UC Irvine UC Irvine Electronic Theses and Dissertations

Title

Mixing Realities: Leveraging Augmented Reality to Support Remote Collaboration for Educational Scenarios, Share Digital Objects Across Distance, and Demonstrate the Physical Nature of Visualizations

Permalink https://escholarship.org/uc/item/3x41x8mc

Author Collins, Max

Publication Date

2021

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA, IRVINE

Mixing Realities: Leveraging Augmented Reality to Support Remote Collaboration for Educational Scenarios, Share Digital Objects Across Distance, and Demonstrate the Physical Nature of Visualizations

DISSERTATION

submitted in partial satisfaction of the requirements for the degree of

DOCTOR OF PHILOSOPHY

in Informatics

by

Max Collins

Thesis Committee: Dr. Kurt Squire, Chair Dr. Theresa Jean Tanenbaum Dr. Alan B. Craig

© 2021 Max Collins

DEDICATION

То

My parents

For always being in my corner

TABLE OF CONTENTS

Dedication		ii
Table of Contents		iii
List of Figures		v
List of Tables		vii
List of Acronyms		viii
Acknowledgements		ix
Curriculum Vitae		xi
Abstract of the Dissertation		XV
Chapter 1. Introduction		1
1.1. Augmented and Virtual Re	eality	4
1.2. Exploring the Relationship	p Between AR and VR	6
Chapter 2. Background and I	Related Work	
2.1. Interaction in XR		
2.1.1. AR Interaction Techn	iiques	
Object Manipulation		
Movement and Navigation		
2.1.2. VR Interaction Techn	iiques	
Object and Scene Manipulati	ion	
Movement and Navigation		
2.2. XR-mediated Collaboratio	on and Communication	
2.2.1. Communication and (Collaboration in AR	
Shared Space Collaborative	AR Applications	
Remote Collaborative AR Ap	oplications	
2.2.2. Shared Virtual Spaces	s and Communication in VR	
2.3. XR for Educational Purpos	ses	
2.3.1. Learning using AR		
Constructivism		
Parallel Distributed Process	ing/Connectionism	
2.3.2. Learning using VR		
2.4. Scientific Visualization		58
Chapter 3. Methodology		60
3.1. System Design & Develop	ment	
3.2. Architecture Design of AR	Applications	
3.3. Data Collection		74
3.3.1. Conditions Examined	1	74
3.3.2. Evaluation Methods		
Questionnaires		
Qualitative Data		

Log D	ata	81
Partic	ipants	81
3.3.3.	Procedure	82
Chapter 4	. Results and Conclusions	
4.1. Cor	idition Rankings	87
4.2. Rer	note AR-mediated Interaction Design Principles	89
4.2.1.	Exploiting Proprioception: Moving Oneself vs. Moving the Object	89
4.2.2.	Promoting User Engagement in Remote Collaborative Environments	92
4.2.3.	Promoting Co-presence in Remote Collaborative Environments	94
Spatia	ıl presence	97
Effect	s of AR	100
Feelir	igs of Connectedness	101
Share	d Objects	
Visua	l Representation of Remote Peer	107
Align	nent between Verbal and Visual Presented Information	109
Chapter 5	. Discussion	110
5.1. Rec	ommendations	110
5.1.1.	Shared (Networked) Objects	112
5.1.2.	Permissions	113
5.1.3.	Method of Manipulation	115
5.1.4. Visual representations of remote peers		
5.1.5. Context-appropriate fiducial markers		
5.1.6.	Customizations and Data Persistence	118
5.2. Lin	litations	119
5.3. Fut	ure Work	120
Reference	¹ S	123
Appendix	A. NASA TLX Questionnaire	138
Appendix	B. MEC Spatial Presence Questionnaire	139
Appendix	C. User Engagement Scale (short form)	141

LIST OF FIGURES

Figure 1 AR giv	es us the ability to pass objects through the screen from one device to another, though physically remote. This screenshot, taken directly from SAROSvis, shows a person in a Zoom call with a virtual background that SAROSvis recognizes as a fiducial marker. SAROSvis then places an AR object, in this case a molecular model, on top of the fiducial marker. Effectively, the person hosting this Zoom call is able to display desired models in AR to remote users of SAROSvis in the remote users' environments, and further interaction can take place from here
Figure 2 Reality	r-Virtuality Continuum (Milgram & Kishino, 1994)(Bambury, n.d.)
Figure 3 WIM in	nterface (Stoakley et al., 1995)28
Figure 4 Shared	l space AR experience34
Figure 6 Screen	shot from XRTB Zoom call where participants are observing a teacher demonstrating experiments and manipulations of the molecular model of ethanol (Sasikumar et al., 2021)62
Figure 7 A user	running SAROSvis on an iPhone observing a presentation through the lens of the application, seeing the rigged AR content attached to the presentation. The model displayed will change depending on the fiducial marker (target image) being displayed on the screen63
Figure 8 A scree	enshot from within the app of a molecular model placed on the target image as displayed on the user's computer screen. The computer screen is displaying a presentation that the remote participant of the teleconferencing call is sharing, thus they are able to choose what the person using the app sees in AR based on which slide they are sharing at a given time
Figure 9 Screen	shot from within the app showing the visual representation of a remote user; in this case, the visual representation is a pencil. The pencil mimics precise transformations and rotations of a remote user who is viewing the same model remotely
Figure 10 Scree	nshot from within the app showing the visual representation of a remote user; in this case, the visual representation is a pencil. The pencil mimics precise transformations and rotations of a remote user who is viewing the same model remotely
Figure 11 Stora	ge capacity vs. network bandwidth, and how SAROSvis compares to pdbARbrowser72
Figure 12 pdbA	Rbrowser Application Architecture72
Figure 13 SARO	Svis Application Architecture. SAROSvis connects clients to the Photon Unity Network (PUN) service so that the position and rotation data is updated in real-time across the network for each participant by tracking them according to their position/rotation relative to the target image. Participants can each scan individual copies of a target image in their own respective spaces and interact with remote partners spatially because of this networked solution

gure 14 In Condition 4, regardless of where the user points the camera, the model is rigged steady in the center of the screen. The camera is opened on even the non-AR conditions for consistency between all conditions76
gure 15 Slide from the presentation presented during data collection. A 2D image of a trigonal planar molecular model is shown
gure 16 Slide from the presentation presented during data collection. A 2D image of a tetrahedral molecular model is shown
gure 17 Slide from the presentation presented during data collection. A target image (stones) is displayed, and when participants view this slide through the lens of their phone or tablet with the application running, the desired 3D objects are displayed rigged to the target image
gure 18 Molecular geometry diagram (<i>Molecular Geometry</i> , n.d.)
gure 19 Condition rankings: participant preferences of each condition tested
gure 20 User Engagement Scale (UES) Results
gure 21 MEC Spatial Presence Questionnaire Results
gure 22 Feelings of Connectedness descriptive statistics chart
gure 23 Sense of Sharing the Same Object amongst Remote Collaborators. Mean scores for each Condition

LIST OF TABLES

Table 1 Primary Research Questions for this work	4
Table 2 Styles of Accomplishing AR (Sherman & Craig, 2018)	13
Table 3 Types of Interaction in AR (Sherman & Craig, 2018)	14
Table 4 Data storage locations with costs and benefits for AR applications	71
Table 5 Conditions used for evaluation Conditions used for evaluation	77
Table 6 MEC Spatial Presence Questionnairie Wilcoxon Signed Ranks Test Statistics	99
Table 7 Feelings of Connectedness Descriptive Statistics	103
Table 8 Feelings of Connectedness Wilcoxon Signed Ranks Test Results	104
Table 9 Sense of sharing an object descriptive statistics	106
Table 10 Feelings of Sharing an Object Wilcoxon Signed Ranks Test between condition significant differences	s with 107
Table 11 Considerations for designing remote AR collaborative experiences	112

LIST OF ACRONYMS

AR	Augmented Reality
AV	Augmented Virtuality
CG	Computer Graphics
EVL	Electronic Visualization Laboratory
HIT Lab	Human Interface Technology Lab
HMD	Head Mounted Display
HMPD	Head Mounted Projection Display
HUD	Head-Up Display
NASA TLX	National Aeronautics and Space Administration Task Load Index
NSF	National Science Foundation
PDB	Protein Data Bank
PDP	Parallel Distributed Processing
PUN	Photon Unity Networking
SAROSvis	Shared Augmented Reality Objects through Screens visualization
SCAPE	Stereoscopic Collaboration in Augmented and Projective Environments
SLAM	Simultaneous Localization and Mapping
ST	See Through
UCI	University of California, Irvine
UES	User Engagement Scale
VE	Virtual Environment
VMD	Visual Molecular Dynamics
VR	Virtual Reality
VW	Virtual World
WIM	Worlds in Miniature
WIMP	Windows, Icons, Menus, Pointer
WoW	Window-on-the-World
XR	"X" acts as a stand in for any of the letters in the initialisms AR, VR, MR, etc. before "Reality"

ACKNOWLEDGEMENTS

I am incredibly grateful to all who helped make this work possible, whether directly or indirectly. My committee, comprised of my advisor, Dr. Kurt Squire, Dr. Theresa Jean Tanenbaum, and Dr. Alan B. Craig, is a tremendous wealth of knowledge. I am extremely grateful for my advancement committee members as well (who joined in addition to those mentioned above): Dr. Constance Steinkuehler and Dr. Deanna Shemek. Both have had profound impacts on my journey at UCI, and I look forward to continued collaboration with them.

Dr. Squire has been the perfect match for me as an advisor. His genuine curiosity about, care for, and determination to improve the world around him are inspiring. I am so glad to have been taken under his academic wing. Dr. Tanenbaum fostered my curiosity in the space from the moment I arrived at UCI, and she gave me space to learn, play, and experiment with AR/VR concepts in some truly transformational classes—and also always made time to have nerdy, sci-fi discussions in her office. Dr. Craig is the reason that I am working in this space in the first place; he has mentored me from the day I stepped foot on a college campus. There are no words to describe the gratitude I have for the countless hours spent working together and learning from a true expert in this space. I am profoundly lucky to have had some of my favorite people sit on my committee.

I would like to thank Shodor for assisting with hosting solutions, particularly Aaron Weeden for being our point person throughout the years. I am also very appreciative of John E. Stone with the Theoretical and Computational Biophysics

ix

Group at University of Illinois at Urbana-Champaign for his tool VMD and the tutorials on how to use it. I also want to thank Dr. Lee Rettberg for his vast knowledge regarding structural biology.

I am thankful for all of the people in the Informatics department that have put up with me randomly dropping by offices to have long chats or just to say hello. I also look back fondly on the immeasurable amount of time spent playing games with folks in the lab—this was time *very* well spent. Lastly, I'd like to thank my family and friends who played a role in this academic journey whether they realized it or not. Thank you to everyone who played a part in this work, I look forward to the future of collaboration!

CURRICULUM VITAE

Max Collins

Website: https://maxacollins.com/

2021	Ph.D. Informatics University of California, Irvine
2020	MS Informatics University of California, Irvine
2017	BS Psychology, Minor in Informatics University of Illinois Urbana-
Champaign	

Publications

1. Collins, M. & Craig, A. (2016). VisMo: Augmented Reality Visualization of Scientific Data and Molecular Structures. JOCSE: Journal of Computational Science Education, 8(1), 12-15. DOI: 10.22369/issn.2153-4136/8/1/3.

 Sasikumar, P., Collins, M., Bai, H., & Billinghurst, M. (2021). XRTB: A Cross Reality Teleconference Bridge to incorporate 3D interactivity to 2D Teleconferencing.
 Extended Abstracts of the 2021 CHI Conference on Human Factors in Computing Systems, 1–4. https://doi.org/10.1145/3411763.3451546

3. Persa, A., Anderson, C. G., Martinez, R., Collins, M., Anderson-Coto, M. J., & Squire, K. 17. Improving Wearable Mindfulness App Through Participatory Design.

Talks

- SIGGRAPH 2021, Immersive Visualization for Science, Research, Art and Education | Mixing Realities: Leveraging Augmented Reality to Support Remote Collaboration in Educational Scenarios and Demonstrate the Physical Nature of Digital Information – August 2021
- XSEDE (The Extreme Science and Engineering Discovery Environment) ECSS (Extended Collaborative Support Services) Symposium | Leveraging Augmented Reality to Enhance Remote Collaboration – April 2021
- XSEDE (The Extreme Science and Engineering Discovery Environment) ECSS (Extended Collaborative Support Services) Symposium | The Development of a Mobile Augmented Reality Application for Visualizing the Protein Data Bank – June 2019
- Academy High School Urbana-Champaign | Presentation on current AR prototyping tools to high school students – February 2020
- Academy High School Urbana-Champaign | Presentation on SparkAR face filter development to high school students – December 2019
- Ingenuity, UCI, Annual showcase event featuring the top student innovations from both the Donald Bren School of Information and Computer Sciences and The Henry Samueli School of Engineering | CrewView AR Theater Set Design — Spring 2018

- Carnegie Mellon University Articulab | Lab Research Presentation regarding findings from Summer research on interactive gameplay with virtual agents — Summer, 2017
- Blue Waters Institute 2017 | Spoke about work on AR visualization and the research around pdbARbrowser (formerly VisMo, VisMe) to the rest of the cohort and affiliates Summer, 2017
- NYU Diversity Summer Student Research Conference | Presented on research on police/public relationships to the rest of the cohort and affiliated faculty and researchers – Summer, 2016
- Associated Colleges of the Chicago Area (ACCA) | Talk regarding scientific visualization work and research around pdbARbrowser (formerly VisMo, VisMe)
 Fall, 2015
- National Conference on Peer Tutoring in Writing | Presentation regarding peer tutoring in the Niles West High School Literacy Center Fall, 2012

Select Research/Work Experience

- Facebook Reality Labs | Research Intern, Human Computer Interaction -Interaction Techniques (PhD)—Fall 2021
- Sony Interactive Entertainment (PlayStation) | Magic Lab Intern Summer 2021
- Empathic Computing Lab | Empathic Computing Lab VIP Intern Fall 2020

- Facebook | UX Researcher Intern Summer 2020
- Facebook Reality Labs (Oculus) | UX Researcher Intern Summer 2019
- Carnegie Mellon University, HCII, Articulab | NSF DREU Intern Summer 2017
- New York University, Social Perception, Action, & Motivation Lab | NSF REU Intern – Summer 2016
- Shodor, Blue Waters Supercomputer, NCSA | Blue Waters Student Intern 2015-16
- University of Illinois at Urbana-Champaign Vision Lab | Research Assistant 2015-17
- University of Illinois at Urbana-Champaign CoLearn Lab | Research Assistant 2016-17

ABSTRACT OF THE DISSERTATION

Mixing Realities: Leveraging Augmented Reality to Support Remote Collaboration for Educational Scenarios, Share Digital Objects Across Distance, and Demonstrate

the Physical Nature of Visualizations

Max Collins Doctor of Philosophy in Informatics University of California, Irvine, 2021 Dr. Kurt Squire, Chair Dr. Theresa Jean Tanenbaum Dr. Alan B. Craig

Augmented Reality (AR) is a medium that gives people the ability to engage with digital information in ways that deviate from more traditional HCI methods (e.g. WIMP user interfaces (Christensson, 2014)). Previous work has shown that AR has useful implications for collaboration, and with remote work experiences and teleconferencing becoming increasingly prevalent and important to people as many are working together over distance (remotely) in higher frequencies, understanding how AR might impact remote work is of particular interest. In this work, I review literature around methods for remote collaboration and using AR (and Virtual Reality (VR)) for collaboration, investigate ways that AR can support efforts to work together across distance, and study how invoking AR may create a sense of joint focus and engagement beyond what traditional remote collaboration tools afford. I explore the design process and development of an AR teleconferencing tool,

xν

SAROSvis, that allows participants to interact with one another around digital assets in AR and pass digital objects through the screen to one another while physically remote. I investigate the use cases of this tool by comparing across curated conditions and evaluating versions of this tool in user studies; I also describe the evaluation methods and user testing results of this system. This work aims to develop a research-based resource for those looking to utilize AR tools for collaboration around 3D entities or environments, with an understanding of the requirements, desires, and limitations of users and scenarios. I propose that there are meaningful ways that people can remotely collaborate in educational spaces using AR to pass (digital) objects to one another through the screen and using AR as a layer on top of the features that existing teleconferencing tools provide can make for more spatially appropriate, collaborative, engaging communication sessions while working remotely.

Chapter 1.

Introduction

Very shortly, the biggest limitation will be our imagination.



-Dr. Alan B. Craig, Understanding Augmented Reality (2013)

Figure 1 AR gives us the ability to pass objects through the screen from one device to another, though physically remote. This screenshot, taken directly from SAROSvis, shows a person in a Zoom call with a virtual background that SAROSvis recognizes as a fiducial marker. SAROSvis then places an AR object, in this case a molecular model, on top of the fiducial marker. Effectively, the person hosting this Zoom call is able to display desired models in AR to remote users of SAROSvis in the remote users' environments, and further interaction can take place from here. In the beginning of 2020, many people around the world experienced a shift to remote work which was (and continues to be) heavily mediated by computers and other technology to facilitate collaboration with peers over distance due to the global pandemic which largely broke out in early 2020. This of course has had an impact on the ways that education can be facilitated, and the means by which educators (and students) approached teaching, learning and interacting with one another. Many people began using videoconferencing tools, Zoom being one of the most popular, to communicate and collaborate with one another (*Video Conferencing, Web Conferencing, Webinars, Screen Sharing - Zoom*, n.d.). Working together in this remote, or virtual, model of instruction was quickly adopted at unprecedented scale, and a lot of uncertainty around the methods for instruction gave rise to exploring new methods for this unorthodox model of collaboration.

Attempting to conform to the traditional methods of communication and instruction that people are used to under normal circumstances was bound to produce some friction under purely remote circumstances, as the situation is markedly different from traditional classroom and school models. Instead, it's imperative that we consider how we can best exploit the affordances that remote communication tools provide when working, learning, and teaching over distance rather than trying to mimic the traditional classroom. This way, we can explore new ways of interacting with one another, and new ways of interacting with the data and information around which we are learning and teaching.

People who typically work with physical entities or 3D environments cannot always easily transition into the videoconferencing workflow to accomplish all of their goals; the transition to remote work has been lossy for many. For example, a chemistry teacher who usually teaches students about the structure of ethanol by building a ball-and-stick model with a molecular model toy set and passing it around the room cannot do this remotely. An art teacher who inspects students' sculptures by walking around and examining their shape and detail within an art classroom cannot do this remotely. The impetus of the work in this dissertation is this barrier that many people now face given certain types of remote work that rely on 3D entities or environments. There are opportunities and novel use cases for existing teleconferencing tools (and exploiting their built-in features and capabilities) that can potentially enhance the experience of working together around 3D material while physically remote.

I propose that Augmented Reality (AR) has immense implications for remote education, specifically when we consider the ways in which AR allows for data visualization, the benefits of exploiting the sense of proprioception (the understanding of how the body is situated and moving in space relative to itself and the surroundings) for educational tasks, the ability to share digital objects between people in interactive ways, and the potential effects on user engagement regarding subject material. In this work, I provide background on AR (and how it relates to Virtual Reality (VR)) and its place in interactive communication and education by examining the relevant literature and work in this space. I then describe the design and development of SAROSvis (Shared Augmented Reality Objects through Screens

visualization), an application made to supplement video calls between remote participants leveraging AR. I outline the methodology used in this dissertation which produced data around the use of AR over distance in educational scenarios, as well as the results from this study and decipher what they could mean, infer future direction based on these results, and posit implications of this work and the possibilities for the future of remote collaboration leveraging AR.

The primary research questions that I focused on in this project are outlined in Table 1. In the next few subsections, I will cover background on AR, VR, and the relationship between the two.

Research question	Method for answering
Does demonstration leveraging AR improve feelings of co-	Comparison between
presence between remote pairs (when compared to	conditions for user testing,
traditional teleconferencing)?	targeted interview questions
Does locus of control over the digital objects/perspective	UES, Comparison between
control make differences in the experience? Should	conditions for user testing,
participants be able to see who is controlling a given object?	targeted interview questions
Can we create a sense of sharing the same object and spatial presence around an object and a remote collaborator using AR? Does this create more workload for users (and if so, is it worth it?)	MEC Spatial Presence Questionnaire, NASA TLX, targeted interview questions

 Table 1 Primary Research Questions for this work

1.1. Augmented and Virtual Reality

To maintain clarity throughout this document, it is important to establish

and maintain a common vocabulary with regard to definitions of AR and VR. Firstly,

I refer to them as forms of media rather than technologies because AR and VR are

methods of consuming information; this is sticking closely to the definition of AR

and VR maintained by Craig (Craig, 2013). Devices that allow for AR and VR consumption are the technologies in this scenario. This is akin to broadcast media; things like TV and radio are forms of broadcast media, and the technologies that allow us to consume them are TVs, phones, and radios. Similarly, there are various technologies that allow users to consume AR and VR content: mobile phones/tablets, head mounted displays (HMDs), CAVES, and HMD glasses. AR and VR are the media that allow users to consume information, and as with other media such as video, audio, and text, there is information that is best consumed with AR and VR and information that is not best consumed using other forms of media.

AR is "a medium in which digital information is overlaid on the physical world that is in both spatial and temporal registration with the physical world and that is interactive in real time" (Craig, 2013). AR is different from VR in that VR provides users with purely digital information shutting out the physical world, whereas AR uses the physical world to place the digital information; AR involves overlaying digital information onto the physical world, and VR involves cutting out the physical world and the user receiving input only from a VR device. In VR, users are fully immersed in the virtual world where all input is digital. An important note about AR and the relationship between the real and digital worlds is that "the virtual overlay [must] be aligned to the real world onto which it is mapped. This is referred to as registration," (Sherman & Craig, 2003). Exploring instances where AR and/or VR can enhance a given scenario rather than adding an unnecessary element and understanding the use cases and interaction techniques for these media will provide a solid foundation of knowledge that will help forecast into the future of information

consumption and digital interaction—interaction with the information, and interaction with one another.

When humans are interacting with virtual environments in VR and virtual objects in the AR, the concept of presence becomes blurry. Sheridan studied presence and how it relates to virtual objects and virtual worlds and came up with a definition that applies well to this context. "Virtual presence" is the "sense of being physically present with visual, auditory, or force displays generated by a computer" (Sheridan, 1992). This definition is applicable because compelling experience will successfully give users a sense of presence, and this is important so that users will overcome the initial novelty of the media and reap the benefits of using the media as a tool.

In the next subsections, more concepts around AR and VR are explored, as well as the similarities and differences between them. In Chapter 2, I will expound upon the descriptions of AR and VR, interactions afforded within them, and their uses for educational purposes.

1.2. Exploring the Relationship Between AR and VR

There are undoubtedly differences in use cases between AR and VR, especially given the current state of each medium, however as the technologies supporting them become more sophisticated it is likely that they may begin to converge as Michael Abrash says and develop a "hand-off" relationship (Abrash, 2017). As AR use becomes more ubiquitous, our perception of digital objects within

the real world will become vastly indistinguishable from perceiving real ones. As Craig points out, living in a world where we accept digital objects as real may not be as far off as we think; people can listen to the radio and feel that they're hearing a band play music, or watch a movie and believe they are seeing a shark that is terrorizing passengers of a boat. The important thing here is that people are accepting computer generated digital creations as though they were real.

Milgram's "reality-virtuality continuum" situates and defines the various realities with which we interact, and describes Mixed Reality (MR) as "a blanket term to describe any experience between the extremes of the continuum," (Milgram & Kishino, 1994). The continuum is broken down into six categories, and both AR and VR are of course found along the continuum as they exist along a spectrum from actual reality to complete virtual reality. All that exists within this spectrum is considered types of MR. Milgram speaks of several classes of MR that are noteworthy and distinguishable from one another (Milgram & Kishino, 1994):

- 1. Monitor-based (non-immersive) AR displays, upon which computer graphic (CG) images are overlaid.
- Same as 1, but using immersive HMD-based displays, rather than WoW monitors.
- 3. HMD-based AR systems, incorporating optical see-through (ST).
- 4. HMD-based AR systems, incorporating video ST.

- Monitor-based AV systems, with CG world substratum, employing superimposed video reality.
- Immersive or partially immersive (e.g. large screen display) AV systems, with CG substratum, employing superimposed video or texture mapped reality.
- Partially immersive AV systems, which allow additional real-object interactions, such as 'reaching in' and 'grabbing' with one's own (real) hand.

As there is copious terminology and variations of terms in this space, it is beneficial to consider Milgram's continuum to provide clarity regarding definitions and establish a common set of definitions for this lexicon.



Reality-Virtuality Continuum Created by Paul Milgram (1994)

Figure 2 Reality-Virtuality Continuum (Milgram & Kishino, 1994)(Bambury, n.d.)

As there is ample jargon used when discussing these topics, and the field is relatively young and heterogeneous, it can quickly become daunting to understand what people are trying to communicate, and to give appropriate labels to applications. While the definitions can be tedious to iron out, it is worthwhile to make sure that those working in this space are all subscribed to similar vocabularies because the absence of consistent terminology impedes meaningful communication.

A term gaining popularity is the all-encompassing "XR," which describes all experiences that lie along this MR continuum when speaking generally about any virtually enhanced experience (*What Is AR, VR, MR, XR, 360?*, n.d.). The "X" acts as a stand in for any of the letters in the initialisms AR, VR, MR, etc. At times for the sake of brevity in this paper, when referring to AR and VR, I may use the term XR; of course, when necessary, I will specify which type of experience I am referencing. These definitions and terms make it possible to navigate the topics covered throughout the remainder of this document with a common vocabulary. In Chapter 2, I will provide a more detailed look into these concepts. Though SAROSvis uses AR visualization, it is worthwhile to understand the background of both AR and VR interactions and both AR and VR educational experiences; it is not uncommon for there to be crossover and similarities in the way that interactions are built and experiences are designed in these spaces.

Chapter 2.

Background and Related Work

2.1. Interaction in XR

2.1.1. AR Interaction Techniques

AR is situated along the reality-virtuality continuum closer to true physical reality than VR, and involves digital content enhancing the real world by means of registering virtual objects to physical-world environments (Milgram et al., 1995). In order to design for the future of AR, it is essential to understand the experiences, applications, and research that currently exist regarding interacting in AR and evaluating these systems.

AR drives interactions in some of the most popular applications used today: Snapchat filters from Snap Inc.; 3D photos and filters on Instagram and Facebook; Facetime and Memoji from Apple; and more. These popular experiences rely on the fact that smartphones are becoming ubiquitous at this point in time with a growing userbase (*Number of Smartphone Users Worldwide 2014-2020*, 2019), and built-in front-facing and rear-facing cameras on these devices allow for relatively cheap AR experiences to be widely tested and adopted.

There are different ways to engage with AR and VR experiences, and as technology to support the media matures, the ways in which we engage with experiences will transform. Craig describes the basic paradigms of AR interaction

into two main categories: subjective and objective views. Subjective view AR involves users viewing the scene through their own perspective in first person. The magic lens is a type of subjective view AR. The magic lens involves devices with cameras that broadcast in real time the view that the user is capturing using the camera. This image is displayed on the screen for the user, along with digital content that is overlaid onto the physical world camera view. This provides an experience where users perceive the physical world augmented with digital objects through the lens of their device. This paradigm is convenient as many people have mobile devices that have cameras. This can also be implemented using AR glasses or seethrough HMDs.

Objective view AR entails the users seeing themselves in the third person and the environment being augmented. The magic mirror is a type of objective view AR, where the user sees themselves being augmented (Craig, 2013). Objective view AR is similar to a security camera in a store: people can see themselves broadcast to a live video feed on a display that is presented to them. With objective view AR, there is digital information overlaid that one sees on the live feed as well.

Head worn displays involve users wearing a head-mounted display or glasses that let them continue seeing the real world while showing the user a digital information overlay simultaneously. Note that this is different than a VR HMD in that users of a VR HMD do not have their real-world surroundings visible to them. One other important type of AR display is projected AR environments, which involve dynamic and interactive projections (as in light projections of imagery) of digital

content onto the real world (such is the case with IllumiRoom (B. R. Jones et al., 2013)).

Styles of Accomplishing AR	Description
Subjective View	Users see the augmented scene directly from their point of view (e.g. magic lens).
Objective View	Users see themselves in the augmented scene (e.g. magic mirror).

Table 2 Styles of Accomplishing AR (Sherman & Craig, 2018)

Interactions in virtual worlds can be put into three basic categories: manipulation, navigation, and communication (Sherman & Craig, 2018). These categories will be the focus of this section on interaction.

Types of Interaction in AR	Description
Manipulation	Users modify the virtual world and objects.
Navigation	Users navigate through the environment.
Communication	Users communicate with other users or agents in the environment.

Table 3 Types of Interaction in AR (Sherman & Craig, 2018)Object Manipulation

Manipulation of digital objects and environmental elements is the most challenging aspect of AR, since users are in the physical world interacting with physical objects as well as digital objects. For manipulation in AR to be successful, users must be able to select virtual objects and then take action on them (Craig, 2013). With AR, there are a rather unique set of affordances for interacting in space. Users are interacting with both the real and virtual worlds, so interactions can, and should, take advantage of elements of both worlds. Solely focusing on virtual *or* real objects would be a missed opportunity and wouldn't take full advantage of what AR has to offer, and arguably wouldn't be true AR; recall that AR involves digital object registration with the real world. A review of work from the International Symposium of Mixed and Augmented Reality (ISMAR) showed that there were three main categories of interaction techniques when it comes to user interfaces: tangible AR, collaborative AR, and hybrid AR interfaces (Feng Zhou et al., 2008). These categories present in ISMAR can be situated within Craig's interaction categories for virtual worlds from Table 2. Tangible AR "bridges the real and virtual world" and lets users manipulate digital objects by manipulating physical ones in the real world. Collaborative AR enhances both face-to-face and remote communication and interaction between multiple users. Hybrid AR is an all-encompassing term for multi-modal experiences that can include various media platforms contributing to an experience. Zhou et al. also remarks on prevalent display mechanisms, which they describe as either seethrough HMDs, projection-based displays, and handheld displays. From this review of work in the field as well as Craig's higher-level descriptions, there exists a working set of interaction categories and display categories.

Oftentimes interactions in AR tend to "mimic real-world interactions," even though users are potentially afforded vastly different interaction techniques than what is possible in the physical world alone, failing to realize the potential of AR (Craig, 2013). Why constrain users to the interaction techniques that are possible in the physical world when there is much more to explore and certain situational advantages of AR? Craig offers the example of summoning objects, and how typically even with virtual objects, users are required to move to the position of an object to retrieve it. This is interesting because there is no reason users should have to move to the location of an object when it could be designed that users may magically

summon objects (or some other interaction beyond moving the body through space). We will likely see new and experimental ways of accomplishing tasks with virtual objects that are native and exclusive to AR as the medium becomes more prevalent.

Craig forecasts that in the future, people will become so accustomed to interactions techniques that AR affords that they will be unsatisfied and possibly even forget that certain interactions are only possible with AR when compared to the bare physical world. Becoming accustomed to interactions within a medium is not uncommon, for example pinching to zoom on touch screen surfaces, and clicking/highlighting words to look up definitions and navigate hyperlinks on computers (Craig, 2013). If someone is accustomed to reading documents on a computer and being able to highlight and have words defined quickly, reading a paper book where this isn't possible may result in them becoming bothered. Likewise, if an AR application provides someone with the ability to see a colleague's name when they might have forgotten this information, when forced to rely solely on their own memory they may become bothered when they don't have the AR overlay available.

Craig continues to give methods for manipulation in AR space such as agent control, where users speak with a virtual assistant to carry out tasks, or aperture selection, which involves users selecting object by viewing the object between their thumb and index finger and then taking action on that object once it is selected (Craig, 2013). These manipulation methods are good for magic lens AR, specifically

when there is no tangible display to touch and interacting with virtual objects must be accomplished without relying on a screen to tap.

SCAPE (Stereoscopic Collaboration in Augmented and Projective Environments) is a platform that utilizes physical components to enhance the virtual reality experience (Brown et al., 2003). This subjective AR experience involves physical objects that are augmented with digital information. SCAPE uses a Head-Mounted Projective Display (HMPD) that allows users to see digital objects and information projected onto physical objects. This AR experience was created to test various interactions by means of physical objects and widgets used within test applications. This is a collaborative experience which involves multiple users who simultaneously receive unique perspectives. To use SCAPE, users are placed into a special environment with is a "workbench" containing physical objects (tracked with camera vision).

The three main interface techniques that the researchers built with SCAPE were the CoCylinder widget, the Magnifier widget, and the CoCube widget. The CoCylinder features a physical cylinder that, when viewed with the HMPD, would ensconce the virtual object that users could move around and perceive from various angels. The Magnifier allows users to magnify digital objects on the workbench, and this also acts as a tool to help users see "underneath" physical objects and markers in the room that my otherwise be occluding objects. The magnifier "implicitly permits the display of graphical avatar "overlays" for physical objects on the workbench having corresponding virtual representations," and this has interesting

implications for how we might link physical and digital objects and dealing with occlusion. The CoCube widget is the most robust, and allows users to select, inspect, and learn about objects with the augmented overlay. The CoCube is a physical cube that users hold up to virtual objects to "capture" them in a virtually transparent box where they can inspect a miniature of the object and furthermore investigate its book of information. Augmenting tangible objects with digital information and transforming them into tools with specific purposes in context greatly expands possibilities of the ways with which we interact with the virtual world.

If an AR experience is using a phone or another tangible device with a screen as the magic lens, users have the option take actions by touching the physical screen to manipulate virtual objects. pdbARbrowser (previously VisMo and VisMe), a molecular visualization tool for smartphones and tablets, gives users the capability to manipulate protein structures in AR (Collins & Craig, 2017a). Being on a tangible device, the interactions that users take to interact with the digital content often takes the form of on-screen actions. There are control buttons on-screen for resizing and moving digital objects and touching to drag on the screen moves objects around. This would not be possible with AR glasses because there would be no tangible interface to accept touches, but this is one way to handle interactions. pdbARbrowser also offers virtual buttons, which can be interacted with using AR glasses. Virtual buttons are digital buttons that are visible through the magic lens and to "click" them, users place their hand onto the area where the button appears. The computer vision algorithm recognizes this action as a touch and thus completes an action.
This leads into other types of computer vision enabled gesture-based interaction techniques. If an experience is built for AR glasses, then it may require hand gestures as input to interact with the virtual world. Piumsomboon et al. tested an AR experience with various gestures that allowed users to control virtual objects and menus (Piumsomboon et al., 2013). They found that this is a viable way to interact with the virtual world "in the air," but it's important to give users some sort of feedback that their actions are taking place, whether that be haptic, visual, or audio feedback. Radkowski and Strizke also tested gesture-based interaction in AR, and found that it was effective for the virtual assembly task they tested with users (Radkowski & Stritzke, 2012).

Burchmann et al. experimented with gestures and computer vision in their FingARtips project in which they gave users gloves that allowed users to interact with virtual objects (Buchmann et al., 2004). Other researchers are developing tools to augment physical touch and sensations from digital objects, which has great implications for sensory feedback in AR (Nojima et al., 2002)(Kawazoe et al., 2019). With compelling enough visual AR and a thorough haptic system, the difference between real physical objects and digital objects will become less distinct in time.

Two well-funded AR glasses projects are the Hololens and Magic Leap (Evans et al., 2017)(*Magic Leap*, 2020). The Magic Leap requires a handheld controller to interact with the virtual objects overlaid onto the physical world, and the Hololens makes use of hand gestures. There are many different ways to interact with the

virtual objects when using AR, and the design of an application is largely dependent on the method of display and in turn what controllers will be available.

There are a multitude of applications for AR, and many different ways that designers can choose to have users interact with systems. The points that interaction can take place is an important aspect that determines the kinds of interactions that are possible in AR. Experiences are either subjective or objective AR, and may or may not involve tangible interface displays that users view the augmented world through (such as a smartphone screen) that allow for input. For now, many experiences involve devices that have tangible, interactive interfaces, so interacting with the virtual world relies on metaphors that users are already used to (clicking, touching, dragging, etc.). This will not always be the case, as not all AR interfaces will have tangible input displays or mechanisms; AR glasses worn by the user won't afford users the ability to tap on a screen. Hand-tracking, voice recognition, other bodily cues, and handheld controllers are experimented with as methods for controlling the digital aspects of an AR scene when a tangible display is not available. Semararo et al. experimented with ways to use fiducial markers to toggle objects, move objects from one fiducial marker to another when in close enough proximity, and other interaction techniques around fiducial markers and their physical proximity (Semararo et al., 2011). Projects like this can help us shift the ways that we think about interactions with more traditional media by experimenting with the affordances of AR.

Movement and Navigation

When interacting in AR, the environment users are operating in is the real world and thus movement is handled by actually moving through the physical world as one would do naturally, though there are augmentations that can be made to aid and assist navigation. Visual and audio cues can assist users by augmenting the real world with navigation information overlaid onto the physical world.

Companies like WayRay and Phiar are developing AR overlays for windshields in cars to assist drivers with navigation and other relevant vehicle information (*WayRay*, 2019)(*Phiar - Artifical Intelligence to Guide You Anywhere*, 2019). These navigation tools are indeed examples of magic lens, but the lens here is the windshield of the car. Rather than using a dashboard to check speed, or looking to a smartphone for directions, AR places digital information that a driver may need strategically (and translucently) onto the windshield.

There are also examples of AR navigation beyond driving assistant overlays. Supplying people on foot with visual cues, GPS and directional information, and directions can be beneficial (Thomas et al., 1998). AR HMD or glasses can provide users with digital visual information as the move through the real world. Reitmayr and Schmalstieg built a collaborative AR navigation system to help people in urban areas (Reitmayr & Schmalstieg, 2004). Their system is interactive and offers users an AR guide to destinations that is responsive to their actions in real time. It is also collaborative, and users can see others who have joined the session and choose to follow, guide, or meet another user (the system calculates a meeting point halfway

between the participants). In addition to these interactions, users can see locationbased information and add annotations to their environment. People using this collaborative navigation and information browser system wear an AR HMD and hold a controller to make selections in the application. This app involves AR interactive objects in the scene registered to the real world, as well as a head-up display (HUD) for information such as compass, current location, and target location.

A head-worn HUD is rigged to the user's sightline, not within the environment, so when a user turns their head or moves, all elements within the HUD remain in the same position with respect to the user as before. To be clear, when I say "rigged," I mean that the AR content is placed in spatial and temporal registration with the world/fiducial marker. AR enables people to have interactive navigation tools that melt into regular activity more seamlessly than current navigation tools. Movement in AR is handled by moving in the physical world, but of course for the AR to work the user must be trackable by the application to register the digital objects and information correctly with the physical world.

2.1.2. VR Interaction Techniques

Interaction is described as "a mutual influence of one thing on another," (Craig, 2013). In VR, there are countless methods for interaction in theory; interactions are only limited by the imagination of the designer and input controller methods of the system. When VR is used to interact, the virtual spaces inhabited are entirely synthesized, thus there are methods of interacting that may be drastically different from real world interactions and the physical limitations by which we are

bound. Many VR interfaces currently include a HMD and a controller held in each hand, and commonly experiences give end-users the familiar experience of using two hands via the handheld controllers, and looking around as a human, or humanlike avatar would. There is definitely a time and place for this as we are, in fact, humans operating in a space; however, there is still plenty of work to be done that involves researchers and designers thinking of how we might interact beyond our human bodies. Exploring previously experimented interaction techniques reveals a lot about how there are many techniques that are incredibly useful and will likely gain traction as VR hardware becomes more robust and compelling, and also that there are countless methods for interaction that will be uncovered as more usecases for VR arise.

In this section I look at various interaction techniques in VR, and some of the ways that researchers have creatively accomplished specific goals within VR space. This will help assess the state of the literature in this sector of XR work and provide ideas for how certain techniques and designs may be applied, combined, and iterated on in future work to achieve desired outcomes. Recall that Craig spoke about the three main aspects of interaction in virtual worlds being manipulation, navigation, and communication; these categories will be the overarching guidelines for this section.

Object and Scene Manipulation

Manipulating the virtual objects and world in VR is a cornerstone of an immersive VR experience, and environments that don't respond at all to a user's

movements or actions are arguably not even VR. Interactions in VR include the various ways that the system responds to user inputs and actions. Designing these interactions thoughtfully is crucial as they will dictate how a user understands the virtual environment and their affordances within it. User interface (UI) design and a particular UI's affordances are so important that "the UI is how we generally envision the computer itself," (Sherman & Craig, 2018).

Sherman and Craig describe four methods of manipulation in VR; three from Mark Mine (Mine, 1995), and one that they add:

- Direct user control: interface gestures that mimic real world interaction
- 2. Physical control: devices that user can physically touch
- 3. Virtual control: devices the user can virtually touch
- Agent control: commands to an entity in the virtual world (Sherman & Craig, 2003)

These manipulation methods are the underlying components to much larger, complex systems and at times they may work in tandem (Ishii & Ullmer, 1997); for example, if someone were to play a virtual shooting game and the controller is a toy gun that they physically interact with while interacting with the virtual model of the gun, they are using physical and virtual control; this is known as a prop (Hinckley et al., 1994). These basic principles can be seen in the interaction designs explored in this section and are good for categorizing the type of experience at hand.

Bimanual interfaces, or methods of interacting with objects with both hands and affecting multiple dimensions of the object (grabbing a cube with both hands and resizing and rotating it), is becoming more common in VR applications (Sherman & Craig, 2018). Controlling the virtual objects and environment in VR can occur in many different ways, but there are seven places where virtual controls can exist for users to interact with (Sherman & Craig, 2018):

- 1. In the world
- 2. In the hand
- 3. In front of the view (head-up)
- 4. Surrounding the user
- 5. On the display
- 6. Through an aperture
- 7. On a panel of interface objects

In a paper about scientific visualization, Sherman and colleagues talk about implementations of VR in scientific visualization work, and weighed the utility of simulating experiments and visualizing data in a CAVE (cave audio-visual experience automatic virtual environment) (Sherman et al., 2004)(Cruz-Neira et al., 1992). Cruz-Neira et al. describe the CAVE as a virtual reality experience contained within a cube, where the participant remains inside the cube and variable projections of the virtual environment are displayed on the walls and change in realtime as the user is tracked and interacts within the environment (Cruz-Neira et al., 1992). The CAVE differs from an HMD in that there is no need for the user to wear gear on their head to shut out the physical world, and the less intrusive nature of the CAVE is an important selling point to many designers; users may need to wear stereo glasses and some device to enable positional tracking.

Two big takeaways from their work are the necessity for VR system to overcome the various immersion issues and the visualization issues that they outline. The immersion issues are field of view, panorama, perspective, body representation, and intrusion. The visualization issues are visual acuity, linearity, look around, progressive refinement, and collaboration. They believe that the CAVE can handle many of these issues uniquely well, and certainly CAVE simulations give way to explore more embodied experiences in VR.

As with all tools, there is a time and place where VR visualization of scientific data may be useful, and instances where 2D visualization models are sufficient. Some of the concerns that Sherman et al. mentioned with regard to VR for scientists who would use it as a simulation tool are fatigue after using an interactive system, utility of being immersed in a visualization, and accuracy. Even with these concerns, as VR continues to advance, scientists are experimenting with using VR to interact with their research in new and innovative ways. "Stereoscopic imagery, wide fields of view, and tracking of the viewer's head enable the scientist to become immersed within the simulation and become a part of the system," and the CAVE has been integral to "transforming virtual reality from a novelty to a useful tool for analysis and display," (Sherman et al., 2004). Much of their work comes from their experiences working with a wide variety of scientists, and Sherman et al. see a

future where VR is an important element of the tool belt of scientists and researchers.

One approach to giving users the power to quickly manipulate the environment in which they are interacting is called WIM, or Worlds in Miniature (Stoakley et al., 1995). This technique involves users manipulating a miniature copy of the virtual world they are in and allows users to make changes to the miniature copy (such as moving and adding objects) that propagate into the full-scale virtual environment. Stoakley, Conway, and Pausch recognize that many VR experiences don't take advantage of the fact that real world constraints don't necessarily need to apply to virtual environments. "We have grown accustomed to these real-world constraints: things we cannot reach, things hidden from view, things beyond our sight and behind us, and things which appear close to each other because they line up along our current line of sight. Our virtual environments should address these constraints and with respect to these issues be "better" than the real world," (Stoaklev et al., 1995). Users can even use the miniature model to move themselves within their environment by moving the avatar that exists in the model, copy objects from one scene to another (think of holding two dollhouses and moving objects between them), and move between scenes themselves.

Stoakley et al. give the example of hanging a picture to illustrate carrying out an action within a system using WIM. This is used as an example because it is typically a two-person task "in which the proximity required to manipulate an object interferes with the desire to see those effects in a larger context. With a WIM,

a single user can stand at a comfortable distance to view the picture in context, while at the same time reaching into the WIM to manipulate it," (Stoakley et al., 1995). The WIM allows users to see objects from various angles to solve interaction issues such as occlusion or range. Stoakley et al. state that having a handheld scale model of the environment provides information with "less cognitive burden" than a system where users are physically flying around in an environment. In a WIM environment, object selection is rather simple. Users can select and manipulate objects in their environment by selecting and manipulating the objects' proxies in the miniature model they are holding.



Figure 3 WIM interface (Stoakley et al., 1995)

WIM is a great tool for VR environments that involve users needing to understand a whole space and move objects that might otherwise be out of reach. An interesting phenomenon in some VR experiences is that it closely mimics the real world; users are required to get up and move to objects in order to interact with them. In some cases, this is a good thing, however there are users that prefer (and work more efficiently) with controllers that give them the power to manipulate desired objects while remaining stationary themselves (Lages & Bowman, 2018). A combination of physically moving around a space in VR (a user walking around, tracked, in the real world to move their avatar and camera view) and using a WIM model to interact with objects that would be much harder to reach is a good compromise that is more powerful than just one or the other. It's important to keep in mind that these interaction techniques are not mutually exclusive, and using ideas from various methods might be the perfect concoction for a particular experience.

A very compelling example of manipulating a VR environment with strong social implications is the Photoportals project. Photoportals use "photography as a unifying metaphor for reference-based interaction techniques," (Kunert et al., 2014). Using this interaction technique, pictures use "visual representations of places, objects and moments in time and to exchange these with others," using photography to allow "the appropriation of inaccessible objects and locations for review and presentations," (Kunert et al., 2014). Users capture photos within an environment to travel to specific spaces by entering the photo, examine the environment by manipulating the photo, move object between scenes, and much more.

Kunert et al. outlined some of the design challenges commonly encountered while developing collaborative systems and compared the techniques that they used to solve them while using Photoportals versus using traditional techniques. For

example, to avoid fragmented visibility, known techniques include having a larger field of view or allowing users to see through objects, but with Photoportals users can share perspectives with other users to reveal parts of the environment to those that may not be able to see due to positioning in an environment. Another issue with collaborative systems is unexpected motion. A traditional approach to avoid this is having multiple users share control, or augmenting group navigation, however with Photoportals users may travel through portals as a replacement for steering through a scene. All of this is possible with Photoportals due to the reference-based interaction techniques built into the system.

In the Photoportals project, multiple users can be in the same physical space engaging with the virtual environment, or they can be remotely located. "For a group of collocated users, each user can independently walk in front of the screen, but navigation through the environment is always performed for the entire group," (Kunert et al., 2014). This is to preserve the feeling of the shared virtual environment. Interaction between users and the environment is also not problematic even when multiple users are interacting with the same objects, for when a Photoportal is captured in scene, it carries a frame, an object or location, viewpoint, and recording time. For example, rather than picking up a plate on a table, a user would capture a Photoportal of the plate and from there they can examine it in 3D, share it with others, copy instances of the plate, or save it for later examination.

Notable manipulation techniques in with Photoportals include methods for bringing objects from scene to scene, remote manipulation, and 3D box mode. "When dragging an object from a scaled portal scene into the main scene, the object retains the scaling of the scene from which it is extracted," (Kunert et al., 2014). Users can snap a picture in one environment, then move to another and open the picture they took previously as a Photoportal and drag (copies of) objects directly from the picture into the current environment. Scaling is also handled in this process, before dragging the object into the new scene, users can zoom in/out in the picture and then place it at the scale they set it to by zooming. Photoportals are viewed in either 2D-box mode or 3D-Box mode, and 3D-Box mode is similar to looking at a scene captured in a glass cube. Users can see it from multiple angles, and this solves an interaction technique of multiple users interacting with the same object; if someone wants to see something from another angle independently, they can snap a picture in 3D-Box mode and do so without needing to drag the whole team to another positioning. With 3D-Box mode, users can perform remote manipulation on objects like the WIM system. They can capture 3D-Box mode Photoportal of some environment, adjust the scale, and move items around in a handheld miniature environment.

Kunert et al. claim Photoportals offer unique affordances to users collaborating in-person or remotely within virtual environments. The researchers evaluated their interface by having design students theorize and design ideas for 3D virtual environments in a UX workshop, and then briefed the students on use of Photoportals and had them interact with and rate the Photoportals system. Kunert

et al. received generally positive feedback, and it seems that using the photograph metaphor for interacting and collaborating in virtual environments proved engaging and also helped to solve various interaction concerns that arise in other systems.

Movement and Navigation

Movement in virtual spaces can be accomplished via physical motion in the real world which is mapped to the virtual world, teleporting, flying driving, etc. (Mine, 1995). There are many ways to handle how a player will control movement, and it's important to keep in mind "speed and direction of motion." Some experiences use keyboards, gamepads, or other controllers that handle things like player motion, selection, and other interactions and users can easily make use of these familiar objects with their hands while in-headset. There are also systems that involve multi-directional treadmills that headset-wearers walk on to move in a virtual environment (Fung et al., 2006).

The critical pieces of information about navigation in VR are the two components that comprise navigation: wayfinding and travel (Sherman & Craig, 2018). Wayfinding components can be additions to the environment such as maps, pathways, and other visual (or audio) aids to help users understand their position in relation to their desired location. Travel is the method that the user can use to get from point A to point B. The ways that travel is handled in VR reflect the manipulation methods Sherman and Craig describe for VR: physical controls, virtual controls, agent controls, and direct user control.

Lages and Bowman compared moving one's body to having a hand controller to move an object, and found that not only does preference depend on users' spatial ability, but previous game experience plays a role as well (Lages & Bowman, 2018). They found that walking to observe an object from different perspectives is more intuitive than learning mechanics of a controller system, and though walking does require spatial ability resources, it is not as taxing as the "cognitive resources being divided between the interface and the task" in the controller case.

Motion in VR is nauseating to some (Regan, 1995), and to combat this many designers implement a teleportation system to allow users to select a spot in their environment and "hop" to it. Getting creative with the methods of interfacing with the virtual world can result in projects that change how we think of controllers, and creativity can make an experience more compelling if the controls are well-integrated into the experience. The VR video game *Lone Echo* takes place in outer space, and users navigate the scene by grabbing onto handholds and pulling themselves around in zero gravity (*Lone Echo*, n.d.). This control scheme not only helps set the theme of the game and feeling of zero gravity control, it also provides a reasonable way to traverse the virtual environment simply with two handheld controllers. There are many different ways to design navigation in VR and understanding the input mechanisms available and the experience itself is essential for designers.

2.2. XR-mediated Collaboration and Communication

2.2.1. Communication and Collaboration in AR

Shared Space Collaborative AR Applications

Describing multi-person AR experiences may sound redundant, as all AR experiences occur by augmenting the real world and other people exist in the real world so communication can be handled as it would normally—by communicating unmediated. There are, however, instances when designers may want to specifically leverage AR to enhance a multi-person interaction or communication system for users that are local to one another (sharing a physical space).



Figure 4 Shared space AR experience

Synchronous (real-time) communication may come in the form of talking in the real world or messaging someone too (physically) distant to communicate verbally with some kind of virtual messaging system. There is also what Craig refers to as "nonreferential AR" which involves augmentations of the real world with enhancements and filters rather than distinct objects and environmental additions (Craig, 2013). This could come in the form of making audio louder (like a hearing aid) or making a room brighter as a visual aid to help users interact.

Asynchronous (not real-time) communication is possible in AR too. Reitmayr and Schmalstieg show that leaving digital notes and annotations around a physical urban area for another person is beneficial for AR-mediated communication (Reitmayr & Schmalstieg, 2004). Nancy Baker Cahill, a Los Angeles based artist, creates digital works of art that require physical presence in a specific location for people to enjoy in AR (Cahill, 2019). Content that can only be seen by users present within a certain context that the developer determines has interesting implications for communication using AR and how we consume and relay information. Written and voice memos are good ways of participating in asynchronous communication, and annotation systems and user-generated content will be further addressed.

Some AR applications are built to support collaboration between users while users are co-located in the same physical space. Resources such as the Placenote SDK which allow developers to create networked AR experiences (*Placenote: Build Spatial Apps*, n.d.). These operate by having users scan a map of their physical surroundings. This map is then uploaded into the cloud where collaborators'

devices download it and from here, users and objects can be tracked in a collaborative way, where each device knows the relative position of each other networked device in the environment. The Placenote SDK is not the only service doing this, nor the only method of AR co-location, but it is one popular example supporting this type of interaction.

Billinghurst and Kato express the importance of seamlessness in face-to-face AR conditions, and even found that participants of a collaborative face-to-face AR experience exhibit behaviors similar to those of interacting with physical objects when interacting with AR (digital) objects collaboratively (Billinghurst & Kato, 2002). This shows that there is opportunity to learn from traditional face-to-face interactions when building for AR experiences. Szalavári et al. describe the five key components of collaborative AR as virtuality (viewing objects that may not be visible in the physical world), augmentation (augmenting real objects with digital information), cooperation (multiple users can cooperate), independence (users can control their viewpoints), and individuality (users can have different experiences) (Szalavári et al., 1998).

Kaufmann and Schmalstieg developed Construct3D to provide a collaborative environment for teaching mathematical concepts (Kaufmann & Schmalstieg, 2002). Construct3D provided a place for students to experiment with geometric models and the system encouraged exploration. They found Construct3D to be beneficial to spatial skills, and suggest that for future success, they would need to continue

working with students to better build for student needs so that an application of this nature could most effectively benefit the target population.

The Studierstube is an early example of collaborative face-to-face AR, in which Szalavári et al. propose their collaborative AR system with findings such as giving users the capability to visualize data in this new way (e.g. moving oneself vs. moving the objects) is surprising at first, but easy to learn (Szalavári et al., 1998). Giving people the ability to visualize any data or object that can be sufficiently modeled with colleagues is a great way to inspire new workflows, types of collaboration and demonstration, and ways of representing information and interacting with data.

Remote Collaborative AR Applications

There are opportunities for AR applications to support remote interactions as well, and these are of particular interest to me as I focus on remote collaborative experiences leveraging AR. These are differentiated from the shared space examples in that the users of a remote AR system are not physically colocated in a shared space. Instead, users are in remote locations from one another, interacting with digital information, objects, and one another by leveraging AR. These types of interactions can be designed for synchronous (same time) or asynchronous (not same time) experiences.

Billinghurst and Kato found that AR can be used to improve sense of copresence while remote conferencing (Billinghurst & Kato, 2002). This exploitation of AR's affordances to improve feelings of co-presence can change the ways that

people perceive colleagues and interact with one another. Synchronous interactions between remote pairs enhanced by AR can involve seeing a remote partner, seeing a remote partner's actions, or sharing perspectives and annotations while visualizing data or objects with a remote partner. Gasques et al. also found that their collaborative remote system for surgical tele-mentoring helped participants facilitate collaboration (Gasques et al., 2021).

Sodhi et al. developed a proof-of-concept system for synchronous AR collaboration (Sodhi et al., 2013). Their system gave users the ability to see a remote partner's hand gestures and manipulations of digital objects in the physical world. This has implications for training, teaching, and demonstration of all sorts. Similarly, Gurevich et al. developed TeleAdvisor which gave remote collaborators the ability to annotate and point out things in the physical world to advise and train from afar (Gurevich et al., 2012). Pointing out aspects of an environment to a remote user can be crucial in some circumstances and proves useful for users whose work lends itself to this (Bauer et al., 1999). The Spinnstube project, aiming to support education efforts through an AR initiative in schools, also allowed for remote viewing of objects (Pemberton & Winter, 2009). Students could join a session and view a digital object in their environment with a remote peer by each wearing special glasses to view the digital material in stereoscopic 3D.

Gauglitz et al. built a system for digital annotations on the physical world that persist, an example of an asynchronous AR collaborative tool (Gauglitz et al., 2014a)(Gauglitz et al., 2014b). Asynchronous AR collaborative tools can be helpful

for people who need to display information to colleagues where doing so in realtime is not necessary (in contrast to the remote collaborative tools that involve one participant demonstrating something to another participant real-time).

Remote videoconferencing supported by AR is another space of interest for synchronous remote interactions. Barakonyi et al. ran user studies for their remote videoconferencing AR tool which gave remote pairs the ability to show AR content to one another during a video call (Barakonyi et al., 2004). They found that latency is extremely important; digital objects should be perceived to be manipulated in real-time, as if manipulating a real object. Digital information should also be attached to something, and not just floating in the air unrealistically, potentially covering the remote collaborators face.

CoVAR is an example of a project combining AR and VR, allowing users to join a remote collaborative session via HMD observing a scene from AR or VR (Piumsomboon et al., 2017). Piumsomboon et al. explored perspective sharing/shifting and sharing of gaze cues amongst participants within a scene. CoVAR along with some of the other work mentioned in this section offer good insights into the area of interest I am focusing on in this dissertation, yet they differ in the ways that the remote collaborators are interacting around the digital objects, and the ways in which the objects are handled between participants. Another example of using AR and VR is the Magic Book, which enables users to see AR content overlaid onto a physical book, much like a pop-up book, and they can subsequently dive into a scene using VR (Billinghurst et al., 2001).

2.2.2. Shared Virtual Spaces and Communication in VR

Understanding shared virtual spaces means taking a look into work in the VR space as well, as co-presence around virtual objects rigged to the physical world absolutely has overlap with co-presence around virtual objects in the virtual world. VR provides an opportunity for those who are physically remote to come together and share the same virtual space/environment, and understanding the needs of as many users as possible means considering various disciplines such as social sciences, engineering, and anthropology to design meaningful experiences that successfully leverage the affordances of the platform (Biocca & Levy, 2013). "A major benefit of virtual shared spaces is that they allow collaboration to take place via computer networks. Thus, not only can two workers share a space while remaining in their offices just down the hallway from one another, but they can also be an ocean away," (Craig et al., 2009). Co-presence online isn't a novel concept; some examples of this include chat rooms online, video games online, and online message boards. Shared experiences leveraging VR are inherently different than those examples however, in that they afford users the ability to have embodied interactions with one another and with their virtual environments.

Couclelis remarks on the ability of modern technology to render distance less important (Couclelis, 1996). The internet, and the devices with which we use it, facilitate remote communication and allow us to work and behave in ways that make distance less of a determining factor in many of our daily activities. Distance will likely become even less relevant as VR is adopted more seamlessly into the lives

of the general public, and as the technologies supporting VR use melt into the periphery and issues such as latency become negligible.

Creating compelling, engaging shared virtual spaces that serve the needs of users as well as foster user interaction is a point of focus for many interaction designers. There's potential to bridge physical gaps with virtual interaction in VR. Oculus Horizon, a massive online virtual space for users to interact, consume media, play games, and more, promises to give users a virtual place to create and explore in shared virtual environments (*Introducing 'Facebook Horizon,' a New Social VR World, Coming to Oculus Quest and the Rift Platform in 2020*, 2019). This project is a very large-scale example of shared spaces in VR, and is the newest of the type, however it is not the first example of shared virtual spaces speaking more generally. When we consider earlier work such as Second Life, a virtual world allowing users to connect, there is plenty to learn about how and why virtual shared spaces can be beneficial to people.

Second Life is a virtual world (VW) that gives users a platform to connect, create, learn and explore (Kaplan & Haenlein, 2009)(Esteves et al., 2011). Note that a VW is not necessarily VR, as a VW doesn't, by definition, need to be an immersive experience but rather an imaginary space and/or objects in a space (Sherman & Craig, 2003). Users, or "residents," do not necessarily have to participate in any goaloriented activities, rather they are given free rein to explore and interact. While one could debate whether or not *Second Life* should be considered a massively multiplayer online role-playing game (MMORPG), there is no doubt that interaction

designers working in VR space can glean insights from *Second Life* and other MMORPGs such as types of information to display, supporting interactions between users, and many more intricacies of online virtual social interaction.

Bigscreen is a successful example of a social VR space centered around media, where users (represented as custom avatars) congregate to consume media, chat, and draw (bigscreen inc., 2016). Users join rooms where other users are hosting and enjoying video content. This is an interesting VR social experiment, as the community works out best practices and guidelines such as "no talking rooms" or thematic rooms. Designers and developers can learn from the ways that users work with a given system to create the experiences that they desire, and from there iterate on and improve systems for users.

Communication in VR can be accomplished with text and voice, and this can be synchronously or asynchronously (Verlinden et al., 1993). There are interesting social relationships to be explored in VR as well. The Painter Project studied creativity and empathy in VR, and experimented with users following along to VR painting mentors while painting a physical work (Gerry, 2017). Gerry suggests this is a new way to teach creativity, as opposed to the watch and recreate model, and that VR has potential to change the ways people communicate "subjective experiences."

Harassment (Blackwell et al., 2019) and safe spaces ("Designing Safe Spaces for Virtual Reality," 2020) in VR are also points of concern. Putting people in an environment with strangers who can remain anonymous could lead to risky

situations. Blackwell et al. study social norm theory to understand how concepts can apply to VR and where challenges lie, and they support experimenting with community-led governance to uphold social norms in VR communities. There are also technical measures that can be taken, such as buffer zones and options to mute/block other players.

2.3. XR for Educational Purposes

2.3.1. Learning using AR

As I am focused on remote collaborative scenarios based in education, understanding the implications of AR and VR for education are important for this work. Collaborative learning is a concept that research has shown to be beneficial, and I believe that the affordances of AR as a medium can help instructors capitalize on unique collaboration methods that can make learning (and teaching) more powerful (Shehab & Mercier, 2019). AR has strong potential to change the ways that education traditionally occurs, specifically education relying on visualizing content that may be 3D in nature. While AR won't be a cure-all solution for all problems that the field of education faces, there are very compelling use cases for AR in education that could benefit teachers and learners alike. It seems that AR has the most potential to be beneficial in education for material that requires visuals, situational awareness, and 3D subject matter, and in this section, we will examine some previous work regarding AR for education.

Traditional classroom instruments that reduce 3D content to a 2D representation on a flat page or textbook strip away some level of meaning from the

material. Without plunging too deep down the rabbit hole of types of learning and effective education models, I will address the work that some researchers have done with AR applications designed for enhancing education and ways that AR can be most effective for improving education.

Kesim and Ozarslan recognize that learning about the physical 3D world in 2D poses challenges, and certain applications of AR could help dissipate those challenges (Kesim & Ozarslan, 2012). They agree that AR can greatly improve certain aspects of education by changing interactions between physical and digital environments, object manipulation, and methods of student collaboration (Billinghurst, 2002). There remains great need for designers to create experiences leveraging AR for education as the substantive work here is relatively scarce.

Wei et al. had great success with using AR for teaching (Wei et al., 2015). They developed a system that let students create AR scenes with visuals and audio without needing programming expertise, and found that students had significantly improved attention, felt that the material was relevant, and had satisfaction with the lessons at hand.

AR can be used to restore imagery from the past as well. Researchers have developed applications to display information and restored versions of cultural sites in physical registration of the true sites (Stricker et al., 2010). Shemek et al. has done this to enhance museum experiences such as the studiolo (study) of Isabelle d'Este and recreate ancient objects (Shemek et al., 2019). Opportunities for using AR to help explore geographically significant locations or support in situ, inquiry-based

learning are abundant, and Klopfer et al. examine these concepts as well as curated narrative-driven experiences (Klopfer et al., 2002).

Chemistry is a frequent target of researchers and developers working with practical applications of AR in education. The inherently physical subject matter begs better representation than text or 2D images in textbooks can provide alone, and therefore AR may offer a better experience than traditional 2D representations. As mentioned in section 2.1.1, pdbARbrowser is an application built for AR visualization of scientifically accurate molecular data produced on high performance supercomputers (Collins & Craig, 2017a). This application allows users to tap into the resources of a database of over 178,000 proteins called the Protein Data Bank (PDB). By typing in a unique identifier code and selecting a representation style, users can visualize these structures via AR using a smartphone or tablet.

Daqri, an AR startup (which is no longer operational) primarily known for AR wearables and industrial visual aids, produced an app that visually augments a physical cube with information about elements in the periodic table, and when users bring certain blocks together they can witness reactions and equation changes (*Elements 4D Interactive Blocks*, n.d.).

Yu-Chien Chen at the HIT Lab studied students interacting with both physical manipulatives and an AR platform using the magic mirror paradigm. Students had the opportunity to work with molecular models made out of real toy balls and sticks in the classroom and AR representations of the same molecular structures (Chen,

2006). Chen used qualitative methods to collect participant information, and found the students reported that the AR visualizations allowed them to represent complex structures without having to worry about fragility of the toy models and limited resources of the physical manipulatives. The AR condition allowed them to compare structures of different amino acids at the same time and manipulate them to understand their physical nature more holistically.

"AR is a visualization tool which can convey either static virtual objects or dynamic animation at the same time. However, technology is not a panacea to all conceptual representation. Instructional designers should be cautious in determining how to integrate AR into a curriculum properly," (Chen, 2006). Careful consideration should take place before deciding the style of implementation of an AR system into a space because the method such as a magic lens (using a device like a see through, video or optical, HMD) or magic mirror will have a great effect on the end user experience. Conversations about costs, benefits, ability, and more are part of good practice, and Chen hopes that studies such as hers be used as a guideline for developing AR utilities for classrooms and managing the use of these utilities.

Nuñez et al. argue that going without AR in a classroom setting while teaching chemistry is a missed opportunity and a failure to the students (Núñez et al., 2008). The researchers supplied an AR system for visualizing subject matter in material sciences, advanced chemistry laboratory, and ceramics and inorganic chemistry and got very positive results from the students they worked with. Students who had trouble with understanding information were able to make sense

of certain concepts because of the shift in representation styles. Beyond the fact that AR systems can elucidate information in the classroom, they can also provide experiences that might not have been possible with other media formats. Nuñez et al. found that "problems that can be easily solved by rotating a structure and analyzing its symmetry properties are almost unsolvable, even for competent students" without the use of AR in this type of education (Núñez et al., 2008).

Copolo and Hounshell represent some opposition to this movement. "Viewing molecules in two dimensions is important since this is the most widely used means of representing molecules in the field of chemistry," (Copolo & Hounshell, 1995). While it is true that 2D representations are the current standard in classrooms (e.g., images in a textbook) and they remain widely available at this moment in education, there can be little expectation for progress if we do not begin to embrace newer methods for visualizing material that is 3D in nature, especially when they are empirically useful.

Representing 3D material as such is crucial. There is information that is potentially lost when optimal visualization techniques are not utilized in certain settings, as seen in the example of Nuñez et al. using AR to help students visualize chemistry material. Looking at a flat, printed image of a protein structure does not adequately represent the true physical nature of how a protein is comprised and situated in space, nor can it be manipulated or perceived from other angles. The folds and contours cannot always be adequately conveyed in an illustration on a page. There may also be portions of a model obscured from view in an image. Using

solely 2D images to show 3D content risks stripping away visual cues and spatial information and potentially jeopardizes the experience. Now is a great time to begin the conversation about when and where to utilize media such as AR in formal education because accomplishing visualizations with AR is becoming more widely available than ever before, cost-wise as well ("HoloKit Is like Google Cardboard for Augmented Reality," 2017).

Azuma speaks of a HMD enhancing an architects work, allowing them to slip a device onto their heads and see a skyscraper outside that isn't there yet, or see the inner workings of a building like an x-ray (Azuma, 1997). Barreiros et al. studied the effects of positive computing and enhancing work spaces using AR and "biophilic" design, or adding nature elements to the work place (Barreiros et al., 2017). Augmenting the work space of workers can increase wellbeing and productivity, and using a HMD as the magic lens frees the hands to perform work with the visual AR assistance. Other work has been done in terms of using AR in the workplace, and with positive results and improved efficiency. There is a promising future for AR for surgeons and other medical fields (Guha et al., 2017). AR for professional training and task assistance should be further explored as it will likely impact a variety of fields, from dentists designing cranial implants (Scharver et al., 2004) to ship building (Yang et al., 2007).

Some educational theories are broached as we explore the examples of manipulatives, illustrations, and AR representations in the classrooms. While there is merit to different representation styles at different times, I am going to briefly

describe why AR offers immense utility for education by considering this concept from a theoretical perspective in the next few subsections.

Constructivism

One theme that presents itself in the conversation surrounding physical manipulatives is constructivism and the ways in which the physical act of doing work can enhance educative experiences. Constructivist approaches to education have proven useful for some teachers in classrooms, and understanding what constructivism means in terms of AR could be useful.

Sometimes the word "constructivist" causes people to falsely accuse the ideology of being things that it is not, as Cobb, Yackel, and Wood point out (Cobb et al., 1992). Constructivism doesn't mean that teachers set students loose to totally organically discover specific concepts that the teacher desires; this approach has too many variables to be effective and there is a risk that the teacher may not achieve the desired outcome. "A teacher who actually took this caricature of constructivism seriously would be abrogating his or her obligations to students, to the school as a social institution, and to wider society," (Cobb et al., 1992). The researchers describe learning as "a process of acculturation," and stress that "the central issue is not whether students are constructing, but the nature or quality of those constructions," (Cobb et al., 1992). Freudenthal likens the constructivist approach to the learning process of humankind repeating history; not as a blank slate, but "as it would have happened if people in the past would have known somewhat like we do now," (Freudenthal, 1981).

This parallelism drawn by Freudenthal leads me to believe that AR's potential for creating environments where principles of constructivism can be realized may be greater than physical manipulatives and illustration representations in certain circumstances. Educators can seed students with information, which was not always available to past students learning the topics, while they build understandings of their world. Students should also be supplied with technologies that have not always been available to "recapitulate the learning process…though in a modified way," (Freudenthal, 1981). Students now have an opportunity to consume information in a way that hasn't been possible before, and if we can give students tools to construct ideologies in terms of the world around them (3D space), learning and expanding knowledge bases could be unprecedented.

Learning by doing in order to construct understanding isn't always necessarily possible. Veterinary students learning about the vascular systems of cows, for example, might not have the opportunity to bring a real cow into their lab and perform procedures to learn them, but AR platforms can potentially help in times like this (*Illinois Cow*, n.d.). In terms of both convenience and ability, implications for what AR can do to education are powerful. Implications for this ideology are strong for VR as well, illustrated well by the NICE project, which allows people to build, collaborate, and interact with the virtual world (and even participants joining via web) (Johnson et al., 1998).

Boudourides writes about learning and affordances that certain technologies offer to users. He argues that technologies "possess interpretive flexibility, as far as

it is void of any objective, fixed properties, but allows for different interpretations by relevant social groups," (Boudourides, 2003). Technology is meaningful because of the way it functions in society, and AR can, and likely will, have different meanings and uses for various groups of people. We can't yet fully understand the effectiveness of AR platforms for education because we haven't allowed teachers and students to utilize existing applications on a large scale. When the education community has more experience with AR, stabilization or the "processes of settling controversies and negotiations among different social groups" can be understood in context (Boudourides, 2003).

Harris and Alexander would like to see constructivist education evolve and "believe that educators must accept the challenge to remain open to emergent, validated approaches that can be integrated into an effective pedagogical repertoire," (Harris & Alexander, 1998). The authors reiterate the point that constructivist education is achieved through guiding students to take agency in their own education. There is no one size fits all approach, but giving students tools to visualize information in a way that could very well be more salient to some types of information is a step in the right direction, especially now that we have AR available to us.

There can't be expectations that all children learn the same way by consuming the same media, and Harris and Alexander explore some of the difficulties experienced in classrooms, even with great teachers. Using AR to allow students to discover and work with concepts that are native to 3D space might draw

them to become more engaged in certain subject material, and "deep, meaningful understanding occurs when children participate fully in their own learning, with previous knowledge and experiences as the starting point for new learning" (Harris & Alexander, 1998).

Parallel Distributed Processing/Connectionism

Parallel distributed processing (PDP), also referred to as connectionism when speaking about human brain and memory, is a theory where the memory in a computer and the human mind are likened architecturally (*Parallel Distributed Processing Models of Memory - Dictionary Definition of Parallel Distributed Processing Models of Memory / Encyclopedia.Com: FREE Online Dictionary*, 2004).

Connectionism specifically looks at memory in terms of a neural network of nodes and connections, and the ways that connections between nodes are strengthened or weakened as they are activated. Networks comprised of nodes in a PDP system are activated in parallel; that is there isn't a set linear procedure of the flow of activation. Connection weights are also updated after activations, and this is how memories are constructed. "Knowledge that governs processing is stored in the connections among the units, for it is these connections that determine what pattern will result from the presentation of an input. Learning occurs through adjustments of connection strengths. Memory storage is just a form of learning, and also occurs by connection weight adjustment," (*Parallel Distributed Processing Models of Memory*, 2004). After connection weights are adjusted and memories formed, when information previously encountered is encountered again, it is with a weighted network that may now activate more effectively and with a more informed view.

Anecdotally, adults often look at children busy interacting with their mobile devices and comment on how much of the real world they are missing. This isn't a completely fair judgement due to the fact that there is undoubtedly a lot of great information that exists for the children within their devices (Craig, 2013). Education using AR can be situated using the PDP model when we consider the way that AR makes it possible to consume the input we receive from our real world, along with the information we must represent digitally hopefully with as little loss as possible.

Consider a student in class who is learning about a protein structure. This student has an illustration of the structure in their textbook, and a ball and stick manipulative model of the structure. Each representation style can be considered a node in a PDP model, along with textual information (another node) about the function of the protein. The connection between these nodes is currently limited by the student's ability to relate these concepts mentally based on their understanding of the representations, no matter how abstract. When arbitrary representations of information are thrown at students, it can be hard for some to make the meaningful connections that we call learning; AR can help with that. Instead of expecting students to cognitively orchestrate methods of making representations meaningful and situate concepts within their real-life implications, AR can be the interactive, visual aid that elucidates scientifically accurate data not typically visible to the naked eye.

AR can act as a steroidal agent that strengthens the connections between nodes. Instead of thinking about how the text describes the images and models she

looks at, she can see the physical nature of the structure by moving her physical body and maintaining her sense of proprioception. She can witness reactions within the protein structure's system, move the structure around, and maintain the scientific integrity of the visualization. The 2D illustrations and many of the limitations (cost, impracticality) of the physical manipulatives are largely eliminated with AR; with physical models, one is limited by the number of toy components to which they have access, whereas the upper bounds of an AR visualization are set by computational power.

Processing the world as we perceive it with our naked eye, enhanced with information either impossible to see with our naked eye alone or difficult to obtain is a form of parallel distributed processing that I venture to say has not been possible before. AR representations of educational material in classrooms can begin to negotiate and reconcile the differences between what we can imagine, and what our physical world limits us to. The spaces between our physical world, the representational models that we use for concepts, and our internal understanding of these concepts leaves room for information to be lost. Seeing the real world in tandem with digital information that can take the form of representations of abstract material, objects that are difficult to bring into classrooms, and more can change the way that young minds learn, and all minds think.

2.3.2. Learning using VR

As stated earlier, there is overlap between AR and VR, so understanding the implications that VR has for education is important to this work as well. VR affords
users unique ways to learn, train, and teach. There are ways of interacting with information that are much different from more traditional 2D interfaces. Here, I outline the ways that VR has helped teachers and learners, and what types of scenarios call for VR applications.

There are a number of ways that 3D visualizations and interactive stereoscopic simulations in VR can lead to greater informational uptake and even help facilitate realizations. Medical educators have shown that training stereoscopically can be significantly more effective than traditional 2D training (Faria et al., 2016). Psotka also wrote about immersive training systems and talked through the various ways that VR could be beneficial to training and education. He assessed VR as a tool through different education ideologies and made substantiated cases for how VR could be meaningfully incorporated (Psotka, 1995).

Some designers take advantage of the fact that VR can take users to places they may not be able to experience in the physical world. A narrative experience like Sky VR: Hold the World gives users a hands-on look at pieces in London's Natural History Museum and allows them to examine and manipulate artifacts with narration from Sir David Attenborough (Sky UK Ltd & Factory 42, 2018). The Apollo 11 experience puts users in the boots of an astronaut on the Apollo 11 mission (Immersive VR Education Ltd., 2016). The Anne Frank House puts users into the shoes of Anne Frank during WWII (Force Field, 2018). VR is an interesting medium for immersive storytelling that throws the user into an embodied experience.

Slightly different than story-driven experiences are free roam, exploratory VR experiences. The Kremer Collection VR Museum (Moyosa Media BV, 2018) and the Isabelle d'Este VR studiolo (study) (Shemek et al., 2019) allow users to explore informational experiences at their own pace. Navigating through venues and examining objects to learn more in VR allows users to visit places that they might not be able to get to physically (Google, n.d.).

Other experiences are more sandbox-esque, such as RoboCo from Filament Games. RoboCo is a VR sandbox game that pushes users to think creatively while building robots that serve various purposes to help "hapless humans in the world of tomorrow," (RoboCo, n.d.). This experience lets users build and experiment with systems using real physics, in a space where resources and danger are not an issue. HoloLAB Champions puts users in a gameshow-like competition where they perform chemistry experiments (Schell Games, 2018), and in the same vein but executed differently, Nanome gives users a place to shrink down to atomic level and manipulate protein structures to understand, and potentially discover new, molecular structures (Nanome, 2018).

Immersion is important to certain aspects of learning. Hu et al. found measurable effects of VR for learning; they found that VR "appears the highest sensitivity on Creative Thinking Instruction" and "interaction in [VR] presents the highest fluency on Creative Thinking Instruction" (Hu et al., 2016). The Magic Book lets people visualize AR content rigged to physical books, and users can choose to enter a scene of the book in VR to be immersed into the book (Billinghurst et al.,

2001). The immersion that VR offers has strong implications for education, and creating an entertaining educational environment. Exploration and interactivity are important to promote "positive learning behaviors," and VR can help facilitate this (Lau & Lee, 2015).

VR learning experiences come in different forms depending on the objective. Some experiences are guided narratives, like immersive movies, that tell a story and walk users through an educational journey. Others are free roam scenes that users are meant to explore and investigate. Some are meant for users to experiment with objects as they would in the real world, but with resources not always afforded in the real world.

Both AR and VR should be in every classroom; not because they can solve all of the issues that educators face, but they can be used as a tool alongside the other traditional methods of teaching. It's a shame that these tools that can give students a place to immerse themselves into experiences, perform experiments and projects, and interact with material in an entirely different way aren't already being used more in education, but hopefully designers and educators begin to focus attention on the overlap of education and XR media. As Sherman and Craig point out, we will need to begin to build a literacy around these media which will propel us into more compelling, interactive experiences (Sherman & Craig, 1995).

2.4. Scientific Visualization

With this dissertation, I am looking at the cross section of AR, education, and visualization—in particular (in most cases, but not strictly limited to) scientific visualization. Sherman et al. look at the rich history and evolution of the field of scientific visualization. They talk about types of representations, how visual cues can be exploited to create more meaningful visualizations (Sherman et al., 2004). Baker and Wickens discuss scientific visualizations as it relates to VR visualizations and interactions and the human factors that are at play in VR visualization scenarios (Baker & Wickens, 1995).

Jones et al. found that understanding complex, rich 3D data such as the information presented around molecular structures is difficult to do, even with traditional visualization software. "The visual subtlety, complexity, and conceptual depth of molecular visualizations present important research and design challenges to chemistry instructors, curriculum developers, and educational researchers," (L. Jones et al., 2005). They point out the opportunity for virtual reality to provide a sense of immersion and more apt interactions around data that is 3D in nature. They also point out the importance of interdisciplinary collaboration for educational scenarios, and the fact that new learners may find difficulty with 2D visualizations. "A chemist can 'see' the theory underlying a visualization. However, cognitive scientists have shown that the visual system has limited neural resources (Trick & Pylyshyn, 1994) therefore, visualizations of abstract molecular concepts may be too complex for learners to process. For students to learn from an image, they must

attend to its relevant characteristics and understand how they demonstrate new concepts. They must also know the scientific conventions (Habraken, 1996)and learn to tune out irrelevant or distracting information (Kozma & Russell, 1997)," (L. Jones et al., 2005).

Keefe and Isenberg forecast that the future of scientific visualization will focus on user interface interaction and 3D computer graphics. Leveraging XR tools to visualize information, especially in contexts where the 3D nature of the data is important to the lesson such as chemistry or physics lessons, can lead to better understanding of concepts and material being presented. "To work effectively with scientific visualizations, users must understand the mapping from data to visual form," (Keefe & Isenberg, 2013).

One limitation of more traditional methods of scientific visualization tools is the 2D display nature of many visualization tools. AR and VR provide for opportunities to engage with scientific data in ways that exploit the 3D aspects of digital objects and give rise to different, potentially more engaging, and possibly more natural interaction techniques when it comes to interacting with the data. Section 2.3.1 explored some compelling the examples of utilizing AR for scientific visualization.

Chapter 3.

Methodology

The purpose of this work is to explore methods of supporting remote collaboration and communication leveraging AR visualization and interactions to support those working around 3D (or spatially relevant) objects/data. Specifically, I explore the concept of giving users the power to pass objects through the screen from one user to another remotely, and the interactions that take place around this scenario. The review of literature shows that there is vested interest in using AR (and VR) to support education, and the field of education is conducive for these types of interventions. The literature also provides good context for the types of interactions and their use cases that have been used in the past for XR applications and experiences.

There are projects working toward understanding how to best collaborate using XR media. A review of work around collaboration in AR proposed that future work in this area should focus on establishing a sense of co-presence, researching "which interaction paradigms will be most effective," and exploring the interaction possibilities around physical objects and the digital overlay (Lukosch et al., 2015). A lot of the previous work regarding remote collaboration and AR focuses on sharing cues from a remote partner (gaze, gesture) and reconstructing a remote environment to see from the perspective of a remote peer. With this work, I propose that we focus specifically on the digital objects being presented in AR, the methods

used to interact with the AR objects being presented, and how these AR objects can be woven into the physical world in order to create an intuitive user experience between remote collaborators.

I am investigating the idea of passing objects through the screen between physically remote participants of teleconferencing video calls by allowing users to visualize these objects using AR. In order to do this, I have developed an application, SAROSvis, around teleconferencing video call applications, in our case Zoom, with various interaction techniques in AR. I designed user studies to evaluate the use of this system and understand what is most salient to participants of a remote call when collaboration around 3D objects is the goal.

For our process, the use of the SAROSvis application is agnostic to the teleconferencing software being used; due to the methods of AR tracking being employed (target images using fiducial markers (Craig, 2013a)), as long as participants are able to share their screen during the video call, the AR objects can be shared between participants of a call using target images. We recruited participants and examined their experiences with the application that we developed for this study, and subsequently reviewed the outcomes including both quantitative and qualitative data. At a high level, this work involved four main steps: review literature, design/build the system, evaluate the system, review evaluation results.

Before moving on, I would like to note that I have indeed considered a VR approach for remote collaboration, specifically as a plugin for teleconferencing tools as well. I took part in the development and evaluation of XRTeleBridge (seen in

Figure 5) in collaboration with the Empathic Computing Lab. This work examines the integration of an MR interface into existing teleconferencing tools (in this case Zoom), and explores the interaction techniques around the camera/view control and manipulation (Sasikumar et al., 2021). This provides for asymmetrical collaborations where participants can join via traditional webcam and/or VR HMD, meaning that participants can demonstrate objects, environments, and topics to others from virtual spaces.



Figure 5 Screenshot from XRTB Zoom call where participants are observing a teacher demonstrating experiments and manipulations of the molecular model of ethanol (Sasikumar et al., 2021)

In the following subsections, I will outline the design and development of the SAROSvis application which I evaluated in this study and explain why certain design decisions were made. I will then review the evaluation methods and procedure employed for the user studies.

3.1. System Design & Development



Figure 6 A user running SAROSvis on an iPhone observing a presentation through the lens of the application, seeing the rigged AR content attached to the presentation. The model displayed will change depending on the fiducial marker (target image) being displayed on the screen

In this section, I will describe the design decisions and development process

for creating SAROSvis. SAROSvis was developed as a means to allow users to ostensibly hand objects to each other through their screens while participating in a video teleconference call. Users can choose which objects to view and hand through the screen, manipulate these objects independently or collaboratively with remote peers, and display a visual representation of their location relative to the object in the physical space of their remote peer.

Providing participants of teleconferencing calls with new ways of interacting with digital assets and data using AR, and in turn new ways of collaborating and communicating with one another, means that we need to have an understanding of existing teleconferencing solutions and how we might develop a solution around them. Zoom provides users with the ability to talk to one another while seeing a live video feed from their webcam. Users can also share their screens and present to the other participants of a call; this can be useful for presentations and showcasing material that may be local to one's computer. Zoom also allows participants to chat via text, create breakout rooms, control video/audio sharing, and more. While this tool is great for remote collaborative work, it does not fully serve users who work with 3D entities or environments; and while sharing the screen can help for certain demonstrations, users are limited to what they can actually display on their screen.

I am particularly interested in seeing if I can accomplish the sense of handing digital, AR assets through the screen from one participant to another by rigging the desired 3D assets to slides of a presentation that would in turn be shared amongst to participants of a teleconferencing call using the screenshare capability of Zoom. Users who participate in the call while using the app can see and interact with relevant 3D material as the host desires by viewing their screen through the lens of the device on which the application is running. The high-level flow of events for this involves a host starting a Zoom call and sharing their screen with a presentation

being displayed. The images in the presentation have AR models tied to them via fiducial markers. The host will ask participants to view their screen through the lens of their smartphone while running the SAROSvis application, and the host can now decide what pops up to the remote participants of the call in AR depending on the slide that they are displaying; they can effectively decide what objects will be displayed to remote participants in their environments. From there, they can manipulate and interact with the AR content in various ways that will be described below.

While designing this system, I kept in mind that this system is intended for use in conjunction with existing teleconferencing tools such as Zoom or Skype. This means that users will be actively focusing their gaze at their computer screens to engage in a call. This is what led to the decision to use target images, or fiducial marker-based AR (Craig, 2013a). Users point their device's camera toward an image target, and the camera observing the scene allows computer vision to place the desired digital model in the correct location; the relevant models will be displayed to the user rigged to the target which, in this case, is a slide or image displayed on the screen during a call.

Figure 7 shows that when a user looks at their screen (through the lens of their smartphone or tablet), they see the digital object (here, a molecular model) overlaid on the target image which is displayed on the screen. The app was built as an extension to pdbARbrowser, which was briefly mentioned in the Background and Related Work section. The application allows users to visualize and manipulate

molecular models from the PDB in AR (Collins & Craig, 2017a). This application makes use of VMD, which renders the molecular models in various representation styles and color schemes, and an architecture diagram demonstrating the data flow is shown in Figure 11 (Humphrey et al., 1996). The pdbARbrowser user visionbased mobile AR (with target image fiducial markers) and was developed in Unity (c#) using the Vuforia vision library(*Vuforia Enterprise Augmented Reality (AR) Software | PTC*, n.d.). Vuforia gave us the ability to use target images to place objects upon the target images, so when the desired image is displayed, the user will see the relevant model pop up on their device's screen.

SAROSvis was created using Unity 3D and the Vuforia plugin as well (Technologies, n.d.)(*Vuforia Enterprise Augmented Reality (AR) Software | PTC*, n.d.). Figure 7 shows the view from the application where a model is rigged to the slide displayed on the user's computer screen, and there is also a floating pencil acting as a pointer in the scene. This pointer is a visual representation of a remote user in the scene whose position and rotation (of their device) is matched by the pencil. The remote participant can maneuver their device and the pencil will mimic their movements accordingly, acting as a remote visual communication tool. This works by taking the position of each in space relative to the target image being views by the device camera and placing the digital pointer at the location of the remote participant so that users can see precisely how and where remote collaborators are operating in the scene. They are, of course, able to point at and touch digital objects which are also placed relative to them in the scene, creating the illusion that remote collaborators are viewing and manipulating a shared digital object though they are

physically remote from one another. The information is networked amongst each user running the application using Photon Unity Networking (PUN) package (*Introduction | Photon Engine*, n.d.). I refer to this feature as remote demonstration, as remote collaborators can perform demonstrations on objects and exhibit visual pointing cues (along with verbal cues through the video call). This architecture is detailed in 3.2.

A participant of a teleconferencing call may choose the models that they would like rigged to a given set of target images (these are variable) and embed these into their presentation slides. While taking part in the teleconferencing call, this participant will share their screen. Other participants will view the shared screen through their device running the app, and thus see the desired objects being passed through the screen to them from their remote partner. When the person driving the presentation changes slides (target images), a new model will be passed to the participants due to the change in target images.

Users can manipulate the 3D objects by using their touch screen device to manipulate and scale assets, as well as shift their perspective by moving their bodies (viewing angle shift). Users can also participate in remote joint demonstration, meaning that multiple participants can manipulate a shared digital asset and see manipulations propagated to all devices.

As the application for this study was built as a branch from pdbARbrowser, the tasks that users took part in during the user studies involved examining 3D molecular structures and categorizing them into geometrical categories. This task

did not require any prerequisite information or training; however, it did require participants to look at the structure being displayed and manipulate their view of the object in some way to know for sure how they should correctly categorize the object. This task was chosen as it is not overly difficult, yet it does require participants to examine and manipulate the object and/or shift their view of the object to complete the goal.



Figure 7 A screenshot from within the app of a molecular model placed on the target image as displayed on the user's computer screen. The computer screen is displaying a presentation that the remote participant of the teleconferencing call is sharing, thus they are able to choose what the person using the app sees in AR based on which slide they are sharing at a given time



Figure 8 Screenshot from within the app showing the visual representation of a remote user; in this case, the visual representation is a pencil. The pencil mimics precise transformations and rotations of a remote user who is viewing the same model remotely



Figure 9 Screenshot from within the app showing the visual representation of a remote user; in this case, the visual representation is a pencil. The pencil mimics precise transformations and rotations of a remote user who is viewing the same model remotely

3.2. Architecture Design of AR Applications

To describe the way that data is handled, stored, and communicated through the network in SAROSvis, I will give a brief overview to some methods for designing system architectures for AR applications. Table 4 shows descriptions, benefits, and costs of possible data storage options for AR applications. It is important to understand that there are different options available when working on visualization projects because assets can quickly become taxing on a systems storage and/or network depending on the architecture, however careful planning depending on the experience can help to avoid potential bottlenecks.

Location	Description	Benefits	Costs
Local	Storage is done on- device, without needing to communicate with a server on a separate remote device.	 Security/ privacy No need for network Very low latency No chance of server being unavailable 	 Requires space on local device Large app to download / install
Remote private	Storage is done on a cloud/server private to the developers, rather than the local device.	 Convenience, availability Potentially large storage available 	 Requires network Maintenance cost
Remote public	Developers have access to a publicly available database, like the RCSB PDB or Google Poly repository.	 Convenience Availability No maintenance cost Storage space not a concern to app developer 	 Requires network Could be taken down Could be unreliable Requires understanding

			•	of how to retrieve (API) Must follow their terms of use
Nowhere/computed	Objects are not stored, but rather computationally (algorithmically) generated. E.g. a random number generator computes random numbers when they are needed and does not store these random numbers beforehand, or a fluid dynamics simulation to create simulated clouds.	 Low/no storage space required 	•	May not be able to create certain tailored assets on the fly Requires computation Requires time to execute the algorithm

 Table 4 Data storage locations with costs and benefits for AR applications



Figure 10 Storage capacity vs. network bandwidth, and how SAROSvis compares to pdbARbrowser



Figure 11 pdbARbrowser Application Architecture

For this work, we are storing locally but since we are building in a collaborative component, we do need to connect to the network. While we are storing models locally for this project due to the fact that we are assessing with a set

number of models, there are two networked aspects: the position and rotation of each model. Figure 11 shows the architecture of pdbARbrowser, which sends an http request to the server to initiate the creation and download of a model, while Figure 12 shows the architecture of SAROSvis where models live entirely local to the application, but each client connects to the server to update information regarding position, rotation and location of shared objects, and position, rotation and location of each participant relative to the others.



Figure 12 SAROSvis Application Architecture. SAROSvis connects clients to the Photon Unity Network (PUN) service so that the position and rotation data is updated in real-time across the network for each participant by tracking them according to their position/rotation relative to the target image. Participants can each scan individual copies of a target image in their own respective spaces and interact with remote partners spatially because of this networked solution.

3.3. Data Collection

The evaluation of this system involved both qualitative and quantitative data collection and analysis. In the following subsections, I will detail the processes carried out to assess participant interaction using the application, the evaluation methods, and reasoning for using said methods.

3.3.1. Conditions Examined

I developed a series of conditions that each participant tested (withinsubjects study design) during the user testing phase of this project. The conditions helped isolate certain interactions within the user experience, and thus allowed us to ask users about their preferences regarding specific interaction techniques. The conditions are outlined in Table 5. The first three conditions are variations on ARenabled versions, and 4-5 are non-AR (closer to a more traditional 3D viewer).

Condition 1 uses AR to rig predetermined models on target images and allows for on-screen user input. Users can tap and swipe on the screen on the digital models to rotate the objects. Users can also pinch two fingers closer together or further apart to rescale the objects. Condition 2 is similar to Condition 1 in that the models are rigged to target, however there are no on-screen manipulations that users can take to manipulate the scale or rotation of the AR objects. That is to say, in order for users to change their perspective of the AR models, they must move themselves physically (or at least the device camera).

Condition 3 is also similar to Condition 1, in that AR is used and on-screen manipulations can be used by the users, however this condition also allows for a networked experience. I call this feature, which allows participants to remotely perform demonstrations and actions on objects, remote demonstration (or remote demo for short). With this networked experience, remote participants can manipulate AR objects and these manipulations propagated to all users in the scene; if one participant moves an AR object, it moves for all users. Users can also see their remote peer represented visually in Condition 3 as a pencil which tracks position and rotation and is updated in real time.

Condition 4 did not use AR and is more of a traditional 3D model viewer. The 3D model is rigged in the center of the user's screen on their device. The user can manipulate the model's rotation and scale, but it's position stays fixed. The camera is still opened to maintain consistency among all conditions regarding the background behind the models being displayed.



Figure 13 In Condition 4, regardless of where the user points the camera, the model is rigged steady in the center of the screen. The camera is opened on even the non-AR conditions for consistency between all conditions

Condition 5 also did not use AR, and is similar to Condition 4, but it is a

networked scene allowing for remote demonstration. The model is still fixed in the center of the screen of the user's device as with Condition 4, but when users manipulate the scale or rotation of the object, these changes are propagated to all users in the scene. The camera is still opened to maintain consistency among all conditions regarding the background behind the models being displayed.

Condition Number	Description		
1	AR with on-screen interactions on screen manipulations (swiping to rotate model, pinching to rescale)		
2	AR with no on-screen interactions (exploiting proprioception)		
3	AR with on-screen interactions (on screen manipulations (swiping to rotate model, pinching to rescale) plus remote demonstration		
4	No AR (like a traditional 3D viewer), on screen manipulations (swiping to rotate model, pinching to rescale)		
5	No AR (like a traditional 3D viewer), on screen manipulations (swiping to rotate model, pinching to rescale) plus remote demonstration		

Table 5 Conditions used	for evaluation	Conditions u	used for	evaluation
-------------------------	----------------	--------------	----------	------------

3.3.2. Evaluation Methods

The data instruments used to evaluate this system involved questionnaires from validated measures, qualitative interviews with participants, and collection of log data surrounding participant task performance. The conditions were examined by using a within-subjects study design, and each participant was presented the conditions in a randomized order (determined using a random number generator). The within-subjects designed was used so that participants could compare their experiences between conditions, especially during the interview sessions. Random condition administration was used in order to counterbalance potential order effects of the conditions that participants saw (*5.2 Experimental Design – Research Methods in Psychology*, n.d.).

To evaluate the system, I chose to use a mix of qualitative and quantitative methods. Leaning on previous work, we used the Nasa TLX, MEC Spatial Presence Questionnaire, User Engagement Scale in conjunction with qualitative (Hart & Staveland, 1988)(O'Brien et al., 2018)(Vorderer et al., n.d.)(Bai et al., 2020). For the results reported in this document, I will be focusing on the findings as supported by both the qualitative themes discovered by thematic analysis, and the quantitative analyses of the questionnaires.

Questionnaires

After participants completed a set of tasks within a condition, they were asked to complete a questionnaire that involved multiple measures to understand the load on participants and their feelings around presence and engagement. The

first portion of the questionnaire included the Nasa Task Load Index (Hart & Staveland, 1988). The participants also completed the MEC SPQ and UES questionnaire (Vorderer et al., n.d.)(O'Brien et al., 2018).

The questionnaires were analyzed using SPSS to produce descriptive statistical analyses and run Friedman Tests, Wilcoxon Signed Ranks Tests, and Bonferroni corrections for the condition comparisons (the Bonferroni correction is a very conservative adjustment to the p-value when running multiple tests/comparisons to reduce the risk of error in finding significance).

Qualitative Data

The questionnaires were not necessarily created for use with AR experiences, so gathering further detail with qualitative interviews was important to understanding the user experience. Before each session, I asked participants questions regarding their job/role, experiences with remote learning/education/working, and any experiences with AR (participants qualified their experiences with examples). The follow-up qualitative interviews took place after the participant tested each of the five conditions. These were semi-structured interviews, and the goal of these sessions was to add more descriptive rich data to the quantitative data and the accuracy data collected from the log files.

The pre-test questions focused on demographic information and gathering information regarding each participant's background. This background is important, as well as qualifying their experiences with related technology, in order to

understand the frame of reference from which the participant may be considering their experience with SAROSvis.

In the following few paragraphs, I will outline the topics covered in the posttest interview as well; this is not an exhaustive list of questions, but rather a detailed look into the guiding themes which drove the questions. In the post-test questions, participants were first shown a slide with a description of each of the five conditions that they tested, and they were asked to rank them from their favorite (or most useful) to least favorite (or least useful) and provide some reasonings for their ranking order. They were then prompted to consider a scenario where their *lowest* ranked condition could be *most* useful, and vice versa. Participants were then asked to consider aloud some contexts they could imagine using an app like SAROSvis or select features from the application.

Participants were asked how certain parts/features of the application affected their experience (AR, on-screen manipulations, moving your body to see around the model, seeing a representation of their remote partner, etc.). They were asked if the experience changed the way that they felt about collaborating with a remote peer, and if they found remote demonstration to be a useful tool. They were asked about feelings of connectedness, and whether or not they felt that they were manipulating/viewing the same object as their remote partner.

Finally, participants were asked about their favorite and least favorite aspects overall, and if there was anything they found confusing about the experience. They were also asked for input on interactions they might add to the

experience to enhance feelings of collaboration, and any other features overall that they thought could be beneficial.

All of these topics were broached to cover as much ground as possible given the areas of interest from the literature and provide a more descriptive reasoning behind data captured from the questionnaires. The results from the qualitative data were then analyzed thematically.

Log Data

The log data collected from participant task performance provided information on the accuracy of the selections. This data was collected in real time as users interacted with the application and made selections on the UI. This allowed for comparison between the correct responses and user given responses. This accuracy data is not of particular focus for this project, especially given the relative ease of the task itself, however it is used here as more of a backup as it is good to collect to see if there are any glaring differences in accuracy between participants.

Participants

The 20 participants in this study were selected from a purposive sample of students and teachers who have experienced learning and/or teaching in a remote education model. This sample provides us with data which allow for constructing hypotheses regarding interaction design techniques in AR, and will allow for future iterations based on usability data and qualitative data (LUBORSKY & RUBINSTEIN, 1995). Participants were located primarily in the US, with some located in New Zealand. Six participants used Android phones, while 14 used iOS devices. Most

participants were either students or instructors at the graduate level, however some were instructors of high school students.

3.3.3. Procedure

Before taking part in the study, participants were asked to install the application on their phone or tablet. The application works on Android and iOS devices, and was installed via TestFlight for iOS and .apk file for Android. Participants joined a Zoom call with the researcher who was facilitating the study, and the researcher shared their screen, displaying a presentation slide deck. The task involved categorizing molecular model structures into one of five categories (seen in Figure 17). This task was chosen as it requires no prerequisite knowledge and is a relatively trivial task, however it also shows an applied use of this application and exploits the necessity of understanding the physical 3D nature of the models presented.

First, the participants were presented with 2D images of molecular models similar to those that they would be viewing throughout the rest of the study. Examples of these can be seen in Figure 14 and Figure 15. These were shown to participants to acclimate them to the subject material, but also to help them understand that 2D imagery of an object may not be sufficient for determining or inferring information about said object. For example, the model in Figure 14 could be a trigonal planar model, however one could also say that there is an occluded bond hidden behind the model making its geometry tetrahedral. Participants were asked to categorize 6 slides worth of 2D images of molecular models. Only a few

participants outright stated that some of the 2D images could be multiple geometries depending on the viewing angle. Regardless of whether or not participants considered this fact, the researcher did impart this information on them before beginning the tasks with the SAROSvis application.



Figure 14 Slide from the presentation presented during data collection. A 2D image of a trigonal planar molecular model is shown



Figure 15 Slide from the presentation presented during data collection. A 2D image of a tetrahedral molecular model is shown



Figure 16 Slide from the presentation presented during data collection. A target image (stones) is displayed, and when participants view this slide through the lens of their phone or tablet with the application running, the desired 3D objects are displayed rigged to the target image Participants were given an introduction to the task, and then asked to open the application. Within the application, they would open a pre-determined, random condition and categorize five different molecular models. After completing the categorization, they were asked to complete a questionnaire on Qualtrics, which involved the Nasa TLX, MEC Spatial Presence Questionnaire, and User Engagement Scale. The participant would then move back to the application to test the next condition. After completing each condition, participants were asked a series of questions in a semi-structured interview where they ranked the conditions, provided advantages and disadvantages, and spoke about what made them feel most co-present, engaged, that we were sharing an object, and more (covered in more detail in the Qualitative Data section).



Figure 17 Molecular geometry diagram (Molecular Geometry, n.d.)

Chapter 4.

Results and Conclusions

I outline the results from the questionnaires and the interviews in the following subsections, organized by the overarching themes that arose from the collected data. These results come from a thematic analysis of the data from user testing, as well as the measures and comparisons from the quantitative data collected.

For data analysis, SPSS Statistics was used to run statistical tests on the questionnaire data collected during user studies. Specifically, Friedman Tests and follow up Wilcoxon Signed Ranks Tests were run. The Friedman Tests allow us to examine whether or not there are differences between the conditions in this study given the within-subjects design (*Friedman Test in SPSS Statistics - How to Run the Procedure, Understand the Output Using a Relevant Example | Laerd Statistics.*, n.d.). The Wilcoxon Signed Ranks test provide information on the significance of differences between specific conditions in the study (*Wilcoxon Signed Rank Test in SPSS Statistics - Procedure, Output and Interpretation of Output Using a Relevant Example.*, n.d.).

I adjusted for potential Type I error (false positive) with a Bonferroni Adjustment to each p-value. This is a conservative method of adjustment to the significance value, so it is very likely that there will not be any false positive results during the analysis of the data. While some may argue that this is overly conservative, it provides a large buffer against potential Type I errors and it is better to be safe.

4.1. Condition Rankings

During the user tests, participants were asked to rank each of the conditions that they tested from their favorite to least favorite condition. They also verbally gave reasons for why they ranked the conditions in the ways that they did. Participants favored using Condition 3 above the other conditions and were vocal in the interviews about why they found this version to be most salient to the remote collaborative scenario when compared to the other conditions. Figure 18 shows the breakdown of rankings by condition for the 20 participants that tested the experience. Participants felt that Condition 3 gave them the best sense of copresence due to the visual representation of the remote participant (the pencil seen in Figure 8 and Figure 9) and liked that it included an AR visualization of the model and ways to interact with it on screen as well as proprioceptively.



Condition Rankings Participant Preferences

Figure 18 Condition rankings: participant preferences of each condition tested

It is promising that participants clearly favored Condition 3, which involved AR + on screen manipulations and remote demonstration, especially given that the conditions were examined in random order for each participant to counterbalance order effects. The majority of participants were vocally excited about the ability to view material in AR with a remote peer synchronously, and all of the participants who participate in teaching described seeing the utility in a tool like this for demonstration and collaboration. This gives good signal toward the idea that AR can enhance these remote collaborative scenarios; in the rest of this chapter, I will dig into the details of how AR might enhance remote collaborations, and furthermore report on where significant results were obtained.

4.2. Remote AR-mediated Interaction Design Principles

Each of the following subsections describes a theme from the data which arose from thematic analysis of the qualitative data and quantitative analyses of the Likert scale items and questionnaires. These results are organized under a heading which best describes the phenomenon, and the body describes what was observed, either qualitatively or quantitatively (or both).

4.2.1. Exploiting Proprioception: Moving Oneself vs. Moving the Object

At first blush, Condition 2, which required participants to move themselves (shift their head/body to change their perspective of the model), was the least favorite of the AR conditions. It was ranked 5th (last place) for 9/20 participants, in the follow-up interviews. The task load may be greater on participants when their only option is to move their heads/bodies to navigate about material/objects displayed using AR, however participants were able to vocalize times and scenarios when this method of visualization could come in handy. After describing the results from the NASA TLX scores, I will detail some of the qualitative data around this method of interaction and weigh some of the costs and benefits of using this technique for educational purposes.

The scores from the NASA TLX scales show that participants reported more workload while working in Condition 2. The raw scores from the NASA TLX scales were used as these raw task load index (RTLX) scores are comparable with and correlate to the traditional TLX scoring mechanism (Byers et al., 1989). A Friedman

Test showed significant differences between conditions $\chi^2(4) = 18.197$, p = .001, and a subsequent Wilcoxon Signed Ranks Test with a Bonferroni Correction (Z = -3.623, p = 0.00 (rounded to three decimal places)) showed that there was a significant difference only between Condition 2 and Condition 4. This means there were significant differences between the condition where users would move themselves to shift their perspective of the content being displayed via AR and the condition where users were purely manipulating the static (position) model with no AR while it was rigged to the center of the screen.

When teaching about something where the physical relationships are critically important, requiring users to shift themselves may help with their grasp of the subject material. If veterinary student must remember precisely where the rumen sits adjacent to the abomasum in a cow's abdomen, then having them examine this at scale while shifting their perspective to observe these organs might be helpful (*Illinois Cow*, n.d.).

It should be noted that for accessibility purposes, having the option for other manipulation techniques (e.g. touch screen, voice, etc.) is beneficial because beyond preference, some users may not be able to perform certain interaction techniques such as body movements. Touch screens may require precise finger movements whereas head/body movements are not as granular. Head/body movements may not be possible for users who cannot move in ways that are seen as more traditional. Precise finger movements may not be possible for users with dexterity issues. The point here is that AR affords multiple methods of interaction to share
and observe objects; providing users with more than one option allows them to use whatever mode of interaction is most salient to them.

The method of interaction highlighted by Condition 2 highlights some very interesting points that may be of use to educators working with spatial concepts and material. There are certainly times when requiring users to move themselves to navigate about material and potentially increasing their workload might be more burdensome than it needs to be; however, this interaction style should not be totally discounted. In fact, it may be the case that as AR becomes more ubiquitous, moving oneself around a digital object feels more natural than tapping and swiping on a screen—especially once the screen is gone and we are no longer interacting with AR content through the lens of a smartphone or tablet (as we move toward wearables such as AR glasses (Rauschnabel & Ro, 2016)). Ironically, this could shift our styles of interacting with objects closer toward the way that we naturally interact with physical objects in the real world. P6 said that while they are more used to touch screen interactions due to the interaction techniques that our phones commonly afford, changing the view is "as easy as just, you know, looking over versus getting my finger and like scrolling through thing...seems more natural."

Of course, with digital objects displayed using AR, users are not limited to the interactions that the physical world affords, such as moving around an object to see the other side or picking it up and turning the object itself. But there may be instances (particularly in educational and training scenarios) where nudging users to rely on the ways in which we interact with physical objects is beneficial.

It should also be noted that in this study, we were using target images (fiducial markers) to rig AR content in a scene, and these target images were displayed on the computer screen of users while participating in the videoconference call. In practice, experiences can still exploit proprioception in AR with printed target images displayed on paper (e.g. worksheets, textbooks, etc.), or using other types of tracking such as SLAM tracking (Liu et al., 2017) to track a user's environment and rig models to the physical space without using a predetermined image. With printed target images, users could pick up and move AR objects as they would with physical objects in the real world by picking up and manipulating the paper (or whatever object that the fiducial markers are printed on) that contains the target image.

4.2.2. Promoting User Engagement in Remote Collaborative Environments

Teleconferencing participant engagement is something that repeatedly came up throughout the pre-test interviews with our participants. Traditional teleconferencing calls while working, teaching, and learning remotely have left many participants feeling bored or feeling that others in their call are not engaged with the content being presented. In interviews, users reported that often in online classes, the majority of participants keep their cameras off and do not engage with the material being presented. All participants interviewed reported that using AR to pass and observe digital objects through the screen has potential for increasing engagement with the subject material.



Figure 19 User Engagement Scale (UES) Results

P13 said that one outcome of using AR during remote classes is that you focus attention to the presentation, and you also tie up the use of the phone so that people aren't texting or distracting themselves in other ways. An activity like this can keep people's attention by insisting focus on the presentation, encouraging user participation, and discouraging use of personal devices for other activities.

P14 noted "I definitely feel more excited if something like this was to appear in a presentation, as opposed to just sitting through all of these Zoom presentations...it could also be fatigue of the fact that I've been sitting in front of Zoom presentations for a year now. This is the first thing that's actually stood out as being different." Participants also voiced some other ideas for more engaging experiences. One point that came up in a small number of the user studies was the importance of the fiducial marker target images. In this study, there was no importance to the visual imagery on the target images, beyond the fact that they provided complex enough imagery to allow the computer vision to work to perform AR tracking. A few participants asked if there was significance to these target images, and it could be useful to curate relevant target images for a given experiences to establish a sense of continuity and immersion (this aligns with findings from Barakonyi et al. (Barakonyi et al., 2004)).

4.2.3. Promoting Co-presence in Remote Collaborative Environments

18/20 participants reported that their experience using SAROSvis changed the way that they felt about collaborating with a remote peer. Participants reported that they had never considered this method of interaction as a means for remote collaboration and were generally excited about the prospects of using a tool like this. Participants who are instructors were excited about this opportunity for their students, and participants who are students felt that this would bring back a level of engagement that has been lost in remote instruction.

P3 said that SAROSvis "could definitely clear up a bunch of misunderstandings... you can look at the same object in Google docs or something, but not really have a complete understanding of what you're both saying, or what you're trying to focus on, and so being able to point at it that way, manipulate it that way is helpful."

Participants shared that being able to see a visual representation of another user within their environment made them feel more co-present (Condition 3). Even though Condition 5 afforded the ability to manipulate an object and see those changes propagated to a remote user, without AR it felt less natural and some users even pointed out that it felt like a glitch, seeing an object moving without them doing anything to it when there is no AR. This was not the case when users could see the AR visual representation of the remote researcher working with them around the AR objects.

P7 said that SAROSvis "makes me want to actually collaborate with a peer... my problem with virtual learning is you can talk about it, or work on a Google Doc together, right, but you can't manipulate the same objects together or see your perspective of it, so I thought it made it [SAROSvis] made it a nicer to work with a peer. It didn't make it seem like such a task... I'll definitely work with someone if we were working like this, even if we were working in separate spaces, I think it'll bring back that interactive quality."

Participants also voiced that there were times when they felt that they were indeed sharing an object with the researcher conducting the experiment, and times when they did not feel this way. To be clear, participants said that there were times that they felt that they were looking at and manipulating the very same object that the remote researcher was manipulating, rather than merely each observing their own copies of the object. The participants overwhelmingly reported feeling this sense of sharing an object in Condition 3 and said that when they could see

manipulations happening in real time remotely while these manipulations were being narrated by the researcher (who was speaking aloud while manipulating the object), as well as the visual representation of the remote researcher, it truly felt that they were sharing an object.

P12 who works with data where the presentation of the physical nature is extremely important noted how great a tool like this would be for remote collaboration, and specifically took interest in the visual representations of each participant's pointer (the pencil in Figure 8). P12 said that seeing how another participant is observing an object is greatly important when working in-person, and this has been lost in remote conditions. The interactions displayed in condition 3 have potential to bring this sense of collaboration and communication cues to the remote workspace.

Participants who had teaching experience said that while the ability to remotely manipulate a shared object is useful, it is important to consider permissions of who can manipulate what, and when; this must be carefully considered to avoid a hectic experience where multiple people try to take interactions at the same time. This is partially a networking consideration, but also a UX consideration as we might for example consider ways to let a user know when they should or should not be manipulating a given model.

Spatial presence

The MEC spatial presence (Spatial Presence: Self Location subscale) questionnaire yielded results showing that there may likely be utility to using AR in remote collaborative scenarios. This scale was created to assess and operationalize spatial presence, and contains nine subscales, each of which aim to gather data around spatial presence and actions within a spatial environment (the full nine subscales include: Attentional Allocation, Spatial Situation model (SSM), Spatial Presence: Self Location, Spatial Presence: Possible Actions, Higher Cognitive Involvement, Suspension of Disbelief, Domain Specific Interest, Visual Spatial Imagery, Absorption) (Vorderer et al., n.d.). While this study was not assessing a VR experience, the Spatial Presence: Self Location subscale was utilized in order to consider how the participants felt about their sense of self relative to the data being presented and the AR objects with which they were interacting.

Running a Friedman test on the data in SPSS showed statistical significance depending on condition: $\chi^2(4) = 89.673$, p = 0.000 (rounded to three decimal places). In order to see the conditions between which this significance lies, a Wilcoxon Signed Ranks Test was performed on the data. Figure 20 shows the average scores from each condition of the spatial presence questionnaire, with a higher score meaning better sense of spatial presence relative to the AR objects. The AR conditions (1-3) outperformed the non-AR conditions (4 and 5), and the significant difference between conditions are reported in Table 6.



Figure 20 MEC Spatial Presence Questionnaire Results

The comparison between conditions using a Wilcoxon Ranked Sign test with

Bonferroni correction to the significance levels is reported in Table 5 below.

Condition	Z	Significance (rounded to 2 decimal places after
Comparison Pair		Bonferroni Correction)
C1-C4	-6.369	.00
C1-C5	-3.531	.00
C2-C3	-3.432	.01
C2-C4	-4.730	.00

C3-C4	-7.223	.00
C3-C5	-5.391	.00
C4-C5	-3.002	.03

Table 6 MEC Spatial Presence Questionnairie Wilcoxon Signed Ranks Test Statistics

The results presented in Table 6 are significant, and meaningful when we consider what each condition that displayed significant results afforded. Condition 4, which did not use AR visualization and did not allow for remote demonstration, essentially a local 3D model viewer with position-static models, showed significant differences in the MEC SPQ responses when compared with every other condition. Condition 4 was significantly ranked lower than all of the rest in terms of spatial presence, based on the results from this questionnaire. Condition 5 was similarly ranked significantly lower than almost all of the other conditions when looking at the data from the Wilcoxon Signed Ranks Test results, except there was no significant difference between Condition 5 and Condition 2. While this may not be particularly surprising, it is important to realize that in this study, AR did indeed afford for a more spatially present experience, and the AR conditions were preferred by the participants.

Participants reported enjoying the ability to see where the remote participant of the session was located in the environment relative to themselves, as well as where the AR object was displayed relative to themselves. Having objects

displayed via AR in the application made participants feel that their experience was more immersive, more interactive, and less like they were just looking at a 3D model using traditional methods (such as on a computer screen in a 3D model viewer).

A few participants talked about making sure that this experience is comfortable for all users and brought up that it could be important to consider power dynamics (teacher vs. student relationship) when designing an experience that shows visual representation of remote participants of said experience in each users' physical space.

Effects of AR

According to the MEC Spatial Presence Questionnaire, AR afforded for a more spatially apt experience and in the follow-up interviews, participants elucidated the aspects of AR which they felt were beneficial for remote collaborative scenarios. P5 talked about how seeing the content in AR rigged to their physical space was interesting: "when it was AR... I felt like it was transforming that little box [the device screen] into a separate plane of existence... where both you are in it and I am in it at the same time."

P11 noted that the fact that there are effects of moving their device camera (their perspective) made them feel more immersed in the experience. They said that they felt that their whole body was engaged, and the digital information is hooked to the actual (physical) environment.

This leads into ideas regarding bodily movements and AR. Some participants specifically stated that moving their body around in order to interact with the AR content felt engaging, and in some ways like a more natural way to interact with 3D data presented in AR. P14 stated that moving their phone around made it feel as though the AR content was actually there in their physical environment and found this quite engaging. They liked the fact that SAROSvis uses the slides as a method for rigging the AR content and driving the experience.

P8 reported that the conditions with AR (1, 2, and 3) were their favorites and felt that AR "was the best part... that you could actually see these objects augmented... and not just a 2D representation." Participants largely felt that having AR deepened immersion, and made the experience feel less like a traditional 3D viewer, and some participants talked about how the AR-enabled interaction of moving oneself around an object felt more natural. P9 felt that AR "deepened immersion to some extent... I think the line between AR and no AR is less important when I'm able to drag with my fingers. Even so, I think that those are intuitive gestures to have... we all know how a touch screen works...I think having to move my phone to a specific space to see the object was interesting. When I could only do that [Condition 2] I liked that more...in general AR did make it feel more interactive."

Feelings of Connectedness

Participants were asked if they felt connected with their partner ("I felt connected with my partner") in a Likert Scale-style question (1- Strongly Disagree to

5- Strongly Agree) after engaging with each condition. The results from this question are unsurprising but telling; each of the conditions that involved remote demonstration (Condition 3 and Condition 5) showed significant difference from conditions which did *not* involve remote demonstration. Table 7 shows the descriptive statistics from each condition with averages from the participant responses. A Friedman test on the data in SPSS showed statistical significance depending on condition: $\chi^2(4) = 51.702$, p = 0.000 (rounded to three decimal places).



Figure 21 Feelings of Connectedness descriptive statistics chart

Condition	N	Mean	Std.	Min.	Max
			Deviation		
1	20	2.5500	1.31689	1	5
2	20	2.8000	1.15166	1	5
3	20	4.5000	.51299	4	5
4	20	2.1500	1.18210	2	5
5	20	4.1500	1.03999	1	5

Table 7 Feelings of Connectedness Descriptive Statistics

Table 8 shows the results from the Wilcoxon Signed Ranks Test to determine the level of significance of the difference between each condition, with Bonferroni correction to the p values.

Condition	Z	Significance (rounded to 2 decimal
Comparison		places after Bonferroni Correction)
Pair		
C3-C1	-3.542	.00
	0.555	
C3-C2	-3.775	.00

C3-C4	-3.850	.00
C5-C1	-3.218	.01
C5-C2	-3.257	.01
C5-C4	-3.745	.00

Table 8 Feelings of Connectedness Wilcoxon Signed Ranks Test Results

These results show us that the conditions where remote demonstration was available may have helped participants feel more connected with their remote peer (the researcher in this scenario). This general finding aligns with information from the literature review on remote AR solutions. Different from most existing solutions however, in this project remote partners were performing manipulations on networked digital objects shared between them through a slide displayed on their screen. Previous work, such as the work of Barakonyi et al., differs in that their ARsupported video call overlayed AR objects onto fiducial markers using the Magic Mirror paradigm, and each participant of a call was displayed a non-networked object (Barakonyi et al., 2004).

Shared Objects

One of the questions that participants answered in the questionnaires after testing each condition asked "I felt as though my partner and I were manipulating one shared object" in a Likert Scale-style question. A Friedman test on the data in SPSS showed statistical significance depending on condition: $\chi^2(4) = 62.967$, p =

0.000 (rounded to three decimal places). A Wilcoxon Signed Ranks Test with Bonferroni correction showed that Condition 3 and Condition 5 felt significantly more like users were sharing an object when compared with any of the other conditions (1, 2, or 4). There was no significant difference between Condition 3 and Condition 5 after the Bonferroni correction for the p value—this is not surprising due to the fact that there is more in common between these two conditions than either of them has in common with any of the other conditions.



Figure 22 Sense of Sharing the Same Object amongst Remote Collaborators. Mean scores for each Condition.

Condition	Ν	Mean	Std.	Min.	Max
			Deviation		
1	20	2.1000	1.11921	1	4
2	20	2.2500	1.20852	1	5
3	20	4.8000	.41039	4	5
4	20	1.8500	1.18219	1	5
5	20	4.4000	.75394	2	5

Table 9 Sense of sharing an object descriptive statistics.

Condition	Z	Significance (rounded to 2 decimal
Comparison		places after Bonferroni Correction)
Pair		
C3-C1	-3.863	0.00
C3-C2	-3.764	0.00
C3-C4	-3.895	0.00
C5-C1	-3.850	0.00

C5-C2	-3.769	0.00
C5-C4	-3.852	0.00

Table 10 Feelings of Sharing an Object Wilcoxon Signed Ranks Test between conditions with significant differences.

This sense of sharing an object is important to collaborative scenarios. During the follow-up interviews, participants were able to elucidate some of the specific things made them feel as though they were sharing an object with their remote collaborator, and how this impacted their experience. P18 said that this is a feature that would come in handy for them right now, as they are dealing with lab fixture set ups while working remotely. P18 said "sometimes I'm saying 'beaker' and he has to go like 'oh what size is your beaker, like what are your dimensions' and I'm like 'I don't know off the top my head...' I have an idea how I should set it up, but it's a little difficult to describe over words really... easier to show that way...when it comes to collaborating and working on the same thing."

Visual Representation of Remote Peer

In the SAROSvis application, the visual representation of the remote peer was represented by a pencil floating in space (as seen in Figure 6). The pencil was chosen due to this experience revolving around educational topics and a writing implement is relevant, but more importantly, the pencil has a point which can indicate direction. The position of the pencil was tracked and constantly updated to the exact location of the remote peer relative to the target image and AR model, and so was the direction that the remote peer was looking toward. These were always updated in real time (at least as quickly as the network would afford).

16/20 participants reported that the pointer affected the experience they had during the study with the SAROSvis application. All 20/20 participants reported that they could see the benefit of using a remote pointer, or some visualization of a remote peer, in remote collaborative environments. The pointer, which was used in Condition 3 to represent the remote participant's location and rotation in space relative to the other participant and the AR objects, was received well by participants. Condition 3 and Condition 5 were good foils to each other, as they both offered shared objects, but Condition 3 used AR and the pointer representation of the remote peer, while Condition 5 did not use AR and had no representation of the remote peer. It seems that this visual representation is necessary for this sense of sharing an object, and this works in conjunction with the verbal cues described in the next subsection as well. P4 said "I feel like that one with a pencil [Condition 3] will help me know that's another person moving it [the AR object]... without that it feels... sometimes I wonder if I'm lagging ... maybe there's some delay and I moved I moved it in somehow it didn't work." P4 is comparing Condition 3, where they feel confident that the remote peer is manipulating the shared object due to the pointer, to Condition 5, where the manipulations taken by the remote peer are seen on the object, which is moving, but there is no representation of the remote peer, so it feels random and disconnected.

A couple participants mentioned that it could be jarring to see another person (or rather, the representation of another person) in one's physical space, if one is not expecting this. It should be made clear exactly what the visual pointer is and how it operates to avoid potential confusion.

Alignment between Verbal and Visual Presented Information

6/20 participants mentioned (without being specifically asked about it) in the interviews that the verbal (the researcher talking whilst manipulating the objects in Conditions 3 and 5) made them feel more connected and like they were sharing an object with the presenter rather than exploring their own copy of the object. While this is not a majority of the participants, this was not asked about specifically in the interview by the researcher, so this frequency with which it did get broached is worth paying attention to.

Participants felt that the verbal information helped solidify the feeling that they were handling the same objects (rather than each having their own copies) as their remote collaborators during the networked conditions (Condition 3 and Condition 5). This aligns with findings from Kim et al. (Kim et al., 2014). Gasques et al. also found that verbal cues along with gestures cues were helpful for collaborative AR scenarios, and even helped account for possible alignment errors of AR content (Gasques et al., 2021).

Chapter 5.

Discussion

5.1. Recommendations

The results from this work show promising insights into the ways that AR can be leveraged to enhance the ways in which we interact with one another while collaborating remotely, and how we can provide a feeling of working together around sharing objects through the screen. The intention of this work is not to extrapolate behaviors to a larger audience, but rather to examine how AR could affect a remote collaborative scenario and in turn, infer design recommendations for future iterations and related work.

I believe that there is a strong case for using AR to support remote collaborations, and for education especially, the benefits of doing so should not be overlooked. In this section, I will outline some points for consideration that should be taken into account when working on projects of this type, as understanding how to optimize for each of the design considerations based on project needs will make for a more effective user experience. In each subsection, I will provide more detail on how one might approach each design consideration in various ways.

Incorporating the ability to pass objects through the screen (virtually) to remote collaborators can potentially help remote collaborators better understand spatially relevant data, engage with subject material and their surroundings in novel

ways, and engage with subject material more enthusiastically. Table 11 gives a title and brief description of the design considerations that I will describe further in the subsequent subsections.

Design consideration	Description
Shared (networked) objects	AR objects appearing in each participants' respective scene which are connected across the network, updating translations and transformations for each party to maintain consistency
Permissions	Determination of actions that each participant can take
Method of manipulation	How the user can manipulate the AR content (viewing angle, translations, rotations, etc.)
Visual representations of remote peers	How the remote participants appear to one another relative to each other and other virtual objects in the scene
Context-appropriate fiducial markers	How closely the material displayed in the target image itself relates to the context and AR objects

Customizations and data	Changes that users can make to AR objects in a scene
persistence	with remote collaborators, and the data that is saved
	for a future session after a session is concluded

Table 11 Considerations for designing remote AR collaborative experiences

5.1.1. Shared (Networked) Objects

Shared objects were a main focus of this project. In this case, the shared objects were models using networking protocols to update the transformations and manipulations of the model for all users in real time. The ability to pass a digital object through the screen from one participant to another using AR in a remote collaborative setting is a powerful metaphor that has not been adequately explored. Allowing users to create and share objects with one another using AR, and subsequently manipulate and customize these objects together, has interesting implications for the future of remote collaboration, and showed compelling results in this study.

When creating a remote collaborative environment leveraging AR, I recommend using this ability to share objects through the screen and manipulate them together in some way. While designing an experience, one can decide what features of the object should be updated (position, rotation, scale, etc.) based on the needs of the scenario. I do believe that it is nontrivial to share objects using target images as well. This not only leans on existing methods of communication that are currently in use (sharing slide decks), but it also gives the driver of an experience immediate control over what objects are being shared through the screen

depending on the target image being displayed. From there, it is technically possible to use other types of tracking (such as SLAM tracking) to rig AR models to a user's environment, not using target images, or using other targets in a user's environment such as pages of a book or printed worksheets. This could be used for more independent study of the AR models, off network; however, for the purposes of remote collaboration in real time, creating networked assets in an experience, making sure that the network is handling the manipulations in a sensible way during the session, and sharing objects through the screen using target images can create an engaging and novel experience for remote collaborators.

Object fidelity should also be considered, as scale and detail can be important in certain cases. For example, in the scenario that P18 described with a lab partner setting up lab fixtures, the size of the beakers was important to their discussion, so making sure that the AR content has accurate dimensions is crucial.

5.1.2. Permissions

Many participants brought up the fact that they would like some sort of permissions handling in a remote collaborative experience, particularly around the shared objects that multiple users can interact with. While for this study, each participant was participating in a scenario with only one remote collaborator (me, the researcher), they saw that it could potentially be confusing if there are more users in an environment and multiple people try to manipulate the same object at the same time. While working with digital objects in remote collaborative scenarios, it is important to consider how to handle multiple participants manipulating the

same object at the same time. P10 said that it could be beneficial to "pass on the control like we do it with Zoom screen sharing, like just one person can do it at a time, so maybe you have some control options, with the object as well, so not anyone can just pick it up."

This issue should be considered from a technical perspective, as well as from a user experience perspective. Technically, this is not a new type of issue; this type of networking reconciliation of data is handled in spaces such as online video games. All players in a scene need to render the same objects so there is consistency and synchronization amongst each instance of the application running. Depending on the client-server model that a systems protocol is operating on, designers can decide how to determine which manipulation is handled first; Lee et al. details one solution, lockstep protocol, for handling multiplayer networking in games to avoid cheating (Lee et al., 2003).

From the user experience perspective, it should be clear to participants of an experience who can manipulate objects, which objects can be manipulated by them, and when they are able to impart manipulations upon objects. Whether this is handled by one participant in the scenario taking a host or master role where they can toggle manipulation abilities for other users, or buttons that each user can press to take turns and pass the proverbial torch to the user who is next in line to demonstrate a manipulation. There are many permutations of ways to consider communicating the permissions to users.

5.1.3. Method of Manipulation

The main methods of manipulation studied in this project were on-screen tapping and swiping, compared with proprioceptive movements of the body to shift perspective of AR content being displayed. When creating AR applications for remote collaboration, I recommend providing the ability to interact with the AR objects and the environment in multiple ways, but also potentially nudging toward certain interactions depending on the use case. To be more specific, it will behoove a designer of an experience to build in multiple methods of interaction due to some methods being more comfortable for some users than others. In this study, some participants reported that moving themselves physically was intriguing and felt like a more natural way to interact with the AR objects, but some participants did not enjoy this and preferred on-screen touch manipulations to manipulate the AR objects. Nudging users toward a specific method of input may be beneficial depending on the scenario. If spatial awareness around the content presented in AR is important to the experience, then telling users to shift themselves about the AR objects may be beneficial—however, for users that do not enjoy this type of interaction or cannot use this method of interaction due to accessibility, having other methods built in is beneficial.

As described in section 2.1 Interaction in XR, there are input mechanisms beyond touch screen input, or physical movements which change the viewing angle. Voice commands, eye tracking, hand gestures, and more are input mechanisms that can be used to interact with AR content. The important aspect is considering desired outcomes of an experience as well as user preferences for interaction technique.

5.1.4. Visual representations of remote peers

The visual representation of remote peers was met with excitement from participants in this study. Participants also had some suggestions for how to make this feature even stronger. In the beginning of remote instruction P17, a high school teacher, said that there was a big push for teachers in their school to use Bitmojis, personalized avatar emojis, to make for a more personalized experience between remote partners who are no longer able to collaborate face to face. Allowing for more customization, even avatars, could be an engaging way to provide represent remote users to others. If users can choose from a list, or even customize their own avatars, they can decide what works best for them in a given scenario. Also, these visual avatars can be used to signify things about users, such as the role that they hold in a given scenario, or information about who hold certain permissions to manipulate AR content.

Craig et al. describes the use of avatars in remote collaborative virtual environments: "when working in a networked collaborative environment, each participant can be represented as a virtual entity ... The mere presence of avatars can greatly improve the ability of the collaborators to communicate through nonverbal means. For example, pointing in a direction, waving an arm, or even just looking in a certain direction can convey valuable information from person to person," (Craig et al., 2009). The representations of the remote peer can be used to indicate certain

things and relay certain information. Avatars can communicate the permissions that a user has, the role that they occupy, name, etc. Various colors, shapes, and objects can be used for avatars to communicate information. Avatar shape may be important to an experience as pointing out parts of a shared object may be important to an experience (which is why for this study, the pointed pencil model was used). There could also be use cases where separate visual objects represent where remote participants are looking and where they are pointing (e.g., a face and a laser pointer).

5.1.5. Context-appropriate fiducial markers

A couple participants brought up the target images containing the fiducial markers and commented on the appearance of these images. P13 noted that sometimes the contrast between the AR objects (colors) and the target images was hard to see. Making sure that the colors of the target images do not clash with the AR content is important.

Some participants also wondered if the target images were related to the content being displayed in AR. In this study, they were not related, however this could be another way to create a more engaging experience. Tying in the target images and making them somehow contextually relevant may be a good way to create cohesion in the experience, and if all relevant information can be presented in the target image itself, this can prevent users from needing to glance at other documents for reference during a session.

5.1.6. Customizations and Data Persistence

14/20 participants mentioned (of their own accord without being specifically asked) that they would like to have the ability to add some type of annotations to the AR content; for example, P7 noted that adding in an annotation system in AR or some visual aid for remote demonstrations would help the visual communication aspect of seeing the visual representation of the remote partner (in Condition 3) become even stronger.

P3 said "more detailed manipulation... if you could have written or like drawn on the object... made a change to it in some way, I think that would increase that feeling [of sharing an object] more... like break parts off and move and focus on that if you wanted to."

A few participants broached the desire to make changes to the AR content, and then save these changes in some way to be revisited later. P17, a high school instructor, said that creating a persisting environment where students could meet to work together remotely and save their work between sessions could enhance the sense of remote collaboration. Being able to collaboratively create, edit, and revisit an object/space may help create a vested interest and sense of camaraderie between remote participants. This can also be a way to facilitate assignments, where students receive a model, enter breakout rooms with smaller groups to work together analyzing, annotating, and manipulating the model in AR, and finally save and export their finished product.

There are also different types of persistence that can be explored. One type involves having an environment that saves state between sessions, and users can revisit the environment which maintains changes from previous sessions. Another method of persistence involves an evolving state, which exists and evolves between sessions as well, while users are *not* in the environment (such as the NICE project (Johnson et al., 1998)). Both of these levels of persistence can be useful depending on the scenario, as they can promote a different type of user interaction and connection with the virtual space/objects.

5.2. Limitations

I am encouraged by the results from the study, but I would like to take some time for the important step of recognizing limitations to this work and understanding how they might have affected the study. This can also shed light on opportunity areas for further research in future work. Firstly, working with more participants would have allowed us to exhaustively test each order combination of the 5 conditions. This would have required quite a few more participants—120 total to be exact. It is for this reason that I used random order counterbalancing with the participants in this study. The data would most likely reach a saturation point well before testing with 120 participants, especially the qualitative portions.

I would have also liked to test the design beyond 1:1 remote collaborative scenarios which were the focus for this study. For the purposes of this study, the researcher-participant dyad was sufficient and provided plenty of rich data, however a compelling future study would examine the conditions presented here with groups to investigate small and large group dynamics around an application like SAROSvis. This could help glean information on potential teacher-student dynamics and remote class structure behaviors around these interaction techniques.

Leading into future work, I would like to iterate further on the designs presented in this work and test these iterations with participants, either longitudinally to assess how recurring participants behave and react to the changes, or with new groups of participants each time.

5.3. Future Work

I am very excited about the future of this work, and the projects that will stem from it. At least two things seem highly likely for the future: remote collaboration isn't going away, and experiences leveraging AR will become more compelling as the technology supporting them continues to advance. As the results from this work are encouraging and have potential to shift the ways in which we think about remote collaboration and communication, I plan to continue to iterate on the designs presented in this work, continue to add interactions and features, and consider ways that an experience such as the one investigated here can be even more accessible and available to educators and students.

It is important to continue to study the needs of remote collaborators in as many scenarios as possible to understand where AR can best benefit them while working remotely. This work provides good information on the types of interactions that can improve a sense of spatial awareness and co-presence around AR content, and further understanding the types of work that people are doing remotely, particularly in education, will help designers create even more targeted applications leveraging AR for remote collaboration. Further study into specific online classrooms, different age groups, and various subjects more is necessary.

I would like to make changes to SAROSvis based on the feedback from this study and continue to collect data from user research studies in order to continue iterating on the design. Some of the features that participants recurringly broached include an annotation system (which is something I wrote about in the paper on pdbARbrowser (Collins & Craig, 2017b)), permissions around shared objects, being able to make changes and customizations to the model being presented in AR with a remote peer, and have these changes persist over time (after closing the app and opening during the next session). These are features that I am actively working on, and plan to study in future sessions.

I am also excited for the future of technologies that will support more methods of AR interaction, such as wearables like AR-enabled glasses that will afford new ways of interacting with AR (Rauschnabel & Ro, 2016). Once things like this are more ubiquitous and easier to access, touch inputs on touch screens to interact with AR content will likely begin to fade away, at least in the ways that we see it today. This is why I think it is important to continue studying proprioceptive bodily movements while interacting with AR objects, as this takes us closer toward the idea that the AR objects can be observed much like physical objects in the physical world.

I am excited by the results from this work, and I think the metaphor of passing objects to one another to manipulate together (or demonstrate on) is very strong. There is a great deal more to explore in this space, but I believe this is a step in a direction that has been relatively underexplored. I eagerly look forward to helping build the future of collaboration, and I think that as remote collaboration continues to mature, leveraging AR to interact with content, and in turn each other, will be of vital importance.

References

5.2 Experimental Design – Research Methods in Psychology. (n.d.). Retrieved April 7, 2021, from https://opentext.wsu.edu/carriecuttler/chapter/experimental-design/

Abrash, M. (2017, October 11). Inventing the Future.

https://www.oculus.com/blog/inventing-the-future/

- Azuma, R. T. (1997). A Survey of Augmented Reality. *Presence: Teleoperators and Virtual Environments*, 6(4), 355–385. https://doi.org/10.1162/pres.1997.6.4.355
- Bai, H., Sasikumar, P., Yang, J., & Billinghurst, M. (2020). A User Study on Mixed
 Reality Remote Collaboration with Eye Gaze and Hand Gesture Sharing.
 Proceedings of the 2020 CHI Conference on Human Factors in Computing
 Systems, 1–13. https://doi.org/10.1145/3313831.3376550
- Baker, M. P., & Wickens, C. D. (1995). *Human Factors in Virtual Environments for the Visual Analysis of Scientific Data*.
- Barakonyi, I., Fahmy, T., & Schmalstieg, D. (2004). *Remote Collaboration Using* Augmented Reality Videoconferencing. 8.

Barreiros, C., Pammer, V., Veas, E., & Oberbichler, U. (2017). What If Factories Looked Like Forests? Redesigning the Manufacturing Industry 4.0 Workplaces with an Augmented Reality Inspired Nature Metaphor. 4.

Bauer, M., Kortuem, G., & Segall, Z. (1999). "Where are you pointing at?" A study of remote collaboration in a wearable videoconference system. *Digest of Papers. Third International Symposium on Wearable Computers*, 151–158.
https://doi.org/10.1109/ISWC.1999.806696

bigscreen inc. (2016). *Bigscreen*. https://www.bigscreenvr.com/ Billinghurst, M. (2002). *Augmented Reality in Education*. 5.

- Billinghurst, M., & Kato, H. (2002). Collaborative augmented reality. *Communications of the ACM*, *45*(7), 64–70. https://doi.org/10.1145/514236.514265
- Billinghurst, M., Kato, H., & Poupyrev, I. (2001). *The MagicBook—Moving Seamlessly between Reality and Virtuality* _____. 3.
- Biocca, F., & Levy, M. R. (2013). *Communication in the Age of Virtual Reality*. Routledge.
- Blackwell, L., Ellison, N., Elliott-Deflo, N., & Schwartz, R. (2019). Harassment in Social Virtual Reality: Challenges for Platform Governance. *Proceedings of the ACM on Human-Computer Interaction*, 3(CSCW), 1–25. https://doi.org/10.1145/3359202
- Boudourides, M. A. (2003). Constructivism, Education, Science, and Technology. *Canadian Journal of Learning and Technology / La Revue Canadienne de l'apprentissage et de La Technologie*, *29*(3). https://www.cjlt.ca/index.php/cjlt/article/view/26537
- Brown, L. D., Hua, H., & Gao, C. (2003). composite augmented visualization. A Widget Framework for Augmented Interaction in SCAPE.
- Buchmann, V., Violich, S., Billinghurst, M., & Cockburn, A. (2004). FingARtips: Gesture based direct manipulation in Augmented Reality. *Proceedings of the 2nd International Conference on Computer Graphics and Interactive Techniques in Austalasia and Southe East Asia GRAPHITE '04*, 212. https://doi.org/10.1145/988834.988871

Byers, J. C., Bittner, A., & Hill, S. (1989). Traditional and raw task load index (TLX) correlations: Are paired comparisons necessary? In A. *Undefined*.
/paper/Traditional-and-raw-task-load-index-(TLX)-Are-In-A-ByersBittner/ae5d6de023e2619bec8edcbd0234ef1777e65f2a

Cahill, N. (2019). Nancy Baker Cahill Artworks. *Nancy Baker Cahill*. https://nancybakercahill.com/artwork Chen, Y.-C. (2006). A Study of Comparing the Use of Augmented Reality and Physical Models in Chemistry Education. Proceedings of the 2006 ACM International Conference on Virtual Reality Continuum and Its Applications, 369–372. https://doi.org/10.1145/1128923.1128990

Christensson, P. (2014, November 7). WIMP (Windows, Icons, Menus, Pointer) Definition. https://techterms.com/definition/wimp

Cobb, P., Yackel, E., & Wood, T. (1992). A Constructivist Alternative to the Representational View of Mind in Mathematics Education. *Journal for Research in Mathematics Education*, 23(1), 2–33. https://doi.org/10.2307/749161

- Collins, M., & Craig, A. B. (2017a). VisMo: Augmented Reality Visualization of Scientific
 Data and Molecular Structures. *The Journal of Computational Science Education*, 8(1), 12–15. https://doi.org/10.22369/issn.2153-4136/8/1/3
- Collins, M., & Craig, A. B. (2017b). VisMo: Augmented Reality Visualization of Scientific
 Data and Molecular Structures. *The Journal of Computational Science Education*, 8(1), 12–15. https://doi.org/10.22369/issn.2153-4136/8/1/3
- Copolo, C. E., & Hounshell, P. B. (1995). Using three-dimensional models to teach molecular structures in high school chemistry. *Journal of Science Education and Technology*, *4*(4), 295–305. https://doi.org/10.1007/BF02211261
- Couclelis, H. (1996). The Death of Distance. *Environment and Planning B: Planning and Design*, 23, 387–389.

Craig, A. B. (2013a). Chapter 2—Augmented Reality Concepts. In Understanding Augmented Reality (pp. 39–67). Morgan Kaufmann. https://doi.org/10.1016/B978-0-240-82408-6.00002-3

Craig, A. B. (2013b). Chapter 6—Interaction in Augmented Reality. In Understanding Augmented Reality (pp. 185–207). Morgan Kaufmann. https://doi.org/10.1016/B978-0-240-82408-6.00006-0 Craig, A. B. (2013c). Understanding Augmented Reality: Concepts and Applications. Newnes.

Craig, A. B., Sherman, W. R., & Will, J. D. (2009). *Developing Virtual Reality Applications: Foundations of Effective Design*. Elsevier. https://doi.org/10.1016/C2009-0-20103-6

Cruz-Neira, C., Sandin, D. J., DeFanti, T. A., Kenyon, R. V., & Hart, J. C. (1992, June).
 The CAVE: audio visual experience automatic virtual environment.
 Communications of the ACM, *35*(6), 64+. Gale Academic Onefile.

Designing Safe Spaces for Virtual Reality. (2020, January 1). *Facebook Research*. https://research.fb.com/publications/designing-safe-spaces-for-virtual-reality/

- *Elements 4D Interactive Blocks.* (n.d.). Kickstarter. Retrieved January 10, 2020, from https://www.kickstarter.com/projects/daqri/elements-4d-interactive-blocks
- Esteves, M., Fonseca, B., Morgado, L., & Martins, P. (2011). Improving teaching and learning of computer programming through the use of the Second Life virtual world. *British Journal of Educational Technology*, *42*(4), 624–637. https://doi.org/10.1111/j.1467-8535.2010.01056.x

Evans, G., Miller, J., Pena, M. I., MacAllister, A., & Winer, E. (2017). Evaluating the Microsoft HoloLens through an augmented reality assembly application. *Degraded Environments: Sensing, Processing, and Display 2017*, 10197, 101970V. https://doi.org/10.1117/12.2262626

Faria, J. W. V. de, Teixeira, M. J., Júnior, L. de M. S., Otoch, J. P., & Figueiredo, E. G. (2016). Virtual and stereoscopic anatomy: When virtual reality meets medical education. *Journal of Neurosurgery*, *125*(5), 1105–1111. https://doi.org/10.3171/2015.8.JNS141563

Feng Zhou, Duh, H. B.-L., & Billinghurst, M. (2008). Trends in augmented reality tracking, interaction and display: A review of ten years of ISMAR. 2008 7th
IEEE/ACM International Symposium on Mixed and Augmented Reality, 193–202. https://doi.org/10.1109/ISMAR.2008.4637362

- Force Field. (2018). Anne Frank House VR on Oculus Go. Oculus. https://www.oculus.com/experiences/go/1596151970428159/
- Freudenthal, H. (1981). Major Problems of Mathematics Education. *Educational Studies in Mathematics*, *12*(2), 133–150.
- Friedman Test in SPSS Statistics—How to run the procedure, understand the output using a relevant example | Laerd Statistics. (n.d.). Retrieved June 2, 2021, from https://statistics.laerd.com/spss-tutorials/friedman-test-using-spss-statistics.php
- Fung, J., Richards, C. L., Malouin, F., McFadyen, B. J., & Lamontagne, A. (2006). A Treadmill and Motion Coupled Virtual Reality System for Gait Training Post-Stroke. *CyberPsychology & Behavior*, 9(2), 157–162. https://doi.org/10.1089/cpb.2006.9.157
- Gasques, D., Johnson, J. G., Sharkey, T., Feng, Y., Wang, R., Xu, Z. R., Zavala, E.,
 Zhang, Y., Xie, W., Zhang, X., Davis, K., Yip, M., & Weibel, N. (2021). ARTEMIS:
 A Collaborative Mixed-Reality System for Immersive Surgical Telementoring. *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*, 1–14. https://doi.org/10.1145/3411764.3445576
- Gauglitz, S., Nuernberger, B., Turk, M., & Höllerer, T. (2014a). In touch with the remote world: Remote collaboration with augmented reality drawings and virtual navigation. *Proceedings of the 20th ACM Symposium on Virtual Reality Software* and Technology, 197–205. https://doi.org/10.1145/2671015.2671016
- Gauglitz, S., Nuernberger, B., Turk, M., & Höllerer, T. (2014b). World-stabilized annotations and virtual scene navigation for remote collaboration. *Proceedings of the 27th Annual ACM Symposium on User Interface Software and Technology*, 449–459. https://doi.org/10.1145/2642918.2647372

Gerry, L. J. (2017). Paint with Me: Stimulating Creativity and Empathy While Painting with a Painter in Virtual Reality. *IEEE Transactions on Visualization and Computer Graphics*, 23(4), 1418–1426. https://doi.org/10.1109/TVCG.2017.2657239

- Google. (n.d.). *Tour Creator*. Tour Creator. Retrieved January 14, 2020, from https://arvr.google.com/tourcreator/
- Guha, D., Alotaibi, N. M., Nguyen, N., Gupta, S., McFaul, C., & Yang, V. X. D. (2017).
 Augmented Reality in Neurosurgery: A Review of Current Concepts and
 Emerging Applications. *Canadian Journal of Neurological Sciences*, *44*(3), 235–245. https://doi.org/10.1017/cjn.2016.443

 Gurevich, P., Lanir, J., Cohen, B., & Stone, R. (2012). TeleAdvisor: A versatile augmented reality tool for remote assistance. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 619–622. https://doi.org/10.1145/2207676.2207763

- Habraken, C. L. (1996). Perceptions of chemistry: Why is the common perception of chemistry, the most visual of sciences, so distorted? *Journal of Science Education and Technology*, *5*(3), 193–201. https://doi.org/10.1007/BF01575303
- Harris, K. R., & Alexander, P. A. (1998). Integrated, Constructivist Education: Challenge and Reality. *Educational Psychology Review*, *10*(2), 115–127. https://doi.org/10.1023/A:1022169018926

Hart, S. G., & Staveland, L. E. (1988). Development of NASA-TLX (Task Load Index):
Results of Empirical and Theoretical Research. In P. A. Hancock & N. Meshkati (Eds.), *Advances in Psychology* (Vol. 52, pp. 139–183). North-Holland.
https://doi.org/10.1016/S0166-4115(08)62386-9

Hinckley, K., Pausch, R., Goble, J. C., & Kassell, N. F. (1994). Passive real-world interface props for neurosurgical visualization. *Proceedings of the SIGCHI* Conference on Human Factors in Computing Systems, 452–458.

https://doi.org/10.1145/191666.191821

- HoloKit is like Google Cardboard for augmented reality. (2017, June 2). *TechCrunch*. https://social.techcrunch.com/2017/06/02/holokit-is-like-google-cardboard-foraugmented-reality/
- Hu, R., Wu, Y.-Y., & Shieh, C.-J. (2016). Effects of Virtual Reality Integrated Creative Thinking Instruction on Students' Creative Thinking Abilities. *EURASIA Journal of Mathematics, Science and Technology Education*, 12. https://doi.org/10.12973/eurasia.2016.1226a
- Humphrey, W., Dalke, A., & Schulten, K. (1996). VMD Visual Molecular Dynamics. Journal of Molecular Graphics, 14, 33–38.
- Illinois Cow. (n.d.). Retrieved June 14, 2021, from https://vetmed.illinois.edu/illinois-cow/

Immersive VR Education Ltd. (2016). Apollo 11 VR on Oculus Rift. Oculus.

https://www.oculus.com/experiences/rift/937027946381272/

Introducing 'Facebook Horizon,' a New Social VR World, Coming to Oculus Quest and the Rift Platform in 2020. (2019). https://www.oculus.com/blog/introducingfacebook-horizon-a-new-social-vr-world-coming-to-oculus-quest-and-the-riftplatform-in-2020/

Introduction | Photon Engine. (n.d.). Retrieved June 13, 2021, from https://doc.photonengine.com/en-us/pun/current/getting-started/pun-intro

- Ishii, H., & Ullmer, B. (1997). Tangible bits: Towards seamless interfaces between people, bits and atoms. Proceedings of the SIGCHI Conference on Human Factors in Computing Systems - CHI '97, 234–241. https://doi.org/10.1145/258549.258715
- Johnson, A., Roussos, M., Leigh, J., Vasilakis, C., Barnes, C., & Moher, T. (1998). The NICE project: Learning together in a virtual world. *Proceedings. IEEE 1998*

Virtual Reality Annual International Symposium (Cat. No.98CB36180), 176–183. https://doi.org/10.1109/VRAIS.1998.658487

- Jones, B. R., Benko, H., Ofek, E., & Wilson, A. D. (2013). IllumiRoom: Peripheral projected illusions for interactive experiences. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 869–878. https://doi.org/10.1145/2470654.2466112
- Jones, L., Jordan, K., & Stillings, N. (2005). Molecular Visualization in Chemistry
 Education: The Role of Multidisciplinary Collaboration. *Chem. Educ. Res. Pract.*,
 6. https://doi.org/10.1039/B5RP90005K
- Kaplan, A. M., & Haenlein, M. (2009). The fairyland of Second Life: Virtual social worlds and how to use them. *Business Horizons*, *52*(6), 563–572. https://doi.org/10.1016/j.bushor.2009.07.002
- Kaufmann, H., & Schmalstieg, D. (2002). Mathematics and geometry education with collaborative augmented reality. ACM SIGGRAPH 2002 Conference Abstracts and Applications, 37–41. https://doi.org/10.1145/1242073.1242086
- Kawazoe, A., Luca, M. D., & Visell, Y. (2019). Tactile Echoes: A Wearable System for Tactile Augmentation of Objects. 2019 IEEE World Haptics Conference (WHC), 359–364. https://doi.org/10.1109/WHC.2019.8816099
- Keefe, D. F., & Isenberg, T. (2013). Reimagining the Scientific Visualization Interaction Paradigm. *Computer*, *46*(5), 51–57. https://doi.org/10.1109/MC.2013.178
- Kesim, M., & Ozarslan, Y. (2012). Augmented Reality in Education: Current Technologies and the Potential for Education. *Procedia - Social and Behavioral Sciences*, 47, 297–302. https://doi.org/10.1016/j.sbspro.2012.06.654
- Kim, S., Lee, G., Sakata, N., & Billinghurst, M. (2014). Improving co-presence with augmented visual communication cues for sharing experience through video

conference. 2014 IEEE International Symposium on Mixed and Augmented Reality (ISMAR), 83–92. https://doi.org/10.1109/ISMAR.2014.6948412

- Klopfer, E., Squire, K., & Jenkins, H. (2002). Environmental Detectives: PDAs as a window into a virtual simulated world. *IEEE International Workshop on Wireless* and Mobile Technologies in Education Proceedings, 95–98. https://doi.org/10.1109/WMTE.2002.1039227
- Kozma, R. B., & Russell, J. (1997). Multimedia and understanding: Expert and novice responses to different representations of chemical phenomena. *Journal of Research in Science Teaching*, *34*(9), 949–968. https://doi.org/10.1002/(SICI)1098-2736(199711)34:9<949::AID-TEA7>3.0.CO;2-U
- Kunert, A., Kulik, A., Beck, S., & Froehlich, B. (2014). Photoportals: Shared references in space and time. Proceedings of the 17th ACM Conference on Computer Supported Cooperative Work & Social Computing - CSCW '14, 1388–1399. https://doi.org/10.1145/2531602.2531727
- Lages, W. S., & Bowman, D. A. (2018). Move the Object or Move Myself? Walking vs. Manipulation for the Examination of 3D Scientific Data. *Frontiers in ICT*, *5*. https://doi.org/10.3389/fict.2018.00015
- Lau, K. W., & Lee, P. Y. (2015). The use of virtual reality for creating unusual environmental stimulation to motivate students to explore creative ideas. *Interactive Learning Environments*, 23(1), 3–18. https://doi.org/10.1080/10494820.2012.745426
- Lee, H., Kozlowski, E., Lenker, S., & Jamin, S. (2003). Multiplayer Game Cheating Prevention with Pipelined Lockstep Protocol. In R. Nakatsu & J. Hoshino (Eds.), *Entertainment Computing: Technologies and Application* (pp. 31–39). Springer US. https://doi.org/10.1007/978-0-387-35660-0_4

Liu, L., Wang, Y., Zhao, L., & Huang, S. (2017). *Evaluation of Different SLAM Algorithms* using Google Tangle Data. https://opus.lib.uts.edu.au/handle/10453/107814

Lone Echo. (n.d.). Lone Echo. Retrieved January 3, 2020, from https://www.echo.games/

LUBORSKY, M. R., & RUBINSTEIN, R. L. (1995). Sampling in Qualitative Research. *Research on Aging*, *17*(1), 89–113. https://doi.org/10.1177/0164027595171005

Lukosch, S., Billinghurst, M., Alem, L., & Kiyokawa, K. (2015). Collaboration in Augmented Reality. *Computer Supported Cooperative Work (CSCW)*, *24*(6), 515–525. https://doi.org/10.1007/s10606-015-9239-0

Magic Leap. (2020). https://www.magicleap.com/

- Milgram, P., & Kishino, F. (1994). A Taxonomy of Mixed Reality Visual Displays. *IEICE Trans. Information Systems*, *E77-D, no. 12*, 1321–1329.
- Milgram, P., Takemura, H., Utsumi, A., & Kishino, F. (1995). Augmented reality: A class of displays on the reality-virtuality continuum (H. Das, Ed.; pp. 282–292). https://doi.org/10.1117/12.197321

Mine, M. R. (1995). Virtual Environment Interaction Techniques.

Molecular Geometry. (n.d.). Retrieved April 14, 2021, from

http://chemistry.elmhurst.edu/vchembook/200geomtypes.html

Moyosa Media BV. (2018). *The Kremer Collection VR Museum on Gear VR*. Oculus. https://www.oculus.com/experiences/gear-vr/2097674173595005/

Nanome. (2018). Nanome. https://nanome.ai/

Nojima, T., Sekiguchi, D., Inami, M., & Tachi, S. (2002). The SmartTool: A system for augmented reality of haptics. *Proceedings IEEE Virtual Reality 2002*, 67–72. https://doi.org/10.1109/VR.2002.996506

- Number of smartphone users worldwide 2014-2020. (2019, November 11). Statista. https://www.statista.com/statistics/330695/number-of-smartphone-usersworldwide/
- Núñez, M., Quirós, R., Núñez, I., Carda, J. B., & Camahort, E. (2008). Collaborative Augmented Reality for Inorganic Chemistry Education. *Proceedings of the 5th WSEAS/IASME International Conference on Engineering Education*, 271–277. http://dl.acm.org/citation.cfm?id=1581120.1581171
- O'Brien, H. L., Cairns, P., & Hall, M. (2018). A practical approach to measuring user engagement with the refined user engagement scale (UES) and new UES short form. *International Journal of Human-Computer Studies*, *112*, 28–39. https://doi.org/10.1016/j.ijhcs.2018.01.004
- Parallel Distributed Processing Models of Memory—Dictionary definition of Parallel Distributed Processing Models of Memory | Encyclopedia.com: FREE online dictionary. (2004). Learning and Memory.

https://www.encyclopedia.com/psychology/encyclopedias-almanacs-transcriptsand-maps/parallel-distributed-processing-models-memory

Pemberton, L., & Winter, M. (2009). Collaborative augmented reality in schools. https://repository.isls.org//handle/1/3024

Phiar-Artifical Intelligence to Guide You Anywhere. (2019). https://www.phiar.net/

- Piumsomboon, T., Clark, A., Billinghurst, M., & Cockburn, A. (2013). User-Defined
 Gestures for Augmented Reality. In P. Kotzé, G. Marsden, G. Lindgaard, J.
 Wesson, & M. Winckler (Eds.), *Human-Computer Interaction INTERACT 2013*(pp. 282–299). Springer. https://doi.org/10.1007/978-3-642-40480-1_18
- Piumsomboon, T., Lee, Y., Lee, G., & Billinghurst, M. (2017). CoVAR: A collaborative virtual and augmented reality system for remote collaboration. SIGGRAPH Asia 2017 Emerging Technologies, 1–2. https://doi.org/10.1145/3132818.3132822

Placenote: Build spatial apps. (n.d.). Retrieved April 28, 2020, from https://placenote.com/

Psotka, J. (1995). Immersive training systems: Virtual reality and education and training. *Instructional Science*, 23(5), 405–431. https://doi.org/10.1007/BF00896880

Radkowski, R., & Stritzke, C. (2012). Interactive Hand Gesture-based Assembly for Augmented Reality Applications.

Rauschnabel, P. A., & Ro, Y. K. (2016). Augmented reality smart glasses: An investigation of technology acceptance drivers. *International Journal of Technology Marketing*, *11*(2), 123–148. https://doi.org/10.1504/IJTMKT.2016.075690

- Regan, C. (1995). An investigation into nausea and other side-effects of head-coupled immersive virtual reality. *Virtual Reality*, 1(1), 17–31. https://doi.org/10.1007/BF02009710
- Reitmayr, G., & Schmalstieg, D. (2004). Collaborative Augmented Reality for Outdoor Navigation and Information Browsing. 11.
- *RoboCo.* (n.d.). Filament Games. Retrieved January 9, 2020, from https://www.filamentgames.com/project/roboco/
- Sasikumar, P., Collins, M., Bai, H., & Billinghurst, M. (2021). XRTB: A Cross Reality
 Teleconference Bridge to incorporate 3D interactivity to 2D Teleconferencing.
 Extended Abstracts of the 2021 CHI Conference on Human Factors in Computing Systems, 1–4. https://doi.org/10.1145/3411763.3451546
- Scharver, C., Evenhouse, R., Johnson, A., & Leigh, J. (2004). Designing cranial implants in a haptic augmented reality environment. *Communications of the ACM*, *47*(8), 32. https://doi.org/10.1145/1012037.1012059

Schell Games. (2018). HoloLAB Champions. Schell Games.

https://www.schellgames.com/games/hololab-champions

- Semararo, L., Craig, A. B., & McGrath, R. E. (2011). *Augmented Reality Demonstrations* (August 2010). https://www.ideals.illinois.edu/handle/2142/18810
- Shehab, S., & Mercier, E. (2019). Implementing collaborative learning in classrooms: Science teachers' perspectives. International Journal of Innovation in Education. https://www.inderscienceonline.com/doi/abs/10.1504/IJIIE.2019.102597
- Shemek, D., Guidazzoli, A., Liguori, M., Bellavia, G., De Luca, D., Verri, L., & Imboden, S. (2019). *Renaissance Remix. Isabella d'Este: Virtual Studiolo. 12*.
- Sheridan, T. B. (1992). Musings on Telepresence and Virtual Presence. Presence: Teleoperators and Virtual Environments, 1(1), 120–126. https://doi.org/10.1162/pres.1992.1.1.120
- Sherman, W. R., & Craig, A. B. (1995). Literacy in virtual reality: A new medium. ACM SIGGRAPH Computer Graphics, 29(4), 37–42. https://doi.org/10.1145/216876.216887
- Sherman, W. R., & Craig, A. B. (2003). Understanding Virtual Reality: Interface, Application, and Design. In Understanding Virtual Reality: Interface, Application, and Design (p. i). Elsevier. https://doi.org/10.1016/B978-1-55860-353-0.50019-7
- Sherman, W. R., & Craig, A. B. (2018). Understanding Virtual Reality: Interface, Application, and Design (2nd ed.). Morgan Kaufmann.
- Sherman, W. R., Craig, A. B., Baker, M. P., & Bushell, C. (2004). Scientific Visualization. In A. B. Tucker (Ed.), *Computer Science Handbook* (2nd ed). Chapman & Hall/CRC.
- Sky UK Ltd, & Factory 42. (2018). Sky VR: Hold the World on Oculus Rift. Oculus. https://www.oculus.com/experiences/rift/2331434793563555/
- Sodhi, R. S., Jones, B. R., Forsyth, D., Bailey, B. P., & Maciocci, G. (2013). BeThere: 3D mobile collaboration with spatial input. *Proceedings of the SIGCHI Conference*

on Human Factors in Computing Systems, 179–188. https://doi.org/10.1145/2470654.2470679

- Stoakley, R., Conway, M. J., & Pausch, R. (1995). Virtual reality on a WIM: Interactive worlds in miniature. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems - CHI* '95, 265–272. https://doi.org/10.1145/223904.223938
- Stricker, D., Pagani, A., & Zoellner, M. (2010). In-Situ Visualization for Cultural Heritage Sites using Novel Augmented Reality Technologies. *Virtual Archaeology Review*, 1(2), 37–41.
- Szalavári, Z., Schmalstieg, D., Fuhrmann, A., & Gervautz, M. (1998). "Studierstube": An environment for collaboration in augmented reality. *Virtual Reality*, *3*(1), 37–48. https://doi.org/10.1007/BF01409796
- Technologies, U. (n.d.). Unity Real-Time Development Platform | 3D, 2D VR & AR Engine. Retrieved May 31, 2021, from https://unity.com/
- Thomas, B., Demczuk, V., Piekarski, W., Hepworth, D., & Gunther, B. (1998). A wearable computer system with augmented reality to support terrestrial navigation. *Digest of Papers. Second International Symposium on Wearable Computers (Cat. No.98EX215)*, 168–171.

https://doi.org/10.1109/ISWC.1998.729549

- Trick, L., & Pylyshyn, Z. (1994). Trick, L. M. & Pylyshyn, Z. W. Why are small and large numbers enumerated differently? A limited-capacity preattentive stage in vision.
 Psychol. Rev. 101, 80-102. *Psychological Review*, *101*, 80–102.
 https://doi.org/10.1037/0033-295X.101.1.80
- Verlinden, J. C., Bolter, J. D., & Van der Mast, C. A. P. G. (1993). Virtual Annotation: Verbal Communication in Virtual Reality [Technical Report]. Georgia Institute of Technology. https://smartech.gatech.edu/handle/1853/3649

- Video Conferencing, Web Conferencing, Webinars, Screen Sharing—Zoom. (n.d.). Retrieved June 14, 2021, from https://zoom.us/
- Vorderer, P., Wirth, W., Gouveia, F. R., Biocca, F., Saari, T., Jäncke, L., Böcking, S.,
 Schramm, H., Gysbers, A., Hartmann, T., Klimmt, C., Ravaja, N., Sacau, A.,
 Baumgartner, T., & Jäncke, P. (n.d.). *MEC Spatial Presence Questionnaire* (*MEC-SPQ*). 14.
- Vuforia Enterprise Augmented Reality (AR) Software | PTC. (n.d.). Retrieved May 31, 2021, from https://www.ptc.com/en/products/vuforia
- WayRay. (2019). https://wayray.com/embedded
- Wei, X., Weng, D., Liu, Y., & Wang, Y. (2015). Teaching based on augmented reality for a technical creative design course. *Computers & Education*, 81, 221–234. https://doi.org/10.1016/j.compedu.2014.10.017
- What is AR, VR, MR, XR, 360? (n.d.). Unity. Retrieved December 14, 2019, from https://unity3d.com/what-is-xr-glossary
- Wilcoxon Signed Rank Test in SPSS Statistics—Procedure, output and interpretation of output using a relevant example. (n.d.). Retrieved June 2, 2021, from https://statistics.laerd.com/spss-tutorials/wilcoxon-signed-rank-test-using-spssstatistics.php#procedure
- Yang, U., Lee, G. A., Shin, S., Hwang, S., & Son, W. (2007). Virtual Reality based Paint Spray Training System. 2007 IEEE Virtual Reality Conference, 289–290. https://doi.org/10.1109/VR.2007.352506

Appendix A.

NASA TLX Questionnaire



Appendix B.

MEC Spatial Presence Questionnaire

1. I had the feeling that I was in the middle of the action rather than merely

observing

Strongly disagree	Somewhat disagree	Neither agree nor disagree	Somewhat agree	Strongly agree
0	0	0	0	0

2. I felt like I was a part of the environment in the presentation

Strongly disagree	Somewhat disagree	Neither agree nor disagree	Somewhat agree	Strongly agree
0	0	0	0	0

3. I felt like I was actually there in the environment of the presentation

Strongly disagree	Somewhat disagree	Neither agree nor disagree	Somewhat agree	Strongly agree
0	0	0	0	0

4. I felt like the objects in the presentation surrounded me

Strongly disagree	Somewhat disagree	Neither agree nor disagree	Somewhat agree	Strongly agree
0	0	0	0	\circ

5. It was as though my true location had shifted into the environment of the

presentation

Strongly disagree	Somewhat disagree	Neither agree nor disagree	Somewhat agree	Strongly agree
0	0	0	0	0

6. It seemed as though my self was present in the environment of the

presentation

Strongly disagree	Somewhat disagree	Neither agree nor disagree	Somewhat agree	Strongly agree
0	0	0	0	0

7. I felt as though I was physically present in the environment of the

presentation

Strongly disagree	Somewhat disagree	Neither agree nor disagree	Somewhat agree	Strongly agree
0	0	0	0	0

8. It seemed as though I actually took part in the action of the presentation

Strongly disagree	Somewhat disagree	Neither agree nor disagree	Somewhat agree	Strongly agree
0	0	0	0	0

Appendix C.

User Engagement Scale (short form)

1. I lost myself in this experience

Strongly disagree	Somewhat disagree	Neither agree nor disagree	Somewhat agree	Strongly agree
0	0	0	0	0

2. The time I spent using the application just slipped away

Strongly disagree	Somewhat disagree	Neither agree nor disagree	Somewhat agree	Strongly agree
0	0	0	0	0

3. I was absorbed in this experience

Strongly disagree	Somewhat disagree	Neither agree nor disagree	Somewhat agree	Strongly agree
0	0	0	0	0

4. I felt frustrated while using this experience

Strongly disagree	Somewhat disagree	Neither agree nor disagree	Somewhat agree	Strongly agree
0	0	0	0	0

5. I found this application confusing to use

Strongly disagree	Somewhat disagree	Neither agree nor disagree	Somewhat agree	Strongly agree
0	\circ	0	0	0

6. Using this application was taxing

Strongly disagree	Somewhat disagree	Neither agree nor disagree	Somewhat agree	Strongly agree
0	0	0	0	0

7. This application was attractive

Strongly disagree	Somewhat disagree	Neither agree nor disagree	Somewhat agree	Strongly agree
0	0	0	0	0

8. This application was aesthetically appealing

Strongly disagree	Somewhat disagree	Neither agree nor disagree	Somewhat agree	Strongly agree
0	0	0	0	0

9. This application appealed to my senses

Strongly disagree	Somewhat disagree	Neither agree nor disagree	Somewhat agree	Strongly agree
0	0	0	0	0

10. Using this application was worthwhile

Strongly disagree	Somewhat disagree	Neither agree nor disagree	Somewhat agree	Strongly agree
0	0	0	0	0

11. My experience was rewarding

Strongly disagree	Somewhat disagree	Neither agree nor disagree	Somewhat agree	Strongly agree
0	0	0	0	0

12. I felt interested in this experience

Strongly disagree	Somewhat disagree	Neither agree nor disagree	Somewhat agree	Strongly agree
0	0	0	0	0