *Equal Sign*, *Subscripts*, *Degree Sign*	Represents meaning specific to a content area(s).		
	→ *Punctuation* → *Fullstop*, *Comma*, *Colon*, *Semicolon*, *Hyphen*, *Aposthrophe*, *Question Mark*, *Bullet*, *Quotation Mark*, *Paranthesis*	Connotates structural relationships between the elements of text. Carries information that is explicit to written language.		

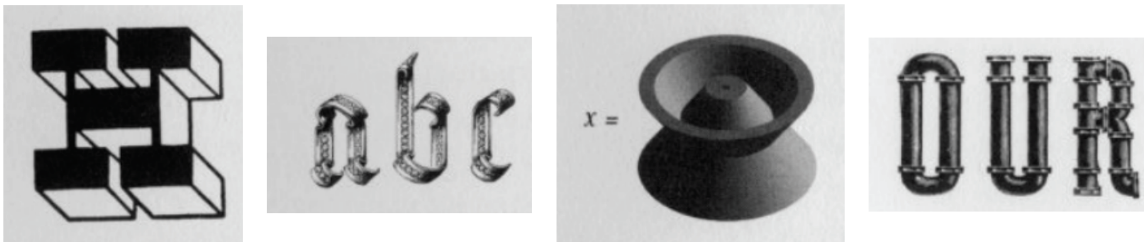
Note. Examination of the typographic variables in the stimulus material. *Typographic Elements* lists the typography related visuals from macro (page) to micro-scale (glyphs). *Function* defines the role each typographic element play in conveying the content. *VR Space* evaluates the relevance of each element in VR space.

Other elements, such as page numbers, chapter names, references in parenthesis, and so on, will be eliminated due to the absence of other parts of the book for these components to work.

Other aspects of screen typography can present designers with new opportunities to enhance the type's visual and editorial power. For example, color in typography may refer to the color of the ink used or to the tonal values of the typeface. The tonal value of a typeface depends on proportion and line weight (Haslam, 2006) i.e., the thicker the lines, the darker the font will appear. It is also affected by the horizontal space between sentences, the use of hyphenations and justification, and vertical space (determined by the combination of x-height and the amount of leading). Darker text blocks tend to look more forward, while lighter text blocks tend to recede in the page (Haslam, 2006). Controlling for this allows for the separation of textual elements. For example, headings may appear darker than the captions, and captions can appear darker than the body copy or vice versa. Because the change in scale in accordance with the user's movement is possible, these values can dynamically be manipulated in VR space to adapt to a user's needs.

In addition, color can be used as an effective tool to emphasize or understate the information on a 2D plane. A word painted red can grab people's attention on a page and therefore become more effective in increasing readers' response rates (Clair & Busic-Snyder, 2012). Space can afford type with different lighting adjustments and surface qualities. Like with real-life objects, surface color, and textures affect light reflection, helping designers make more prominent elements of typographic emphasis. Figure 3.2 and Figure 3.3 illustrate some examples and possible directions.

Figure 3.2: Stylistic Alternatives to Glyhps in VR



Note. Stylistic alternatives to letterforms in 3D Space from Miller (1996). In order; *Extrusion, Curling Ribbons, Rotation, Tubing.*

The position of the typographic elements also reinforces the hierarchy. Running headers and footers, for example, are usually quiet on the page. Being positioned at the margins, those repetitive elements are usually not the most prominent elements on the page. On the other hand, headlines and subheads are expected to be positioned at the page's beginning. Positioning of the elements on the page reveals relational aspects through the alignment of the text. Four basic alignments of text are: left-aligned, right-aligned, centered, and justified (Haslam, 2006). Each supports reading different information or the page's visual appearance using different strengths. Book readers have been customized with certain alignment styles over time, some being stylistic outcomes, others being functional concerns.

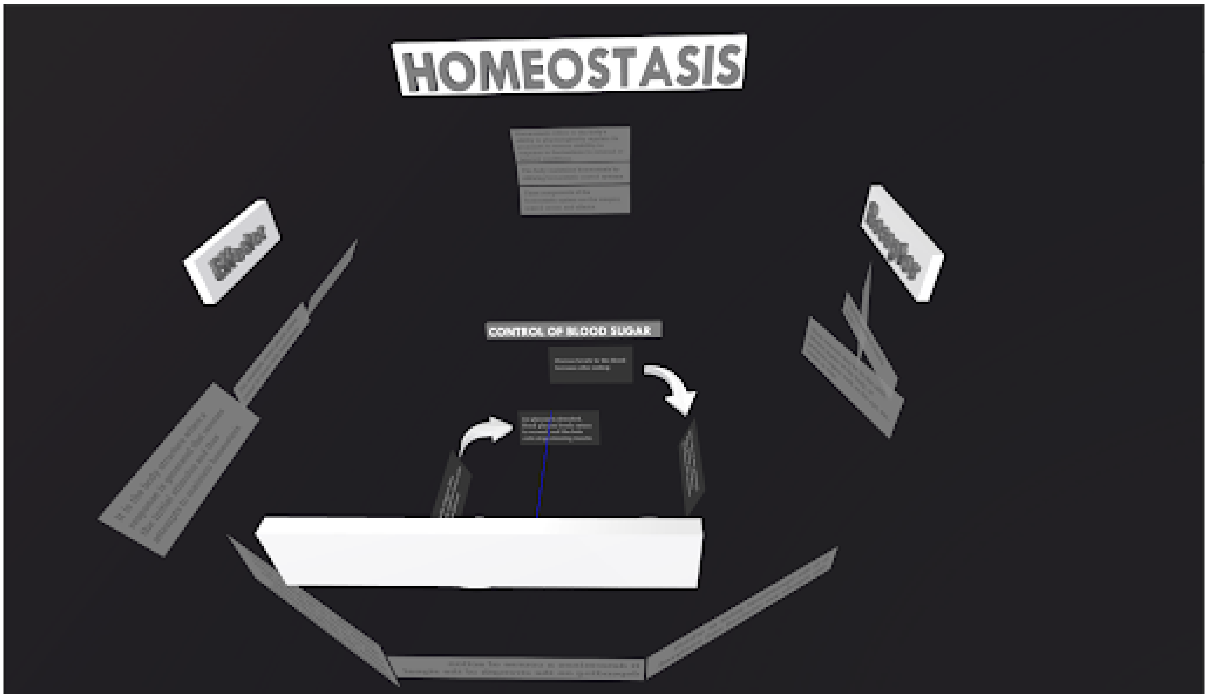
Figure 3.3: Other Editorial Treatments for Text in VR

FORM	LIGHTING	SURFACE QUALITY
extrusion	area light	glass
rotation	point light	glossy
tubing	spot light	refraction
shadowing	sun light	transparent
sewing	beam shape	velvet
molecular construction		toon
modular construction		subsurface
bloating		different textures

Note. Possible editorial treatments for text in VR with varying light and surface qualities.

And finally, decisions about size affect the hierarchy of the page. The readership, content, purpose, and format usually inform such decisions. Most text in different media uses a range of type sizes. The careful selection of additional sizes in relation to grid, typeface, and weight determines the typographic hierarchy. The relative visual importance of typographic elements can be signaled as larger and often bold text. Still, an editorial decision can be made by using smaller headings and establishing hierarchy through the use of position and/or color.

Figure 3.4: Spatial Prose with Textual-Diagrams in VR



Note. A screenshot from the experiment environment displaying prose and the diagram.

Most of these possible treatments will not be implemented in the design of spatial text. Instead, micro and macro-level typographic elements will stay identical between the conditions to only reflect a change in layout. In other words, the main information hierarchy, placements, type size, editorial applications, as well as local and global boldness of type will remain the same, except for prose; text will fit in the immediate field of view, but without requiring scrolling (simultaneous global and local access to text), and for textual-diagrams labels will surround the user (Figure 3.4).

3.3 Experimental Design

To measure the differences in lexical recognition and cued-recall of sentence location performances between flat and spatially distributed layouts in VR, two testing conditions were implemented:

3.3.1 Control Condition

Aims to gauge participants' baseline memory performances using a layout similar to a Portable Document Format (PDF) but delivered in Virtual Reality (Figure 3.5b). Participants used the right-hand controller to scroll through the text.

3.3.2 Experimental Condition

Textual elements were spatially distributed to surround the user (Figure 3.5a). These elements were stationary, and users followed the reading sequence by rotating their bodies 360 degrees clockwise to read sequentially.

Figure 3.5: Images From the Two Conditions

(a) Spatial Layout Condition　　(b) Flat layout Condition

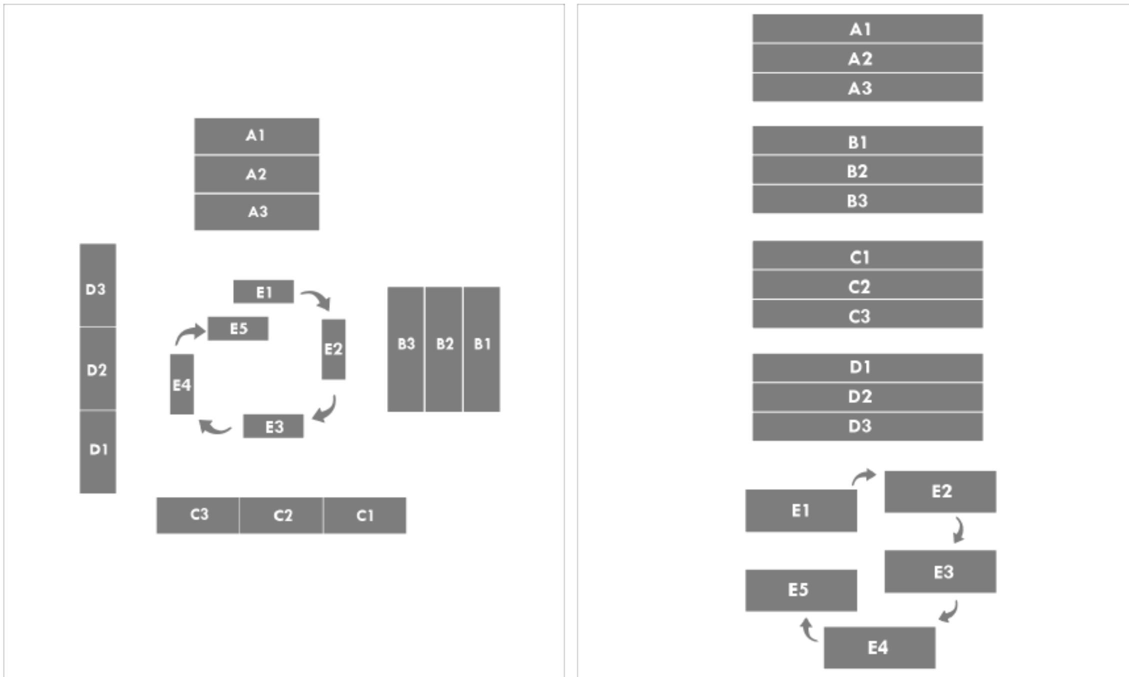
Note. In the spatial condition participants are placed in the middle of the scene, and in the flat condition, they are faced to the text panel.

Students read about the two physiological concepts *homeostasis* and *labor in humans.* The text comprised paragraphs and a diagram displaying an example of the concept. The material is reproduced in Appendix A. The task was to remember the text as completely as possible. Each participant was timed and limited to a 7-minute maximum reading span. There were four possible conditions, and subjects were randomly assigned to one (homeostasis flat, homeostasis spatially distributed, labor flat, labor spatially distributed). Each subject had an equal chance to receive any treatments under the study. Although not presented in this analysis, some additional data was collected for future studies.

After reading, participants first took a lexical recognition test to identify previously read sentences among several distractors. There were two kinds of distractors; (1) identical, (2) paraphrased, and (3) absent. For prose, there were 19 statements shown in total (7 identical, 7 paraphrased, and 5 absent); for the diagram, there were 13 statements (5 identical, 5 paraphrased, and 3 absent). All were shown in random order. The maximum time given to the participants for this test was 15 minutes.

Following that, a cued-recall test asked participants to place each of the 17 sentences in the locations where they read them. Cued-recall involves retrieval of content with the help of cues (Lorch et al., 1995). Participants were given an outline of the text and sorted the location of sentences based on that outline (Figure 3.6). Each sentence was shown in randomized order. The maximum time given to the participants for this test was 7 minutes.

Figure 3.6: The Outline of the Text for Spatial and Flat Layouts



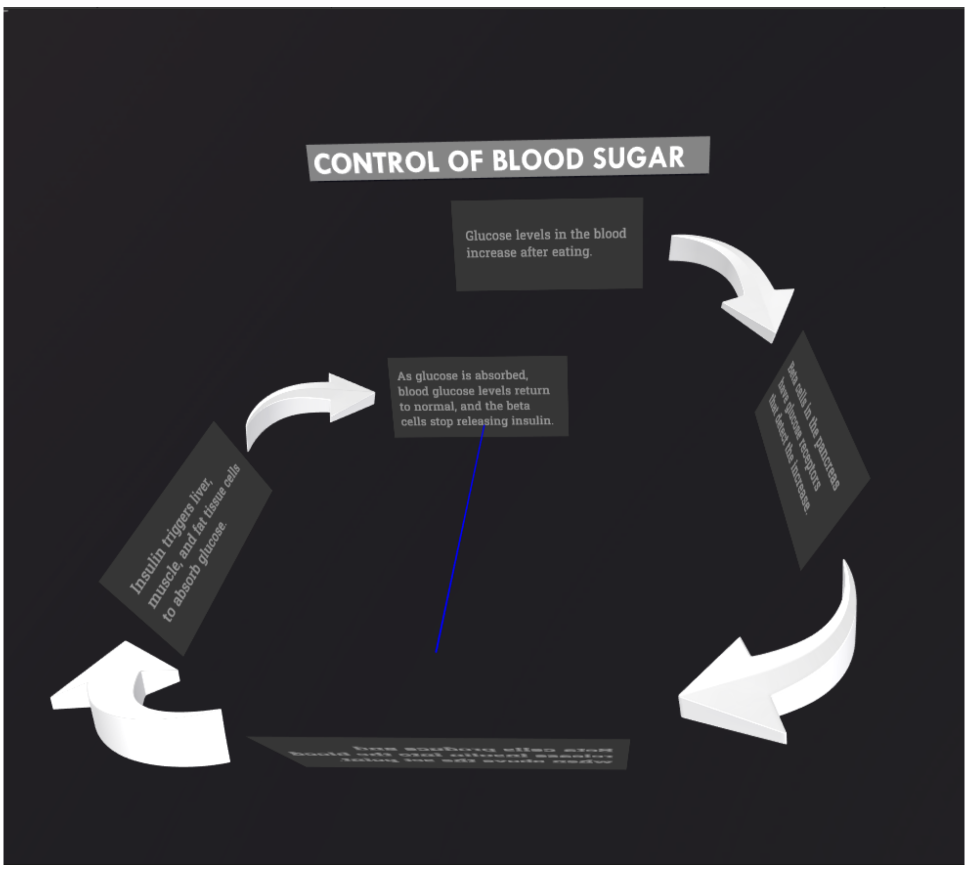
3.3.3 Improving Comparability

The preliminary way to improve comparability between the two conditions was to show the control condition in VR instead of a computer display. Reading from a scrolling PDF in VR provides two things for the experiment design. First, it eliminates the possibility of a difference in performance time and accuracy due to familiarity with reading text from a computer display. Secondly, it removes the effect of enhanced computer graphics hardware, contributing to more effective text rendering. VR presents new challenges regarding reading comfort due to low-resolution and pixel density measures compared to computer screens, which are known to be more exhausting and headache-inducing, especially for long-term use (Li & Wong, 2021).

A secondary means to enhance comparability between the two conditions was in relation to the text difficulty and length. Both texts were measured regarding text difficulty using several different readability tests. Ultimately, both texts had a Flesh-Kincaid Grade Level of 10.2 (Flesch, 2007) and Flesch Reading Ease Score of 42.7 (Kincaid et al., 1975). The average number of words per sentence was 11, and the average number of syllables per word was 1.8. Each text had 23 sentences and 270 words and an estimated average reading time of 2.7 minutes.

Finally, both readings had the same text structure, which shows the order in which the scientific concept occurs. Two kinds of textual elements in both readings, prose, and diagram, also had the same features. The prose consisted of four paragraphs, and the diagram had five panels that gave examples of the concepts explained in the prose (Figure 3.7).

Figure 3.7: Diagrams in Spatial Layout



3.4 Design of the Experiment Space

3.4.1 Content

Prose text started with a paragraph that briefly described the main concept and introduced the topic of the text. Following that, three other paragraphs described the components of homeostasis control systems; receptor, control center, and effector, or the stages of labor; dilation, expulsion, and placental for the second text. Each paragraph described the concept by its function, explained its structure, and gave an example of the body part responsible for fulfilling that function or events taking place in each stage, respectively (see Appendix A).

Diagrams explained the reactions produced by each concept. Homeostasis describes the process of regulating blood sugar in humans, and labor describes the role of oxytocin in labor.

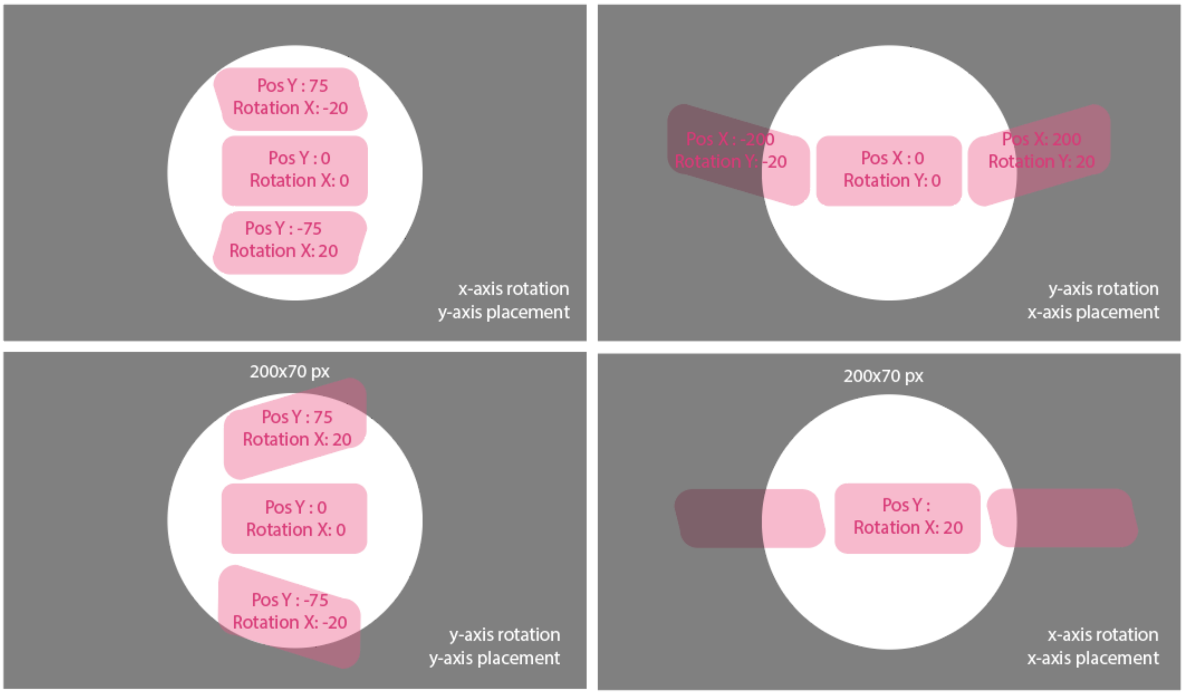
3.4.2 Text

For both environments, Roboto Regular was used for the body text. It was chosen because of its elevated x-height, open counters, augmented weight, width, and medium stroke contrast. These factors are expected to increase legibility. Titles are sans serif, and they are derived from TW Cen MT Bold. The contrast ratio between foreground and background color is 4.5:1, which aims to increase readability using #EEEEE color space. (#5F5F5F for the background, #6D6D6D text.)

3.4.3 Dimensions

Refers to the dimensional treatments of the text elements and the layout. In the flat layout condition, elements only had length and width without any noticeable thickness, similar to Portable Document Format (pdf) on a computer. The text was placed on a tablet, facing toward the participants at all times. Only 1/3 of the text was visible before scrolling. The sentences were linearly placed from top to bottom, the diagram being the last element. Participants were stationary throughout the reading. Scrolling through the text was almost intuitively understood and easily performed.

Figure 3.8: Four Screens and Text Arrangements



Note. For each paragraph of prose the sentences are displayed at various angles in X and Y axis and orientations.

In the spatial layout condition, each paragraph extended to the Z-axis (depth) by varying the rotation values of sentences (Figure 3.8). The first and the last sentence of each paragraph had a rotation of 20 degrees in the x- or y-axis and enclosed the middle sentence, which was located in the middle. Two of the four paragraphs had horizontal, and the other two had a vertical arrangement of sentences.

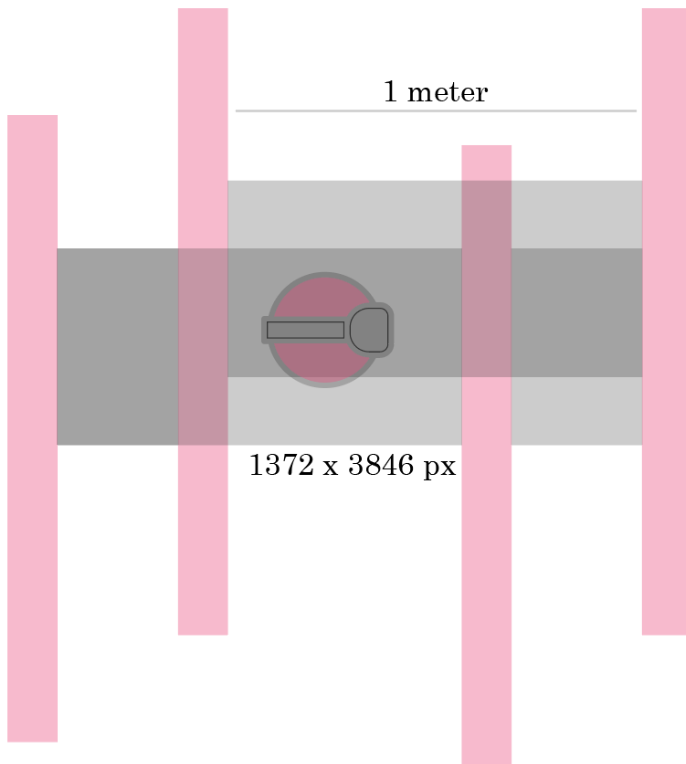
Each sentence was placed on a panel with 200 pixels in width and either 35, 50, or 70 pixels in height based on the length of the sentence. Diagram panels were slightly smaller at 150 pixels and had 70 pixels in height.

All letters in the main title, including the subtitles, were three-dimensional. They were extruded to have a visible thickness. A spotlight lighted them with an appropriate focus cone. The title was bigger in size and was placed slightly above the subtitles. The main purpose was to signal the starting point of the reading visually.

3.4.4 Posts

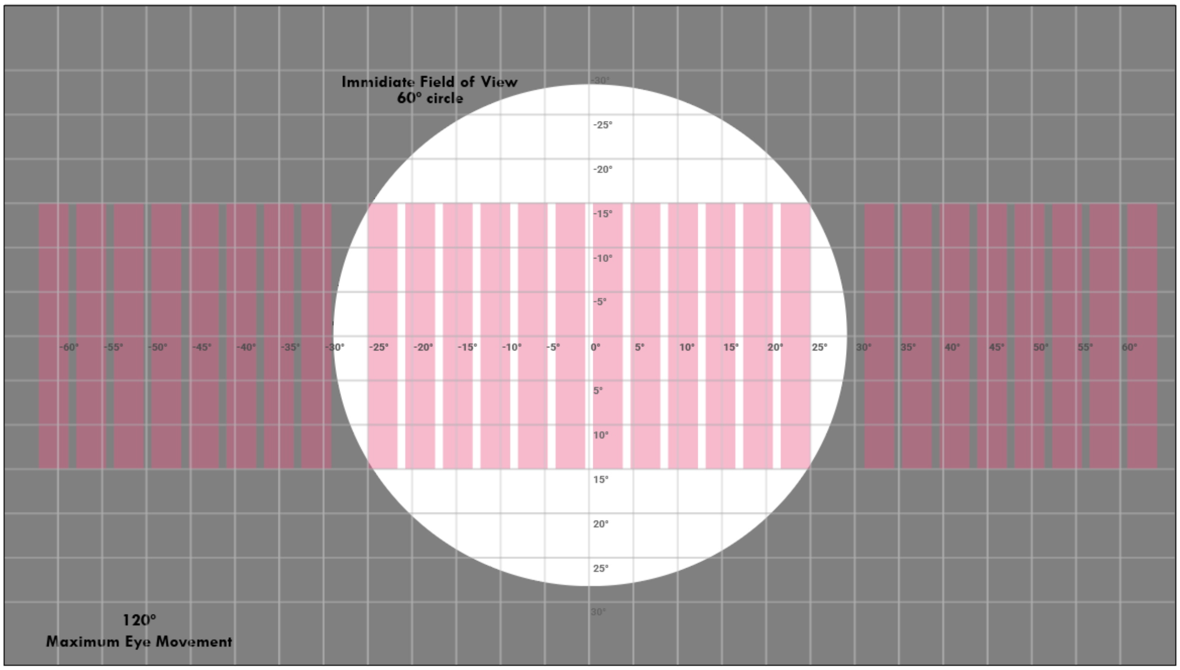
Posts were used to divide the 360 space into areas and help organize content, including type (Figure 3.9). The intended viewing distance for text was one meter in the world space so that participants could have more space to move. Therefore, posts were placed 1 meter apart from one another and formed four screens placed in cardinal coordinates. Each paragraph was placed between the two posts using a screen size of 1372 x 3846 pixels. Initially, the camera's height was 170 cm from the floor to keep the body text in the eye's comfort zone. This measurement was rearranged per each participant's height.

Figure 3.9: The Posts Used in Placing the Text Panels



Each screen uses two grid systems (Figure **??**). One is the angular grid which divides the screen into visual angles of 5 degrees. On top of it, a twelve-column grid is placed inside the visual angle of 60 degrees, also called the eye comfort zone (Bowman et al., 2003). The center of the body text is placed 6 degrees lower than the mathematical center (Alger, 2015). The flat screen with a one-meter distance corresponds to 53.1 degrees in visual angle.

Figure 3.10: Spherical and 12-column Grid



The reading direction was from left to right, requiring two full 360 rotations to finish reading the entire text, including the diagram (Figure 3.10). Because this is the reading direction of English, participants followed this rotation order for the initial reading, then continued to rotate and walk around the text as they preferred. Diagram was placed half of these distances (0.5 meters in the distance, 85 cm in height) and enclosed the reader fully. To complete reading the whole diagram, subjects had to turn 360 degrees. This rotation was signaled with arrows between the panels.

3.5 Study Design

3.5.1 Practice Session

The practice session familiarized participants with the VR headset and the tests they would take in the main session. During practice, the researcher helped participants to properly wear the VR headset to provide optimal reading experience and physical comfort. This included adjusting the side and top straps and firm arms on either side of the headset until the headset was resting lightly on the participant's face and the text appeared clearly. Another measure was Inter-pupillary distance (IPD), the distance between the two pupils. This measure ensures a comfortable spacing between the two lenses to achieve the best image clarity. This was an especially important measure for comfortable reading because otherwise, the text appears blurred. The Oculus Quest 2 comes with three different presets; 61mm or smaller (narrowest), 61-66mm (middle), and 66mm or larger (widest). The researcher provided that measure by having the subjects fixate on the dot on the wall and using a ruler to determine the distance between two pupils.

Following that, participants were introduced to the controller functions. For the whole experiment process, only the right-hand controller was used. It functioned in two ways; the first function was scrolling the text in the flat condition, and the second was a ray cast constantly from the controller so that participants could follow the lines of text while reading.

The reading used in the practice session was about *Greenhouse Effect*. It was not a college-level text and was shorter than the main session's text. It had 5 sentences, a diagram with 5 panels, and a spatially distributed layout. Participants had three minutes to finish reading the text and then took the two memory tests. The lexical recognition test included 4 items, and the cued-recall test had 9 items to be placed on the text outline. After finishing the tests, the researcher reviewed each item in the lexical recognition test and explained what was expected in the main session.

Before continuing the main session, participants watched a short

video to clear their short-term memory. Following that, they were introduced to the main task by the researcher.

3.5.2 Main Session

After participants wore the headset comfortably, the timer started on their marks. They were told earlier that they would read about the physiological concept of homeostasis. Once the text was revealed, participants were told to look around and see the text structures before starting the reading or practice scrolling through the text, depending on their assigned condition. After familiarizing themselves with the environment, they were prompted to memorize the text as completely as possible while reading.

3.5.3 Lexical Recognition

After completing the reading, participants first took the lexical recognition test. In this test, they were asked to read 32 sentences in total. For each sentence, they were asked to decide if the sentence was *identical, paraphrased, or absent*.

1. **Identical:** Describes the sentences read in the text word by word, i.e., direct extractions from the text.

2. **Paraphrased:** Defines the sentences which carried the same meaning as the identical sentences but used synonyms and/or changed the word order. As a result, sentences still carried the same meaning but had little or no lexical overlap with sentences in the original text.

   It is the body structure that detects changes in a variable, which is a substance or process that is regulated. **(identical)**

   A variable is an element or activity that is regulated, and it is the body structure that notices the differences in that variable. **(paraphrased)**

3. **absent:** Specifies the sentences that did NOT appear in the text. Participants were warned that the accuracy of the statements should not influence their responses. In other words, whether the statement is true or false if their judgment says it was not in the text, they select *absent*. Some sentences presented false information regarding the topic (i.e., invalid) but nonetheless shared notable lexical overlap with the verbatim sentences.

For example, sweat glands release sweat to `decrease` *body temperature.* **(identical)**

For example, sweat glands release sweat to `increase` *body temperature.* **(invalid identical)**

Some statements were appropriate to the topic. Nevertheless, they did not appear in the text at all. Subjects were expected to mark them as *absent*. For example, the following sentence is generally valid but not in the text.

Regulation of blood glucose is largely done through the endocrine regarding hormones. **(valid but absent)**

Participants who already know certain information regarding the topic of the text are expected to confuse some of the new valid statements (with or without lexical overlap) with statements taken from the text. On the other hand, they should be able to recognize statements that are not true of the text.

3.5.4 Cued Recall of Sentence Location

In this test, participants were given a total of 17 sentences (12 for the prose and 5 for the diagram) and were asked to locate them using a provided layout of the text.

Notice that the texts described the components of each concept under the subtitles (e.g., Receptor or Dilation) but never explicitly mentioned the terms on the sentence level. The absence of the referents in the sentences intentionally prevented participants from locating sentences directly under the subtitles, which is known to help with recall accuracy substantially (Winograd & Lynn, 1979).

3.6 Experimental Materials

3.6.1 Apparatus

An Oculus Quest 2 (2020 model) was used for the experiment. A room-scale guardian was set, which defined a safe play boundary that was 2.4m (7 ft 10.5 in) by 2.1m (6 ft 10.6 in). The experiment took place in a research room at the library which provided a silent reading environment. All participants used the same headset, and the base station remained identical for all participants.

3.6.2 Participants

40 participants (23 Female, 17 Male) were recruited from the North Carolina State University campus. Their age ranged from 18 to 24 (30%), 25 to 34 (62.5%), and 35 to 44 (7.5%). Each participant received a $20 compensation for their time. Participants had to be older than 18 years old, currently enrolled as students (undergrad=17.5%, masters=35%, and doctorate=42.5%), and without a chronic major medical illness or recent hospitalization. All participants had native (60%) or professional (40%) proficiency in English. They had 9 majors, the most common being Master of Graphic and Experience Design students with 12 participants (30%). Most participants reported previous VR experience (75%), with 3 (7.5%) of which noted regular use over 1 hour per session.

3.7 Measures

This research borrows methods from Experimental (Hegarty & Just, 1993) and Educational Psychology (Lorch et al., 1995). Therefore the measures are defined as follows;

- **Lexical Recognition:** It is defined as the ability to recognize previously encountered verbatim sentences amongst distractor sentences from a lengthy prose text (Terry & Mason, 1982). It was measured as a function of the spatial layout treatment. The test consisted of 32 sentences, 13 of which were for diagrams.

  For prose, 7 sentences in each set were verbatim sentences from the text. From each of the identical sentences, a paraphrased sentence was constructed. The paraphrased sentences involved word substitutions and changed word sequences from the original sentences. For diagrams, there were 5 identical and 5 paraphrased sentences. The paraphrasing method was adopted from Kintsch & Bates (1997). Finally, there were 5 distractors generated for prose and 3 for diagram questions. The distractors contained statements that were factually false in regard to what was given in the text, although there were lexical overlaps between the sentences from which they were derived. This type of recognition memory test is referred forced-choice test where one of the items is usually a target item, and the others are distractor items (Tajika, 2001).

- **Cued-Recall of Sentence Location** Cued-recall offers cues that are related to the previously presented information during the retrieval of the memory (Moult, 2011). In this test, participants were asked to locate each sentence they read in the correct slot using an outline of the text. The visual outline had a box for each sentence and space between sentences when a paragraph was formed. Sentences were shown randomly, and participants were told to create the right reading sequence using the text's visual outline. The partial-credit scoring method was used similarly to the lexical recognition test.

3.7.1 Scoring

Both tests adopted a partial-credit scoring method instead of correct/incorrect scoring (Figure 3.1). Despite being commonly used in memory tests, all-or-nothing scoring has low discriminating power. It produces a limited range of scores in which most information about individuals is lost (Gonthier, 2022).

Table 3.1: Partial Scoring for lexical recognition and cued-recall tests.

	Lexical recognition scoring		
	Identical	Paraphrased	Absent
Identical	3	2	1
Paraphrased	2	3	2
Absent	1	1	3
	Cued-recall of location scoring		
	Sequence and Serial Position	Sequence	Serial Position
Points	3	2	1

For the lexical recognition test, subjects who recognize an identical sentence as *identical*, or paraphrased sentence as *paraphrased*, or a sentence that was not in the text as *absent* get the highest point of 3 for each correct answer. However, because a subject who recognizes an identical sentence as *paraphrased* instead of *absent* certainly demonstrates better memory and therefore gets partial credit (2 points) for choosing *paraphrased* or gets the lowest score (1 point) for choosing *absent*. Partial credit scoring contains the same basic information as all-or-nothing scoring. Still, it increases discriminating power due to the inclusion of information from trials that were not totally correct (Ostrow et al., 2015).

Cued-recall of sentence location test also used partial credit scoring. Therefore, if a participant accurately recalled the sequence of the text (order implied by the subtitles) and the correct serial position (sentence order), they got the highest point (3 points)(e.g., marks B1 for a sentence that appeared in the second panel and was the first sentence). If the stimuli were recalled in the correct sequence, participants got 2 points (e.g., marks C2 for a sentence that appeared in C1), and if the sentence was placed under the correct serial position, they got 1 point partial credit. In other words, the omission of the serial position of a sentence does not invalidate the whole sequence. The subject still gets points for placing sentences in the correct paragraphs. Table 3.1 summarizes the partial-credit scoring method used in the lexical recognition and cued-recall tests.

CHAPTER

4

RESULTS

4.1 Research Questions

The two research questions which led the analysis were:

RQ1. How do the spatially distributed layouts in VR impact lexical recognition and cued-recall of sentence location performances of prose scientific text?

Because diagrams differ from sentential representations in their spatial and structural connections and the content (Larkin & Simon, 1987), they are hypothesized to show a different impact on memory performances. Therefore the secondary research question was:

RQ2. How does the impact of spatially distributed layouts in VR change the lexical recognition and cued-recall of sentence location performances based on prose vs. diagram structures of the scientific text?

To answer these two questions, multiple regression analysis was used. It quantified the differences between the two layout conditions.

4.1.1 Model Specifications

This research uses a standard Ordinary Least Squares regression (OLS) model to predict the relationship between lexical recognition, cued-recall of sentence location scores, and layout design. This section looks at the relationship between lexical recognition scores and layout design. In a typical multiple regression model, all independent variables are simultaneously entered into the regression equation, and each independent variable's individual predictability of the dependent variable is evaluated (Fidell et al., 1996).

Before collecting the data, potential causal relationships were established. The regressand in the model was *LexP* (aggregated verbal recognition scores for prose). The primary regressor of interest was *Layout* and operationalized as a binary variable (spatially distributed=1, flat=0). The regression model includes the following control variables; **(1)** *Labor* (reading topic) which was operationalized as a binary variable *(labor=1 for labor, labor=0 for homeostasis)*, **(2)** *Male* (gender) was recorded as a categorical variable *(male=1, female=0)*, **(3)** *Age* was collected as a ratio variable *(18-25, 25-30, and 30-35)*, **(4)** *Edu* (education level) was operationalized as the highest degree completed *(High school degree or equivalent (e.g. GED), Bachelor's degree (e.g. BA, BS), Master's degree (e.g. MA, MS, MEd), Doctorate or professional degree (e.g. MD, DDS, PhD))*, **(5)** *Design* was a categorical variable for designers and non-designers (*designer=1, non-designer=0)*, **(6)** *Visuo* (visuospatial working memory ability), and **(7)** *TextRec* (text recall ability) which was sampled by asking participants to self-rate their abilities for each at four levels; *poor, fair, average, good, and excellent*. The answers were mapped to 1,2,3,4,5 for analysis. **(8)** *TopInt* (topic interest) and **(9)** *VRper* (VR perception) were measured using Likert-scales to evaluate six statements about the topic and the perception of VR technology by the individuals *(Strongly Disagree, Disagree, Agree, Strongly Agree)*. The answers were mapped to 1,2,3,4 for analysis. **(10)** *VRfam* (VR familiarity) was operationalized as a categorical variable on three levels *(Yes, I have used VR before, No, I have never used VR before, and Yes, I use VR very often)* and mapped to 0, 1, 2. Finally, **(11)** *Eng* (Language proficiency) was collected as another categorical variable marking participant's English language proficiency on three levels *(Native or Bilingual Proficiency, Full Professional Proficiency, and Limited Working Proficiency)*. The answers were mapped to 1,2,3.

Covariates are relevant variables predicted to have a relationship with spatial layout and can also be determinants of memory scores. For example, individuals with a high capacity for visuospatial working memory will likely have better scores because their visuospatial working memory is activated through a three-dimensional layout design. Also, VR familiarity can contribute to increased reading processing, leading to higher memory scores. VR perception measures participants' view of VR reading regarding its usefulness, attractiveness, and easiness. A negative perception of VR reading might result in lower memory scores due to the prejudices against VR as a platform to read from. Additionally, gender-based differences in spatial abilities can correlate with the perception of three-dimensional objects and might contribute to memory scores. Additionally, high interest in the text topic can result in better memory scores through previous knowledge or high motivation. Text recall abilities can positively affect the relationship between layout design and memory score for someone who rates this ability high. We can also expect the relationship between the predictor and explanatory variable based on the individual's language proficiency. Moreover, major changes in memory occur across the human lifespan. Thus, including age might allow controlling for potential relationships where young age leads individuals to obtain better verbal recognition scores.

The multiple regression equation for lexical recognition scores for prose is as follows;

$$\begin{aligned} LexP &= \beta_0 + \beta_1 Layout + \beta_2 Labor + \beta_3 Male + \beta_4 Age + \beta_5 Edu + \beta_6 Design \\ &+ \beta_7 Visuo + \beta_8 TextRec + \beta_9 TopInt + \beta_{10} VRper + \beta_{11} VRfam + \beta_{12} Eng + \epsilon_i \end{aligned} \tag{4.1}$$

While the preferred model specification is displayed through equation 4.1, for this analysis, a bivariate model that includes only *LexP* will be performed to understand how the relationship changes as significant covariates are controlled.

$$LexP = \beta_0 + \beta_1 Layout + \epsilon_i \tag{4.2}$$

A similar multiple regression model assesses the relationship between cued-recall scores and the layout design. The main regressand in the model was *RecP*, which represents the aggregated cued recall scores for prose. Therefore the fitted regression model was;

$$\begin{aligned} RecP &= \beta_0 + \beta_1 Layout + \beta_2 Labor + \beta_3 Male + \beta_4 Age + \beta_5 Edu + \beta_6 Design \\ &+ \beta_7 Visuo + \beta_8 TextRec + \beta_9 TopInt + \beta_{10} VRper + \beta_{11} VRfam + \beta_{12} Eng + \epsilon_i \end{aligned} \tag{4.3}$$

Cued-recall, like recognition, is a kind of memory test, but differently; it involves cues to aid the memory. After receiving the cue (Lovelace & Southall, 1983), the participant tries to recall the item. All the aspects of the previous model specification are expected to affect the cued-recall of sentence location performance. Therefore, they are included in both models.

The secondary research question of this experiment looks at the differences in memory performances based on two kinds of textual elements: diagrams and prose. For that, *LexD* and *RecD* will be used to run another regression model.

4.1.2 Assumptions of Multiple Regression

For conducting a regression analysis, some assumptions have to meet. The following are four relevant assumptions of multiple regression analysis.

In this experiment, a random assignment was used to allocate participants to each group. The primary function of random selection is estimating an unbiased average treatment effect (Stock et al., 2003). In this context, "unbiased" means that, under repeated measures, the participant's characteristics or any other variable will not be related to the assignment to the conditions. It is essential because although an individual experiment may have baseline differences between the states, differences are expected to be canceled out in either direction over repeated measures. Another vital purpose of random assignment is ensuring that the potential confounders are evenly distributed (Goldberg, 2019). This provides that observed differences between the experimental conditions do not differ on one or more confounding variables, and only the independent variable causes the observed differences.

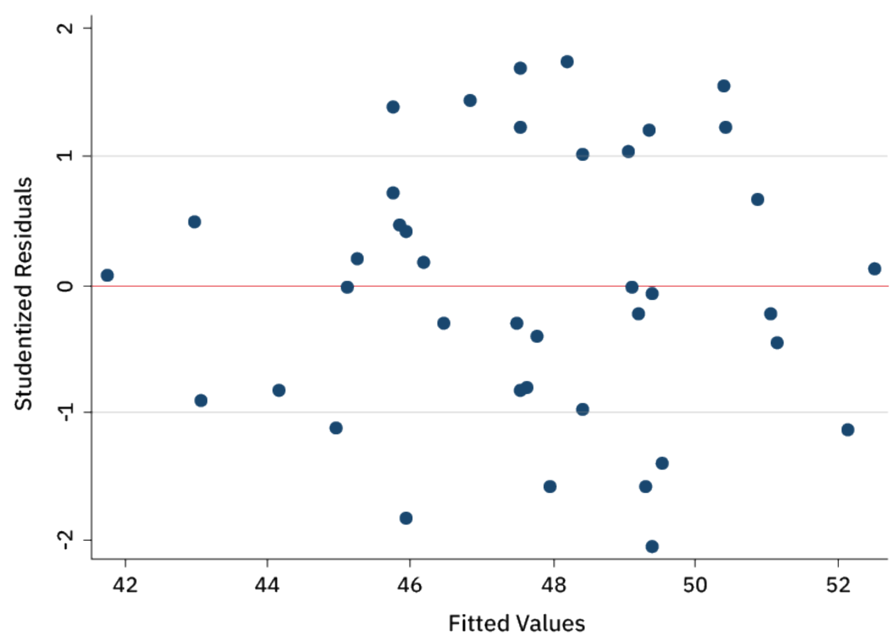
1. **E(u|X=x)=0:** This assumption relates to the conditional distribution of error terms given each observation of the independent variable has a mean of 0 (Stock et al., 2003). Because this is a randomized controlled experiment with a binary treatment, participants are randomly assigned to each group, independently of their characteristics, including those that determine lexical recognition and cued-recall of sentence location scores. Because of this random assignment, the conditional mean of u given X is 0, and the first condition of multiple regression is met (Stock et al., 2003).

2. **(Xi, Yi),i=1,...,n, are i.i.d.:** The model assumes that the observations are independently and identically distributed (Stock et al., 2003). For this experiment, a subset of

participants was randomly selected from a single population; NC State University students. Therefore, a sample of 40 students drawn from this population has the same distribution and is independent of one observation to the next due to randomization (Stock et al., 2003).

3. **Large Outliers Are Unlikely:** Multiple regression models are sensitive to outliers which are known to distort the statistics (Stock et al., 2003). An observation with a studentized residual that exceeds -3.3 or 3.3 is defined as an outlier (Smith, 1988). The studentized residuals plot shows no data points that fall outside the range of -3 to +3 standard deviation units (Figure 4.1).

Figure 4.1: Scatter Plot for Standardized Residuals



4. **No Perfect Multicollinearity:** The fourth assumption, multicollinearity, is caused by two independent variables that are highly correlated(Alin, 2010). To test for multicollinearity, variance inflation factors (VIFs) were calculated for the independent variables in the model. VIF scores significantly higher than 10 can be a source of

multicollinearity. Figure 4.1 VIFs are within the normal range for each predictor in the model. Therefore multicollinearity assumption is met.

Table 4.1: Variance Inflation Factor (VIF); The Measure of the Amount of Multicollinearity.

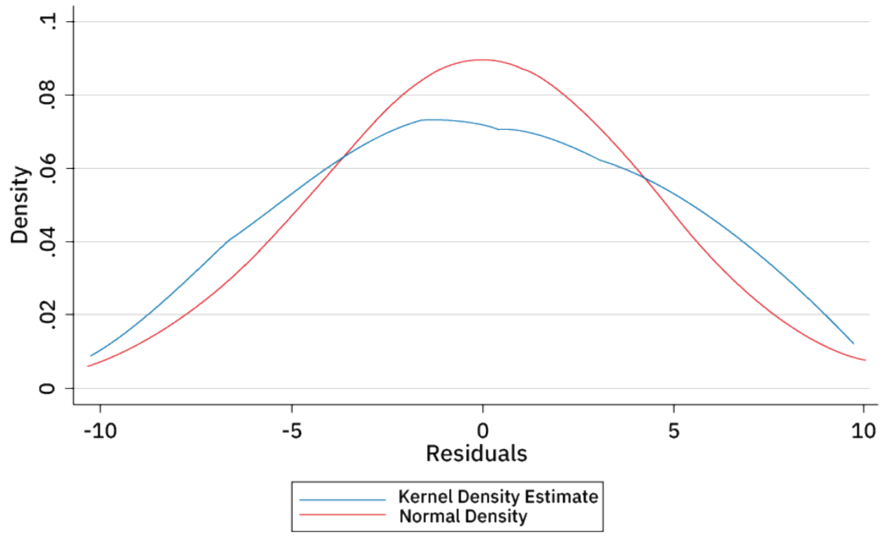
Variable	VIF	1/VIF
Age	2.53	0.39
Profession	2.29	0.43
Visuospatial Ability	2.19	0.45
Text Recall Ability	2.13	0.47
Education	2.03	0.49
VR Perception	1.73	0.57
Gender	1.71	0.58
Layout	1.69	0.59
VR Experience	1.57	0.63
Topic Interest	1.47	0.68
Mean VIF	1.93	

Normality and homoscedasticity are other assumptions of multiple regression (Stock et al., 2003). Normality assumes that the outcome and the error terms are normally distributed. Kernel density estimates of residuals in figure 4.2 follow a possible deviation in the upper and lower tail but show no indications of non-normality. Homeoscedasticity suggests that the variance of the dependent variable is the same for all independent variable values. BreuschePagan/Cooke Weisberg test for heteroskedasticity showed that the estimated model met the homoscedasticity assumption (p-$value$= 0.7, α= .05).

4.1.3 Descriptive Statistics

Table 4.2 shows descriptive statistics on the variables of interest. On average, participants in this sample had 47.7 points on the lexical recognition test for prose text with a range of 39 to 57 points and 27.98 points on the cued-recall test with ranging values of 15 to 36. The majority of the participants were studying design-related majors (.65). VR perception

Figure 4.2: Kernel Density for Normality Assumption



was rated at .54, and topic interest at .57, with a similar range of .24 to .86 and .24 to .86, respectively. The average participant in this experiment had a positive view of the VR technology (.54) but rated his/her text recall ability lower (.33) than visuospatial working memory abilities (.4).

VR familiarity was a categorical variable with three levels and showed that most participants had experience using VR before (68%) while only 7.5% reported regular use. This suggests that the average person in this sample was familiar with VR; however, the extent to which they know about VR remains unclear without additional information about the characteristics of their experience. Moreover, this implies that the results are only generalizable to populations with more VR experience which does not represent the general population. 43% of participants were male. Most participants (62%) were between the ages of 25-34. 42% of the participants were masters, and 43% were Ph.D. students. Finally, 68% of the students reported bilingual or native proficiency, 33% reported full proficiency, and 5% reported limited working proficiency.

Table 4.2: Descriptive Statistics for Prose.

	Categories	Mean	Standard Dev.	Min.	Max.
Lexical Recognition Scores		47.7	5.12	39	57
Cued-Recall Scores		27.98	5.7	15	36
Gender		.43	.5	0	1
Designer		.65	.48	0	1
Age	18-24	.3	.46	0	1
	25-34	.62	.49	0	1
	34-44	.075	.3	0	1
Education	Undergrad	.15	.36	0	1
	Masters	.42	.5	0	1
	PhD	.43	.5	0	1
VR Familiarity	used before	.68	.47	0	1
	used often	.075	.26	0	1
	never used	.25	.43	0	1
Language Proficiency	Native/Bilingual	.63	.49	0	1
	Full Professional	.33	.47	0	1
	Limited Working	.05	.22	0	1
VR Perception		.54	.15	.24	.86
Topic Interest		.57	.13	.1	.77
Visuospatial WM		.4	.23	0	.67
Text Recall Ability		.33	.23	0	.67

4.2 RQ1: Memory Scores for Prose

Table 4.3 shows results for regressing lexical recognition scores for prose on binary variable text layout. Model (1) includes only lexical recognition scores for prose *LexP*, and Model (2) includes aggregated cued-recall of sentence location scores for prose *RecP* as the bivariate regression with only *Layout* as the regressor. The preferred specifications are Model (3) for lexical recognition and Model (4) for the cued-recall of sentence location scores which includes the full set of covariates listed in equations 4.1 and 4.3.

Model (1). Indicates outcomes from the regression of aggregated lexical recognition scores for prose on categorical variable text layout type. The results do not provide statistically significant evidence that spatially distributed layout improved the lexical recognition test scores of the participants.

Model (2). Shows the results from the regression of aggregated cued-recall scores for prose on binary variable text layout type. The results indicate that the spatially distributed layout did not explain significant variance in cued-recall memory for the sentence location test scores.

Model (3). Looking at the unique individual contributions of the predictors, the results show that, on average, students who read about labor scored 4.45 points higher (p=0.021) compared to those who read about homeostasis. Furthermore, a unit increase in students' perception of VR scores resulted in 7.2 points less on the lexical recognition test (p= 0.009). Finally, on average, people who used VR often increased their lexical score by 11.13 points (p= 0.003) after adjusting for all covariates in 4.1.

Model (4). Looking at the unique individual contributions of the predictors, the results show that on average relative to age groups 18-24 and 35-44, age group 25-34 (p= 0.027) have scores that are 7.1 points higher.

4.3 RQ2: Comparing Memory Scores for Prose vs. Diagram

To compare the aggregated lexical recognition and cued-recall of the sentence location scores, a standard Ordinary Least Squares (OLS) regression model is used. Table 4.4 shows;

Model (1). The bivariate regression model with only lexical recognition for diagrams. (*LexD).*

Model (2). The biavariate model for cued-recall (*RecD)* scores for diagram and *Layout* as the main regressor.

Model (3) and Model (4). are the preferred models with the full set of covariates listed above in the model specification section (4.3). The covariates are identical for diagrams.

Model (1), Model (2), and Model (3) do not provide statistically significant evidence that spatially distributed layout improved the lexical recognition or cued-recall tests. Model (4) suggests that, on average, students who read about labor scored 3.27 points less (p= 0.021).

Table 4.5 compares the participants' answers in percentages for identical and paraphrased sentences based on the layout condition and prose vs. diagram. Using partial scoring, except for identical questions from prose text with a spatial layout (69% for 3D and 66% for 2D), higher percentages of the participants who read from the flat layout choose the correct answer with the highest score for both prose and diagram. For the same identical questions, participants who read it from the flat layout chose "paraphrased" for identical sentences with a higher percentage (26%) compared to spatial layout (17%). In contrast, "absent" was chosen by 14% and 8% for spatial and flat layouts, respectively.

For the diagram, the correct answer with the highest score was chosen with the flat layout for paraphrased (72%) and identical (69%) questions. For the same identical questions with spatial layout, 32% of the participants chose paraphrased, and 8% chose absent, whereas 19% of the participants chose paraphrased for the same identical questions, and 9% selected absent. For paraphrased questions, the same number of participants selected identical (17%), and a higher percentage (21%) selected absent for the same questions.

The table 4.6 compares cued-recall results based on the layout condition. When asked about the location of the sentence, 60% of the participants accurately recalled the sequence of the text (order implied by the subtitles) and the correct serial position (sentence order) for prose text with the spatial layout, and 57% for flat layout condition. Those who only accurately recalled the serial position of prose sentences had similar percentages for both conditions (15% for spatial layout and 16% for flat layout). Likewise, the percentages of participants who missed both the serial position and sequence were 25% for spatial and 27% for flat layout.

For the diagram, 68% of participants accurately identified the serial position and the sequence of the sentences for the spatial layout, and 61% for the flat layout. Those who only marked the serial position correctly were 19% of the participants for the spatial layout and 13% for the flat layout. However, 26% of the participants failed to recall both the serial position and the sequence for the flat layout, in comparison to 13% for the spatial layout.

4.4 Thematic Analysis

A thematic analysis (Braun & Clarke, 2006) was conducted to understand factors that played a role in participants reading experiences and performances in VR (Table 4.7).

At the end of the main session, four questions were asked:

- **Q1.** "Did you think the content for the texts was equally difficult (homeostasis and labor)? Why?"

- **Q2.** "What strategies did you use to remember the text for each test (cued-recall and lexical recognition)?"
- **Q3.** "Did you think reading in VR was uncomfortable (for spatially distributed and flat each)?"
- **Q4.** "What else should be done to improve the environment to increase memory rates/functions?"

From the 40 participant interviews, one recording was broken and excluded from the analysis. Table 4.7 summarizes the data analysis processes for the interviews. First, the researcher transcribed the voice recordings verbatim. As patterns emerged, initial codes were identified, and initial themes were formed from codes. After a review of the themes and their meanings, each theme was named and reported.

Three themes were identified, with a total of 9 subthemes (Table 4.8). **Semantics** refers to the meaning-related aspects of the recall and recognition process. *Environment* refers to the reactions to environmental design and/or technology-specific limitations of the environment. *Desirable Features* emerged as inputs by the user, highlighting the desirable components in the reading environment that may increase memory functions.

Text-related cues are described as factors relating to meaning in language that assisted participants' memory performance. Three subthemes that emerged from the initial coding were; topic familiarity, text structure, and use of keywords.

One of the most mentioned factors during the process of interviews was topic familiarity. Familiarity resulted from a personal connection to the topic or previous education. For example, one participant reported his expectation to perform better in labor reading due to his wife's pregnancy, while another said the topics were familiar from high school.

Text structures refer to the organization of the text that represents meaning in terms of semantic and visual structure. Most participants talked about using the logical order of the concepts to remember the information. One participant explained;

> "A common strategy was utilizing how the sections were broken up. Paragraphs go from the definition to function to the body parts. Consistency of text structure helped choose the right slots for each sentence."

Seven participants also reported the use of visual structure in helping them follow the text structure better. When asked about the differences in text difficulty, five out of those read from a flat layout described the easier text as *"linear"*, *"easier to follow"*, or *"a clear process"*. The two participants who read from spatial layout expressed that the "image of the text" and "ability to create a narrative" was stronger in the spatial layout.

Five participants reported that they used titles as mnemonic devices. One participant reported:

> "I looked at the heading and got a general idea of what the text will be about."

Some participants employed a conjoint use of more than two strategies to remember the text. The most common one was using keywords and their position in relation to the logical order of the paragraphs.

> "My first strategy was to look at a number of sentences for paragraph and diagram. Then, I tried to remember keywords."

With 45 total excerpts, text-related cues were the most mentioned theme amongst the other two.

Environment refers to the effect of design choices given the technology-specific limitations of the environment. Kinetic memory, egocentric point of view, and technical limitations were three subthemes from the data.

Spatial layout conditions allowed participants to walk around the text blocks. A preference for reading while sitting was manifested by some participants, while others appreciated the opportunity to develop muscle memory in relation to the text.

> "I learned somewhere that movement contributed to a more active form of learning. That is why I expect to perform better in spatial reading."

A by-product of walking in the environment was the ego-centric point of view. Nine participants explained that they used their bodies as reference points and split the text into chunks based on their position to remember the locations of general themes.

> "The spatial arrangement was easy to remember because I could picture the positions of the text elements. I used my body as a point of reference."

However, three participants expressed a loss of orientation as the duration and complexity of the text increased.

> "I felt a little lost in the beginning. I needed to turn and look around to reorient. It added extra stress. I needed to remember the placement of words in space and the content."

Apart from the design choices, some technicalities about the state of VR technology were found to affect reading comfort. One that recurred often was using the Oculus Link cable, which provided a better performance with fewer disconnects, reduced latency, and less lag. However, three participants mentioned the line extending from the headset as an obstacle to moving freely in the environment. Another factor was the weight of the headset. Five participants found the headset heavy and ill-fitting overall. Finally, the resolution of the screens affected the reading comfort for some participants expressing that the text was often blurry if the distance and the visual angle were not parallel to the head position, which also caused awkward head movements. Nevertheless, seven participants expressed their experience as *"surprisingly comfortable"* with VR reading.

> "Reading in VR was pretty comfortable. It is strange because it is not how we read."

For two participants, an interesting factor that created the most unpleasant spatial reading experience had astigmatism despite corrective lenses. These participants declared fatigue and nausea by the end of the experiment.

Desirable features include input by the participants to improve reading performance in VR. The most represented concept in this theme was interactions. Twelve participants wanted flexibility in rearranging and interacting with the text elements. The two believed that a seated reading experience with the ability to shuffle text using the controllers would create an ideal interaction method for reading.

> "I would like to be able to grab text and rearrange it in a way that would make more sense to me, personally."

Participants indicated that features like note-taking, color coding, highlighting words, adding annotations, and additional imagery as desirable aids for their memory.

4.5 Other Observations

It was observed that participants used other strategies and behaviors during the experiment. Although there was no limitation and enough space to walk around in the spatial layout, only a few participants did. Those who moved used their bodies to crouch and walk away and towards the text rather than just rotating. Moreover, they created *a path in their heads*, which was used as a mental route where the key concepts were located, reinforcing an egocentric frame of reference and connection between words and loci (i.e., method of loci).

In the flat condition, boxes around the sentences had different sizes and were ordered linearly. Some participants noted that they used the dimensions and the ordering of the packets but failed to use this to improve their cued-recall performance because the boxes in the test were all in equal measure. This is expressed as creating "an image of the text" by some participants. Similarly, a participant commented that these boxes made up "chunks," which were grouped as three sentences with different sizes per paragraph. This structure does not generally prevail in traditional prose. Still, it was implemented in the experiment setting to control for the boxes in the spatial condition, which increased the text's legibility.

In the flat layout, the text had a serial order, which the participants saw as more *habitual*. On the other hand, spatial arrangement suggested parallel reading order in which jumping between the text was much more *intuitively* done. One participant defined the second system as a "mental map." For some, it led them to lose the order of reading.

Two participants read the text out loud and reported this as something they used to do when memorizing the text. None of the participants reported any other specific memory techniques they used before the study. However, one participant was involved in the Rubik's cube competitions for the last five years, and memorizing algorithms that translate to various cube movements was commonplace for him. Though he reported not using these techniques in the experiment setting, he also said he located specific keywords from the text concerning his body.

Table 4.3: Regression results for lexical recognition and cued-recall of sentence location on layout design and covariates for prose.

Regressors	**Model 1** *(LexP)*	**Model 2** *(RecP)*	**Model 3** *(LexP+Covariates)*	**Model 4** *(RecP+Covariates)*
Spatial Layout	-.75	.55	-3.5	-1.3
	(1.63)	*(1.8)*	*(2.25)*	*(2.38)*
Labor			**4.45***	-4.7
			(1.79)	*(2.57)*
Male[a]			-.73	-1.3
			(2.28)	*(1.92)*
Age 18-24[b]			7.05	3.35
			(3.77)	*(4.30)*
Age 25-34			4.45	**7.1***
			(2.54)	***(3.02)***
High school Graduates[c]			-.97	4.17
			(2.56)	*(3.33)*
Bachelor Graduate			1.5	2.04
			(1.69)	*(3.05)*
Designer[d]			-.11	-4.12
			(1.88)	*(2.08)*
Visuospatial Memory			-.74	6.33
			(4.56)	*(3.5)*
Text Recall			3.04	5.89
			(4.04)	*(4.25)*
Topic Interest			6.11	1.45
			(6.48)	*(6.89)*
VR Perception			**-7.2***	-2.44
			(2.5)	*(4.09)*
Used VR before[d]			1.78	2.79
			(2.93)	*(2.3)*
Used VR often			**11.13***	1.55
			(3.4)	*(3.66)*
Native/Bilingual[e]			-2.72	-7.59
			(5.32)	*(6.45)*
Full Professional			-2.53	-10.11
			(4.62)	*(5.6)*
Observations	40	40	40	40
R^2	(0.005)	(0.0024)	(0.47)	(0.49)

Note. The outcomes for Models (1) and (3) show lexical recognition scores, and Models (2) and (4) show results for cued-recall of sentence location scores. All standard errors (in parentheses) are heteroscedasticity-robust (* $p < 0.05$). Reference groups for; (a) female,(b) 35-44 age group, (c) doctorate or professional degree, (d) non-designers, (e) limited working proficiency.

Table 4.4: Regression results for lexical recognition and cued-recall of sentence location on layout design and covariates for diagram.

Regressors	Model 1 *(LexD)*	Model 2 *(RecD)*	Model 3 *(LexD+Covariates)*	Model 4 *(RecD+Covariates)*
Spatial Layout	-1.15	1.1	-2.3	1.33
	(.83)	*(.95)*	*(1.58)*	*(1.24)*
Labor			-.51	**-3.27***
			(1.2)	***(1.24)***
Male[a]			.39	-0.55
			(0.8)	*(1.03)*
Age 18-24[b]			1.96	2.6
			(2.24)	*(1.93)*
Age 25-34			2.29	2.46
			(1.48)	*(1.35)*
High school Graduates[c]			.89	2.86
			(2.86)	*(2.1)*
Bachelor Graduate			.35	-.25
			(1.52)	*(1.15)*
Designer[d]			.59	.06
			(1.0)	*(.99)*
Visuospatial Memory			2.22	1.2
			(1.78)	*(1.88)*
Text Recall			2.68	1.88
			(1.88)	*(2.34)*
Topic Interest			5.23	-2.01
			(3.41)	*(4.18)*
VR Perception			-2.22	-2.01
			(1.9)	*(1.44)*
Used VR before			-.99	1.6
			(1.6)	*(1.46)*
Used VR often			1.92	1.81
			(1.7)	*(1.92)*
Native/Bilingual[e]			-2.31	1.95
			(2.7)	*(2.28)*
Full Professional			-2.28	2.43
			(2.62)	*(2.24)*
Observations	40	40	40	40
R^2	(0.04)	(0.03)	(0.41)	(0.46)

*Note.*The outcomes for Models (1) and (3) are lexical recognition scores, and Models (2) and (4) are for cued-recall of sentence location scores for diagrams. All standard errors (in parentheses) are heteroscedasticity-robust. * $p < 0.05$. Reference groups for; (a) female,(b) 35-44 age group, (c) doctorate or professional degree, (d) non-designers, (e) limited working proficiency.

Table 4.5: The comparison of identical and paraphrased sentences based on layout condition, in percentages.

		Identical Questions			Paraphrased Questions	
		3D(%)	**2D(%)**		**3D(%)**	**2D(%)**
PROSE	Identical(*)	69	66	Identical	23	23
	Paraphrased	17	26	Paraphrased(*)	53	58
	Absent	14	8	Absent	24	19
DIAGRAM	Identical (*)	60	72	Identical	17	17
	Paraphrased	32	19	Paraphrased (*)	62	69
	Absent	8	9	Absent	21	14

Note. 3D refers to the spatial layout, and 2D refers to the flat layout. The correct answer is marked with (*) for the two columns.

Table 4.6: The comparison of cued-recall results based on the layout condition.

		3D (%)	**2D(%)**
PROSE	Serial Position+Sequence	60	57
	Serial Position	15	16
	Incorrect Answers	25	27
DIAGRAM	Serial Position+Sequence	68	61
	Serial Position	19	13
	Incorrect Answers	13	26

Note. 3D refers to the spatial layout, and 2D refers to the flat layout. Serial Position+Sequence is the correct answer, and Serial Position refers to the partially correct answers for 3D and 2D columns.

Analytic steps suggested by (Braun & Clarke, 2006)	**Analytic steps of the current study**
Familiarizing with the data	Interviews are transcribed systematically and verbatim.
Generating initial codes	Patterns that are repeated in participants' expressions are identified as initial thoughts and reflections. The main focus is placed on the statements rather than the study's goal.
Searching for themes	Codes are divided into themes and subthemes. The differences between the flat and spatially distributed layouts are critically reviewed. The emerging themes are evaluated under the study's goals.
Reviewing themes	Themes are reviewed once more to solidify core meanings. Early themes are reviewed to join the main themes when related.
Naming and defining the emerging themes	relevant names were applied to describe the core themes.
Producing the report	Themes are described and reported by the researcher.

Table 4.7: The description of analytic steps.

THEME 1			**THEME 2**			**THEME 3**	
Text-Related Cues			Environment			Desirable Features	
Subthemes		Frequency	**Subthemes**		Frequency	**Subthemes**	Frequency
Topic Familiarity		11	Kinetic Memory		5	Interactions	13
	Title	5	Egocentric POV		12	Color	11
Text Structure	Visual Structure	7		Cord	3	Image	7
	Logical Order	12	Technical Limitations	Heavy HMD	3		
Use of Keywords		10		Blurry/Unfocused Text	6		

Table 4.8: Frequency of occurrence for subthemes and themes in VR reading experience.

CHAPTER

5

DISCUSSION

The present research looks at the role of a spatial layout design on participants' memory for scientific prose and diagrams. To summarize, two testing conditions (flat and spatially distributed layout for text) were used to measure the scientific text's lexical recognition and cued-recall performances. A secondary analysis is also performed to see if the same memory tests differ in their scores based on the diagram and prose formats.

Overall, the spatial layout did not help participants' memory performances for diagrams and prose. For prose and the diagram, different topics affected participants' scores. The maximum score for lexical recognition was 57 points for prose and 39 points for diagrams. Participants in the experiment scored an average of 3 points higher, reading about homeostasis in diagrams on the test in comparison to text about labor, and scored almost 5 points more when they read about labor in prose in comparison to homeostasis, suggesting that an effect of the topic and the text type (prose vs. diagram) exists. This might suggest that having texts that are not just in the same subject area but also more closely related for both diagram and prose could be beneficial.

Participants' VR perception and familiarity were decisive factors that affected their memory scores. A unit increase in students' perception of VR scores resulted in 7.2 points less on the lexical recognition test, and people who used VR often scored 11.13 points more on the lexical recognition test. First of all, everyday people's perception of VR is likely shaped by the external narratives surrounding the technology. VR has emerged in popular culture as a futuristic and highly immersive tool that can replace real-world sensory information and create, as Bowman & McMahan (2007) put it, "...oohs and ahs and a unique user experience". Historically, this expectation was at its highest in the early '80s when VR was seen as the next big thing. Still, the day's technology did not meet the expectations, leading to "VR Winter." (Foxman, 2018). Today, these high expectations are closest to being delivered, especially after the increased interest of big tech companies such as Meta. However, VR is neither the center of the technology ecosystem nor the driving force for it (Evans 2020). Undoubtedly, the technology is unique and exciting, but practitioners still use relatively few immersive VR systems in the real world (Bowman & McMahan, 2007), let alone everyday users. Therefore, the contrast between the positive perception of VR and lower test scores may be explained by a mismatch between expectations created around the technology and the humble typographic setting they used to read. This might disappoint participants as reading a lengthy text about science is not what they think they would do in VR. Secondly, the experiment environment presented a novel reading pattern drastically different from what humans are used to. Given that reading is habitual and ancient, offering an alternative to how we usually read is difficult to adapt. However, this alternative may not replace how we read but can be a fruitful choice based on the specific reading task, especially for people who are more spatial thinkers or are more familiar with VR technology.

Complementary to this view, people who are more familiar with VR and use it regularly performed significantly better than those who used VR once before or for the first time. Because those who have used virtual reality before tend to feel more in control than those who are experiencing it for the first time, this might lead to frustration, which can impact memory functions. During the interview, some of these issues were raised; the most frequently noted were the hefty HMD, the long cord that extends from it, and the blurry/unfocused lettering. Experienced users had an easier time managing environmental and technology-related conditions. Moreover, from the participant's perspective, the lack of schemata for spatial reading increases effort. The concept of effort is well-enunciated in psychology. James (1890) notes, "If you ask how many ideas or things we can attend to at once, the answer is not very easily more than one, unless the processes are very habitual."

An important function of habituation is automation which reduces the effort of performing (Wickens, 2014), which was lacking for those less familiar with VR. Before automation, there is *germane load* which refers to the mental resources dedicated to formulating and automating schemata in long-term memory. Sweller et al. (1998) observed that appropriate instructional designs reduce extraneous load but increase germane cognitive load. A fundamental assumption about human cognition is that mobility and spatial ability act through physiological and behavioral mechanisms developed throughout human evolution, and they are automated (Cashdan & Gaulin, 2016). This assumption was used to represent and navigate text in space, suggesting a high intrinsic cognitive load like discovering the natural environment and architectural structures. However, this reading format might increase the user's overall cognitive load without enhancing memory due to using additional mental resources to build and automate a schema. This may suggest that the text is not processed as part of spatial cognition unless this environment has aspects of the physical world, such as objects.

Despite these setbacks, VR is rapidly evolving. In order to make VR a learning tool for reading, research should offer users a cognitive benefit that occasionally surpasses the tradition of reading from flat screens. Therefore, it can be argued perhaps the broader adoption of new reading patterns is not the result of the fanaticism of tradition or the difficulty of creating a schema for it but rather a lack of rigorous research which can provide a measurable benefit for the reader, and hence adaptation of a new reading pattern can prevail over habits. This may prove stronger for subgroups with more VR experience and a positive VR perception.

Today, many applications are looking to improve reading efficiency by manipulating typographic elements. For example, "BeeLine" is an extension for Chrome that adds colored gradients to the sentences to guide the reading pattern. The "Bionic Reading" is another API extension from Swiss designer Renato Casutt that highlights letters in each word to guide the eye through the text. Both of these applications allow users to customize specific parameters of the system. Neither BeeLine nor Bionic Reading provides rigorous research to justify their claims except for one unpublished, ad-hoc study with no references for Bionic Reading. In typography research, the results of the empirical studies do not always override what typographers think or do and vice versa (Dyson & Kipping, 1997; Dyson, 2014; Hartley & Harris, 2014). Lonsdale (2014) states that identifying the level of agreement between science and practice can inform and complement one another. For example, typographic tradition can affect the design of the experiment environment when science has inconclusive answers. This may hold true for two-dimensional typography that is fed by an already existing body of research and years of design practice. However, when there is a lack of tradition and theoretical research to draw from, designers have to gain new perspectives in the making process. Moreover, researchers must keep expanding until a shift in approach or underlying assumptions of typography occurs through empirical evidence. VR reading is in its pre-paradigmatic stage and needs more research to make more people believe it has potential. On the way to this, rigorous experiments and design sketches should be produced.

Participants reported conflicting views on the ego-centric reference frame used in the experiment. As part of the design, participants were asked to walk and rotate to see text that surrounded them 360 degrees around. This was an analogy to scrolling, obscuring the parts of the passage at any given time. This resulted in a dispersed view of the text content, which talks about the parts of the same topic. For instance, seeing *receptor* and *effector* simultaneously as parts of homeostatic systems was impossible. Wickens & Carswell (1995) describe the principle in display design called *proximity compatibility* as one guideline to use in determining where a display should be located, given its relatedness to other displays. The two dimensions of proximity determine the layout design significantly. First, *perceptual proximity* defines how close the task-related information lies in the user's perceptual space (e.g., using the same physical dimensions), and secondly *processing proximity* defines the extent to which multiple sources are used as part of the same task, i.e., the sources that are integrated should have close proximity. The compatibility of these two dimensions affects user performance. In this context, the obscured view of the same content may prove incompatible due to close processing proximity but distant perceptual proximity since parts of the text do not require independent processing. Including a global overview of the text that shares the same topic and offering zooming in on demand may prove more functional for some participants. For instance, homeostasis and its aspects can be in the same field of view but in different groupings for various groupings of information.

Moreover, this navigation method led to getting lost in space and losing the reading order for five participants. The current hierarchy in the spatial layout aims to use levels and type sizes to help participants orient themselves. However, some participants still struggled with clearly understanding the relationships, their present location, and deciding where to look next within the system. When it comes to paper documents, there are usually some organizational standards. For instance, contents (at the very beginning) and indices (at the very end) pages in books provide guidance on where specific items are located within the text's main body. In the text, relative position concepts like "before" and "after" have actual physical correlates and relatively fixed vertical orientations. In a spatial layout, the text you once faced can be "behind," "in front," or "next to" the user as they move. This unusual interaction may account for differences in memory performance. Including aspects directly copied from a book structure can prove useful for future applications. For example, a more linear and permanent positioning of main titles may provide a static orientation, i.e., an anchor point, while still having subtitles expand in space. As a result, participants are less likely to feel overwhelmed by what they encounter.

On the other hand, other participants claimed that physical movements, similar to going around a house's rooms, caused better processing of the environment. However, walking and rotating around is a physically more effortful activity than sitting and turning the pages with your fingertips. While some participants used their bodies to their fullest and moved a lot, others expressed their wish to sit while reading. Learning styles in *Experiential Learning Theory* explain how learners may dwell in some parts of the learning cycle while ignoring others (Kolb et al., 2014). According to the theory, individual learning styles are not fixed traits but dynamic states resulting from continual learning experiences. One person's embodied adaptive stance toward the world shows how these habits influence the person cognitively, perceptually, and behaviorally (Peterson et al., 2015). For some individuals, although it was effortful, movement awareness enabled greater retention of information in accordance with their preferred approach to learning.

Text-related cues were reported as an integral part of text memorization. The most recurring factor cited by participants as contributing to higher results on the lexical recognition test was topic familiarity. Text memory is often guided by semantic memory, which facilitates the creation and retrieval of meaningful information (Tulving, 1972). While semantic memory mainly relies on text-based processing (Glenberg et al., 1987b), for some participants, knowledge-based processing was elicited when reading about the topic they are familiar with. The common source for the knowledge was a personal experience with the process described in the text. In other words, the reading process simultaneously elicited parts of episodic memory—memory for individually experienced events— for some participants. The distinction between episodic memory and semantic memory—general knowledge of the world—is a key concept in memory study (Tulving, 1972). Episodic memory supports the retrieval of knowledge by helping relate new information to previously stored information more readily when the content is familiar (Yonelinas, 2002), such as when drawing upon a personal memory of pregnancy to remember information about labor in women. For instance, one of the participants who went through his wife's first pregnancy reported a deep understanding of the processes, conditions, and biological factors experienced firsthand and very recently, while for others, the topic was processed differently. Moreover, research shows that readers' pre-existing cognitive frameworks influence what they pay attention to and remember in the text (Spyridakis & Wenger, 1991). More specifically, according to the Selective Attention Strategy, three steps are taken to process and learn textual information (Reynolds, 1992). Initially, information is processed at a minimal level, and the reader rates the importance of different segments, and consequently, she focuses her attention selectively on the more important parts of the text. Due to the allocation of attention, these parts are learned better than the less important ones. According to Reynolds (1992) importance of the text should be referred to as *salience*, meaning the property of standing out. Salience predicts that regardless of the causes or origins of it, salient text segments are learned better because selectivity puts more attention on them (Hidi, 1995). Nevertheless, it should be noted that this blanket term does not distinguish between objective importance or personal and subjective evaluations of the text. For example, a second source of familiarity was a previous class about the topic participants read about. One participant explained that he learned about the concept of homeostasis in a first-year college-level biology course and found it important to understand how our bodies work; therefore, he expected to perform well above average on the test.

Participants assigned the sentences to the appropriate locations in the outline using tactics related to the text structure and keyword memory for the cued-recall of sentence location test. The text structure was memorized using titles, visual structure, and the logical order of paragraphs. The text structure is among many factors that may contribute to difficulty recalling (Taylor, 1982). Expository texts are usually arranged using a hierarchy of key concepts and supporting details. Generally speaking, many explanatory text structures do not follow a well-defined and conventional structure that is known to help with memorization (Taylor, 1982). However, in the stimulus material, both texts described a biological process in sequential order, and this structure was consistent throughout the two topics. It was mainly helpful for participants to place sentences in the correct order and location based on macro- and micro-level structures. Because readers cannot remember the whole content, they must create a *macrostructure*, or a gist, of the important ideas in the text Kintsch & Bates (1977). The macrostructure is also used as a retrieval aid to recall more detailed text information, even though some micropropositions—representations of local propositions as words and sentences—may be specifically remembered Kintsch & Bates (1977). When accessible, a reader's schema for a particular text genre, such as a schema for narrative text, regulates how macro rules are produced (Taylor, 1982). A cause-and-effect structure was strictly followed between the two topics in the stimulus material, which produced a consistent text structure for the processes described. Often simultaneously, titles were memorized as micropropositions, anchoring parts of the content into the correct locations and contributing to the microstructure formation as well. Inherently, this structure creates a logical order of concepts that are used as aids to improve memory of the text.

5.1 Limitations and Future Work

Four major limitations of this study could be addressed in future research. The first limitation is caused by the significant estimates of covariates in the preferred models. This indicates that the random assignment failed to balance confounders, which caused experimental conditions to non-trivially differ on variables such as VR familiarity and age (Goldberg, 2019). Overall, the statistical significance found in covariates explained the observed differences that are not the same on average between the two groups. It is possible that the treatment condition had, by chance, more participants who use VR often than the control condition. It is also possible that, simply by chance, more participants in the treatment condition were familiar with VR technology and therefore had a more positive perception of it. Research indicates that the chances of this shortcoming of random assignment (differences in confounders) occurring are less likely as sample sizes get larger (Stock et al., 2003; Goldberg, 2019). Future research should include a larger sample size to keep the biased random assignment at a negligible level.

The second limitation relates to the sampling method or more specifically participant recruitment used in the research. This research uses unrestricted and self-selected convenience sampling, which hinders generalization. Most participants in this sample had prior experience with VR, though the extent is unclear; this is not representative of the students in the US. Therefore, results are likely only generalizable to populations with more VR experience. This study should be repeated with a larger population and use a probability sampling method in the future.

The third limitation relates to typographic research. The classical research on typography has mostly been concerned with the legibility of printed or digital texts at the micro level (at the level of characters) (Hartley & Burnhill, 1976) and the organizational problems at higher levels such as layout designs (Lonsdale, 2014). Micro and macro-level elements are interrelated regarding reading performance (Lonsdale et al., 2006). For instance, layout designs are affected by parameters like type size, line length, and interlinear space. Much of the knowledge about micro-level aspects generated from reading has been undertaken by psychologists and opticians whose main aim is to understand cognition and behavior broadly (Thiessen et al., 2020). As a result, experimental work on reading has little to no reference to a coherent set of principles related to typographic decision-making (Hartley & Burnhill, 1976). The interrelated elements are mostly studied in isolation (Hartley & Burnhill, 1976; Lonsdale et al., 2006). This trend continues for three-dimensional typography. The current VR reading research landscape is occupied by micro-level investigations such as text readability (Larson et al., 2000; Knaack et al., 2019; Qing & Edara, 2022; Jankowski et al., 2010; Rau et al., 2018; Kojic et al., 2022) and macro-level investigations such as layout and screen size (Ragan et al., 2012; Polys & Bowman, 2004; Polys et al., 2011) but their interaction are rarely measured. Here, a distinction in VR reading should be made between object-filled, immersive environments and purely textual worlds. It is plausible to say in all VR environments with text; they are treated either as labels, as part of UI elements, or as annotations in a predominantly object-filled environment. It is worth noting that principles for VR typography that the practitioners produce are still just anecdotal evidence with no theoretical base, especially on the micro level. Notwithstanding the state-of-the-art overview of VR reading herein, this study remains limited to that created by the desktop paradigm and isolated research done by other disciplines. Future research should discover the possibilities in either direction.

The fourth limitation relates to the state of VR technology and its effect on readability. A preliminary examination of the experiment environment in VR indicated an effect called *judder* (Smit et al., 2009). In a VR experience, judder can refer to a variety of visual flaws, such as shaking, blurring, or any combination of these. The two most occurring reasons for judder are *smearing* and *strobing*. Smearing is defined as the perception of motion blur that reduces the sharpness and detail of the image in VR. Smearing occurs when a pixel is lit while it moves across the retina. Strobing is the perception of multiple copies of an image at the same time in VR (Smit et al., 2009). It is caused by a low refresh rate or image persistence of the display and becomes more apparent when smearing is eliminated. Image persistence causes images on the screen to remain unchanged on the screen for seconds to minutes. Judder was persistent during reading in VR due to smearing and strobing in the letterforms. Judder increases due to rapid turns of the head and eye and is known to cause motion sickness (Bowman et al., 2003). They were more sensible due to the rapid eye movement reading required. Two individuals disclosed that they had *astigmatism*, which is an anomaly in the eye brought on by a deviation from spherical curvature and resulting in distorted images because light rays are prevented from converging at a single focus (Harris, 2000). They attributed the increased judder effect to their condition, which may prevent those participants from performing the task. Future research should consider the quality of vision during the recruitment process and use high-quality headsets, which can mitigate the adverse effect of low-quality displays.

Some participants had strong suggestions about the features that they wished the environment had to improve their memory. The use of color to differentiate chunks of information and the use of images to support the text were two that appeared in the participants' feedback. The imagery, including color that is displayed simultaneously with the text content, establishes two different but linked memory traces (Clark, 1987). It goes without saying that storing memory in two different functional regions as opposed to simply one greatly increases the likelihood that it will be retained and recalled (Clark, 1987). Imagery code includes the representation of purely visual properties such as color or brightness and spatial and spatiotemporal relationships. Therefore, adding image and color in the experiment environment may hinder memory performance because the same code interferes strongly with one another (spatial and visual) because they call upon the same representational and processing resources (Clark, 1987). However, because most text content includes such elements, readers still rely on them.

The most recurring theme in the interviews was the participants' desire to interact with the individual text elements. The main idea was to customize text elements in order to create a personal organization of the text. This desire parallels the core idea of *method of loci*. The two factors that the method of loci builds off are the imposition of the organization on a previously disorganized list, which was not the case in this experiment, and the linkages between the loci and the items, which force people to digest the information meaningfully and elaborately. Therefore, the reader can create a mental map of spatial text to serve as an association anchor for other items that must be remembered. This idea which operates on a natural human tendency to memorize and recall information, can improve the effectiveness and longevity of text memory. Moreover, the customization may include scale, distance, and rotation variations which may help to differentiate between various kinds of information and to reorganize the concepts to generate additional meaning and, in turn, improve recall. For example, a group of keywords stacked together after the initial read can mean that they share a common theme, and the form they make can act as a spatial label. In some cases, body text can be attached to each keyword at an opposite angle, providing no vision occlusion but still having the option to keep the related concepts together. The options can be numerous. Because desktop computers rely on 2D displays that are limited in space, VR can afford an offloading cognition onto the environment using all three dimensions to distribute information. Finally, the ability to move closer to the text for readability or focus and step back to see an overview of the complete information can be a part of the interaction system, which aids easy navigation and control over visible space. For example, in prose, a new paragraph is formed whenever there is a change in a particular point or idea. Then paragraphs make the main themes labeled by the titles. In flat conditions, these relationships are represented linearly. They cannot be accessed simultaneously within the same field of view once they exceed the page text area. On the contrary, in VR conditions, text can provide a global overview of the content afforded by the vast real estate. By creating an outline that is accessible at any given time during the reading, participants may be able to allocate attention to the content instead of managing largely separate views by scrolling.

CHAPTER

6

CONCLUSION

This chapter will conclude the study by summarizing the important research findings in relation to the research aims and questions, as well as discussing their worth and contribution.

This study aimed to investigate the effect of spatially distributed layouts on college students' recall and recognition performances. The central questions for this research were as follows:

RQ1 How do the spatially distributed layouts in VR impact lexical recognition and cued-recall of sentence location performances of prose scientific text?

RQ2 How does the impact of spatially distributed layouts in VR change the lexical recognition and cued-recall of sentence location performances based on prose vs. diagram structures of the scientific text?

Quantitative analysis revealed that spatial representation of the content did not improve memory functions; recall and recognition of the information. Further findings showed that VR perception and familiarity affected participants' scores in the lexical recognition test. Moreover, the qualitative analysis of the interviews revealed that the text-related cues were the primary ways to decide between paraphrased, identical, and absent sentences in the lexical recognition test, while the egocentric point of view accompanied text-related cues to locate sentences in the correct positions.

While the results are not representative of the students in the US, this approach provides new insights into the effect of spatial typography on human memory for those who are designing for future VR typography. This research instigates an idea for the practitioners that typography could break into two-dimensional planes. Because while participants did not perform better in the spatial condition, the fact that they did not perform worse proves that VR typography can convey information while retaining the context and clues that allow one to travel across it. This research positions itself as one of the few precedents of VR typography, especially for long text, but it also raises questions about the initial constraints of the display, type position, and visual delight that require further attention from the practitioners. Many more design sketches with varying typographic forms at any size, position, and orientation in three-dimensional space should be produced to start the conventions that will work for recall, recognition, and learning. Based on the conclusions, practitioners should consider (1) adopting a stationary position for the users to reduce effort, (2) expanding on the types of interactions that affect the text by allowing the customization of single textual elements, which may increase memory, (3) keeping the variations in text angle to a narrow range in order to increase consistency, and finally (4) improve the hierarchy for the visual elements while allowing the movement of various displays.

Further research is needed to determine the effect of acceptance or resistance toward VR on learning, specifically for science learning. A more comprehensive survey of VR perception with all of its aspects, such as ease of use, level of resistance, and exposure to new technologies and innovation, can help researchers better understand their implications on memory and other aspects of cognition, such as attention, intelligence, spatial ability, and long-term memory in relation to spatial typography. In relation to this, a detailed description of the frequency of use and the scope of use, ranging from professional to entertainment, should be comprehensively investigated solely and in relation to VR perception.

REFERENCES

Abrash, M. (2021). Creating the future: Augmented reality, the next human-machine interface. In *2021 IEEE International Electron Devices Meeting (IEDM)*, pages 1–2. IEEE.

Adams, M. & Collins, A. (1977). A schema-theoretic view of reading. technical report no. 32.

Ahir, K.,Govani, K.,Gajera, R., & Shah, M. (2020). Application on virtual reality for enhanced education learning, military training and sports. *Augmented Human Research*, 5:1–9.

Al-Diban, S. (2012). *Mental Models*, pages 2200–2204. Springer US, Boston, MA.

Alderson, J. C. (2000). *Assessing Reading*. Cambridge Language Assessment. Cambridge University Press.

Alesandrini, K. L. (1984). Pictures and adult learning. *Instructional Science*, 13(1):63–77.

Alger, M. (2015). Visual Design Methods for Virtual Reality. *Personal Website*, (September):98.

Alin, A. (2010). Multicollinearity. *Wiley interdisciplinary reviews: computational statistics*, 2(3):370–374.

Allport, D.,Rennison, E., & Strausfeld, L. (1995). Issues of gestural navigation in abstract information spaces. *Conference on Human Factors in Computing Systems - Proceedings*, 2:206–207.

Alvermann, D. E. & Boothby, P. R. (1983). A preliminary investigation of the differences in children's retention of "inconsiderate" text. *Reading Psychology*, 4(3):237–246.

An, S. (2013). Schema theory in reading. *Theory & Practice in Language Studies*, 3(1).

Anderson, D. E.,Vogel, E. K., & Awh, E. (2013). A Common Discrete Resource for Visual Working Memory and Visual Search. *Psychological Science*, 24(6):929–938.

Anderson, R. C. & Pearson, P. D. (1984). A schema-theoretic view of basic processes in reading comprehension. technical report no. 306.

Anderson, R. C. & Pearson, P. D. (1988). A schema-theoretic view of basic processes in reading comprehension. *Interactive approaches to second language reading*, pages 37–55.

Andrade, J. & Meudell, P. (1993). Is spatial information encoded automatically in memory? *The Quarterly Journal of Experimental Psychology*, 46(2):365–375.

Andrews, C.,Endert, A., & North, C. (2010). Space to think: Large, high-resolution displays for sensemaking. *Conference on Human Factors in Computing Systems - Proceedings*, 1:55–64.

Atkinson, R. C. & Shiffrin, R. M. (1968). Human memory: A proposed system and its control processes. In *Psychology of learning and motivation*, volume 2, pages 89–195. Elsevier.

Baceviciute, S.,Mottelson, A.,Terkildsen, T., & Makransky, G. (2020). Investigating representation of text and audio in educational vr using learning outcomes and eeg. *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*.

Baddeley, A. (1982). Reading and working memory.

Baddeley, A. (1998). Recent developments in working memory. *Current opinion in neurobiology*, 8(2):234–238.

Baddeley, A. (2003). Working memory: Looking back and looking forward. *Nature Reviews Neuroscience*, 4(10):829–839.

Baddeley, A. (2012). Working memory: Theories, models, and controversies. *Annual Review of Psychology*, 63:1–29.

Baddeley, A. & Wilson, B. A. (2002). Prose recall and amnesia: Implications for the structure of working memory. *Neuropsychologia*, 40(10):1737–1743.

Baddeley, A. D. & Hitch, G. (1974). Working memory. In *Psychology of learning and motivation*, volume 8, pages 47–89. Elsevier.

Ball, K. K.,Beard, B. L.,Roenker, D. L.,Miller, R. L., & Griggs, D. S. (1988). Age and visual search: Expanding the useful field of view. *JOSA A*, 5(12):2210–2219.

Barfield, W.,Rosenberg, C.,Han, S.-H., & Furness, T. (1993). Spatial situational awareness as a function of frame of reference, virtual eyepoint elevation, and geometric field of view. In *SID International Symposium Digest of Technical Papers*, volume 24, pages 107–107. SOCIETY FOR INFORMATION DISPLAY.

Bartlett, F. C. (1995). *Remembering: A study in experimental and social psychology*. Cambridge University Press.

Baumann, J. F. (1984). The effectiveness of a direct instruction paradigm for teaching main idea comprehension. *Reading Research Quarterly*, 20(1):93–115.

Bellezza, F. S.,Cheesman, F. L., & Reddy, B. G. (1977). Organization and semantic elaboration in free recall. *Journal of Experimental Psychology: Human Learning and Memory*, 3(5):539.

Berkowitz, S. J. (1986). Effects of instruction in text organization on sixth-grade students' memory for expository reading. *Reading Research Quarterly*, 21(2):161–178.

Berlucchi, G. & Aglioti, S. (1997). The body in the brain: neural bases of corporeal awareness. *Trends in neurosciences*, 20(12):560–564.

Bjork, E. L. & Bjork, R. A. (1996). *Memory*. Elsevier.

Borgman, C. L. (1986). The user's mental model of an information retrieval system: an experiment on a prototype online catalog. *International Journal of Man-Machine Studies*, 24(1):47–64.

Bower, G. (1970). Analysis of a mnemonic device. *American Scientist*, 58:496–510.

Bower, G. H.,Black, J. B., & Turner, T. J. (1979). Scripts in memory for text. *Cognitive psychology*, 11(2):177–220.

Bowman, D. A.,Hodges, L. F., & Bolter, J. (1998). The virtual venue: User-computer interaction in information-rich virtual environments. *Presence: Teleoperators and Virtual Environments*, 7(5):478–493.

Bowman, D. A. & McMahan, R. P. (2007). Virtual reality: How much immersion is enough? *Computer*, 40(7):36–43.

Bowman, D. A.,North, C.,Chen, J.,Polys, N. F.,Pyla, P. S., & Yilmaz, U. (2003). Information-rich virtual environments: Theory, tools, and research agenda. *Proceedings of the ACM Symposium on Virtual Reality Software and Technology, VRST*, Part F1290:81–90.

Brandt, T.,Dichgans, J., & Koenig, E. (1973). Differential effects of central versus peripheral vision on egocentric and exocentric motion perception. *Experimental brain research*, 16(5):476–491.

Bransford, J.,Franks, J.,Morris, C., & Stein, B. (2014). Some general constraints on learning and memory research. *Levels of Processing in Human Memory (PLE: Memory)*, page 331.

Braun, V. & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative research in psychology*, 3(2):77–101.

Brewer, W. F. (1988). Postscript: Imagery and text genre. *Text - Interdisciplinary Journal for the Study of Discourse*, 8(4):431–438.

Brewer, W. F. & Treyens, J. C. (1981). Role of schemata in memory for places. *Cognitive psychology*, 13(2):207–230.

Bringhurst, R. (1996). *The Elements of Typographic Style*. Elements of Typographic Style. Hartley Marks.

Brod, G. (2021). Toward an understanding of when prior knowledge helps or hinders learning. *npj Science of Learning*, 6(1):24.

Brunyé, T. T.,Taylor, H. A., & Rapp, D. N. (2008). Repetition and dual coding in procedural multimedia presentations. *Applied Cognitive Psychology*, 22:877–895.

Butcher, K. R. (2006). Learning from text with diagrams: Promoting mental model development and inference generation. *Journal of Educational Psychology*, 98(1):182 – 197.

Cashdan, E. & Gaulin, S. J. C. (2016). Why go there? evolution of mobility and spatial cognition in women and men. *Human Nature*, 27:1–15.

Chandler, P. & Sweller, J. (1991). Cognitive load theory and the format of instruction. *Cognition and Instruction*, 8(4):293–332.

Chaudhry, M. (2021). Creating effective virtual reality learning experiences: Lessons learned. In *Education and Training in Optics and Photonics*, pages Th4A–1. Optica Publishing Group.

Christie, J. M. & Just, M. A. (1976). Remembering the location and content of sentences in a prose passage. *Journal of Educational Psychology*, 68(6):702.

Christoffels, I. K.,De Groot, A. M., & Kroll, J. F. (2006). Memory and language skills in simultaneous interpreters: The role of expertise and language proficiency. *Journal of Memory and Language*, 54(3):324–345.

Chung, S. T. (2002). The effect of letter spacing on reading speed in central and peripheral vision. *Investigative Ophthalmology & Visual Science*, 43(4):1270–1276.

Clair, K. & Busic-Snyder, C. (2012). *A typographic workbook: A primer to history, techniques, and artistry*. John Wiley & Sons.

Clark, James M Paivio, A. (1987). A dual coding perspective on encoding processes. *Imagery and related mnemonic processes: Theories, individual differences, and applications*, pages 5–33.

Clinton, V.,Taylor, T.,Bajpayee, S.,Davison, M. L.,Carlson, S. E., & Seipel, B. (2020). Inferential comprehension differences between narrative and expository texts: a systematic review and meta-analysis. *Reading and Writing*, 33(9):2223–2248.

Cooper, M. (1989). Computers and design. *Design Quarterly*, pages 1–31.

Cowan, N. (2014). Working Memory Underpins Cognitive Development, Learning, and Education. *Educational Psychology Review*, 26(2):197–223.

Craik, F. I. & Lockhart, R. S. (1972). Levels of processing: A framework for memory research. *Journal of verbal learning and verbal behavior*, 11(6):671–684.

Daller, H.,Van Hout, R., & Treffers-Daller, J. (2003). Lexical richness in the spontaneous speech of bilinguals. *Applied linguistics*, 24(2):197–222.

Dasilva, M.,Laidlaw, D.,Greenberg, B.,Basser, P.,Pierpaoli, C.,Chiocca, E., & Deisboeck, T. (2001). An immersive virtual environment for dt-mri volume visualization applications: a case study.

Delogu, F.,Raffone, A., & Belardinelli, M. (2009). Semantic encoding in working memory: Is there a (multi)modality effect?. *Memory*, 17(6):655 – 663.

Denis, M.,Beschin, N.,Logie, R. H., & Sala, S. D. (2002). Visual perception and verbal descriptions as sources for generating mental representations: Evidence from representational neglect. *Cognitive Neuropsychology*, 19(2):97–112.

Donaldson, W. (1992). Measuring recognition memory. *Journal of Experimental Psychology: General*, 121(3):275.

Donaldson, W. & Tulving, E. (1972). *Organization of memory*. Academic Press.

Dyson, M. C. (2004). How physical text layout affects reading from screen. *Behaviour & information technology*, 23(6):377–393.

Dyson, M. C. (2014). 9. Applying psychological theory to typography: is how we perceive letterforms special? *Visual Communication*, (2014).

Dyson, M. C. & Haselgrove, M. (2001). The influence of reading speed and line length on the effectiveness of reading from screen. *International Journal of Human-Computer Studies*, 54(4):585–612.

Dyson, M. C. & Kipping, G. J. (1997). The legibility of screen formats: are three columns better than one? *Computers & Graphics*, 21(6):703–712.

Estes, W. K. (1977). On the interaction of perception and memory in reading. *Basic Processes in Reading*, 17:1–24.

Eysenck, M. W. (1978). Levels of processing: A critique. *British Journal of Psychology*, 69(2):157–169.

Fang, Z. (2005). Scientific literacy: A systemic functional linguistics perspective. *Science education*, 89(2):335–347.

Fang, Z. (2006a). The language demands of science reading in middle school. *International Journal of Science Education*, 28(5):491–520.

Fang, Z. (2006b). The language demands of science reading in middle school. *International Journal of Science Education*, 28(5):491–520.

Fastrich, G. M.,Kerr, T.,Castel, A. D., & Murayama, K. (2018). The role of interest in memory for trivia questions: An investigation with a large-scale database. *Motivation science*, 4(3):227.

Fidell, S.,Silvati, L.,Howe, R.,Pearsons, K. S.,Tabachnick, B.,Knopf, R. C.,Gramann, J., & Buchanan, T. (1996). Effects of aircraft overflights on wilderness recreationists. *The Journal of the Acoustical Society of America*, 100(5):2909–2918.

Flesch, R. (2007). Flesch-kincaid readability test. *Retrieved October*, 26(3):2007.

Foxman, M. H. (2018). *Playing with virtual reality: early adopters of commercial immersive technology*. PhD thesis, Columbia University.

Frantz, R. S.,Starr, L. E., & Bailey, A. L. (2015). Syntactic complexity as an aspect of text complexity. *Educational Researcher*, 44(7):387–393.

Frase, L. T. & Kammann, R. (1974). Effects of search criterion upon unanticipated free recall of categorically related words. *Memory & cognition*, 2:181–184.

Frase, L. T. & Schwartz, B. J. (1975). Effect of question production and answering on prose recall. *Journal of Educational Psychology*, 67(5):628.

Geller, J.,Davis, S. D., & Peterson, D. J. (2020). Sans forgetica is not desirable for learning. *Memory*, 28(8):957–967.

Gibson, E. J. (1969). Principles of perceptual learning and development.

Giulia Cataldo, M. & Oakhill, J. (2000). Why are poor comprehenders inefficient searchers? an investigation into the effects of text representation and spatial memory on the ability to locate information in text. *Journal of educational psychology*, 92(4):791.

Glenberg, A. M.,Meyer, M., & Lindem, K. (1987a). Mental models contribute to foregrounding during text comprehension. *Journal of Memory and Language*, 26(1):69–83.

Glenberg, A. M.,Meyer, M., & Lindem, K. (1987b). Mental models contribute to foregrounding during text comprehension. *Journal of Memory and Language*, 26(1):69–83.

Goldberg, M. H. (2019). How often does random assignment fail? estimates and recommendations. *Journal of Environmental Psychology*, 66:101351.

Gonthier, C. (2022). An easy way to improve scoring of memory span tasks: The edit distance, beyond "correct recall in the correct serial position". *Behavior Research Methods*, pages 1–16.

Goodman, N. (1976). *Languages of art : an approach to a theory of symbols.* Hackett, Indianapolis.

Gough, P. B. (1984). Word recognition. *Handbook of reading research*, 1:225–253.

Gough, P. B.,Kavanagh, J., & Mattingly, I. (1972). One second of reading. *Cambridge: MIT Press*, pages 331–358.

Graesser, A. C.,McNamara, D. S., & Kulikowich, J. M. (2011). Coh-metrix: Providing multilevel analyses of text characteristics. *Educational Researcher*, 40(5):223–234.

Graesser, A. C.,McNamara, D. S.,Louwerse, M. M., & Cai, Z. (2004). Coh-metrix: Analysis of text on cohesion and language. *Behavior research methods, instruments, & computers*, 36(2):193–202.

Grech, A. M.,Nakamura, J. P., & Hill, R. A. (2018). The Importance of Distinguishing Allocentric and Egocentric Search Strategies in Rodent Hippocampal-Dependent Spatial Memory Paradigms: Getting More Out of Your Data. *The Hippocampus - Plasticity and Functions*, (July).

Guri-Rozenblit, S. (1989). Effects of a tree diagram on students' comprehension of main ideas in an expository text with multiple themes. *Reading Research Quarterly*, 24(2):236–247.

Guthrie, J. T. & Kirsch, I. S. (1987). Distinctions Between Reading Comprehension and Locating Information in Text. *Journal of Educational Psychology*, 79(3):220–227.

Gyselinck, V.,Ehrlich, M. F.,Cornoldi, C.,De Beni, R., & Dubois, V. (2000). Visuospatial working memory in learning from multimedia systems. *Journal of Computer Assisted Learning*, 16(2):166–176.

Hampson, P. J. & Morris, P. E. (1996). *Understanding cognition.* Blackwell.

Harris, R. J.,Cady, E. T., & Tran, T. Q. (2013). *Comprehension and memory.* Routledge.

Harris, W. (2000). Astigmatism. *Ophthalmic and physiological Optics*, 20(1):11–30.

Hartley, J. & Burnhill, P. (1976). Explorations in space: A critique of the typography of bps publications. *Bulletin of the British Psychological Society*.

Hartley, J. & Harris, J. L. (2014). *Reading the Typography of Text*. Routledge.

Hartley, J. & Trueman, M. (1983). The effects of headings in text on recall, search and retrieval. *British journal of educational psychology*, 53(2):205–214.

Hasher, L. & Zacks, R. T. (1979). Automatic and effortful processes in memory. *Journal of experimental psychology: General*, 108(3):356.

Haslam, A. (2006). *Book design*. Laurence King Publishing.

Hegarty, M.,Carpenter, P. A., & Just, M. A. (1991). Diagrams in the comprehension of scientific texts.

Hegarty, M. & Just, M. A. (1993). Constructing mental models of machines from text and diagrams. *Journal of memory and language*, 32(6):717–742.

Heim, M. (1995). The design of virtual reality. *Body and Society*, 1(3-4):65–77.

Hidi, S. (2001). Interest, reading, and learning: Theoretical and practical considerations. *Educational psychology review*, pages 191–209.

Hidi, S. E. (1995). A reexamination of the role of attention in learning from text. *Educational Psychology Review*, 7(4):323.

Hillner, M. (2009). *virtual typography*. AVA Publishing SA.

Hine, K. & Tasaki, H. (2019). Active view and passive view in virtual reality have different impacts on memory and impression. *Frontiers in Psychology*, 10:2416.

Horn, R. E. (1998). Visual language. *MacroVu Inc. Washington*.

Huttenlocher, J. & Presson, C. C. (1973). Mental rotation and the perspective problem. *Cognitive Psychology*, 4(2):277–299.

Hyde, T. S. & Jenkins, J. J. (1969). Differential effects of incidental tasks on the organization of recall of a list of highly associated words. *Journal of Experimental Psychology*, 82(3):472.

Irwin, J. W. (2021). Teaching reading comprehension processes.

Ishizaki, S. & Mitchell, W. J. (1996). Typographic performance. *Technology*.

James, W. (1890). *The principles of psychology, Vol II*. Henry Holt and Company, NY, US.

Jankowski, J.,Samp, K.,Irzynska, I.,Jozwowicz, M., & Decker, S. (2010). Integrating text with video and 3d graphics: The effects of text drawing styles on text readability. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '10, page 1321–1330, New York, NY, USA. Association for Computing Machinery.

Jefferies, E.,Ralph, M. A. L., & Baddeley, A. D. (2004). Automatic and controlled processing in sentence recall: The role of long-term and working memory. *Journal of memory and language*, 51(4):623–643.

Jerald, J. (2015). *The VR book: Human-centered design for virtual reality*. Morgan & Claypool.

Jian, Y. C. & Wu, C. J. (2015). Using Eye Tracking to Investigate Semantic and Spatial Representations of Scientific Diagrams During Text-Diagram Integration. *Journal of Science Education and Technology*, 24(1):43–55.

Jones, M. N.,Willits, J., & Dennis, S. (2015). Models of semantic memory.

Jury, D. (2009). History of typography. In *Handbook of Research on Writing*, pages 41–76. Routledge.

Just, M. A. & Carpenter, P. A. (1980). A theory of reading: From eye fixations to comprehension. *Psychological review*, 87(4):329.

Kahneman, D. (1970). Remarks on attention control. *Acta Psychologica*, 33:118–131.

Kapur, S.,Craik, F. I.,Tulving, E.,Wilson, A. A.,Houle, S., & Brown, G. M. (1994). Neuroanatomical correlates of encoding in episodic memory: levels of processing effect. *Proceedings of the National Academy of Sciences*, 91(6):2008–2011.

Kilteni, K.,Groten, R., & Slater, M. (2012). The sense of embodiment in virtual reality. *Presence: Teleoperators and Virtual Environments*, 21(4):373–387.

Kincaid, J. P.,Fishburne Jr, R. P.,Rogers, R. L., & Chissom, B. S. (1975). Derivation of new readability formulas (automated readability index, fog count and flesch reading ease formula) for navy enlisted personnel. Technical report, Naval Technical Training Command Millington TN Research Branch.

Kintsch, W. (1982). Memory for text. In *Advances in Psychology*, volume 8, pages 186–204. Elsevier.

Kintsch, W. (1994). Text comprehension, memory, and learning. *American psychologist*, 49(4):294.

Kintsch, W. & Bates, E. (1977). Recognition memory for statements from a classroom lecture. *Journal of Experimental Psychology: Human Learning and Memory*, 3(2):150.

Kintsch, W.,Patel, V., & Ericsson, K. (1999). The role of long-term working memory in text comprehension. *Psychologia*, 42:186–198.

Kintsch, W. & Van Dijk, T. A. (1978). Toward a model of text comprehension and production. *Psychological review*, 85(5):363.

Kintsch, W. & Walter Kintsch, C. (1998). *Comprehension: A paradigm for cognition.* Cambridge university press.

Knaack, L.,Lache, A.-K.,Preikszas, O.,Reinhold, S., & Teistler, M. (2019). Improving readability of text in realistic virtual reality scenarios: Visual magnification without restricting user interactions. In *Proceedings of Mensch und Computer 2019*, pages 749–753.

Knight, C. & Munro, M. (1999). Comprehension with[in] virtual environment visualisations. In *Proceedings Seventh International Workshop on Program Comprehension*, pages 4–11.

Kojic, T.,Vergari, M.,Möller, S., & Voigt-Antons, J.-N. (2022). Assessing user experience of text readability with eye tracking in virtual reality. In *Virtual, Augmented and Mixed Reality: Design and Development: 14th International Conference, VAMR 2022, Held as Part of the 24th HCI International Conference, HCII 2022, Virtual Event, June 26–July 1, 2022, Proceedings, Part I*, pages 199–211. Springer.

Kolb, D. A.,Boyatzis, R. E., & Mainemelis, C. (2014). Experiential learning theory: Previous research and new directions. In *Perspectives on thinking, learning, and cognitive styles*, pages 227–248. Routledge.

Kolers, P. A. (1973). Remembering operations. *Memory & Cognition*, 1:347–355.

Kolers, P. A. (1976). Reading a year later. *Journal of Experimental Psychology: Human Learning and Memory*, 2(5):554.

Kolers, P. A. (1979). Reading and knowing. *Canadian Journal of Psychology/Revue canadienne de psychologie*, 33(2):106.

Kolers, P. A. & Ostry, D. J. (1974). Time course of loss of information regarding pattern analyzing operations. *Journal of Verbal Learning and Verbal Behavior*, 13(6):599–612.

Kong, Y.,Seo, Y. S., & Zhai, L. (2018). Comparison of reading performance on screen and on paper: A meta-analysis. *Computers & Education*, 123:138–149.

Kruley, P.,Sciama, S. C., & Glenberg, A. M. (1994). On-line processing of textual illustrations in the visuospatial sketchpad: Evidence from dual-task studies. *Memory & Cognition*, 22(3):261–272.

Köhler, S.,Moscovitch, M., & Melo, B. (2001). Episodic memory for object location versus episodic memory for object identity: Do they rely on distinct encoding processes? *Memory & Cognition*, 29:948–959.

Lanier, J. (1992). Virtual reality: The promise of the future. *Interactive Learning International*, 8(4):275–79.

Larkin, J. H. & Simon, H. A. (1987). Why a Diagram is (Sometimes) Worth Ten Thousand Words. *Cognitive Science*, 11(1):65–100.

Larson, K.,Dantzich, M.,Czerwinski, M., & Robertson, G. (2000). Text in 3d: some legibility results. pages 145–146.

LaViola Jr, J. J.,Kruijff, E.,McMahan, R. P.,Bowman, D., & Poupyrev, I. P. (2017). *3D user interfaces: theory and practice*. Addison-Wesley Professional.

Levie, W. H. & Lentz, R. (1982). Effects of text illustrations: A review of research. *Ectj*, 30(4):195–232.

Li, K. C. & Wong, B. T.-M. (2021). A literature review of augmented reality, virtual reality, and mixed reality in language learning. *International Journal of Mobile Learning and Organisation*, 15(2):164–178.

Linderholm, T. (2006). Reading with purpose. *Journal of college reading and learning*, 36(2):70–80.

Linderholm, T. & van den Broek, P. (2002). The effects of reading purpose and working memory capacity on the processing of expository text. *Journal of educational psychology*, 94(4):778.

Liu, F. (2010). A short analysis of the nature of reading. *English Language Teaching*, 3(3):152–157.

Liu, J.,Prouzeau, A.,Ens, B., & Dwyer, T. (2020). Design and Evaluation of Interactive Small Multiples Data Visualisation in Immersive Spaces. (1):588–597.

Lockhart, R. S. (2002). Levels of processing, transfer-appropriate processing, and the concept of robust encoding. *Memory*, 10(5-6):397–403.

Loftus, E. F. (1992). When a lie becomes memory's truth: Memory distortion after exposure to misinformation. *Current Directions in Psychological Science*, 1(4):121–123.

Lonsdale, M. (2014). Typographic features of text. outcomes from research and practice. *Visible Language*, 48(3):29–67.

Lonsdale, M. d. S.,Dyson, M. C., & Reynolds, L. (2006). Reading in examination-type situations: the effects of text layout on performance. *Journal of Research in Reading*, 29(4):433–453.

Lopez, C.,Halje, P., & Blanke, O. (2008). Body ownership and embodiment: vestibular and multisensory mechanisms. *Neurophysiologie Clinique/Clinical Neurophysiology*, 38(3):149–161.

Lorch, E. P.,Klusewitz, M. A., et al. (1995). Effects of typographical cues on reading and recall of text. *Contemporary educational psychology*, 20(1):51–64.

Lorch, R. F. (1989). Text-signaling devices and their effects on reading and memory processes. *Educational psychology review*, 1:209–234.

Lovelace, E. A. & Southall, S. D. (1983). Memory for words in prose and their locations on the page. *Memory & cognition*, 11(5):429–434.

Lupton, E. & Miller, A. (1999). *Design writing research.* Phaidon Press London and New York.

Lévy, P. (1998). Becoming virtual, reality in the digital age. *Plenum Trade, New York.*

Mayer, R. E. (2004). Designing multimedia technology that supports human learning. *The design of instruction and evaluation: Affordances of using media and technology*, pages 33–49.

Mayer, R. E. & Gallini, J. K. (1990). When is an illustration worth ten thousand words? *Journal of educational psychology*, 82(4):715.

McCormick, E. P. & Wickens, C. D. (1995). Virtual reality features of frame of reference and display dimensionality with stereopsis: Their effects on scientific visualization. Technical report, AIR FORCE INST OF TECH WRIGHT-PATTERSON AFB OH.

McKinley, M. P.,O'Loughlin, V. D., & Bidle, T. S. (2019). *Anatomy and physiology : an integrative approach.*

McKinley, M. P.,O'Loughlin, V. D.,Bidle, T. S., & York, J. G. (2013). *Anatomy & physiology: An integrative approach.* McGraw-Hill New York, NY, USA.

McKoon, G. & Ratcliff, R. (1992). Inference during reading. *Psychological Review*, 99(3):440 – 466.

McVee, M. B.,Dunsmore, K., & Gavelek, J. R. (2005). Schema theory revisited. *Review of educational research*, 75(4):531–566.

Merwin, D. H.,Vincow, M. A., & Wickens, C. D. (1994). Visual analysis of scientific data: Comparison of 3d-topographic, color and gray scale displays in a feature detection task. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, volume 38, pages 240–244. SAGE Publications Sage CA: Los Angeles, CA.

Meyer, B. J. (1975). Identification of the structure of prose and its implications for the study of reading and memory. *Journal of Literacy Research*, 7(1):7–47.

Meyer, B. J. & Rice, G. E. (1982). The interaction of reader strategies and the organization of text. *Text-Interdisciplinary Journal for the Study of Discourse*, 2(1-3):155–192.

Meyer, B. J. & Rice, G. E. (1992). 12 prose processing in adulthood: The text, the reader, and the task. *Everyday cognition in adulthood and late life*, page 157.

Mikropoulos, T. A. & Natsis, A. (2011). Educational virtual environments: A ten-year review of empirical research (1999–2009). *Computers and Education*, 56(3):769–780.

Miller, J. A. (1996). *Dimensional Typography, No1: Words in Space Kiosk Report.* Princeton University Press, USA.

Mishra, P. & Brewer, W. F. (2003). Theories as a form of mental representation and their role in the recall of text information. *Contemporary Educational Psychology*, 28(3):277–303.

Moult, M. (2011). *Cued Recall*, pages 751–752. Springer New York, New York, NY.

Muir, L. & Hawes, G. (2013). The Case for e-Book Literacy: Undergraduate Students' Experience with e-Books for Course Work. *Journal of Academic Librarianship*, 39(3):260–274.

Nagels, A.,Chatterjee, A.,Kircher, T., & Straube, B. (2013). The role of semantic abstractness and perceptual category in processing speech accompanied by gestures. *Frontiers in Behavioral Neuroscience*, 7:181.

Newcombe, N. S. & Frick, A. (2010). Early education for spatial intelligence: Why, what, and how. *Mind, Brain, and Education*, 4(3):102–111.

Newport, R.,Pearce, R., & Preston, C. (2010). Fake hands in action: embodiment and control of supernumerary limbs. *Experimental brain research*, 204(3):385–395.

Ni, T.,Bowman, D. A., & Chen, J. (2006). Increased display size and resolution improve task performance in information-rich virtual environments. *Proceedings - Graphics Interface*, 2006:139–146.

Nichols, S. & Patel, H. (2002). Health and safety implications of virtual reality: a review of empirical evidence. *Applied ergonomics*, 33(3):251–271.

Norman, D. A. (1981). Categorization of action slips. *Psychological review*, 88(1):1.

Norman, D. A. & Bobrow, D. G. (1979). Descriptions: An intermediate stage in memory retrieval. *Cognitive Psychology*, 11(1):107–123.

Norman, D. A. & Verganti, R. (2014). Incremental and radical innovation: Design research vs. technology and meaning change. *Design issues*, 30(1):78–96.

Nunberg, G. (1990). The linguistics of punctuation.

Ostrow, K.,Donnelly, C.,Adjei, S., & Heffernan, N. (2015). Improving student modeling through partial credit and problem difficulty. In *Proceedings of the Second (2015) ACM Conference on Learning@ Scale*, pages 11–20.

Palladino, P.,Cornoldi, C.,De Beni, R., & Pazzaglia, F. (2001). Working memory and updating processes in reading comprehension. *Memory and Cognition*, 29(2):344–354.

Panel, N. R.,of Child Health, N. I.,Development, H.,Initiative, N. R. E.,for Literacy (US), N. I.,Service, U. S. P. H., & of Health, U. S. D. (2000). *Report of the National Reading Panel: Teaching children to read: An evidence-based assessment of the scientific research literature on reading and its implications for reading instruction: Reports of the subgroups.* National Institute of Child Health and Human Development.

Parong, J. & Mayer, R. E. (2018). Learning science in immersive virtual reality. *Journal of Educational Psychology*, 110(6):785.

Pausch, R.,Proffitt, D., & Williams, G. (1997). Quantifying immersion in virtual reality. In *Proceedings of the 24th annual conference on Computer graphics and interactive techniques*, pages 13–18.

Perfetti, C. A. (1985). Reading skills. *Psychiatry*, 50:1125–1129.

Peterson, K.,DeCato, L., & Kolb, D. A. (2015). Moving and learning: Expanding style and increasing flexibility. *Journal of Experiential Education*, 38(3):228–244.

Piaget, J. (2003). Part i: Cognitive development in children–piaget development and learning. *Journal of research in science teaching*, 40.

Piaget, J. & Cook, M. T. (1952). The origins of intelligence in children.

Polys, N. F. & Bowman, D. A. (2004). Design and display of enhancing information in desktop information-rich virtual environments: Challenges and techniques. *Virtual Reality*, 8(1):41–54.

Polys, N. F.,Bowman, D. A., & North, C. (2011). The role of Depth and Gestalt cues in information-rich virtual environments. *International Journal of Human Computer Studies*, 69(1-2):30–51.

Press, T. M. I. T. & Issues, D. (1998). Review Reviewed Work (s): The Visible Word : Experimental Typography and Modern Art 1909- 1923 by Johanna Drucker Review by : Brian Donnelly Published by : The MIT Press Stable URL : http://www.jstor.com/stable/1511830. 14(1):81–84.

Qing, Z. & Edara, P. (2022). Human vision vs. computer vision: A readability study in a virtual reality environment. In *2022 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW)*, pages 61–64. IEEE.

Ragan, E. D.,Bowman, D. A., & Huber, K. J. (2012). Supporting cognitive processing with spatial information presentations in virtual environments. *Virtual Reality*, 16:301–314.

Raja, D.,Bowman, D. A.,Lucas, J., & North, C. (2004). Exploring the Benefits of Immersion in Abstract Information Visualization. *Proceedings of the 8th Immersive Projection Technology Workshop*, pages 61–69.

Rau, P.-L. P.,Zheng, J.,Guo, Z., & Li, J. (2018). Speed reading on virtual reality and augmented reality. *Computers and Education*, 125:240–245.

Rayner, K. & Bertera, J. H. (1979). Reading without a fovea. *Science*, 206(4417):468–469.

Reicher, G. M. (1969). Perceptual recognition as a function of meaningfulness of stimulus material. *Journal of experimental psychology*, 81(2):275.

Reynolds, R. E. (1992). Selective attention and prose learning: Theoretical and empirical research. *Educational psychology review*, 4:345–391.

Richgels, D. J.,McGee, L. M.,Lomax, R. G., & Sheard, C. (1987). Awareness of four text structures: Effects on recall of expository text. *Reading Research Quarterly*, pages 177–196.

Rivas Méndez, D.,Armstrong, C.,Stubbs, J.,Stiles, M., & Kearney, G. (2018). Practical recording techniques for music production with six-degrees of freedom virtual reality. In *Audio Engineering Society Convention 145*. Audio Engineering Society.

Robertson, G. G.,Card, S. K., & Mackinlay, J. D. (1993). Information visualization using 3D interactive animation. *Communications of the ACM*, 36(4):57–71.

Roediger, H. L. & Blaxton, T. A. (1987). Effects of varying modality, surface features, and retention interval on priming in word-fragment completion. *Memory & cognition*, 15(5):379–388.

Roehling, J. V.,Hebert, M.,Nelson, J. R., & Bohaty, J. J. (2017). Text Structure Strategies for Improving Expository Reading Comprehension. *Reading Teacher*, 71(1):71–82.

Ruggiero, G.,Ruotolo, F., & Iachini, T. (2009). The role of vision in egocentric and allocentric spatial frames of reference. *Cognitive processing*, 10(2):283–285.

Rumelhart, D. E. (1977). Toward an interactive model of reading. In *Attention and performance VI*, pages 573–603. Routledge.

Rumelhart, D. E. & McClelland, J. L. (2017). Interactive processing through spreading activation. In *Interactive processes in reading*, pages 37–60. Routledge.

Rundus, D. (1971). Analysis of rehearsal processes in free recall. *Journal of experimental psychology*, 89(1):63.

Rydning, D. R.-J. G.-J.,Reinsel, J., & Gantz, J. (2018). The digitization of the world from edge to core. *Framingham: International Data Corporation*, 16.

Sachs, J. S. (1974). Memory in reading and listening to discourse. *Memory and Cognition*, 2(1):95–100.

Sadeghi, K. (2007). The key for successful reader-writer interaction: Factors affecting reading comprehension in l2 revisited. *Asian EFL journal*, 9(3):198–220.

Salzman, M. C.,Dede, C., & Loftin, R. B. (1999). Vr's frames of reference: A visualization technique for mastering abstract multidimensional information. In *Proceedings of the SIGCHI conference on Human Factors in Computing Systems*, pages 489–495.

Sanders, A. E.,Holtzer, R.,Lipton, R. B.,Hall, C., & Verghese, J. (2008). Egocentric and exocentric navigation skills in older adults. *The Journals of Gerontology Series A: Biological Sciences and Medical Sciences*, 63(12):1356–1363.

Scaife, M. & Rogers, Y. (1996). External cognition: How do graphical representations work? *International Journal of Human Computer Studies*, 45(2):185–213.

Schank, R. (1975). Conceptual dependency theory. conceptual information processing.

Scott, D. & Power, R. (2002). Generating textual diagrams and diagrammatic texts.

Shen, Y. (2008). An exploration of schema theory in intensive reading. *English Language Teaching*, 1(2):104–107.

Shiffrin, R. M. & Schneider, W. (1984). Automatic and controlled processing revisited.

Shing, Y. L. & Brod, G. (2016). Effects of prior knowledge on memory: Implications for education. *Mind, Brain, and Education*, 10(3):153–161.

Simpson, M.,Wallgrün, J. O.,Klippel, A.,Yang, L.,Garner, G.,Keller, K.,Oprean, D., & Bansal, S. (2016). Immersive analytics for multi-objective dynamic integrated climate-economy (dice) models. In *Proceedings of the 2016 ACM companion on interactive surfaces and spaces*, pages 99–105.

Skupin, A. & Fabrikant, S. I. (2003). Spatialization Methods: A Cartographic Research Agenda for Non-geographic Information Visualization. *Cartography and Geographic Information Science*, 30(2):99–119.

Slater, M. (1999). Measuring presence: A response to the witmer and singer presence questionnaire. *Presence*, 8(5):560–565.

Slater, M. (2003). A note on presence. *Presence Connect*, 3(January 2003):1–5.

Slater, M. & Wilbur, S. (1997). A framework for immersive virtual environments (five): Speculations on the role of presence in virtual environments. *Presence: Teleoperators & Virtual Environments*, 6(6):603–616.

Small, D.,Ishizaki, S., & Cooper, M. (1994). Typographic space. *Conference on Human Factors in Computing Systems - Proceedings*, 1994-April:437–438.

Small, D. L. (1999). Rethinking the Book I1. (1987).

Smit, F.,van Liere, R., & Froehlich, B. (2009). A programmable display layer for virtual reality system architectures. *IEEE Transactions on Visualization and Computer Graphics*, 16(1):28–42.

Smith, E. E.,Jonides, J.,Koeppe, R. A.,Awh, E.,Schumacher, E. H., & Minoshima, S. (1995). Spatial versus object working memory: Pet investigations. *Journal of cognitive neuroscience*, 7(3):337–356.

Smith, R. M. (1988). The distributional properties of rasch standardized residuals. *Educational and Psychological measurement*, 48(3):657–667.

Solnyshkina, M.,Zamaletdinov, R.,Gorodetskaya, L., & Gabitov, A. (2017). Evaluating text complexity and flesch-kincaid grade level. *Journal of social studies education research*, 8(3):238–248.

Sperling, G. (1963). A model for visual memory tasks. *Human factors*, 5(1):19–31.

Spyridakis, J. H. & Wenger, M. J. (1991). An empirical method of assessing topic familiarity in reading comprehension research. *British Educational Research Journal*, 17(4):353–360.

Stanovich, K. E. (1980). Toward an interactive-compensatory model of individual differences in the development of reading fluency. *Reading research quarterly*, pages 32–71.

Stein, B.,Littlefield, J.,Bransford, J., & Persampieri, M. (1984). Elaboration and knowledge acquisition. *Memory and Cognition*, 12:522–529.

Stock, J. H.,Watson, M. W., et al. (2003). *Introduction to econometrics*, volume 104. Addison Wesley Boston.

Swanson, H. L. & Berninger, V. (1995). The role of working memory in skilled and less skilled readers' comprehension. *Intelligence*, 21(1):83–108.

Swanson, H. L.,Cochran, K. F., & Ewers, C. A. (1989). Working memory in skilled and less skilled readers. *Journal of Abnormal Child Psychology*, 17:145–156.

Sweller, J.,van Merrienboer, J. J. G., & Paas, F. G. W. C. (1998). Cognitive Architecture and Instructional Design. *Educational Psychology Review*, 10(3):251–296.

Tajika, H. (2001). Recognition Memory, Psychology of. In Smelser, N. J. & Baltes, P. B., editors, *International Encyclopedia of the Social Behavioral Sciences*, pages 12832–12835. Pergamon, Oxford.

Tardif, T. & Craik, F. I. (1989). Reading a week later: Perceptual and conceptual factors. *Journal of Memory and Language*, 28(1):107–125.

Taylor, A.,Sanson, M.,Burnell, R.,Wade, K. A., & Garry, M. (2020). Disfluent difficulties are not desirable difficulties: The (lack of) effect of sans forgetica on memory. *Memory*, 28(7):850–857.

Taylor, B. M. (1982). Text structure and children's comprehension and memory for expository material. *Journal of educational psychology*, 74(3):323.

Terry, W. S. & Mason, H. M. (1982). Recognition of sentences from prose. *Bulletin of the Psychonomic Society*, 19(1):7–10.

Thiessen, M.,Beier, S., & Keage, H. (2020). A Review of the Cognitive Effects of Disfluent Typography on Functional Reading. *The Design Journal*, 23(5):797–815.

Treisman, A. (1986). Features and objects in visual processing. *Scientific American*, 255(5):114B–125.

Treisman, A. M. (1969). Strategies and models of selective attention. *Psychological review*, 76(3):282.

Tulving, E. (1972). Organization of memory. *Episodic and semantic memory*.

Tulving, E. (1974). Cue-dependent forgetting: When we forget something we once knew, it does not necessarily mean that the memory trace has been lost; it may only be inaccessible. *American Scientist*, 62(1):74–82.

Tutt, D. & Harty, C. (2013). Journeys through the cave: The use of 3d immersive environments for client engagement practices in hospital design.

Verhoeven, L. & Perfetti, C. (2008). Advances in text comprehension: Model, process and development. applied cognitive psychology, 22, 293-301. *Applied Cognitive Psychology*, 22:293 – 301.

Waller, R. (1979). Typographic access structures for educational texts.

Waller, R. (1982). Text as diagram: Using typography to improve access and understanding. *The Technology of Text*, 1.

Waller, R. (1987). The typographic contribution to language towards a model of typographic genres and their underlying structures.

Wang, A. Y. & Thomas, M. H. (2000). Looking for long-term mnemonic effects on serial recall: The legacy of simonides. *The American journal of psychology*, 113(3):331.

Warde, B. (1955). *The crystal goblet*. Sylvan Press.

Webster, R. (2016). Declarative knowledge acquisition in immersive virtual learning environments. *Interactive Learning Environments*, 24(6):1319–1333.

Webster, R. D. (2014). Corrosion prevention and control training in an immersive virtual learning environment. In *CORROSION 2014*. OnePetro.

Weinstein, C. E. & Mayer, R. E. (1983). The teaching of learning strategies. In *Innovation abstracts*, volume 5, page n32. ERIC.

Wickens, C. D. (2014). Effort in human factors performance and decision making. *Human Factors*, 56(8):1329–1336.

Wickens, C. D. & Carswell, C. M. (1995). The proximity compatibility principle: Its psychological foundation and relevance to display design. *Human factors*, 37(3):473–494.

Wilson, J. R. & Rutherford, A. (1989). Mental models: Theory and application in human factors. *Human Factors*, 31(6):617–634.

Winograd, E. & Lynn, D. S. (1979). Role of contextual imagery in associative recall. *Memory & Cognition*, 7:29–34.

Wise, J. A.,Thomas, J. J.,Pennock, K.,Lantrip, D.,Pottier, M.,Schur, A., & Crow, V. (1995). Visualizing the non-visual: spatial analysis and interaction with information from text documents. *Proceedings of the Information Visualization Conference*, pages 51–58.

Wong, Y. Y. (1996). Temporal typography a proposal to enrich written expression. *Conference on Human Factors in Computing Systems - Proceedings*, pages 408–409.

Yang, F.,Qian, J.,Novotny, J.,Badre, D.,Jackson, C. D., & Laidlaw, D. H. (2020). A virtual reality memory palace variant aids knowledge retrieval from scholarly articles. *IEEE Transactions on Visualization and Computer Graphics*, 27(12):4359–4373.

Yeh, Y.-Y. & Silverstein, L. D. (1990). Limits of fusion and depth judgment in stereoscopic color displays. *Human factors*, 32(1):45–60.

Yonelinas, A. P. (2002). The nature of recollection and familiarity: A review of 30 years of research. *Journal of memory and language*, 46(3):441–517.

APPENDICES

STIMULUS MATERIALS

Homeostasis

Homeostasis refers to the body's ability to physiologically regulate its processes to ensure stability in response to fluctuations in external or internal conditions. The body maintains homeostasis by utilizing homeostatic control systems. Three components of the homeostatic system are the receptor, control center, and effector.

Receptor

It is the body structure that detects changes in a variable, which is a substance or process that is regulated. It typically consists of sensory neurons. These neurons may be in the skin, internal organs of the body, or particular organs such as the eye, ear, tongue, or nose.

Control Center

It is the operational point at which the signals are received and analyzed. It determines a course of action depending on the strength of the signal. Generally, a portion of the nervous system such as the brain or spinal cord, or an organ of the endocrine system like the thyroid gland is responsible for choosing an action.

Effector

It is the body structure where a response is generated that reverses the initial stimulus and thus attempts to maintain homeostasis. For example, sweat glands release sweat to lower body temperature. Most body structures can serve as one, although muscles and exocrine glands are often parts of it.

CONTROL OF BLOOD SUGAR

1. Glucose levels in the blood increase after eating.
2. Beta cells in the pancreas have glucose receptors that detect the increase.
3. Beta cells produce and release insulin into the blood when above the set point.
4. Insulin triggers liver, muscle, and fat tissue cells to absorb glucose.
5. As glucose is absorbed, blood glucose levels return to normal, and the beta cells stop releasing insulin.

True Labor

True labor is defined as uterine contractions that increase in intensity and regularity, and that result in changes to the cervix. Contractions typically are weak and irregular, but as levels of oxytocin continue to rise in the later stages, they become more intense and frequent. Labor has three physiological stages: the dilation, expulsion, and placental stage.

Dilation

This stage begins with the onset of regular uterine contractions and ends when the cervix is thinned and dilated to ten centimeters in diameter. This is the longest of the three stages and is marked by the greatest variability in duration. Those who have never given birth generally experience this stage longer, lasting from eight to twenty-four hours, than those who have given birth previously.

Expulsion

It begins with the complete dilation of the cervix and ends with the delivery of the fetus from the mother's body. This stage may last as little as several minutes but typically takes thirty minutes to several hours. The umbilical cord is clamped and tied off when the baby's body is fully expelled.

Placental Stage

After the baby is expelled, the uterus continues to contract. These contractions help compress uterine blood vessels and displace the remaining tissues from the uterine wall. It is typically completed within thirty minutes.

ROLE OF OXYTOCIN IN LABOR

1. Both mother and the baby secrete oxytocin.
2. Oxytocin stimulates the uterus to contract.
3. Uterine contractions cause the baby's head to push against the cervix, making the cervix stretch.
4. Dilating the cervix initiates nerve signals, which cause more oxytocin to be released.
5. The process escalates until the fetus is expelled from the uterus.