UNDERSTANDING USER ENGAGEMENT IN IMMERSIVE AND INTERACTIVE STORIES

A Dissertation
Presented to
The Academic Faculty

by

Steven P. Dow

In Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy in the
School of Interactive Computing
College of Computing

Georgia Institute of Technology
December 2008

UNDERSTANDING USER ENGAGEMENT IN IMMERSIVE AND INTERACTIVE STORIES

Approved by:

Dr. Blair MacIntyre, Advisor
School of Interactive Computing
Georgia Institute of Technology

Dr. Jay Bolter
School of Literature, Communication, and Culture
Georgia Institute of Technology

Dr. Mark Guzdial
School of Interactive Computing
Georgia Institute of Technology

Dr. Elizabeth Mynatt
School of Interactive Computing
Georgia Institute of Technology

Dr. Michael Mateas
Computer Science Department
University of California Santa Cruz

Date Approved: August 19, 2008

ACKNOWLEDGEMENTS

Most of all I would like to thank my family who have always supported me with love and humor, and who taught me the value of hard work.

I am especially grateful to my advisor, Blair MacIntyre, for believing in me and helping me develop as a researcher. Between our zany brainstorming sessions to our late-night demo preparations, it has been an excellent adventure I will soon not forget.

I want to acknowledge my committee members—Jay Bolter, Mark Guzdial, Michael Mateas, and Beth Mynatt—for posing thoughtful questions and exposing me to new ideas.

I have been extremely fortunate to part of the collaborative environment in the School of Interactive Computing and the GVU Center at Georgia Tech. I cannot think of any other organization where we could create a science project of such magnitude without a clear “problem” being addressed. All of my projects throughout my PhD have been collaborative—nothing is more satisfying than being part of something that a single person could never create. To that extent, I would like to thank my friends and colleagues for challenging and encouraging me through their own hard work and dedication: Gregory Abowd, Aaron Bobick, Marshini Chetty, Jill Coffin, Josh Cothran, Sauvik Das, Irfan Essa, James Foley, Andrea Forte, Maribeth Gandy, Beki Grinter, Diane Gromala, Robert Hamburger, Ellie Harmon, Giovanni Iachello, Scott Klemmer, Janet Kolodner, James Landay, Annie Lausier, Jaemin Lee, Yang Li, Brian Magerko, Ali Mazalek, Manish Mehta, Nancy Nercessian, Chris Oezbek, Zach Pausman, Sha Xin Wei, Michael Terry, Geb Thomas, Amy, Voida, Steve Voida, Noah Wardrip-Fruin, Susan Wyche, Yan Xu, Jose Zagal, and to everyone in the Augmented Environments Lab and the Human-Centered Computing program.

I want to acknowledge the National Science Foundation and the GVU Center for sponsoring the installations of *AR Façade*, and the Beall Center for Art and Technology and GrandTextAuto.org for hosting the second deployment. A special thanks goes to Manish Metha for his invaluable contributions to the *AR Façade* project and to all my participants across both installations.

I want to recognize Andrew Stern and Michael Mateas for their brilliant work on *Façade*—my dissertation would be decidedly different if it were not for your bold venture into a "wicked problem space."

Finally, I dedicate my dissertation to Randy Pausch. I am constantly inspired, not only by his pioneering work on immersive entertainment experiences, but by the life he led and the dreams he aspired to.

TABLE OF CONTENTS

	Page
ACKNOWLEDGEMENTS	iii
LIST OF TABLES	viii
LIST OF FIGURES	ix
LIST OF ABBREVIATIONS	x
SUMMARY	xi
CHAPTER	
1. Introduction	1
1.1 Introduction to Immersive and Interactive Stories	1
1.2 The Medium and the Pleasures of the Medium	3
1.3 Thesis Statement	6
1.4 Motivation	8
1.5 Contributions	10
1.6 Exclusions	14
1.7 Summary	15
2. A Theoretical Framework for Embodied Narrative Engagement	17
2.1 Framing Embodied Narrative Engagement	18
2.2 Three Realms of Pleasures and the Material Properties	20
2.3 Combining Pleasures	40
2.4 Defining Embodied Narrative Engagement	58
2.5 Other Experiential Concepts	64
2.6 Chapter Discussion	74
3. Technical Description of AR Façade	77
3.1 Making Façade "Embodied"	78
3.2 Design Challenges for Mixing the Physical and Virtual	85
3.3 An Eleven-Week Gallery Deployment of AR Façade	93

3.4 Experiences of Non-Player Participants 103
3.5 Chapter Discussion 113
4. Methodology and Overall Impressions 117
4.1 Background on User Study Methodologies for Experience Design 118
4.2 User Studies of Desktop-based Façade 120
4.3 Data Gathering for AR Façade 123
4.4 Analysis Methods for AR Façade 129
4.5 Influences and Overall Impressions of AR Façade 137
4.6 Summary 154
5. Player Demographics and Styles of Play 155
5.1 Player Demographics and Episode Statistics 157
5.2 Episode Analysis 162
5.3 Styles of Play (Case Studies) 174
5.4 Towards a Measure of Play Style 191
5.5 Chapter Discussion 198
6. The Effects of Immersive Interfaces on Agency 201
6.1 Immersed and Unconstrained 202
6.2 Diminished Agency 209
6.3 Players Strategies for Achieving Agency 215
6.4 Trading Presence for Agency 225
6.5 Chapter Discussion 231
7. The Effects of Immersive Interfaces on Dramatic Involvement 235
7.1 Second-Person Narrative Voice and the Player-Character 236
7.2 First-Person Interfaces and the Player-Character 243
7.3 Emotional Involvement 251
7.4 Player Tactics for Maintaining Distance 258
7.5 Interface Effects on Emotional Distance 273
7.6 Chapter Discussion 288

8. Future Research for Immersive and Interactive Stories 291

8.1 Constructing Future Immersive and Interactive Stories 293

8.2 Modeling Play Style 303

8.3 Constraining Loosely-Constrained Immersive Interfaces 307

8.4 Managing Media Distance 312

8.5 Exploring the Potential and the Limitations 318

8.6 Summary and Final Thoughts 324

APPENDIX A: Online Survey of Façade (SurveyMonkey.com) 326

APPENDIX B: Online Survey of Façade (Results) 330

APPENDIX C: Interface Comparison Study (Game Instructions) 335

APPENDIX D: Interface Comparison Study (Questionnaire) 336

APPENDIX E: Interface Comparison Study (Interview Guides) 338

APPENDIX F: Interface Comparison Study (Results) 340

APPENDIX G: Eleven-Week Gallery Deployment (Optional Paper-based Survey) 342

APPENDIX H: Eleven-Week Gallery Deployment (Survey Results and Episode Data) 343

APPENDIX I: Two-Week Player Investigation (Interview Guide) 346

APPENDIX J: Two-Week Player Investigation (Demographic Questions) 347

APPENDIX K: Two-Week Player Investigation (Questionnaire) 350

APPENDIX L: Episode Video Coding Analysis (Reference Sheet) 352

APPENDIX M: Episode Video Coding Analysis (Kappa Statistic Summary for Inter-Coder Reliability) 353

APPENDIX N: Episode Video Coding Analysis (Player Visualizations) 355

APPENDIX O: Possible Correlations Across Both Installations 358

APPENDIX P: Myers-Briggs Personality Tests 360

APPENDIX Q: Interpersonal Distance 361

APPENDIX R: Wizard Docent Investigations (Interview Guides) 363

REFERENCES 364

LIST OF TABLES

Table 3.1: Total number of text characters erased across three versions 107

Table 4.1: Summary of research studies of AR Façade 124

Table 4.2: AR Façade players (N=45) self-report use of technology 143

Table 5.1: Player averages for key episode data across both installations (N=45) 160

Table 5.2: Player 17's opening exchange with Trip 167

Table 5.3: Player 8 exchange near the end of her episode 169

Table 5.4: Player 3 exchange near the end of her episode 171

Table 5.5: Player 29 exchange near the end of her episode 172

Table 5.6: Player 18 exchange after encountering an AI crash 173

Table 5.7: Player 32 exchanges exemplifying the engager play style 175

Table 5.8: Summary of other players who exhibited an "engager" style of play 178

Table 5.9: Player 16 exchanges exemplifying the performer play style 179

Table 5.10: Summary of other players who exhibited a "performer" style of play 181

Table 5.11: Player 21 exchanges exemplifying the tinkerer play style 182

Table 5.12: Summary of other players who exhibited a "tinkerer" style of play 185

Table 5.13: Player 19 exchanges exemplifying the observer play style 186

Table 5.14: Summary of other players who exhibited a "observer" style of play 188

Table 5.15: Player 43 exchanges exemplifying the partaker play style 189

Table 5.16: Clustering study participants (N=45) into one of the five play styles 195

Table 5.17: Play styles and episode statistics (averages/standard deviations) 196

Table 6.1: Player 25 speaking a continuous 187-character-long utterance 207

Table 6.2: Player 6 demonstrating his "keyword" strategy for communication 219

Table 6.3: Player 14 demonstrates “fluid” conversation 221

Table 6.4: Player activity across three versions in the interface comparison study 226

Table 7.1: Player 38 filling in backstory for the player-character role 242

Table 7.2: Player 25 reacting to the dramatic moment 249

Table 7.3: Player 26 reacting emotionally 252

Table 7.4: Player 35 reacting emotionally 254

Table 7.5: Rating differences between engagers and partakers 263

Table 7.6: Interpersonal distance across two installations 275

Table 7.7: Interpersonal distance in the keyboard version across two FOV values 276

Table 8.1: Hypothetical run-time indicators of potential play styles 305

Table F.1: Key survey questions from the interface comparison study 340

Table F.2: Key episode statistics from the interface comparison study 340

Table F.3: Paired-samples T-test statistics for the interface comparison study 341

Table H.1: General demographics of 40 respondents to the optional paper survey 343

Table H.2: T-test comparing KB and AR for DialogPerMin at Beall Center (N=232) 344

Table H.3: T-test comparing KB and AR for StmtLength at Beall Center (N=232) 344

Table H.4: T-test comparing KB and AR for GesturePerMin at Beall Center (N=232) 345

Table H.5: T-test comparing KB and AR for TimeOfPlay at Beall Center (N=232) 345

Table O.1: Pearson Correlation analysis for all 45 players 358

Table Q.1: Paired-samples T-test for IPD in three versions of Façade (N=12) 361

Table Q.2: Independent-samples T-test for IPD at the Beall 361

Table Q.3: Independent-samples T-test for IPD (FOV change) 362

Table Q.4: Independent-samples T-test for IPD across genders 362

LIST OF FIGURES

Figure 2.1: Embodied Narrative Engagement (ENE) Framework	19
Figure 2.2: The interrelationship between agency and dramatic involvement	41
Figure 2.3: The interrelationship between agency and presence	47
Figure 2.4: The gulfs of execution and evaluation from direct manipulation theory	48
Figure 2.5: The interrelationship between presence and dramatic involvement	53
Figure 2.6: The intersection of presence, agency, and dramatic involvement	59
Figure 2.7: Taxonomy of examples of immersive and interactive stories	61
Figure 3.1: Screenshot of Façade, with the characters Trip and Grace	79
Figure 3.2: Original Façade's keyboard and mouse interaction	80
Figure 3.3: Conceptual architecture diagram for desktop Façade	81
Figure 3.4: The interface for (a) desktop Façade (b) AR Façade	82
Figure 3.5: Hardware for the Atlanta AR Façade installation	83
Figure 3.6: Creating a physical replica of the virtual apartment of Facade	84
Figure 3.7: Wizard-of-Oz interface in the Atlanta installation of AR Façade	85
Figure 3.8: GrandTextAuto's Beall Center exhibit	93
Figure 3.9: Beall Center layout during the GrandTextAuto exhibit	95
Figure 3.10: AR Façade at the Beall Center hardware and stage	98
Figure 3.11: Physical setup at the Beall Center exhibit of AR Façade	99
Figure 3.12: AR Façade conceptual architecture with two wizard variations	100
Figure 3.13: Two different wizard interfaces used at the Beall exhibit	101
Figure 3.14: The audience experience in Atlanta versus the Beall Center	104
Figure 3.15: Wizard activity during the eleven-week installation	109

Figure 3.16: The wizard station occluded from the audience view 110

Figure 4.1: Flowchart of data gathering and analysis process for AR Façade 130

Figure 4.2: An excerpt from the interview transcripts with researcher notes 131

Figure 4.3: Sketches of the affinity diagram analysis method 131

Figure 4.4: Player 2's experience of the registration error 142

Figure 5.1: Key player demographics and episode data (N=45) 158

Figure 5.2: Distribution of episode endings (N=45 players) for AR Façade 159

Figure 5.3: Player 38 directed to look at the Italy photo 161

Figure 5.4: Legend for game episode visualization 164

Figure 5.5: Player 26 shows normative behavior throughout the episode 165

Figure 5.6: Player 42 exemplifies transitioning between different styles of play 165

Figure 5.7: Players greeting the characters at the beginning 166

Figure 5.8: Player 17 had interaction problems throughout and quit early 167

Figure 5.9: A technical disturbance where Trip appears to walk through the wall 168

Figure 5.10: Player 8 acted socially appropriately, but quit the experience 168

Figure 5.11: Player 8 by the door trying to politely excuse herself 169

Figure 5.12: Player 7 style of play changed when the fighting started 170

Figure 5.13: Player 9 spoke less as the experience carried on 170

Figure 5.14: Player 3 quickly moving to get in front of Trip before he leaves 171

Figure 5.15: Player 29 running towards Grace and shying away from Trip 172

Figure 5.16: P29 visualization showing her perform divergently towards the end 172

Figure 5.17: Player 18 hugging Grace after the AI engine crashes 173

Figure 5.18: Images of Player 32 exhibiting a "engager" style of play 176

Figure 5.19: Visualization for engager Player 32 176

Figure 5.20: Images of Player 16 exhibiting a "performer" style of play 180

Figure 5.21: Visualization for performer P16 180

Figure 5.22: Images of Player 21 exhibiting a "tinkerer" style of play 183

Figure 5.23: Visualization for tinkerer P21 183

Figure 5.24: Images of Player 19 exhibiting an "observer" style of play 186

Figure 5.25: Visualization of observer P19 187

Figure 5.26: Images of Player 43 exhibiting a "partaker" style of play 190

Figure 5.27: Visualization of partaker P43 190

Figure 5.29: Endings across the five play styles for all participants (N=45) 197

Figure 6.1: Player 4 trying to drag the characters together 207

Figure 6.2: Survey results for comparative study (N=12) 209

Figure 6.3: Player 14 having a fluid conversation with Trip and Grace 222

Figure 6.4: Interaction ratings across play styles (N=45) 224

Figure 7.1: Images of Player 25 showing her dramatic involvement 250

Figure 7.2: Images of Player 26 showing her emotional involvement 253

Figure 7.3: Images of Player 35 showing her emotions change 254

Figure 7.4: Difference between engagers and partakers for preferred character 264

Figure 7.5: Difference between players' genders for preferred character 265

Figure 7.6: Number of females and males classified as engagers and partakers 266

Figure 7.7: Average episode time for men and women 267

Figure 7.8: Interpersonal distance across genders 270

Figure 7.9: Interpersonal distance in meters among men and women players 271

Figure 7.10: Interpersonal distance across two installations 274

Figure 8.1: Theoretical "delay curves" for the two wizard methods 299

Figure B.1: Ethnicity and gender of online survey respondents 330

Figure B.2: Self-reported typing ability of online survey respondents 331

Figure B.3: Reports of playing quantity by online survey respondents 331

Figure B.4: Gameplay strategies used by online survey respondents 332

Figure B.5: Subjective ratings about conversation by online survey respondents 332

Figure B.6: Reasons given by online survey respondents for breakdowns 333

Figure B.7: Strategies used by online survey respondents to adapt to breakdowns 333

Figure H.1: Players relating AR Façade to other media experiences (opt. survey) 343

Figure O.1: Players' overall rating vs. players' age (N=45) 359

Figure O.2: Players' overall rating vs. players' episode ending (N=45) 359

Figure P.1: Personality test results from Beall study (N=33) 360

LIST OF ABBREVIATIONS

AR	Augmented Reality
ENE	Embodied Narrative Engagement
HMD	Head-Mounted Display
IIS	Immersive and Interactive Stories
KB	Keyboard-Based Interaction
SB	Speech-Based Interaction
VR	Virtual Reality
WOz	Wizard-of-Oz methods

SUMMARY

Popular science fiction often proffers the *Holodeck* vision for future immersive media: a seamless and transparent interface connecting users to a virtual world where they transform into a story character and influence unfolding events. In my dissertation, I offer empirical observations on a discussion that has been largely theoretical to this point. I explore the psychological concept of user engagement through an immersive and interactive story experience called *AR Façade*. In the experience, the “player” interfaces with an animated married couple through an augmented reality (AR) interface that allows for unconstrained body movement and speech communication. The player finds herself in the middle of a marital conflict and can influence how the social scenario plays out through her actions and statements. I have studied the *AR Façade* experience from the user perspective by conducting mixed-method investigations in two instantiations: our proof of concept lab demo at Georgia Tech and an eleven-week gallery installation at the Beall Center for Art and Technology in Irvine, CA.

My thesis challenges the assumptions ingrained in the *Holodeck* vision by offering empirical evidence that immersive display technology both supports and counteracts the experiential pleasures sought by proponents of the *Holodeck* medium. Focusing on the experiential aspects of the human-computer interface, I examine how a media experience changes when going from traditional desktop interaction to immersive augmented reality. While the goal of many presence researchers is to strive for an “illusion of non-mediation,” I conclude that explicit mediation may be required for reaching embodied narrative engagement. An immersive interface should be mediated to provide clear mechanics to support player agency (the feeling of empowerment over events) and allow the player to manage their distance from the designated character role. In the process of presenting evidence for my claim, I clarify the terminology across the

stakeholder disciplines and present an empirical case study that spans media theory, design practice, and computer science.

CHAPTER 1

INTRODUCTION

Think of the computer, not as a tool, but as a medium. —Brenda Laurel, Computers as Theatre (p126)

1.1 Introduction to Immersive and Interactive Stories

In defining and shaping computational media as Immersive and Interactive Stories, new experiential pleasures arise at the intersection of presence, agency, and dramatic involvement. The combined effects of these three psychological notions—each a unique, dedicated area of study and media practice—allows us to explore a human fascination with *embodied narrative engagement*. The science fiction fantasy of Star Trek's *Holodeck*[1] embodies the concept of an ultra-realistic simulator that can activate all the senses and allow the ship's crew to interact with story scenarios for recreation and training. On the television show, the experience of the *Holodeck* can be seamlessly depicted by cutting between shots of the Enterprise ship and the simulated environment. The science-fiction is further reinforced by explaining the "technology" of the *Holodeck*—a combination of replicated matter, tractor beams, shaped force fields, and holographic images. The *Holodeck* is by all means a contrived medium, but it has captured the fascination of media theorists and computationalists because of the experiential pleasures and expressive possibilities it has come to represent. It is a

1 http://en.wikipedia.org/wiki/Holodeck (accessed on 9/22/08)

seductive thought: to provoke people to feel transported into a fictitious world, transformed into a story character, and able to influence the unfolding events.

In my dissertation, I offer an empirical, practice-based approach to a discussion that has been largely theoretical to this point. Working with my colleagues at Georgia Tech, and building upon the groundbreaking interactive drama, *Façade*[2], we have constructed a fully-functional exemplar of the *Holodeck* vision. *AR Façade* is an Immersive and Interactive Story experience where a user/player[3] interfaces with an animated married couple through an augmented reality (AR) interface that allows for unconstrained bodily movement and speech communication (Dow et al., 2006). The player finds herself in the middle of a marital conflict and can influence how the social scenario plays out through her actions and statements. I have studied the *AR Façade* experience from the player perspective by conducting qualitative investigations in two instantiations: our proof of concept built in a lab at Georgia Tech's GVU Center in Atlanta, GA and an eleven-week gallery installation at the Beall Center for Art and Technology in Irvine, CA[4].

My thesis challenges the assumptions ingrained in the *Holodeck* vision by offering empirical evidence that immersive display technology both supports and counteracts the experiential pleasures sought by proponents of the *Holodeck* medium. In the process of presenting evidence for my claim, I clarify the terminology across the stakeholder disciplines and present an empirical case study that spans media theory, design practice, and computer science.

---

2 http://www.interactivestory.net/ (accessed on 9/22/08)

3 Throughout my dissertation, I use player and user interchangeably to refer to the first-person experience of *AR Façade.*

4 http://beallcenter.uci.edu/ (accessed on 9/22/08)

1.2 The Medium and the Pleasures of the Medium

I draw an important distinction between the material features of a medium and the pleasurable affects invoked by the medium. A medium is the material by which "messages" are carried; both the messages and the material properties of the medium can have experiential effects on an interactor/observer. Experiential pleasures refer to psychological states in the observer's mind, and need not be positive to be pleasurable (e.g., one can take pleasure in pain). Designers of any medium attempt to create content that is culturally and personally interesting, and to tap into the broad palette of human desires.

Although digital media designers explore many forms and configurations of new media technologies, this dissertation focuses on computational media defined by the combination of immersive interfaces, interactivity, and narrative structure. **Immersive and Interactive Stories** (IIS) are first-person experiences that embed a user as character in a story and allow the user to enact the role designated to them. Immersive technologies—such as head-mounted displays, haptic devices, surround sound, olfactory displays, spatial tracking, motion capture systems, and the like—aim to exploit advances in sensing and displays to simulate (or partially simulate) an environment and to empower the full-range of human physicality. An interactive medium reacts to actions by a participant; the underlying procedural logic determines how a participant's input changes the state of the medium. Finally, within the context of story experiences, a narrative structure defines how a story is presented through the characters, plot, and point of view. Immersive interfaces, interactivity, and narrative structure are three material properties that can be manipulated and engineered to create desired effects. *AR Façade* is one of the first media experiences to combine an immersive augmented reality interface with an underlying narrative structure that responds interactively based on speech and physical actions of the participant.

Immersive and interactive stories have the potential to trigger a psychological state of **embodied narrative engagement** (ENE): the combination of the feeling of being in a story world (**presence**), the feeling of empowerment over unfolding events (**agency**), and the feeling of being caught up in the plot and characters of a story (**dramatic involvement**). Embodied narrative engagement varies from individual to individual and gets shaped through the formulation of the medium and the content. The psychological notions of presence, agency and dramatic involvement are discussed in more detail in Chapter 2, and are similar to the experiential notions Murray describes in *Hamlet on the Holodeck* (1997). To help mitigate some of the confusion over the meanings across communities, I adopt slightly different terminology than Murray and tighten the definitions.

Presence refers to the subjective feeling of being within a world, whether virtual, physical or some combination thereof. One of the most cited definitions of presence is Lombard and Ditton's "the perceptual illusion of non-mediation" (1997), the goal for many researchers working with virtual reality (VR) technologies. A community of researchers from psychology, computer science, and philosophy has converged on this concept to explore telemedicine, VR therapy and rehabilitation, telerobotics, VR education, and entertainment[5].

The sense of **agency** refers to feeling empowerment over events within a medium and having the motivation to take action. I distinguish between different sources of agency. One can feel agency through the ability to navigate a world (“movement” agency) or through the ability to modify the world (“object” agency). Most relevant to

[5] Presence community resources: http://www.presence-research.org/ and http://vr.coe.ecu.edu/vpbib.html (accessed on 9/22/08)

my thesis is a notion of “narrative” agency—the ability to effect the course of unfolding plot events and the outcome of a story. Agency is only felt when a user takes some action within an environment and believes that their actions impact events within that environment.

The notion of dramatic involvement has a longer history of discourse, perhaps making it more difficult to define succinctly. An individual can experience **dramatic involvement** when they are interested and caught up in the characters and plot of a story. Storytellers have explored various narrative points of view and plot structures, from the earliest days of oral communication up to modern forms of media. Aristotle first dissected the dramatic form of narrative, and put forth the idea of structuring plot events to build and release tensions (335 B.C.).

Just as immersive interfaces, interactivity and narrative structure combine as material properties of *immersive and interactive stories*, the experiential pleasures of presence, agency, and dramatic involvement coalesce to bring about the possibility of *embodied narrative engagement*. Each of the psychological states requires a degree of “willing suspension of disbelief” (Coleridge, 1817) on behalf of the participants to see past the imperfections of the immersive interface, interactivity, and narrative structure.

My research questions and my thesis statement focus on the paired relationships between presence, agency and dramatic involvement. Do immersive interfaces increase presence and result in more embodied narrative engagement? Does the feeling of presence lead to feeling more agency? Does the feeling of presence lead to feeling more dramatic involvement? In my dissertation, I seek to illustrate the tensions that arise when trying to simulateously maximize the three experiential pleasures. Some have assumed that maximizing presence by creating a “transparent” or non-mediated interface will automatically lead to the goal of embodied narrative engagement. My analysis of prior theory and the empirical evidence from *AR Façade* argues that non-mediated immersive interfaces will decrease some players' sense of agency and will not provide players a means to explicitly manage their distance from dramatic content.

1.3 Thesis Statement

Through my empirical investigations of *AR Façade* I have formed a thesis that describes a relationship between immersive interfaces and the desired effects: presence, agency, dramatic involvement, and the combined notion of embodied narrative engagement. While it is generally accepted in the VR/presence research community that improving the immersive qualities of the human-computer interface can increase the sense of “being there” (Heeter, 1992), little (if any) presence research focuses on immersive and interactive stories where a user’s body is placed within a story environment and given a designated character role. While the goal of many presence researchers is to maximize “the perceptual illusion of non-mediation” (Lombard and Ditton, 1997), for designers of immersive and interactive stories, presence is a means to an end.

For the purposes of this work, I assume that the goal of combining immersive interfaces, interactivity, and narrative structure is to trigger a psychological sense of embodied narrative engagement (the feeling of being in a story world, able to influence unfolding events, and dramatically involved in the plot and characters). I argue that to increase the overall sense of embodied narrative engagement in immersive and interactive stories, the immersive interface should include explicit mediation—clear mechanics to support player agency and the means for players to manage their distance from the designated character role. Thus my thesis statement claims:

> *The strategy of perceptually immersing the user and minimizing mediation to create the sense of presence does not necessarily maximize the sense of embodied narrative engagement, because complete transparency hides interaction mechanisms that strengthen the sense of agency and does not provide sufficient means for users to manage their distance from dramatic content.*

My thesis assumes that a purely immersive medium would not employ non-realistic interaction affordances as they would take away from the "perceptual illusion of non-mediation". Mediation in its many forms—from non-reality-like interface mechanisms to multiple narrative points of view—runs counter to the spirit of non-mediated presence. My argument in this dissertation is that explicit mediation is required to reach the goal of embodied narrative engagement—even if it sacrifices presence.

I also acknowledge the ability of users to adjust to the imperfections of the interface and to reach embodied narrative engagement by mentally constructing a suspension of disbelief. In my case studies of *AR Façade*, I will show that some players overcame the lack of sufficient mediation, and were able to feel agency and maintain enough psychological distance to become dramatically involved. However, most players—even those that did feel some sense of embodied narrative engagement—would have benefited from having more mediation, not less. My investigation concludes that an immersive interface to an interactive story should include different forms of mediation so that players can create and maintain their preferred level of embodied narrative engagement.

I acknowledge that some media designers will likely have different goals for immersive experiences and for what constitutes "optimal" engagement. These designers may choose valid ulterior goals, divergent from the *Holodeck* notion of "living in the story”, for example intentionally depriving the player of agency or infringing their “safety zone,” all aimed at a different type of engagement. Such work is beyond the scope of this dissertation. Similarly, although the tradeoffs between agency and dramatic involvement may be of interest to authors of interactive narratives, this work focuses on perceptually immersive interfaces.

1.4 Motivation

One way to understand a phenomenon is to isolate it through the manipulation of the material world. I seek to better understand the experiential pleasures of this new medium by actually entering the design space. My collaborators and I created *AR Façade* not only to make a breakthrough in interactive entertainment, but also to exercise an ethnographic-style approach to empirically investigate concepts that have only previously been subject to “thought” experiments. In the process of designing and studying the experience, I am positioned to submit strongly supported arguments to the theoretical discussion and to contribute practical design considerations to guide the development of immersive and interactive stories.

While this dissertation centers on the entertainment aspects of immersive and interactive stories, it also speaks to broader issues of computer-mediated communication and education. Computers have become one of the primary means by which humans communicate, express themselves, and understand each other. The advancements to computational media are not exclusively about one-way communication from artists/ authors to audiences, but can have implications for how individuals will come to interrelate within the interconnected world, particularly with the rise in popularity of 3D online communities, such as Second Life. I am particularly motivated by the belief that much of the world wide web (and the “sensed” physical world) will be absorbed into and represented within shared virtual environments, combining elements of Second Life and Google Earth, similar to the vision posited by Roush (2007).

Immersive and interactive stories can provide new pedagogical tools. The triumvirate of experiential pleasures each has roots in learning science. Presence researchers have developed immersive VR interfaces for rehabilitation and motor skill training, a sort of "learning by exposure" (Hodges et al., 1995; Hodges et al., 1996), largely employed by the military. Likewise, interactive computer-based learning experiences—including games—have been integrated into educational curricula, because of the benefits of "learning by doing" (Roussou, 2004). Furthermore, case-based learning scientists and cognitive scientists—who claim humans arrange knowledge (and reorganize it) as a set of scripts in our mind (Kolodner, 1993; Shank et al., 1994)—have explored the power of narrative towards a theme of "learning by scenario". When taken together as one pedagogical medium, the three strategies seek to emulate the situated learning encountered in everyday life (Lave and Wegner, 1991). *AR Façade* creates a social scenario that draws upon participants' cultural and interpersonal knowledge and compels them to physically enact the role set out for them.

Based on our work thus far, I cannot speak about transfer of knowledge back to the real world, but existing theory suggests that the situational, performative nature of immersive and interactive stories will allow lessons to really "sink in". This medium not only seeks to exploit human biology—by structuring the interface so that our bodies "intuitively understand"—it seeks to build on social and cultural knowledge. Interactive and immersive stories have the potential to contribute to cultural training and other socially complex lessons where other media endeavors have not been as successful. The overarching motivation is to explore the emotional and cognitive "pleasures" and to extract design considerations for the medium that can speak to authors and educators alike.

1.5 Contributions

As contributions to the research community, my dissertation work offers a theoretical framework for Embodied Narrative Engagement (ENE), a design artifact of an Immersive and Interactive Story (IIS), research strategies for investigating participant behavior/opinions in this new medium, and empirical evidence that expands our understanding of user engagement and challenges some of the common assumptions in presence research.

1.5.1 Theoretical Framework for Embodied Narrative Engagement

Building on Murray's notions of experiential pleasures in *Hamlet on the Holodeck* (1997) and Mateas' *Preliminary Poetics for Interactive Drama* (2001), in Chapter 2 I propose a theoretical framework to communicate about the intersection of presence (a feeling of being within an environment), agency (a feeling of empowerment over events), and dramatic involvement (a feeling of being caught up in the plot and characters of a story). This trifecta—embodied narrative engagement—describes the premise behind the *Holodeck*, a potential illusion of feeling transported into a simulated fictitious world, transformed into a story character, and able to influence the unfolding events.

My theoretical framework goes further to describe the paired relationships that arise between each of the three experiential pleasures. My analysis of related literature in HCI, media theory, and presence theory suggests that the particular interrelationships are mutually reinforced to a point, but that tensions can arise. I put forth concrete media examples that exemplify each experiential notion and each paired relationship. I also briefly present some of my earlier collaborations in immersive entertainment: the *Voices* *of Oakland* outdoor location-based audio tour (Dow et al., 2005), the *Four Angry Men* augmented reality jury drama (MacIntyre et al., 2003), and the tangible installation work done in the Topological Media Lab (Sha et al., 2001). None of my earlier projects go as far as *AR Façade* to effectively combine all three material properties in the same experience.

1.5.2 The Design and Implementation of AR Façade

In 2005, I collaborated to build *AR Façade* (Dow et al., 2006), an augmented reality version of *Façade*, the critically acclaimed interactive drama created by Mateas and Stern (2000; 2003). Technically, we achieved what few entertainment experiences have accomplished—combining interactive virtual characters, non-linear narrative, and unconstrained embodied interaction. On the surface, *Façade* was a natural candidate for conversion to AR: it takes place in a small world (two rooms in a small apartment), requires no fast-paced interaction that would be challenging with current hardware, and supports rich interaction between the player, the characters and the world. However, the conversion of *Façade* to *AR Façade* sheds light on a range of architectural, interaction and content issues that will be relevant to any experience that immerses a player in a mixed physical/virtual space. In Chapter 3, I describe the details of *Façade* and the basic design for *AR Façade*, including our use of “Wizard of Oz” methods (Dahlback et al., 1993; Dow et al., 2005) employed to overcome the most challenging interaction issues. I also discuss the infrastructural changes required to prepare *AR Façade* for a gallery installation at the Beall Center for Art and Technology in Irvine, CA.

1.5.3 Research Strategies for Investigating Immersive and Interactive Stories

AR Façade and the original *Façade* provide a unique experience for players, and so I have had to rethink the traditional notions of user evaluation. Games and interactive entertainment—especially those that push on narrative and social interaction—are situated experiences that call for ethnographic-style analyses to explain players' interpretations and behaviors. My focus has not been on strict evaluation as much as understanding the user experience and investigating larger questions of embodiment in interactive entertainment. In Chapter 4, I explain my mixed-method approach that combines qualitative observational methods, raw data collection, retrospective interviews, and open-ended interviews exercised across the two installations of *AR Façade*. I also created a preliminary video analysis technique for immersive and interactive stories that codes different types of player speech utterances, gestures, and technical anomalies for the course of an episode. My data gathering and analysis methods contribute to a descriptive account of player behaviors and interpretations and provide evidence for my thesis claims.

1.5.4 A Descriptive Account of Player Behaviors and Interpretations

In Chapter 5, I will present a descriptive account of the range of behaviors and interpretation from players of *AR Façade* to illustrate the diversity and ambiguity of embodied narrative engagement observed at the two *AR Façade* installations. I contribute a description of play styles in interactive and immersive stories—demonstrated through five case studies—similar to Bartle's discussion of a player taxonomy in MUDs (1996). Through my video analysis of game episodes I illustrate how players can shift between styles during the course of an episode and I analyze the possible correlations that emerge with respect to episode data and individual player demographics when bucketing individuals into one of the particular player types.

1.5.5 An Empirical Argument about the Effect of Immersive Interfaces

The empirical data collected in a series of studies about *AR Façade* suggests that the AR interface induces a strong sense of presence, creating a more "immediate" experience for players. The unconstrained AR interface did not inhibit players from acting and speaking freely. The first-person immersive video-see-though HMD reinforced the second-person narrative point of view and the player's character role within the script.

In Chapter 6, I demonstrate through game episodes how participants acted "naturally" in the interface, speaking quite verbosely at times and attempting actions that fell outside of the three prescribed actions (hug, comfort, kiss). The interviews revealed that this naturalness in the interface raised user expectations for how the environment should respond. The AR versus desktop contrast helped to elucidate experiential differences; the desktop environment provides much clearer affordances for what objects can be manipulated and how much text input is understood at one time. I reveal tactics used by players attempting to obtain a sense of agency in the AR version, and show how other players held on to an illusion of agency despite inadequate affordances. I argue that designers seeking the overall goal of embodied narrative engagement should manage expectations for interactivity by explicitly mediating immersive interfaces.

In Chapter 7, I demonstrate that many players transform into the character role, treat the characters as believable social partners, and act out within the dramatic moments. My analysis describes the raw emotions elicited by the scenario and the tactics used by players to seek psychological distance from their dramatic involvement in the story. The AR/desktop contrast revealed that some players preferred the less intimate desktop experience, precisely because it supported their desire for a degree of emotional distance. I argue that designers seeking the overall goal of embodied narrative engagement should allow players to manage their emotional distance by supporting mediation in immersive interfaces.

1.5.6 Future Research for Immersive and Interactive Stories

In Chapter 8, I will discuss the overarching assumption of non-mediation, the premise of popular depictions of immersive technology such as the *Holodeck*. I propose that mediation will provide designers more control over the interrelationships between presence, agency and dramatic involvement. While a move against pure immediacy would be seen as counterproductive by many presence researchers, mediating the experience—by explicitly building in constraints, structuring manipulation hooks, and allowing for story experimentation—facilitates more authorial control over users' degree of agency and dramatic involvement. I pull in additional comments and anecdotes from the *AR Façade* studies, along with other general ideas, to present design considerations and to raise research questions for future work in this area.

1.6 Exclusions

In my dissertation I will not be speculating about broad societal effects of immersive and interactive storytelling. I employ an empirical, mixed-method approach for understanding one instance of this type of media experience. Due to this empirical focus, I make notably few references to contemporary video games or films to support my positions. While this is not a media theory thesis, I do hope to make sense of my findings within prior theoretical work, particularly in Chapter 2. Likewise, despite the fact that I employ empirical methods, I do not follow traditional HCI or psychology research in testing hypotheses about interface improvements. My collaborators and I built an example immersive and interactive story, and I performed ethnography-inspired research. The research findings resemble social science (more so than laboratory-style research) because they reveal how specific people behaved in this one media experience over a definitive timeframe.

I have decided to investigate an assumption that pervades contemporary culture: that a totally immersive, transparent interface is the ultimate media form—the premise of the *Holodeck*. My intention is to deconstruct the popular portrayal of virtual reality as a perfectly seamless medium, whether it serves as the ultimate tool for society (i.e., *Star Trek*) or the ultimate weapon for entrapping the masses (i.e., *The Matrix*). My investigation embarks with the assumption that many media designers will adopt the goal of embodied narrative engagement. This is admittedly narrow, since many designers will explicitly strive for a different aesthetic and reject the goal of catharsis—the emotional release that presupposes total uncritical involvement. I will not be addressing the avant-garde, counter tradition adopted by some media designers. Rather, I argue that even if the goal is ultimate embodied narrative engagement, the way forward is not necessarily through pure non-mediation, but through a measured balance of the material properties. Thus, my dissertation is intended for computational scientists and media designers who seek to create "mainstream" immersive and interactive story experiences for entertainment and education. However, I expect artists seeking to advance the counter tradition in new media will also learn from and exploit my structured conceptualization of immersive and interactive stories to deliberately distort the balance of material properties.

1.7 Summary

In this chapter, I have introduced my dissertation topic through the familiar guise of the *Holodeck*, the fictional perfection of a promising medium I refer to as Immersive and Interactive Stories. I differentiate between the material properties of immersive and interactive stories (immersive displays, interactivity, and narrative structures) and the new experiential pleasures that open up at their intersection. I suggest the phrase "embodied narrative engagement" to describe the potential illusion of being in a story world (presence), feeling of empowerment over unfolding events (agency), and feeling caught up in the plot and characters of a story (dramatic involvement).

My thesis statement unpacks the effect of first-person immersive display technology on a user's sense of presence, agency, and dramatic involvement. I highlight tradeoffs that potentially arise when striving for presence (the illusion of non-mediation) and advocate a broader goal of embodied narrative engagement (which takes on many forms and requires explicit mediation). Outside of the immediate implications for entertainment technology, my work is motivated by the potential contributions to computer-mediated communication and education. As a starting point for understanding the immersive and interactive story media, my work is a descriptive account of one such experience and does not speculate on the myriad of media creations that will follow or the broader social impact. Finally, my dissertation contributions and chapters can be summarized as follows:

- A theoretical framework for "Embodied Narrative Engagement" (Chapter 2)
- An exemplar immersive and interactive story design artifact, *AR Façade*. (Chapter 3)
- Research strategies for capturing and visualizing participation in interactive and immersive stories (Chapter 4)
- A descriptive account of player behaviors and interpretations of the *AR Façade* experience (Chapter 5)
- Empirical evidence about the effects of first-person immersive interfaces in interactive stories (Chapters 6 and 7)
- Design considerations and open research questions for future immersive and interactive stories (Chapter 8)

CHAPTER 2

A THEORETICAL FRAMEWORK FOR EMBODIED NARRATIVE ENGAGEMENT

> *Media are continually commenting on, reproducing, and replacing each other, and this process is integral to media.* —Jay David Bolter and Richard Grusin, Remediation: Understanding New Media (p55)

In this chapter, I turn to media theory, HCI, and philosophy to disassemble and examine the constructs behind the *Holodeck* premise. I define key terminology for the medium of immersive and interactive stories and outline a framework that explores several core experiential states and their influence on each other. I outline three experiential pleasures at the core of immersive and interactive stories: presence, agency, and dramatic involvement. I clarify what they mean, drawing from the literature and distinguishing them from other closely related concepts, such as Murray's notions of immersion and transformation (1997).

I describe the theoretical effect of combining these aesthetics, aimed at creating a state of embodied narrative engagement, and highlight some of the intellectual debates that arise at the paired interrelationships between presence, agency and dramatic involvement. In particular, I highlight discussions in support of my thesis: that complete interface transparency hides interaction mechanisms that strengthen the sense of agency and does not provide sufficient means for users to manage their distance from dramatic content. A state of embodied narrative engagement cannot be maximized without explicit mediation in the interface, even if this means sacrificing a sense of presence.

I consider the concept of embodied narrative engagement in light of other important experiential theories: play, performance, flow, and phenomenological philosophy. Finally, I summarize the chapter reflecting on how my prior work falls within the framework, and motivating the *AR Façade* endeavor, which has become the empirical focus of my dissertation work.

2.1 Framing Embodied Narrative Engagement

I propose a framework that draws into focus three interdependent, widely-studied psychological states that can be experienced by participants in immersive and interactive stories. The framework uses a Venn diagram to discuss relationships on two levels, both the material features of the medium and the pleasurable effects invoked by the medium (see Figure 2.1). Generally speaking, creators attempt to manipulate the material properties of a medium to strive for experiential affects or "pleasures" (emotions that need not be positive to be pleasurable). Storytellers attempt to create content that is culturally and personally interesting and to tap into human desires for adventure, fear, arousal, drama, love, and hate, to name a few. For immersive and interactive stories, authors not only utilize traditional narrative strategies (by creating compelling characters, twisting plots, story arcs, etc.), they will also manipulate the viewer's perceptual experience (with immersive interfaces) and design ways for viewers to participate (through interactivity).

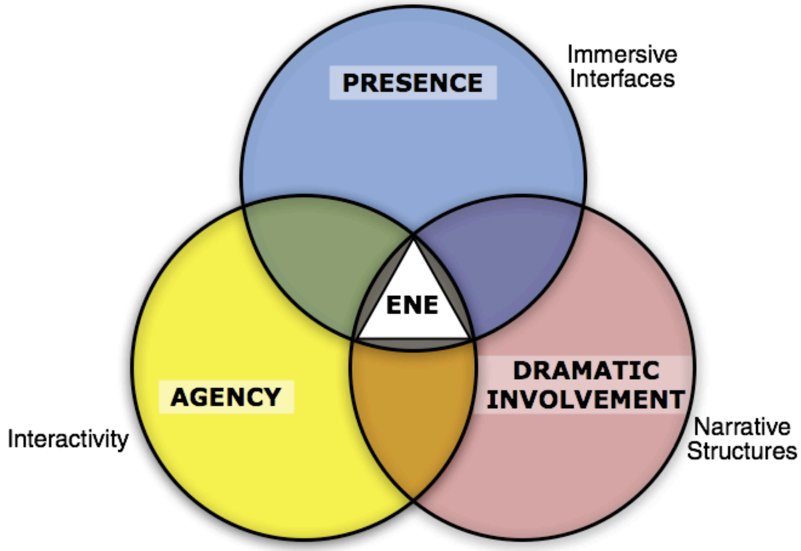


Figure 2.1: Embodied Narrative Engagement (ENE) Framework: the combination of three *experiential pleasure*s: presence, agency, and dramatic involvement. The *material properties* (perceptually immersive interfaces, interactivity, and narrative structures) are features of the medium that can be manipulated by creators.

At the center of this Venn framework is the experiential possibility of embodied narrative engagement (ENE), a phrase I use to describe the combined pleasures: *presence* (a feeling of being within an environment), *agency* (a feeling of empowerment over events), and *dramatic involvement* (a feeling of being caught up in the plot and characters of a story). All three components could refer to experiences outside of any medium, but for this discussion I assume these are pleasures derived from media experiences.

I arrived at this framework partly based on Murray's three aesthetics (1997), but primarily out of my own experiences creating the immersive and interactive story *AR Façade* (described in Chapter 3) and my studies of the player experience detailed throughout the dissertation. My choice to elevate the three experiential pleasures does not rule out other important experiential pleasures that may contribute to the overall experience, but these are the most relevant for my discussion of immersive and interactive stories. Moreover, I do not dismiss the efforts of media designers who choose not to emphasize these aesthetics or who pursue an overall goal other than ENE.

As psychological states in a participant's mind, each of the three experiential pleasures require a degree of complicity on behalf of the participant. Laurel describes the cognitive and emotional capacities humans muster to "give ourselves" to a medium (Laurel, 1993, p112). This notion of complicity is expressed in Coleridge's notion of the "willing suspension of disbelief" where an audience member can accept as true the premises of a theatrical play, even if they are fantastical or impossible (Coleridge, 1817). The "willing suspension of disbelief" usually refers to passive narrative experiences (plays, films, novels, etc.), where the spectator/reader does not necessarily have a narrative role or agency in the course of events, and the goal of the playwright is to minimize unintentional disturbances. The concept applies equally to presence and agency, as participants must allow themselves to feel part of a simulated world and to feel as if their actions are making an impact. In the design of immersive and interactive stories, the material properties (narrative structure, interactivity, and perceptual immersion) can only be manipulated to a point. Participants achieve the full illusion by cognitively closing the gap. When the illusions of these three psychological states succeed, an opportunity exists for embodied narrative engagement, marked by improvisational enactment and deep catharsis, which I describe in more detail in Section 2.4.

2.2 Three Realms of Pleasures and the Material Properties

I will now define each region of the framework and discuss the related literature for both the experiential pleasure (i.e. presence) as well as the underlying material property (i.e. perceptually immersive interfaces). For each realm of pleasure I discuss the idealized notion and then present counter arguments in the literature that indicate maximizing the experiential pleasure may not always be optimal.

2.2.1 Presence and Perceptually Immersive Interfaces

The concept of presence is central to an extensive body of theoretical, psychological and technical work (e.g., Short et al., 1976; Heeter, 1992; Biocca, 1997; Lombard and Ditton, 1997; Meehan et al., 2001; Meehan et al., 2002) that extends over half a century and includes a yearly conference and a respected research journal. Presence refers to the subjective feeling that one's body is within a world, whether virtual, physical or some combination thereof. Individuals are said to feel presence in a medium when they feel as if they are "there" (Heeter, 1992), or in some sort of "direct" (Lombard and Ditton, 1997), "immediate" (Bolter and Grusin, 1999), "natural" (Lombard and Ditton, 1997). or "post-symbolic" (Lanier and Biocca, 1992, p160) connection to a medium. Presence is felt when the medium seems "real" (Rheingold, 1991), "transparent or invisible" (Marsh et al., 2001), or "disappearing" (Pimentel and Teixeira, 1993). In this definition, presence speaks to a tradition in HCI of talking about the "invisible" computer (Norman, 1998) or "seamless" interfaces (Weiser, 1994)[6]. One of the most cited definitions of presence is Lombard and Ditton's "the perceptual illusion of non-mediation" (1997), the notion that perceptually submerging a participant in an artificial media environment can result in the individual failing to perceive or acknowledge the existence of the medium.

Presence is often used synonymously with immersion, but I want to call out an important distinction highlighted in Murray's definition of immersion: "the sensation of being surrounded by a completely other reality... (taking over) our whole perceptual apparatus ...and learning to do the things the new environment makes possible." (1997) Murray's humanities perspective includes two notions: the feeling of being present or

---

6 Although Weiser draws a pronounced contrast between VR and ubiquitous computing, stating "the opposition between the notion of a virtual reality and ubiquitous, invisible computing is strong..." (Weiser, 1991)

submerging into the "perceptual apparatus," and the feeling of engaging in and being captivated by the action. The first part adequately describes the notion of presence as defined by the presence community; the second part goes beyond presence, towards other subjective constructs, such as agency and dramatic involvement. Ryan also offers a more expansive notion when referring to immersion, noting that the word applies to any absorbing activity, such as being "immersed in a crossword puzzle" (Ryan, 2001, p14). Backing up Murray, Ryan states that "immersion requires an active engagement..." (p15), but she also says "immersion is a corporeal experience" (p21).

The Virtual Reality (VR) community also struggles to describe the relationships of these experiential notions with the same ambiguous terminology. In Witmer and Singer's discussion of presence, they first define involvement as "focusing one's energy and attention on a coherent set of stimuli" and immersion as "perceiving oneself to be enveloped by and included in an environment" (1998). They then claim that "involvement and immersion are necessary for experiencing presence" (Witmer and Singer, 1998). From this VR perspective, immersion is one component of presence; from Murray's humanities perspective, presence is one component of immersion. In both communities, the terms presence and immersion become convoluted and interchangeable. I do not discredit Murray's or Ryan's definitions of immersion, nor do I refute the existence of such a construct. Rather, I will use the word *engagement,* as I have defined it below (in Section 2.4) as a potential illusion of feeling within a story environment, of feeling of empowerment over events, and of feeling caught up in the plot and characters of a story. I will reserve the word *immersion* (or at least the descriptor immersive) to refer to features or qualities of media technology that create sensory impact for the user. In the rest of this section, I examine the literature to unpack this more narrow and material sense of the word immersion.

Media that surrounds more of a user's perceptual system is more immersive, for example, encircling a significant portion of the spectator's visual field of view. According to Biocca and Delaney, perceptual immersion is "the degree to which a virtual environment submerges the perceptual system of the user" (1995). This term refers to how a virtual world encases a human's perceptual organs and motor systems. Biocca provides a detailed breakdown of perceptual embodiment in virtual environments, including a discussion of the human sensory/motor channels (Biocca, 1997). Many sensory and motor channels are integrated together requiring a tightly coupled directly manipulatable system of inputs and outputs to create realistic perceptual immersion, such as simulating the sense of touch with haptic technology[7], which I discuss further in Section 2.3.2.

Perceptual immersion can describe the human sensory bandwidth (e.g., spectators experience more immersion watching a film in a IMAX theater than on an iPhone). It can also refer to more physical forms of interfaces that take advantage of users' bodies (e.g., players experience more perceptual immersion with a Wiimote than with a classic game controller). The inputs and outputs to the game remain the same; the player is still controlling an avatar or shooting a weapon, but the interface is more perceptually immersive.

A number of currently successful commercial endeavors utilize research in immersive technology, including surround sound and large-screens (e.g., IMAX theaters), computer graphics (e.g., Pixar movies), physical input devices (e.g, the Wii), and virtual reality rides (e.g. Disney amusement parks). The evolution of immersive interfaces can be traced back to nineteenth century stereoscopic imaging inventions, such as the

7 http://haptic.mech.northwestern.edu/HapticResearch.html (accessed 9/22/08)

Stereoscope and the Stereographoscope[8]. One of the first ventures in full perceptual immersion came from Morton Heilig and his 1955 patent for the Sensorama[9], a device composed of mechanical parts that creates a sensory simulation. As Robinett points out, the content for Sensorama did not go beyond non-narrative simulations, such as riding a motorcycle through the streets of Brooklyn (Robinett, 1994). During the 80s and 90s, the vision of Virtual Reality (VR) received popular attention, partly due to the pop-cultural hype built up in movies such as *The Lawnmower Man* and books such as Rheingold's *Virtual Reality* (Rheingold, 1991). Virtual reality research continues today with advances in Head-Mounted Displays (HMD) and with numerous projection-based Computer-Aided Virtual Environments (CAVEs) around the world, as well as in military flight simulators[10] and driving simulators[11].

Augmented Reality (AR) is similar to VR, but rather than create a fully simulated world, the physical world itself is augmented with virtual content. Sutherland created the first "see-through" AR head-mounted display in 1968, allowing low-fidelity line graphics and text to be overlaid on the surrounding space (Sutherland, 1968). AR technology was developed primarily for its potential to augment practical work situations, such as equipment repair (Feiner et al., 1993), minimally invasive surgeries (State et al., 1996), and outdoor information services (Feiner et al., 1997). Since MacIntyre et al. pressed the idea of exploring AR as a medium (2001), there have been numerous entertainment endeavors, including outdoor games in immersive AR (e.g., ARQuake (Thomas et al., 2000) and Human Pacman (Cheok et al., 2004)), AR card games on handheld devices

---

8 http://cprr.org/Museum/Ephemera/Stereo-Viewers.html (accessed 9/22/08)

9 http://en.wikipedia.org/wiki/Sensorama (accessed 9/22/08)

10 http://en.wikipedia.org/wiki/Flight_simulator (accessed 9/22/08)

11 http://www.nads-sc.uiowa.edu/(accessed 9/22/08)

(Wagner et al., 2005), and some narrative-based AR experiences (Moreno et al., 2001; MacIntyre et al., 2003). In Chapter 3, I discuss details of the design issues we faced when creating a narrative-based AR experience for entertainment.

Others have focused less on simulation, and more on exploring the expressive power of our sensory palette. Utterback describes the physical computing movement as an exploration of "the interface between physical bodies and various representational systems" or a focus on the materiality of the "connective tissue between our bodies and the codes represented in our machines" (Utterback, 2001). Krueger's early artistic creations of "artificial reality" conceptually explored the human body as a medium (1990) (along with Stelarc[12]).

In the presence community, many researchers have worked on isolating and quantifying the effects of particular variables of the interface, the so-called "immersion factors", looking at frame rate and passive haptics, (Meehan et al., 2002); aural effects, (Kramer, 1995); system responsiveness (Steuer, 1995); interface allowances for body movement (Slater, 1998); and others (Heeter, 1992; Sheridan, 1992; Witmer and Singer, 1998; Meehan et al., 2002). This focus on engineering the immersiveness of various display technologies has led to efforts to measure presence and the effects of the immersion factors. Presence is typically measured in one of three ways: objective measures of reflexive or physiological response (Meehan et al., 2002), subjective measures such as questionnaires and observation (Witmer and Singer, 1998), and assessments of the user's ability to differentiate between the physical and virtual world (Barfield, 1995).

One of the most notable presence experiments takes place over a VR "pit" at the University of North Carolina (Meehan et al., 2002). Meehan and his collaborators

12 http://en.wikipedia.org/wiki/Stelarc (accessed 9/22/08)

collected physiological data and subjective measures to prove that frame rate positively impacts presence. Except for the fact that participants were standing next to a big virtual hole, the VR pit environment was intentionally "dry". Adding narrative structure to the environment would have confounded the physiological data. One could argue that a non-immersive, engaging narrative experience would solicit stronger physiological responses than a believably immersive, non-narrative experience. That is not to say that the "pit" methods are not valuable, but that a broader experiential notion of engagement likely trumps presence when it comes to creating an emotional impact. Many presence researchers recognize that physiological measures of presence are less useful when the user is interacting with real content. Presence questionnaires (PQ) present an alternative, but introduce subjectivity and ambiguity. The Witmer and Singer PQ originally developed for VR asks the user if the "virtual environment seems consistent with your real-world experiences?" (1998). In this question, "consistency" is ambiguous, since a VR or AR experience can be perceptually consistent with the real world, or the content can be consistent with the real world (i.e., realism vs. fantastical or surreal content). Also, it is unclear the degree to which the PQ applies to AR, where media is integrated into a real environment.

Some theorists delineate the presence construct into different types. For example, Biocca draws out a distinction between physical presence (the sense of "being there"), self-presence (the sense of "feeling one's own body"), and social presence (the sense of "being with another body") (1997). Biocca's division between "being there" and having a sense "of one's own body" highlights one primary difference between VR and AR: in AR, the user already interacts in a real, physical place, removing the need for an interface for controlling a virtual body. A study by Tang et al. comparing VR and AR showed that the mismatch between actual body and virtual body in VR can lead to a decrease in presence, and that AR allows for more natural body movements (Tang et al., 2004). It is unclear how the concept of presence can be understood in AR, where being "there" in an artificial place may not be applicable. A similar issue arises when considering the so-called Mixed Reality (MR) spectrum (Milgram, 1994) and the various incarnations of physical computing (e.g. Ishii et al., 1997) where the user's physical presence is rarely considered. In the broader range of MR media the presence construct may need to be reconsidered.

Social presence taps into a thread of research focused on the "intimacy" or the "richness of interaction" with other social actors (Short et al., 1976; Rice, 1992), both human and non-human. In everyday situations (in "unmediated" reality), a compelling personality might have a strong "presence" or a good actor may have "stage presence." In relation to media studies, "social presence" research studies the effect of immersion factors on communication between two physically separated people (Lombard and Ditton, 1997). Other "social presence" work focuses on identifying and creating humanistic attributions and emotions for artificial agents (e.g., Cassell et al. 2000; Lee and Nass, 2003) and robots (Thomaz, et al., 2006), which I return to in Section 2.2.2.

The word presence also surfaces in the context of explaining one's involvement in dramatic content. The concept of "dramatic presence", originally presented in Kelso et al.'s study of staged actors, refers to a user's sense of "being in a dramatic situation" due to the culmination of sequential events (Kelso et al, 1992). Likewise, Schubert and Crusius discuss the "book problem" and argue that the psychological phenomenon of presence is the same in books and film as it is in virtual reality (2002). While they do concede that VR seems most strongly related to a sense of actually being there, Shubert and Crusius claim that the "transportation" in literature (Gerrig, 1993) and the "diegetic effect" in film (Tan, 1996) are used to produce presence through the power of narrative. I would categorize these notions under the broader definition of immersion (as defined by

Murray and Ryan), because they elude to a notion of being attentive or captive or giving oneself to the experience. I return to the concept of dramatic presence when I talk about the interrelationship between presence and dramatic involvement.

Again for clarification, I draw a distinction between presence (feeling "within" an environment) and engagement (interested in the content). Many researchers either talk about immersion—encompassing both notions—or conclude that greater presence leads to greater engagement. Lombard and Ditton claim "Presence implies a direct and natural experience rather than just the processing of symbolic data and is therefore likely to be more compelling" (1997). Baker et al. also support this link between VR immersion leading to presence and presence to engagement: "Immersion can contribute to a sense of presence, embodiment, and engagement with the virtual world that is rarely experienced at the desktop" (1998). Marsh et al. also argue that "transparency (in VR) keeps users in the 'flow' of their activities and consequently enhances experiences in users" (2001). In response to this often-implicit goal of greater transparency, media theorists Bolter and Gromala have argued that true transparency is a myth and point to interactive art as a counter tradition of compelling non-transparent forms of mediation (2003). Other presence researchers have questioned the mantle of "maximizing" presence, such as Ijsselsteijn et al. who suggest that for certain tasks "changes that diminish presence may enhance performance" (2000). My thesis also argues for a more nuanced view of presence in light of a broader goal of embodied narrative engagement, where presence may need to be sacrificed to produce the desired effect.

2.2.2 Agency and Interactivity

The dictionary broadly defines agency as "the condition of being in action" (Dictionary.com), but in this dissertation agency will refer to a feeling of empowerment over events within a medium and the motivation to take action. The agency concept has been described by technologists and media theorists, with some variability, but certainly less ambiguity than presence. According to Murray, “agency is the satisfying power to take meaningful action and see the results of our decisions and choices.” (1997) In a later publication Murray states that we experience agency when “the world responds expressively and coherently to our engagement with it” (2004).

In order for agency to occur, a media experience must provide a means for user input and must result in some impact on the output—a feedback loop I will refer to as interactivity. Any application that allows user manipulation—including word processing programs—arguably provide a sense of agency. As Mateas explains, it takes more than mere interface buttons and knobs to twiddle, if “all this twiddling has little effect on the experience, there is no agency” (2001). Mateas’ definition of agency posits the notion of a “balance between the material and formal constraints,” where the material constraints amount to interactive elements in the interface and formal constraints refer to high-level motivation for action afforded by the activity (2001). Agency only occurs with action.

Without an explicit interaction loop, prior media such as film and radio have not been able to provide a sense of agency. Some have argued that media can create a sense of agency without interactivity, for example Perlin, who describes the feeling of experiencing a character’s agency (2004). Perlin concedes that in a linear narrative spectators give up their own agency, but rather experience the story characters’ agency: “to read Harry Potter is to experience his agency” (Perlin, 2004). As I mention earlier, all three experiential pleasures I talk about in this dissertation require a degree of suspension of disbelief; thus, it is possible for a viewer to experience some degree of agency without interaction. However, the most powerful experience of agency only occurs when the medium provides interactivity.

Some researchers have attempted to characterize interactivity. Laurel defines interactivity in terms of a system's provisions for enabling a user to "act within a representation" and characterizes it through three variables: 1) frequency (how often you could interact) (2) range (how many choices available) (3) significance (how much the choices really affected matters) (Laurel, 1993, p20). Laurel claims that "significance" is a key aspect of agency, particularly in interactive storytelling, where "the effect of a player's actions on the plot has to be substantial" (1993, p142). Ryan discusses possibilities for interactivity in terms of an exploratory-ontological dichotomy, where on one end "the user is free to move around the database...but does not impact the destiny of the virtual world" (exploratory) and on the other, the user can "send the virtual world on different forking paths" (ontological) (Ryan, 2001; Aarseth, 1997). Indeed the question of how interactivity works within narrative structure is the center of debate between "narratologists" and "ludologists," as I discuss later in the chapter.

Again from Dictionary.com, interactivity is "the extent to which a computer program and human being may have a dialog." For Shneiderman and others who developed the concept of "direct manipulation," the dialog should be a two-way, reciprocal flow of information with three properties: (1) continuous representation of the object of interest (2) physical actions or label button presses instead of complex syntax (3) rapid, incremental, reversible operation whose impact on the object of interest is immediately visible (Shneiderman, 1983). Some aspects of direct manipulation, such as visibility and representation, connote properties of perceptually immersive interfaces. However, Shneiderman's description of rapid, incremental, and reversible actions is one of empowerment over the system towards a stronger sense of agency. The focus is on creating a tight coupling between a user's actions and the feedback provided by the system, such as demonstrated in recent research on tangible interfaces (Ishii and Ulmer, 1997; Patton and Ishii, 2007; Ishii, 2008). I refer to direct manipulation again when I discuss the intersection of presence and agency.

In another interpretation of interactivity, Artificial Intelligence (AI) researchers are developing intelligent software agents that would be able to interact with humans (Agre, 1997; Cassell et al., 1999; Cassell et al., 2000; Mateas and Stern, 2003), although AI-based agents do not necessarily require explicit user input to act on behalf of a user (Isbell, 2004). Mateas' discussion of interactive drama describes how an autonomous character can provide rich material and formal constraints at the level of language and thought, thus enabling high-agency experiences with strong character narrative content (2001). AI techniques are used most commonly for character agents in modern video games, especially for non-narrative functions such as adjusting character positions based on movements by players (Cox and Fu, 2005). Early research on software agents for everyday office tasks—such as Crowston and Malone's email sorting agent—have often manifested in an anthropomorphic form (Crowston and Malone, 1988; Xiao, 2001). Researchers of intelligent agents grappled with how the "dialog" or interaction would occur with these agents and how best to represent them.

The distinction between intelligent software agents and user-controlled direct manipulation interfaces sparked a contentious debate among researchers at the IUI and CHI conferences in 1997 (Schneiderman and Maes, 1997). Maes argued for users to give up some control, to "delegate" to software agents, while Schneiderman advocated giving user's full control so they "can be responsible for the decisions they make" (1997). Shneiderman also argued that to suggest computers can think amounts to "deception" and that "the user may feel poorly treated" (1988).

Designing humanistic characteristics into the digital agents has the advantage of providing lifelikeness, although some would argue it only provides an unwarranted illusion of personification. Suchman, for example, states “the personification of the machine is reinforced by the ways in which its inner workings are a mystery and its behavior at times surprises us” (Suchman, 1987). Anthropomorphizing a software agent can set high expectations for interaction. According to Laurel (1993), many people criticized the interactive agent “Phil” in The Knowledge Navigator 1987 promotional video for Apple Computer[13]. She reflected that many people regarded the character as unintelligent; Phil had human-like traits (a real actor), but did not have human language capabilities (because they tried to simulate actual dialog interaction). Laurel reports that a later version of the video used a cartoon representation and people supposedly found the new version much more likable because the character was more consistent and appropriate to the action (Laurel, 1993, p62). This notion relates to Mori’s theory of the “uncanny valley” where humans would actually respond negatively to robots that are near facsimiles of humans, but not quite (Mori, 2005).

Reeves and Nass provided evidence—through a long series of psychological lab experiments—that “individuals’ interactions with computers, television, and new media are fundamentally social and natural, just like interactions in real life” (1996). People do not need to perceive a humanistic form to attribute human qualities, as shown in studies of Roomba robot owners that revealed personification occurring within a non human-like form (Sung et al., 2008). Mateas’ exploration of “alien presence” is recent example of non-anthropomorphic notions of intelligent agents (Romero and Mateas, 2005).

Returning to the topic of agency, I believe a user’s sense of agency in any storytelling medium is likely distributed between their own sense of control and the empowerment of story characters. In immersive and interactive stories, I would distinguish between different sources of agency. One can feel agency through the ability

13 http://en.wikipedia.org/wiki/Knowledge_navigator (accessed 9/22/08)

to navigate a world ("movement" agency) or through the ability to modify the world ("object" agency) or through interaction with characters ("character" agency). Most relevant to my thesis is a notion of "narrative" agency—the ability to effect the course of events and outcome of a story. Agency is only felt when a user takes some action within an environment and believes that their actions impact events within that environment.

Even within this brief discussion, it is not clear that maximum agency (complete control through direct and reversible manipulation) can provide the optimal experience in all contexts. Relinquishing control gives ground to external forces, such as software agents that can work on behalf of the user to accomplish tasks. In narrative worlds, autonomous characters create opportunities for richer global agency by providing new material and formal affordances for action (and, since the basis of drama is conflict, the formal affordances will be communicated as conflict, even if it limits the user's sense of empowerment). Moreover, the powerful notion of "rapid incremental reversible operations" from direct manipulation theory appears in conflict with the notion of "non-mediation" to create a sense of presence. One could argue that less immediate experiences like web surfing could provide better content control (more agency). I will look more closely at specific interrelationships later.

2.2.3 Dramatic Involvement and Narrative Structure

There is a long history of discourse—going back to early Greek philosophy—of dissecting the poetics of storytelling and describing the emotional experience of becoming "involved" in a drama. Aristotle believed that imitation or "mimesis"[14] comes naturally to human beings, that our propensity to imitate and learn through imitation from a young age differentiates us from animals (Aristotle, 330 BC). By imitating life in

14 Mimesis is the Greek word meaning 'imitation' or the representation of nature.

drama[15], a mode of fiction represented in performance, humans are able to experience a catharsis[16] or a climax of emotions (Aristotle, 335 BC). I propose a simple definition for the experience of drama: an individual can experience dramatic involvement when they are interested and caught up in the characters and plot of a story.

Other literary theorists also talk about a mental state that moderates textual representations of stories, such as Kendall Walton's theory of fiction as make-believe and his concept of "mental simulation" (Walton, 1990). Others describe the feeling of being "transported" (Gerrig, 1993), "emotionally involved" (Green and Brock, 2000), or "lost" (Nell, 1988) in text. The notion of transportation—also used by Green and Brock—emphasizes the role of attention: "transportation is a convergent process, where all of the person's mental systems and capacities become focused on the events occurring in the narrative" (Green and Brock, 2000). According to Laurel, spectators can become so "engrossed in (movies), you forget about the projector, and you may even lose awareness of your own body" (1993, p16). Douglas and Hargadon (2000) state "the pleasure of immersion stems from our being completely absorbed within the ebb and flow of a familiar narrative schema... engagement[17] lies in our ability to call upon a range of schemas ... and to venture in the direction of authorial intention." Not surprisingly, the word "immersion" surfaces again as an appropriate description for a state of absorption in various media including books.

---

15 The word drama is derived from the greek word δράμα, meaning "action".

16 Catharsis is a greek word meaning "purification" or "cleansing." It refers to a sudden emotional climax that evokes overwhelming feelings of sorrow or laughter.

17 Douglas and Hargadon use the word engagement to refer to a reader's ability to reflect and critique the narrative as it is being read. My allocation is completely different as I state below, engagement refers to the reader's interest and involvement in the story.

One of the key experiential pleasures of drama is the ability to empathize with the characters in a story or to "vicariously experience what the characters in the action seem to be feeling" (Laurel, 1993, p16). Perlin states that when we read a novel or listen to a storyteller tell a story about a guy and a gal, "by some transference process we become that guy or gal during the duration of the story" (Perlin, 2004, p13). According to Perlin, the transference process only occurs "because we agree to give over our choice-making power and to passively allow the narrative to lead us" (2004, p12).

With the emergence of digital forms of narrative, some theorists have sought to expand the aesthetics, to strive for novel experiential pleasures. Murray's description of transformation expounds on the newfound possibility to "shape-shift" or to "masquerade… as a cowboy or space fighter" (1997). She then goes further, proclaiming "our capacity to imagine life from multiple points of view" (Murray, 1997, p161), and describing an interactor's ability to mutate, to take on various skins, or to mosaic patterns of thought.

The dramatic involvement that people feel (or that they allow themselves to feel) likely has as much to do with personal interest with story events and the specific content than anything else. Rather than proclaim what makes content compelling, I will now turn towards some of the underlying narrative structures that authors and storytellers have adopted over time to achieve the desired affects.

In Aristotle's Poetics he theorizes drama as six hierarchical categories: plot, character, reasoning, diction, lyric poetry, and spectacle (Aristotle, 335 BC). Laurel argues that formal cause flows all the way down through the hierarchy, and material cause all the way up. So, plot is the formal cause of character, character the formal cause of reasoning, etc., while spectacle is the material cause of poetry, up through character is the material cause of plot (1993).

The resounding points that remain relevant today are that authors construct a plot to explicate some theme (formal cause). Characters enact the plot through their thought processes, language, and patterns of actions (material cause). As Mateas explains, the audience only views the material cause (2001). If the drama is successful the audience can "recapitulate the chain of formal causation" and understand the motivation of the characters and the point of the plot (Mateas, 2001). Aristotle also described the features of a 'good tragedy', including completeness and unity[18], that allow an audience member to feel a sense of closure. Closure is a term now used in psychology to describe an individual's desire for an answer as opposed to prolonged ambiguity.

As Laurel explains, "the stuff of narrative is description, while the stuff of drama is action" (1993, p94). Drama goes beyond mere narrative because it is enacted through characters, intensified by manipulating time, and strives for unity of action (completeness). According to nineteenth century dramatist, a good drama follows a dramatic arc with five parts: exposition, rising action, climax, falling action, and dénouement (Freytag, 1863). By structuring narrative as such, the author attempts to build up emotional responses to characters and the events that take place.

Another strategy available to authors is choice of narrative voice or point of view. The different forms of narrative point of view are long-standing techniques from literature. The author can create a different effect depending on how the story is told: first-person through the voice of a protagonist character ("I sit"), second-person with the reader as the main character ("You sit"), and third-person through the voice of an external narrator ("Frank sits"). Second-person narrative voice is not typically used in film, but authors of interactive fiction have more recently employed this voice to relay to the

18 Definition of Unity derived from Aristotelian aesthetics according to Dictionary.com: represent action as taking place in one day (unity of time), as occurring within one place (unity of place), and as having a single plot with a beginning, middle, and end (unity of action).

reader what she is currently encountering. Authors have long been experimenting with different "voices" to situate the reader. The variations on literary point-of-view are more intricate, particular when paired with different tenses[19]. Authors and filmmakers also manipulate the time order of events (e.g. *Pulp Fiction* and *Memento*), playing with a notion of non-linearity, but generally the viewer is still escorted through the story in a way that maintains dramatic involvement.

With the advent of film and video games, it becomes necessary to draw a distinction between the literary concept of point-of-view (narrative voice) and the camera notion of point-of-view (camera viewpoint). In film and video games 'first-person' and 'third-person' are often used to describe the two most common camera viewpoints—from a character's vantage point or from an outside camera showing the characters. Filmmakers typically shift the camera viewpoint frequently within any particular scene (Katz, 1991) and most audience members never notice. In the course of over a century of filmmaking, directors have come to understand what each camera viewpoint provides to rhetorically effect the experience of a story. Camera shifting is used less in video games, which usually adopt a consistent camera viewpoint (e.g. first-person in *Quake*, third-person in *SIMS*), although some games give the player an option to switch views (e.g. *Mario Cart*). Ryan draws out a spectrum between internal and external perspectives (Ryan, 1997)[20]. In an internal viewpoint, the user "identifies with an avatar, or apprehends the virtual world from a first person perspective," as one does in a first-person shooter. An external viewpoint situates the user "as a god who controls the fictional world from above," as one does in a SIMS-like game. Camera viewpoint shifts

---

19 http://en.wikipedia.org/wiki/Point_of_view_(literature) (accessed 9/22/08)

20 Ryan points out that these axes are adopted from the typology of user functions and perspectives in *Cybertexts* (Aarseth, 1997)

do happen automatically in video games during non-controllable "cinematic" sequences, which, are actually more like short films.

Despite specific camera techniques and viewpoint shifting, most movies and TV shows employ a first-person narrative voice where the main characters act out the plot and speak as if events are happening to them. Occasionally, a third-person narrator's voice is employed to relay parts of a story as characters act in the backdrop (e.g., in *The Shawshank Redemption,* when Morgan Freeman talks about Tim Robbins' character "Andy"). Rarely (if ever) do films employ second-person narrative voice where characters turn and speak to the camera as if you are a character. Second-person gets used in television shows like *American Idol* when the moderator talks to "you" the viewing audience, although the audience is not an explicit character. The American version of *The Office* uses the technique as a play on reality-TV, where the characters turn and speak to the camera that documents the office antics. Video games typically employ second-person narrative voice during the game play and third-person during cinematic sequences, although there are exceptions.

The narratologist Ryan describes the different semiotic manifestations of narratives from diegetic narration (verbally telling a story), to mimetic narration (enacting a story through gesture and dialogue), to simulative narration[21] (events unfold within an artificial world that can be manipulated) (Ryan, 2001a). The emergence of interactivity leaves authors grappling with how to maintain dramatic involvement when the "reader" transforms into a story character. In Murray's discussion of transformation, she points out the difference between understanding a theme by becoming someone else versus exploring variations on a theme (1997). Mateas refers to these two stances as "transformation as masquerade" and "transformation as variety", respectively (2001). On

21 Simulative narrative was presented during a lecture at the Digital Media department at Georgia Tech on October 30, 2007

one hand Murray discusses the experiential potential of "morphing" into a character's shape and becoming deeply attached to the character's emotional state. Murray writes about the power of enactment as an "agent of personal transformation" (1997, p170), very much a cathartic emotional impact. On the other hand, transformation offers "new ways of mastering fragmentation" (Murray, 1997, p156), choosing between a multiplicity of perspectives. Murray claims that a sense of closure comes, not from reaching the end of a story, but in understanding the "work's structure" through exhausting the possibilities (1997).

Murray's discussion of transformation highlights a counter argument offered by some narrative theorists. The playwright Brecht pursued a philosophy that would undermine Aristotelian poetics, because of the tendency for spectators get immersed in the stories and lose their critical distance (Brecht and Willet, 1964). The approach practiced by Brecht in his Epic Theatre and developed further by Boal in his Theatre of the Oppressed intentionally pushes against suspension of disbelief to highlight the artificiality of the experience (Boal, 1979). Brecht and Boal were part of an ideological movement known as postmodernism[22]—a critique of the bourgeois, elitist culture. On the Grand Text Auto blog[23], reflecting on her book Narratives as Virtual Reality, Ryan contrasts the "realist school" of print narratologists who had "perfected the art of immersivity" with "digital texts... influenced by postmodern aesthetics, whose insistence on self-reflexivity is hostile to immersion." Then she poses a fascinating question: "whether this resistance to immersivity is a matter of ideological position, or whether it is inherent to the medium?" A reflection by Bolter—in early days of hypertext—seemingly supports the latter, that hypertext is a "vindication of postmodern theory" (Bolter, 1992).

22 Postmodernism refers to a philosophical movement that started around the 1950s as a rejection and reevaluation of common assumptions in architecture, art, literature, music, film, and theatre.

23 http://grandtextauto.org/2003/05/26/narrative-as-virtual-reality/ (accessed 9/22/08)

If self-reflexivity is indeed inherent to interactive media, it raises the possible contradiction with efforts to keep the medium (uncritically) immersive.

As another counter argument to the notion of "transportation as masquerade" some theorists point to the distance required to feel empathy with a story's characters. "Empathy is subject to the same emotional safety net as engagement—we experience the character' emotions as if they were our own, but not quite; the elements of 'real' fear and pain are absent" (Laurel, 1993, p120). This raised the question as to whether emotional distance is necessary to empathize with characters in a dramatic situation. For some people, participation in dramatic action (even non-immersive, non-interactive) may require a degree of detachment—distance that allows them to absorb the story themes, empathize with the characters, and to be critical of representations put forth by the author.

2.3 Combining Pleasures

In Murray's book, where she outlines three aesthetics of the *Holodeck*, she states that "the combination of pleasures, like the combination of the digital medium itself, is completely novel" (Murray, p181). In this section, I seek to examine the combination of pleasures made possible in immersive and interactive stories, starting with the specific interrelationships between presence, agency and dramatic involvement. Dissecting the Venn diagram for the embodied narrative engagement framework, I look at the relationships and the potential contradictions that arise as these pleasures intermix. Throughout this section, I also present prominent examples from computing and the media arts.

2.3.1 The interrelationship between dramatic involvement and agency

The juxtaposition of dramatic involvement and agency lies at the center of a contentious debate between narratologists and ludologists[24] and continues to influence how narrative intermingles with video games. The fundamental tension is nicely summarized by Costikyan's reflection: "Divergence from a story's path is likely to make a for a less satisfying story; restricting a player's freedom of action is likely to make a less satisfying game" (Costikyan, 2000). Adams says "interactivity is almost the opposite of narrative; narrative flows under the direction of the author, while interactivity depends on the player for motive power" (Adams, 1999). This notion of being able to control the flow of narrative is rooted in Aristotelian poetics. Authorial control over narrative linearity allows for the "intensification" of story events to manufacture conflict. Authors build up tension that can later be resolved, leading to the ultimate feelings of catharsis and closure.

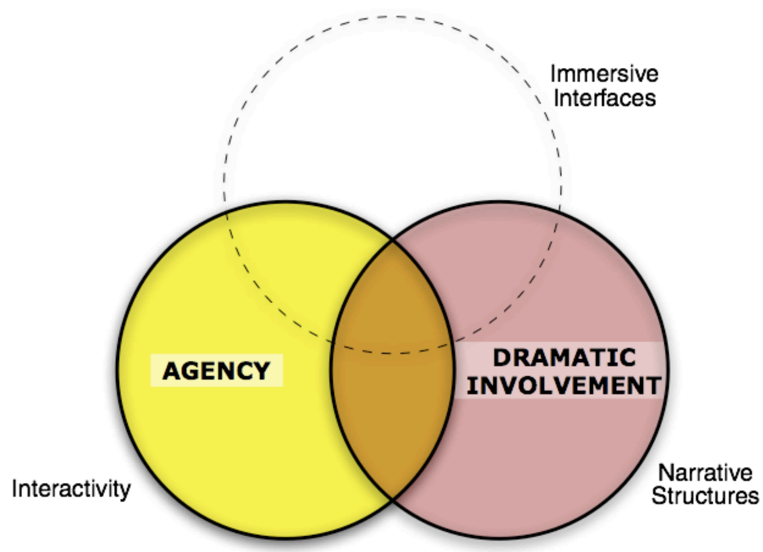


Figure 2.2: The interrelationship between agency and dramatic involvement

---

24 Where narratology is a long-standing discipline of narrative; the term "ludology" was put forth by Espen Aarseth (1997) to advocate a new field of study focused on games and game play, rather than framed in narrative discourse.

One could argue the very thing that gives linear storytelling its dramatic power is the fact that participants do *not* have agency. A spectator feels empathy for the characters and succumbs to the fate of dramatic events. Spectators are invested without having power to act. As Laurel points out, interactive narrative writers have long struggled with relinquishing control of plot lines and story direction while maintaining a dramatic effect (Laurel, 1993).

Towards merging narrative and interactivity, some theorists would argue that interactive video games—like all non-digital games—have elements of narrative. Murray says “games are always stories, even abstract games such as checkers or Tetris” in the sense that they “cast the player as the opponent-battling or environment-battling hero” (Murray, 2004). Pearce also argues from the narratologist perspective “a good game, even one without an obvious ‘storyline’ ... will tend to follow something that resembles the emotional curve of a dramatic arc” (2004). Although, she states “narrative has a profoundly different function in games than it does in other narrative-based media” (Pearce, 2004). While this could be a semantic argument over the definition of narrative, the self-proclaimed ludologist Aarseth argues “the pleasure of games is quite different from the pleasures of the novel” (2004), suggesting a fundamental move away from the classic pleasures of dramatic involvement.

Where video games designers think about infusing narrative into interactive simulations, authors of interactive fiction seek to bring procedural logic to stories. In her discussion of the poetics of interactivity, Ryan explores a range of interactive structures for narrative, such as ‘the network’, ‘the branching tree’, ‘the vector with side branches’, ‘the maze’, ‘the flow chart’, and ‘the braided plot’ (Ryan, 2001). But as many theorists have argued, including Ryan, the number of narrative paths quickly becomes combinatorially overwhelming. As Adams points out, an interactive narrative cannot “be all things to all players" (2006). As a player's interaction possibilities increase, the system requirements also become impossibly elaborate.

In an effort to conquer the "demands of interactivity," the Oz Project at CMU sought "to bring together writers, artists, and artificial intelligence researchers to produce fundamental technology that can support this new form of art and entertainment" (Bates et al., 1992), a form that has become known as interactive drama (Mateas, 2002). The group's research agenda included constructing computational methods and theory for interactive drama and developing autonomous agents that would integrate elements of perception, cognition, emotion, action, and language. Moreover, the group attempted to reposition the "reader" as a participant in a story, and to create artificial, yet believable, conversational partners—like *Eliza*. Weizenbaum's *Eliza* program was one of the first "successful" implementations of a conversational agent because it constrained the conversation topic to the format of a psychiatric interview (1966). The situational assumptions built into the psychiatrist's reflexive model (e.g., "Well, why are you unhappy?") provided the minimal structure needed for some conversational flow. Other researchers have produced believable conversational agents (Cassell, 1999; Johnson, 2000; Corradini, 2005), but generally their behavior is not integrated into a complete dramatic performance.

Ryan asserts that, "the central idea of interactive drama is to abolish the difference between author, spectator, actor, and character" (1997). In an early exploration of this reformulation of the player role, Kelso et al. experimented with the dramatic possibilities by structuring interaction with real actors where there was no computation involved (Kelso, 1992). An uninitiated "player" would enter a set with very little direction and begin interacting with actors who were free to enact their character's personality within a loosely-defined script. They found that players acted out within the dramatic moment—

what they refer to as dramatic presence—which I believe fits my notion of embodied narrative engagement (see Section 2.4). The player was “there”, involved in the drama, and able to influence the course of events.

Born out of the Oz Project, Mateas posits a theoretical framework to integrate agency into Aristotle’s poetics of drama. Making a case for interactive drama, Mateas says that when a player is integrated within the story as a first-person character, the player decides what she can and cannot do, and might possibly do, based on material and formal constraints (Mateas, 2001). Mateas places the “player character” in the Aristotelean formal/material causal chain at the level of the character and then explains that player agency can be experienced when there is a balance or a “sweet spot” between formal constraints (or motivations for action from the plot) and material constraints (conventions for action made available through the language and patterns in the medium) (2001). Mateas’ theory motivated his own research agenda to build *Façade*, “a first-person, interactive dramatic work … with a strong sense of agency.” (Stern and Mateas, 2005)

Whether *Façade* succeeds as both a story and a game is open to debate, but after years of studying the user experience of *Façade*, I believe it fails to entirely resolve the tensions that arise in the opposing natures of agency and dramatic involvement. As I describe in Chapter 5, many players in my in-depth episode analysis tended to change their behavior throughout the three phases of *Façade’s* narrative structure. As the tension mounted and the characters worked to resolve their marital difficulties, the player is often left feeling like they are simple watching scripted drama unfold. Moreover, in our large-scale online survey of players (see Appendix A and B), we found that a majority of players felt they had the least influence on Façade’s story during the later phases. A terse summary of my analysis would be: Façade is more like a game for the first half and more like a story during the latter half.

Others have theorized different solutions to this tension. Hammond, Pain and Smith suggest looking at Brectian and Boalian approaches in theatre where “the player both acts and reflects on narrative form from an outside perspective...” with the goal to “provide the player with the narrative construction kit most productive of player agency” (Hammond et al., 2007). The approach would provide players with high agency by giving reflective distance from the narrative. More recently, Murray has argued for the “replay story”, like *Groundhog Day*, where the protagonist can “do-over” events and experience parallel worlds without privileging any one of them (Murray, 2004). Murray cites the example of Cooper’s *Reliving Last Night*, an interactive video where the spectator can relive a woman’s previous evening, but can adjust three readable parameters—what she wears, what she serves, and what music she plays—that “taken as whole they present a fuller understanding...of the intriguingly rich space of possibilities” (Murray, 2004, p7). Mateas acknowledges the potential conflict between agency and Murray’s “transformation as variety” and argues that *Façade* supports this notion of variety through replay, because “multiple run-throughs have different, unitary plot structures” (2001).

Other examples that explore some combination of interaction, character, and narrative include *Virtual Babyz*, *Dogz and Catz*[25], *Winchester's Nightmare*[26], *Thespian* (Si et al., 2005), and *FearNot!* (Paiva et al., 2004), to name a few. Others have focused more on the underlying technology, such as story generation (Meehan, 1976; Riedl et al., 2003;

25 http://petz.us.ubi.com/ (accessed 9/22/08)

26 http://nickm.com/if/winchester.html (accessed 9/22/08)

Riedl and Young, 2004), believable characters (Cassell et al., 1999; Cassell et al, 2000), and intelligent drama management (Magerko, 2006).

I do not deeply examine the tension between agency and dramatic involvement in my dissertation. However, presence and the effect of immersive interfaces have deep interplay with both dramatic involvement and agency, and thus have some impact on this tension. The two subsequent sections discuss the paired interrelationships with presence, and thus are more relevant to my overall thesis.

2.3.2 The interrelationship between presence and agency

Laurel explores the connection between presence and agency with her observation that "immersion and agency are deeply interrelated. Without agency, we are simply absorbing images, or, as in the case of motion-platform rides, having ourselves shaken around by some other agency" (Laurel, 2004). It can be an uncanny experience to be physically immersed and to be swept about in a fantastical theme park ride, but without any control over one's body movement. Those experiences often circumvent the desire for interactivity by sensibly constraining the audience (e.g. a train car or a submarine boat), where the action typically happens outside the reach of the participants, giving them no sense of agency.

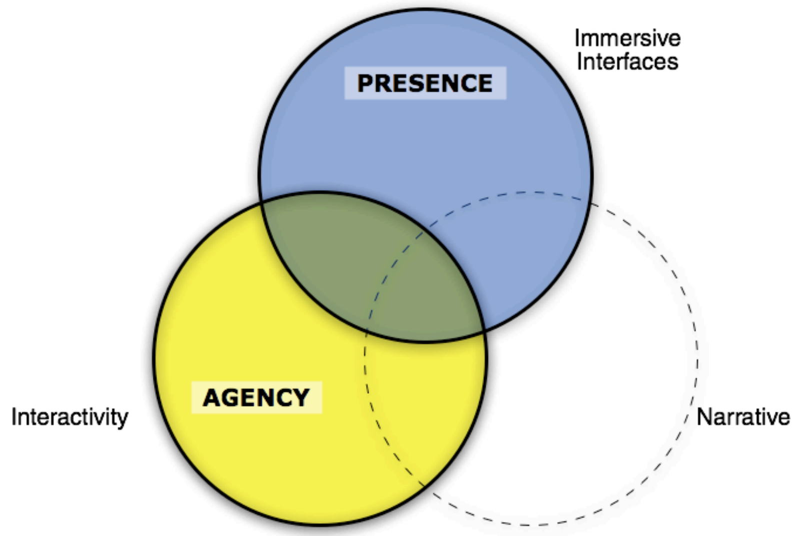


Figure 2.3: The interrelationship between agency and presence

Immersive virtual reality experiences—which induce a strong sense of presence—beg for more interactivity. Being "there" invokes a thirst for more agency, which is why questions about participant agency appear on presence questionnaires. For example, the Witmer and Singer presence questionnaire includes questions about control factors, asking users if they felt "able to control events" and if they felt the environment was "responsive to actions that you initiated" (1988). These questions are actually asking about the experiential pleasure of agency. They suggest that the interactivity that contributes to a sense of agency can also help users feel a connection to the media environment, and hence increase their sense of presence.

This mutual reinforcement of presence and agency is manifest in the concept of direct manipulation, originally coined and exemplified in several information visualization applications by Shneiderman (1983). Prior to the theory of direct manipulation, Sutherland demonstrated Sketchpad, the first example of directly-manipulatable pen-based graphical user interfaces (Sutherland, 1968). Direct manipulation can be used as a lens to chronicle the transition from command line interfaces to windowing systems to more recent research efforts in tangible computing

(Ishii, 2008; Jacob, 2008). These paradigm shifts are marked by improvements in user empowerment and increased visibility of information. As I discussed previously, direct manipulation captures the presence notion of visibility and representation, as well as the agency notion of rapid, incremental, and reversible actions.

While presence and agency can reinforce each other, this relationship quickly becomes problematic. Laurel warns about the "vague way agency is often handled in human-computer activity" citing two common violations: "uncomfortable holes in the mimetic context" and "vague forces ... of supplication rather than cooperation." (1993, p142). In the former, user agency is destroyed by a lack of awareness of system operations; in the latter, the user's agency is undermined by external forces bossing the user around so that she feels no control. Hutchins, Hollan and Norman provide a detailed cognitive account of direct manipulation in user interfaces, examining "the cognitive effort it takes to manipulate and evaluate a system" (Hutchins et al., 1986). They write about the "distance" between a system's interface and a user's goals. This distance can be bridged by minimizing the "Gulf of Execution" (or the user's difficulty in translating a goal into a physical action in the system) and the "Gulf of Evaluation" (or the user's difficulty in determining whether the system's response met the desired goal) (Hutchins et al., 1986).

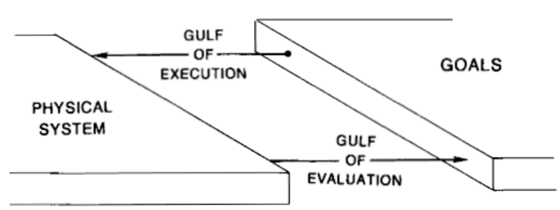


Figure 2.4: The gulfs of execution and evaluation from direct manipulation theory (Hutchins et al., 1986)

Specifically the 'semantic' distance deals with questions of whether it's possible to concisely convey what one wants to say and how much the user must translate or "process" system output to make sense of it. Hutchins et al. express the desire to reduce the semantic distance through system design—especially "high-level languages" that mediate "structure between intentions and expressions"—although they assert the user's ability to adapt to the system (1986).

Azuma discusses the possibility for direct manipulation in virtual reality interfaces and suggests that solving registration errors—or properly aligning and synchronizing virtual cues with real world events—will help minimize interface distance. He says "the power of direct manipulation comes from leveraging the user's previous experience with the real world" (Azuma, 1997). While this makes sense, it only solves half of the problem.

Immersive environments provide reality-like representations and invite users to rely on their instincts to form and execute goals. By instinct, users know what they can do by first assuming the rules of reality. They know what the action should accomplish and how they execute it (e.g., users know how to grasp for a mug). Initially users "don't have to think" because the interface builds on reality-like interaction, and thus users experience a narrow gulf of execution. However, non-mediated AR/VR interfaces often do not fulfill expectations, and thus players experience a large gulf of evaluation (trying to understand what the system understands). Their sense of agency is only destroyed when the system does not behave as expected. Without infinite computational power, the system will not be able respond to all possible user actions in emulation of reality (e.g., users expectations are not met if the mug is not programmed to respond). Without communicating the constraints, the gulf of evaluation is unnecessarily wide. Of course, when a users' actions are not met the first time, they close the gap cognitively by adjusting their expectations and goals. When Bolter and Gromala warn about "the dangers of transparency," they are referring to the problems that arise due to lack of proper affordances: “the interface will mask the operation of the system exactly when the user needs to see and understand what the system is doing” (2003, p55).

A potential conflict exists between the apparent lack of mediation in immersive interfaces and the constraints and affordances that support interactivity. As Laurel says, “people must be constrained” (1993, p99). The way forward has been explored extensively by proponents of tangible interfaces where the goal is “to bridge the gaps between cyberspace and the physical world” by providing “seamless couplings” between digital data and mechanical objects (Ishii and Ullmer, 1997; Patton and Ishii, 2007). For example, the MetaDesk attempted to physically embody many of the GUI desktop metaphors (e.g. windows, icons, handles as physical objects) (Ullmer and Ishii, 1997). In tangible interfaces, designers seek to take advantage of existing mechanical constraints and affordances, and include innumerable special cases that would be difficult to capture in a virtual experience. Haptic devices—which simulate the sense of touch on the computer through a directly manipulable force-feedback apparatus—also seek to create this tight coupling between natural movements and system responsiveness. Both tangible table-top interfaces and haptic devices have strict mechanical limitations, and thus are generally only applied to specialized applications (Thomas et al., 2001; Ryokai et al., 2004).

Other researchers have looked at the affordances of clothing as an interactive platform, such as Sha’s “topological media” spaces that continuously sense movements from garments and alter the artistic audio-visual environment (Sha et al., 2003). Advances in conductive textiles[27], tiny sensors platforms[28], “smart-skin” surfaces (Rekimoto, 2002; Rekimoto, 2008), and organic user interfaces (Holman and Vertegaal,

---

27 http://www.lumigram.com/catalog/index.php?language=en (accessed 9/22/08)

28 http://www.tinyos.net/ (accessed 9/22/08)

2008) are moving the field of ubiquitous computing towards a vision of computation in any object and interaction on any surface (Weiser, 1991).

While the efforts to tightly couple physical gestures to digital environments has been explored extensively, an open question remains over how and whether it is even possible to setup similar constraints with speech. It is an open question about how to effectively close the gulf of evaluation in speech interfaces. The immersiveness of the physical interface and the unbridled, ‘natural’ interaction raises user expectations for the system. This is particularly evident when anthropomorphizing the interaction, as expressed by the linguist Brennen, “people’s expectations about human/computer interaction are often inherited from what they expect from human/human interaction” (Brennan, 1990, in Laurel, 1993, p151). Speech interfaces, in particular, illustrate this dissonance as users enjoy using natural speech but then struggle to understand how the system understands their utterances. Many speech systems enforce some sort of language constraints, limiting the vocabulary the user can use to communicate commands (Cohen et al., 2000) or guiding the dialogue down a constrained path (Walker, et al., 1998). Others have explored multi-modal methods, such as the “tap-to-speak” (Oviatt et al., 1994; Oviatt et al., 1996), to signify the beginning of a speech utterance, resulting in better recognition accuracy. The problem is that most speech application domains involve a degree of complexity that require a wide range of language, which further complicates the problem of informing the user what language can be used and when. Brennen does offer a saving grace by noting “the fundamental ability of human beings to adapt to their conversational partners makes the whole human/computer enterprise possible” (Brennen, 1990, in Laurel 1993, p151).

Still other media endeavors seek to exploit the immersiveness of physical environments, rather than the physicality of the body. Digital media tours of historic sites and museums are popular and have been implemented in various ways (Cheverst et al., 2000; Vlahakis et al., 2001; Aoki et al., 2002; Dow et al., 2005). In these examples, users are perceptually immersed in a culturally meaningful space, but their interaction with the "information space" is conducted primarily through handheld button interfaces. The ARCHEOGUIDE project, for example, provides an index to contextual information based on location, closely emulating a mobile encyclopedia (Vlahakis et al., 2001).

The interrelationship of presence and agency highlights a tradeoff between natural expressiveness and actual empowerment. This notion has also been explored recently by Jacob et al with their concept of "reality-based interaction" where they emphasize "building on our pre-existing knowledge of the everyday, non-digital world", but acknowledges that reality might have to be "traded" in return for expressive power, efficiency, versatility, ergonomics, accessibility and practicality. (Jacob et al., 2008). The resounding question moving forward is whether providing explicit mediation—as affordances different than everyday life—diminished the sense of presence? More importantly, can designers strike a balance between the "reality-based" interfaces that give rise to presence and the "non-real" constraints/affordances that clearly communicate interactivity and support agency? As my thesis states, I argue that providing mediation within an immersive and interactive story experience increases a user's overall combined sense of presence, agency, and dramatic involvement (embodied narrative engagement).

2.3.3 The interrelationship between presence and dramatic involvement

When media environments combine first-person immersive interfaces and second-person narrative voice, a mutually-supportive interrelationship emerges. Many narrative theorists would argue that a good story connects to readers and draws them into a fictitious world. Some go as far to say readers are mentally immersed or "transported" into the story (Green, 2008), but I would argue that this is different than the perceptual notion of presence, as I've defined it here. Someone can be "there", but have no interest in the content. For example, in most VR presence experiments, the environments are designed as "dry" as possible so that small changes in physiological responses can be attributed to interface manipulations, not dramatic content. These two configurations of media (books and "boring" VR worlds) demonstrate the independence of presence and dramatic involvement.

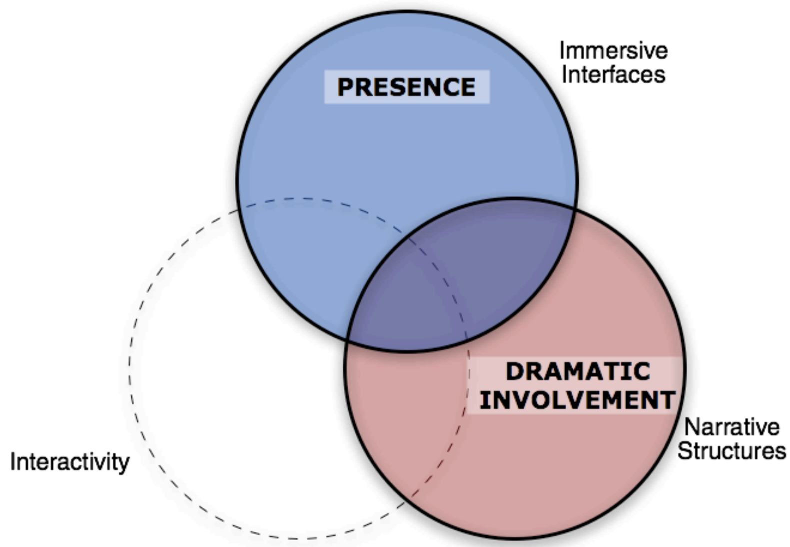


Figure 2.5: The interrelationship between presence and dramatic involvement

So combining perceptually immersive interfaces with a deeply involving story, especially one that uses second-person narrative voice (i.e., "you sit"), brings about the notion of "transportation as masquerade," as Mataes described in paraphrasing Murray (2001). Increasing a sense of presence through the use of immersive interfaces is widely believed to intensify spectators' dramatic involvement in content. Media richness theory states that richer, "higher-bandwidth" mediums are more effective at communication, because they can "convey equivocal information" (Rice, 1992). For example, Short et al. conducted a series of social psychology experiments in the seventies to show that visual media communicates more "warmth" and "sociability" over audio media, and likewise audio over text (1976). The thrust of most presence research is built on the premise that improving features like image size and quality, color depth, frame rate, stereo surround sound, network speeds, etc will inevitably improve the experience (Kramer, 1995; Meehan et al., 2003; Slater et al., 1998; Steuer, 1995). One can point to the continual increase in sales of consumer electronics as evidence of society's thirst for improved quality and richness in media[29].

In addition to increasing the sense of presence through the use of "first-person" immersive interfaces, the "player-character" role in the script can amplify the "transformation as masquerade" into narrative worlds. In fully simulated VR environments, a user can experience different "shapes", as Bricken illustrates "You can be the mad hatter or you can be the teapot" (Laurel, 1991, p372). In a narrative sense, the action is directed towards "you" in second-person narrative voice. As I discussed above there are many valid narrative voices available in writing and filmmaking, but immersive interfaces designed to induce a sense of "being there" may not provide as much flexibility.

According to Zettl, filmmakers utilize subjective camera shots in film, transferring the viewer from an event-spectator to an event-participant and encouraging them to "participate in an event psychologically" (Zettl, 1990). Horror films often explore the intimacy of subjective camera views to build dramatic tension and to increase the sense of presence, but rarely hold that viewpoint throughout (*The Blair Witch Project* might be an exception). Camera viewpoint shifting is commonplace in film, but in HMD-based or CAVE-like VR (and in AR, especially), the camera viewpoint is always a subjective camera shot. In a fully-immersive display, unless the participant is given no movement

---

29 As a personal anecdote, I have a number of friends who have large flat screens or projection surfaces with surround sound who play first-person shooter video games with the lights off. They report feeling extremely present and engaged in the gaming experience because they have blocked out everything in the environment.

agency at all, the user's head and body movements typically determine their viewpoint into the world. User-controlled subjective cameras in virtual and augmented environments likely increase the sense of presence.

While virtual reality opens the possibilities for greater flexibly in shape-shifting (because a user's movement can be mapped to a fictitious avatar's) and for more built-in abilities (flying, falling, etc.), AR is bound to real-world constraints. In VR the participant can see herself (or her avatar) externally; in first-person immersive AR, the user's body image remains bound to the real world and the user can never leaves his or her body. Technically, first-person camera viewpoints are not required in augmented reality, although HMD-based AR usually places a small forward-facing camera(s) on the user's head. Chastine et al. experimented with camera shifting in AR for collaborative referencing tasks where one participant views an augmented world through a collaborating user's camera viewpoint (Chastine et al., 2007). In general, little research has been conducted on the effect of camera shifting and "out of body" perspectives in HMD-based AR and VR and how they effect presence.

The *Four Angry Men* augmented reality experience—based on the American courtroom drama *Twelve Angry Men*—explores a clever twist on the constraint of first-person display with second-person narrative voice (MacIntyre and Bolter, 2003). Participants see video-based virtual jurors through a head-mounted display, and at any time they can switch between different points of view by physically moving between chairs. The user witnesses the drama from the viewpoint of one of four jurors, and her perception of the scene in any particular chair reflects the expectations, beliefs, and prejudices of that juror.

Other seminal examples of media endeavors that explore the intersection of immersive interfaces and narrative include *Placeholder*, where visitors enter a VR environment and listen to bits of spoken narrative left by other visitors to the virtual environment (Laurel et al., 1994). Pausch et al. worked with Disney on a production of *VR Aladdin* for the EPCOT Center in Orlando, Florida (1996). In this VR experience, visitors fly a magic carpet though an ancient city marketplace, hearing stories from various characters during the flight (e.g., a parrot shopkeeper, etc.). In the “mixed reality” experience *Desert Rain,* visitors experienced three stages of a narrative through different media: a physical briefing room, a virtual world projected onto a water curtain, and a final physical space, staged as a motel room. Actors guide visitors to the water curtain space where the visitors are expected to navigate the fictional world to find a special virtual door. Rather than pass through the door virtually, players walk through a water curtain to find the final physical stage (Koleva et al., 2001).

Narrative-based audio tours of historic sites also operate at the intersection of narrative and perceptual immersion, although the display mechanisms are not particularly immersive. The Alcatraz audio tour[30], for example, uses authentic voices and sound effects to heighten the dramatic effect for visitors. Similarly, the *Voices of Oakland* is an outdoor location-based audio experience for an historic cemetery told in first person voice from the perspective of the deceased (Dow et al., 2005).

In many examples of AR and VR, the narrative voice is 3rd person, and holds a “god-like” or external view of a narrative world. The *MagicBook* technology employs a god-like perspective where characters graphically pop-out of the page and might even “talk” to the reader, but the artificial world is in the book, not surrounding the user (Billinghurst et al., 2001; Grasset et al., 2007). Other physical manifestations of story worlds, like Tangible Viewpoints (Mazalek et al., 2002), place the players above the action like most board games. The dramatic action of the narrative world all happens at a

30 http://www.alcatraz.cc/ (accessed 9/22/08)

distance, potentially sacrificing some sense of presence in the fictitious world. A strict notion of immersiveness (first-person display with second-person narrative voice) might be required to achieve believable presence in a fictitious world, but many authors may find it acceptable to sacrifice presence and, in turn, get more options for narrative voice.

Another counter argument to a strict view of presence is that increased intensity combined with a believable masquerade will only be acceptable to a point. Murray warns about the potential of continually enacting destructive patterns (Murray, 1997, p173). Ryan also raises a moral question about the possibility of the *Holodeck* and whether it crosses a threshold that prior media could not:

> *Any attempt to turn empathy, which relies on mental simulation, into first-person, genuinely felt emotion would in the vast majority of cases trespass the fragile boundary that separates pleasure from pain (Ryan, 2001a)*

The ability to empathize, or to really put oneself in another's shoes, is compromised by not allowing distance from the emotional drama. The worry is that rather than empathizing with a character's emotions from a distance, the emotions are problematically directed at the user. Similarly, building from Brecht's critique of Aristotelian poetics, Pinchbeck has argued that players of first-person shooter games are "steered towards an uncritical relationship with the affordances of the experience" and that "successful immersion implies, by definition, an acceptance of the rules of the artificial representation" (Pinchbeck et al., 2006, p7). These arguments proclaim that immersed spectators could lose critical distance and that emotions traditionally classified as empathetic could dangerously turn into first-person emotions. The question is whether viewers desire critical distance and whether designers would like to provide that distance to their audiences. Using less immersive devices or diverging from 2nd-person narrative voice could allow for more removed ways of telling the story. Such strategies might detract from an optimal sense of presence, but they would perhaps allow for more critical distance.

2.4 Defining Embodied Narrative Engagement

Building on the definition of engagement as discussed in the literature, I introduce the specific phrase embodied narrative engagement to refer to the feeling of being physically transported into a fictitious world, transformed into a story character, and able to influence the unfolding events. In the general sense of the word, engagement refers to a person's involvement or interest in the content or the activity of any experience that, by design, captivates people: a fabulous cuisine, an architectural wonder, a rock concert, etc. What is personally interesting can vary from individual to individual. In a broad sense, someone may be engaged in an experience for various reasons including curiosity, adventure, fear, arousal, self-reflectiveness, or perversity—so even an unpleasant experience can be engaging for some people.

Turning to the literature, what do others consider the ultimate goal of media experiences? In Computers as Theatre, Laurel holds up engagement as a "desirable—even essential—human response to computer mediated activities" (1993, p112). She states that "engagement has cognitive components, but it is primarily understood as an emotion" and "[Engagement] is the state of mind that we must attain in order to enjoy a representation of an action" (Laurel, 1993, p112-3). The word engagement emerges in core HCI theory when discussing direct manipulation. Hutchins et al. describe direct engagement as "a feeling of involvement directly with a world of objects in a domain" (1986). Bolter and Grusin use the phrase "authentic emotional experience" to describe the ultimate goal of any narrative media (1999). Murray goes further to suggest a societal desire for deeper expressiveness—that "every age seeks out the appropriate medium in which to confront the unanswerable questions of human existence" (Murray, 1997, p280).

Engagement can occur in any medium; one does not need to feel agency or presence to feel engrossed in content. One can be engaged in a novel, in the sense of relating to a character or being intrigued by the plot. Webster and Ho studied audience engagement in multimedia presentations by looking at user attention, curiosity, and intrinsic interest (1997). In the context of interactive media, Turkle refers to the "holding power" of video games (1995). Likewise, McMahan proclaims that engagement occurs when "a player (in a game) reaches a level of near-obsessiveness... sometimes referred to as deep play" (2003). Seif el Nasr offers further definitions of engagement and attempts to bring performance theories into the discussion of developing interactive narrative (2007). In the context of immersive media, Ryan's discussion of the ultimate goal of entertainment makes use of an interesting metaphor: "like taking a dip in a Jacuzzi: it is easy to get in, but you cannot stay in very long, and you feel tired when you get out" (Ryan, 2001b, p11).

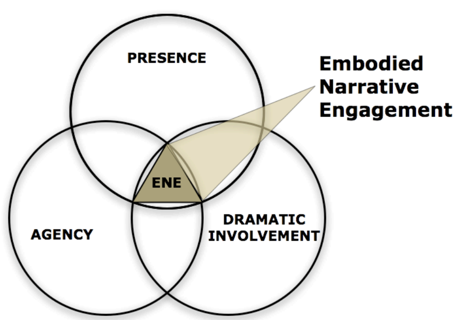


Figure 2.6: The intersection of presence, agency, and dramatic involvement: the theoretical psychological construct, Embodied Narrative Engagement

I conceptualize embodied narrative engagement at the intersection of presence (the feeling of being within an environment), agency (the feeling of empowerment of actions), and dramatic involvement (the feeling of being caught up in the plot and characters of a story). For this dissertation, I will focus on engagement as it pertains to media experiences, specifically narratives (to distance it from a host of non-mediated everyday activities such as fine dining as well as non-narrative media experiences such as video games). And, since "narrative engagement" would apply to any narrative media, even books, I have added the "embodied" descriptor to encompass the sense of one's body and empowerment, of being in the fictitious world and in the moment of action. Embodied narrative engagement is intended to describe the experiential pleasures during the media experience, not before (expectations. anticipation, hype, etc.) nor after (satisfaction, recollection, etc.), although certainly a good Immersive and Interactive Story will likely lead to strong feelings of satisfaction and memories of the experience. Embodied narrative engagement is not experienced uniformly by all people, but impacted by cultural context and individual factors, much like books, film, web content, and all other media.

The material independence of immersive interfaces, interactivity, and narrative structure is evident. Media experiences can be a combination of two or three of these elements and completely void of others. As Ryan states "there is no compelling reason for a VR application to be both narrative and interactive. In fact it could be neither" (1997, p3). The experience of "flying through" (or rather being flown through) the Grand Canyon in a 360° theme park theatre can be quite exhilarating. Likewise, a book is neither interactive or immersive, and a word processing program is neither narrative nor particularly immersive.

The framework for Embodied Narrative Engagement can serve as an organizing taxonomy for prior work from various research communities.

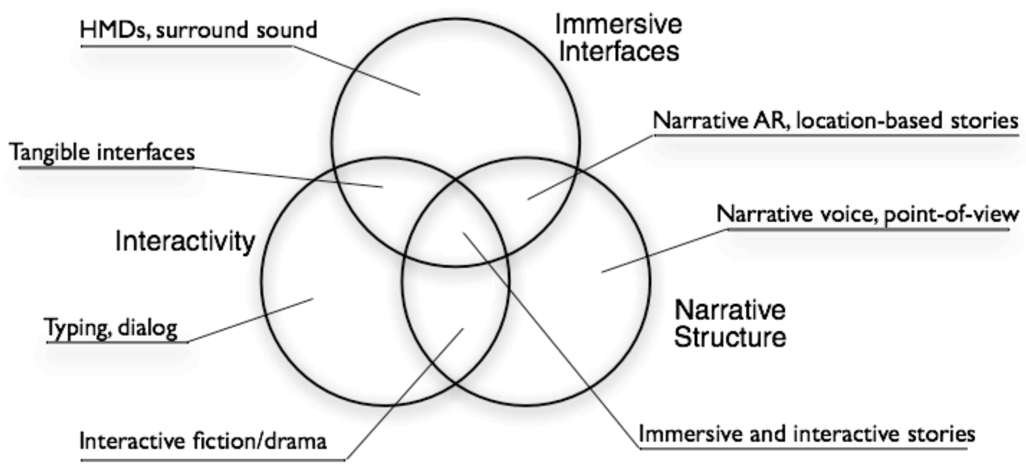


Figure 2.7: Taxonomy of examples of immersive and interactive stories

The *AR Façade* experience—the immersive and interactive story created as part of my dissertation and detailed in Chapter 3—falls into the center of this taxonomy as does a number of notable examples from research. In the KidsRoom project, researchers created a physical children's playspace equipped with cameras to sense the movements of participants (Bobick et al., 2000). The researchers designed a simple, linear narrative that sufficiently constrained the space of possible interactions so that participants could enter the experience unencumbered by sensors. The narrative was not adaptive, but it managed to offer kids a sense of agency by requiring them to physically interact.

In the "Mad Tea-Party" AR experience created by Moreno et al., the user plays the character Alice and sits at a table where she can see video of the Mad Hatter, Dormouse, and March Hare embedded in the space (2001). The experience explores perceptual immersion and true first-person narrative point of view, and it attempts to give the user a sense of agency by procedurally launching into video sequences based on user gestures, such as splashing tea with a physical teacup. However, narrative structure is not particularly complex as it simply cycles through media segments.

Disney's *Turtle Talk with Crush*[31] is an interactive "living creature" amusement park attraction based on a *Finding Nemo* character. Audiences sit down in front of large display (an animated view of the ocean) and interact with a 3D computer-generated model of Crush. The "magic" of this attraction is that a hidden actor views the audience through hidden cameras and speaks the voice of Crush. Crush's animation synchronizes to the actor's voice in real-time and audiences feel like Crush is truly alive, especially when Crush calls directly on guests to interact with him. (e.g, "Hey little dude with the bright red shell!"). Although the experience does not attempt true first-person immersion (guests are not "in" the water), it cleverly exploits the affordance of a fish tank so that it feels like Crush is "there," just on the other side of the glass.

Projects at USC's Institute for Creative Technology (Hill et al., 2001) and UCF's Media Convergence Lab (Hughes et al., 2005) have also attempted to bring together mixed reality, natural speech, gestural interaction, and AI-based story engines. Many of these experiences have been created for learning or training purposes, such as military reconnaissance or phobia treatment (Traum et al., 2007; Parsons and Rizzo, 2008), although other experiences, such as MR Time Portal, have strong entertainment value as well (Stapleton and Hughes, 2003).

For such examples of immersive and interactive stories to be successful, and to possibly reach that state of embodied narrative engagement, there must be a degree of suspension of disbelief (Coleridge, 1817). When interactive and immersive stories succeed in creating that illusion, an opportunity exists for improvised enactment and deep catharsis. However, interactive and immersive stories can go too far, potentially infringing on the individual, and not allowing for "transformation as variety" (Mateas,

31 http://en.wikipedia.org/wiki/Turtle_Talk_with_Crush (accessed 9/22/08)

2001) or more fragmented or "hypermediated" (Bolter and Grusin, 1999) ways of approaching the medium.

Murray often solicits critique for her fantastical hopes for the new medium, such as proclaiming that "the *Holodeck*...provides a safe space in which to confront disturbing feelings we would otherwise suppress; it allows us to recognize our most threatening fantasies without becoming paralyzed by them" (1997, p25). However, Murray does heed warnings for the potentially frightening consequences of engaging fully the thoughts and emotions of simulated roles. Her recommendation to harbingers of the new medium: "the more fully constructivist the story environment, the more opportunities it will offer to move beyond the enactment of destructive patterns" (1997, p173).

Ryan works to dispel the myth of the *Holodeck*: "Even if the hardware and software problems could be resolved, an important question remains. What kind of gratification will the experiencer receive from becoming a character in a story?" (Ryan, 2001a). She argues that people may not want this formulation of the medium, even if technologists could solve the vast challenges. These prognostications from media theorists delight in the possibilities and forewarn the dangers. Similarly, my conceptual framework outlines potential contradictions—beyond any technical challenge—that must be considered in the design of interactive and immersive stories.

While the three experiential pleasures of presence, agency, and dramatic involvement mutually reinforce each other, as aesthetics with independent goals, they can also compete in mutual opposition. Specifically, I claim that increasing the sense of presence does not maximize the overall sense of embodied narrative engagement, because complete transparency hides interaction mechanisms that strengthen the sense of agency, and does not provide sufficient means for users to manage their distance from dramatic content.

2.5 Other Experiential Concepts

I would be remiss if I did not discuss my notion of embodied narrative engagement with respect to play, performance, flow, and phenomenological theories. In this section, I discuss how ENE differentiates and overlaps with these other broad areas of theory.

2.5.1 Play Theory

Is embodied narrative engagement a form of play? ENE certainly overlaps if we accept Ellis' definition of play as "behavior motivated by the need to avoid boredom and maintain arousal" (Ellis, 1973, p17). However, as Aarseth implies, play seems to serve a function different than dramatic involvement (1997). According to Zimmerman "play is the free space of movement within a more rigid structure. Play exists both because of and also despite the more rigid structures of a system" (2004). This dry definition applies as well to the mechanics of games as it does to the "play in a steering wheel" (Zimmerman, 2004). Play is more associated with games than with stories, since typically narratives do not leave room for play.

Loizos, who studied play in higher primates, noted that behavior patterns from contexts with immediate and obvious intentions are applied to play situations where they "seem to be divorced from their original motivation and are qualitatively distinct from the same patterns appearing in their originally motivated contexts" (Loizos, 1969). We transfer behavior patterns from "serious" context into play contexts. Certainly, in immersive and interactive stories that imitate real-world contexts, behavioral patterns are likely to emulate the real context.

In Sutton-Smith's theoretical discussion of play, he states "there is no notion more characteristic of human achievement than the creation of illusory cultural and theoretical worlds" and that our "participation in such play worlds can be seen not as a defect... but rather as participation in a major central preoccupation of humankind" (Sutton-Smith, 1997, p54). Aristotle would agree that the creation of simulated worlds is part of being human. Schechner poses the question, "does viewing tragedies or playing video games dull people to pain or train them to administer it" (2002)? Is play used as a way to escape reality or a means to deal with it?

Huizinga defines the notion of the "magic circle" as a necessary and clear boundary for play. The game world is complete and separate from everyday life, and requires players to adopt a lusury attitude when stepping away from everyday life into a game space (Huizinga, 1971). Human play environments (e.g. most games and sports) generally provide a means for calling "time-out." Some contemporary game designers have experimented with blurring that boundary, to create games and play environments that become interwoven into economic, educational, and other social contexts. (e.g. "serious" games[32], alternate reality gaming[33], pervasive games[34], live action role-playing (LARP) games[35]).

Similar to embodied narrative engagement, play involves bringing known behavioral patterns into a simulated world and it may require an explicit need for stepping away from the simulated world. Play differs from embodied narrative engagement, in that it occurs in animals and can be as simple as a momentary diversion. The unique human ability to tell and engage in stories sets ENE apart from the more general idea of play.

---

32 http://www.seriousgames.org/ (accessed 9/22/08)

33 http://en.wikipedia.org/wiki/Alternate_reality_game (accessed 9/22/08)

34 http://www.pervasive-gaming.org/ (accessed 9/22/08)

35 http://en.wikipedia.org/wiki/Live_action_role-playing_game (accessed 9/22/08)

2.5.2 Performance Theory

Is the experiential pleasure of embodied narrative engagement similar to acting? When actors act, they indeed enact physical gestures and speech as part of a narrative. Actors play the role of a character within a drama with other characters, potentially getting “wrapped up” in the emotions to pull off more believable performance. A number of performance theories have been developed to explain how humans can pull of the ability to manipulate their own capacities. Early 20th century Russian director Stanislavski was motivated to analyze and codify performances after observing theatre actors over several nights, where sometimes the actors were completely “being a role” and other nights their performances were less inspired. Stanislavski created a “system” where an actor analyzes deeply the motivations and emotions of the character by asking themselves “magic if” questions to understand the emotions they would feel if they were in the character’s situation (Stanislavski, 1936). In the 40s and 50, Strasberg advanced a notion of acting called “The Method,” where an actor recalls emotions from his or her own life and uses them to identify with the character (1988). Professional actors seek to take on emotions that match the character being portrayed.

In many ways, embodied narrative engagement seeks to situate a participant’s mind similarly to deeply engaged actors. However, in most performance forums, actors are following a script and have emotionally prepared for the situation ahead of time. The sense of agency or empowerment over the unfolding events is irrelevant since the actors are simple executing the authors intentions. Improvisation, on the other hand, bears an even closer resemblance to ENE because actors perform in the moment and in response to unfolding events from the environment. Improv theatre often takes the form of comic performance[36]. The improv practiced in “Theatre of the Oppressed”, operated by Boal,

36 Chicago’s Second City, the TV show “Whose Line Is It Anyway?”, the traveling act Theatresports, to name a few.

was meant to be serious rather than comedic, but it achieved its goals by intentionally disrupting the dramatic flow (1979). Actors would disregard the "4th wall" and solicit suggestions from the audience for how a character on stage should overcome some oppression (Boal, 1979).

Much like games and play environments, there is a clear divide between the character role and a social role in everyday life. According to the performance theorist Schechner, "most people know the difference between enacting a social role and playing a role onstage — wearing the clothes, making the gestures, uttering the words, maybe even feeling the emotions of characters in a drama" (2002, p171). In Schechner's broad view, "performances are actions" (2002, p1). This "everyday" notion of performance is shared by Goffman who defines performance as "all the activity of a given participant on a given occasion which serves to influence in any way any of the other participants" (1956, p15). Similarly, de Certeau sought to understand everyday practice by looking at the "ways of operating" and articulating the tactics and strategies used by people to negotiate a culture of production and consumption, including media representations (de Certeau, 2002). One question is how authentic are the behaviors and emotions exhibited by participants in Immersive and Interactive Stories? During these simulations, do participants perform "mock" behaviors or genuine behaviors borrowed from everyday life? Do participants pre-mediate their actions or do they react extemporaneously?

Schechner stresses that "every performance is unique" and must be considered within the cultural/historical context (2002, p30). Benjamin also wrote about the "unique existence" and "aura" of original artwork (1936). He rationalized that mechanical reproductions of art—even the most faithful reproductions—lacked "presence in time and space" (Benjamin, 1936). Reproductions continued—driving a transfer of power from the elite to the masses—into the digital age where Benjamin's concept continues to be influential. Bolter et al. apply the concept of aura to new media, and mixed reality in particular, arguing that digital authors now have opportunities to build on the authenticity of physical spaces, such as historic cemeteries (2004). One could argue that the introduction of interactivity into digital art places less emphasis on the aura of the creation itself, as it does each performance of the creation. In that sense, our culture would place more auratic emphasis on witnessing each live musical performance, rather than a studio recorded performance. A user's performance within a simulation or game would be more auratic than the simulation itself. However, Auslander argues that media technology has encroached on live events—hyper-mediated through amplification and amalgamation—to the extent that they can hardly be called "live" performance (1999).

Looking at the potential experiential affects of simulation media, some theorists have noted the metamorphic effect of performance. Schechner differentiates between "transportation" (temporarily moved or touched) and "transformation" (a permanent change to who people are) (2002, p72). Penny contemplates the ethics of simulation because of the potential to "desensitize" participants to physical actions citing the example of phobia patients who enter VR worlds (2004). The simulated worlds are used for this serious purpose "precisely because they have impact on people's lives" (Penny, 2004, p74). Penny raises ethical questions about the consequences of violent video games that "hardwire" young people and tunes their minds for shooting at humans (2004). Baudrillard goes further, warning about the difference between acting and simulating. He notes that "to simulate is to feign to have what one hasn't" (Baudrillard, 1981, p5). His argument is that feigning something one does not have can actually produce the presence of those authentic emotions. Feigning illness results in actual sickness. Baudrillard argues that simulation threatens the difference between the "real" and "imaginary" (1981).

Mixing reality also leaves open the risk of distorting and masking reality, and then ultimately creating new realities (Baudrillard, 1981).

From specific performance methods, to the experiential pleasure of acting, to the ethical questions, performance studies can inform the design of immersive and interactive stories. Laurel has argued for new media adaptations of conventions from theatre and film, such as the theatrical notion of exits and entrances and the "panoramic close-up" from film (Laurel, 1993). Montola's exposition of live-action role-playing (LARP) games (e.g. Dungeons and Dragons) could provide an useful framing for some first-person, interactive digital experiences (2007). Although LARP experiences require no digital medium, it will be useful to understand how they script the narrative environment and structure interaction around power hierarchies and personified character constructs (Montola, 2007).

Some examples of media experiences allow trained actors to perform through the medium, such as Disney's Turtle Talk with Crush[37] and the Quasi robot[38]. The idea of using masks or puppets as a performative technique is not unlike the Wizard of Oz, where an elaborative illusion is pulled off by a man behind the curtain (Baum, 1900). Developing "behind the curtain mechanics" can lead to more actor-driven animated characters and robots.

2.5.3 Flow Theory

Is embodied narrative engagement a form of flow? The concept of flow, defined and studied by the psychologist Mihaly Csikszentmihalyi, describes, "the holistic experience that people feel when they act with total involvement" (1975, p36). In contemporary terms, this is the state of being "in the zone". The theory of flow is best

---

37 http://en.wikipedia.org/wiki/Turtle_Talk_with_Crush

38 http://www.interbots.com/

known for the “channel” diagram that maps out challenges against skills (Csikszentmihalyi, 1990). To achieve a flow state, a balance must be struck between the challenges of the task and the skills of the individual. If the task is too easy, a person can become bored. If the task is too difficult, the person might experience frustration. This aspect of flow closely resembles Vygotsky's theory of proximal development in learning situations, that we are only able to grasp the concepts just ahead of our development path (1978). Csikszentmihalyi’s flow theory has been instructive to many domains of life including religion, sports, and education. In the field of human-computer interaction, researchers have looked to flow theory to model web navigation (Hoffman, 2002; Pace, 2004), video game interaction (Johnson and Wiles, 2003; Sweetser and Wyeth, 2005; Chen, 2007), and direct manipulation interfaces (Bederson, 2004).

According to Csikszentmihalyi, the following are characteristics of flow (1975, p.72):

- Clear goals (expectations and rules are discernible)
- Concentrating and focusing, a high degree of concentration on a limited field of attention (a person engaged in the activity will have the opportunity to focus and to delve deeply into it)
- A loss of the feeling of self-consciousness, the merging of action and awareness
- Distorted sense of time - one's subjective experience of time is altered
- Direct and immediate feedback (successes and failures in the course of the activity are apparent, so that behavior can be adjusted as needed)
- Balance between ability level and challenge (the activity is neither too easy nor too difficult)
- A sense of personal control over the situation or activity
- The activity is intrinsically rewarding, so there is an effortlessness of action

- When in the flow state, people become absorbed in their activity, and the focus of awareness is narrowed down to the activity itself, merging action and awareness.

Several of the characteristics of flow also apply to my notion of embodied narrative engagement: a distorted sense of time, a loss of self-consciousness, and being absorbed in the situation. While flow reflects much of the language I use to define ENE (attention, absorption, involvement), it differs in several dimensions. Embodied narrative engagement does not require clear goals, direct and immediate feedback, personal control, or a balance between abilities and challenges. For most narrative media, it is not clear to me how to describe the experience as flow. For example, how is flow experienced when watching TV?

Csikszentmihalyi conducted his research on flow through a series of ethnographic shadowings, interviews, and surveys. One of his methods was to survey his participants at random times during the day about their activities and feelings at the moment. Others have attempted to operationalize and measure flow, particularly as it related to interface design (Bederson, 2004) and learning activities (Shin, 2006). Although many of the metrics of flow are unrelated to narrative engagement, measures of time distortion appear to have promise (Weybrew, 1984; Czerwinski et al., 2001). In a study of task performance with varying levels of challenge and assistance, Czerwinski et al. found that the more difficult the task, the longer the participants perceived the activity relative to the actual time (2001), providing evidence for Csikszentmihalyi's theory that people in a state of flow do not perceive as much passage of time.

The flow concept could prove to be very useful for certain types of immersive and interactive stories, particularly when the content is framed with explicit goals, or where the user is granted a strong element of control and feedback, as with most video games.

Csikszentmihalyi points out that not all the characteristics are required for flow (1990), so it might be that flow theory encapsulates embodied narrative engagement, but for now I am assuming a fine distinction.

2.5.4 Phenomenological Theory

Is embodied narrative engagement a subset of phenomenological theory? The language of phenomenology is incredibly similar to my discussion of embodied narrative engagement, particularly presence and agency. The philosophy of phenomenology developed by Husserl, Heidegger, Merleau-Ponty and others seeks to understand our consciousness of phenomena. According to Dourish, phenomenologists "analyze how we perceive and experience the phenomena of the everyday world" (2001, p105). Husserl founded phenomenological theory as a "study of the essence of consciousness as experienced from the first-person point of view" (Smith, 2006). Husserl's student, Heidegger went an important step beyond Husserl's ideas to reject the long-standing doctrine of mind-body Cartesian dualism held by Descartes (e.g. I think, therefore I am) (Dourish, 2001). Heidegger introduced the concept of "being-in-world" (*Dasein*) to emphasize how thinking and being are inherently linked, not disconnected, as if some homunculus runs the show in our minds (Dourish, 2001).

The phenomenologist Merleau-Ponty pays particular attention to the role of the body and our perceptual system (2005). Building on the notion that consciousness is always embodied, Merleau-Pointy believes the body is "our general medium for having a world" (2005, p147). Some contemporary philosophers have contemplated the "brain in a vat" thought experiment (as represented in *The Matrix* films)—that our minds are completely disembodied and perceive a simulated world, and thus how can we really take for granted anything we "know" to be true. This idea has been refuted by Putnam, Dennett and others on various grounds, strengthening the phenomenological claims of the embodied nature of our existence (Dennett, 1978; Putnum, 1981). Ryan suggests that "virtual reality systems act as a reminder of the productive implication of the body in the phenomenal world... VR offers a dramatization of phenomenological doctrine" (2001, p72). Our bodies are necessary for our existence, in physical and virtual realities.

Schutz developed phenomenological theory as it applied to the social world, arguing that our ability to understand each other starts with our "own lived experience" (Schutz, 1932, p13). He proposed an interpretative model of the social world that relies on the rational mind to see actions in the world within patterns of intentions (Schutz, 1932). Schutz's sociological orientation of phenomenology influenced prominent theorists in sociology, including Garfinkel in his concept of ethnomethodology (1967), and later, Suchman (1987).

Dourish, seeking to elucidate both social and physical phenomenological foundations for HCI practice, states that "embodied interaction... is an approach to the design and analysis of interaction that takes embodiment to be central to, even constitutive of, the whole phenomenon" (2001, p102). He points at trends in HCI, how tangible computing seeks to capitalize on "our familiarity with real world objects" and how social computing is concerned with with "situated" nature of actions in everyday social situations (Dourish, 2001). Related to my definition of agency, Dourish writes about the concept of coupling as a process by which actions are made effective in technological systems. Users manage this coupling and seek to turn objects of inquiry into useful tools (Dourish, 2001).

While phenomenological theory shares much of the same language as embodied narrative engagement, the main difference is that embodied narrative engagement is a theoretical psychological state and phenomenology is a philosophical position about our existence in the world. While virtual worlds are indeed artificial, they are experienced phenomenologically just as we encounter the non-mediated world. Towards achieving the illusion of embodied narrative engagement, immersive and interactive stories will seek to capitalize on physical skills as well as social and cultural “capital” used to negotiate and maintain our relationships to each other. Ultimately, it will be the users who will come to manage their coupling with the physical interface and with other social characters, both remote computer-mediated humans and artificial personas.

2.5.5 Summary of Related Theory

In this section I introduced play theory, performance theory, flow theory, and phenomenological theory as fields of thought closely related to my concept of embodied narrative engagement. Each can have important influences on the development of immersive and interactive stories, just as play and flow theory have influenced modern game designers (Chen, 2007). The concept of embodied narrative engagement can be differentiated from play and flow theory primarily by the notion of dramatic involvement (seemingly at odds with a sense of active participation). Performers may not experience the same level of agency as in embodied narrative engagement, unless they work without a script, as is the case with improvisational acting. Phenomenology speaks to the broadest notion of our existence in the world and encompasses the gamut of life experiences, from playing a game, to acting on stage, to climbing a mountain, to engaging an immersive and interactive story.

2.6 Chapter Discussion

In this chapter I have proposed a framework for relating the experiential pleasures and the material properties of the *Holodeck*. I discussed the experiential notion of embodied narrative engagement as an overlap of presence (a feeling of being within an environment), agency (a feeling of empowerment over events), and dramatic involvement (a feeling of being caught up in the plot and characters of a story). In exploring the literature around each concept, I revealed the potential conflicts that arise when the experiential pleasures are taken to their idealized logical end.

The interrelationships between the concepts suggest potential tradeoffs and important questions. If an author relinquishes control to a user, does it take away from their ability to intensify the occurrence of events, and in turn diminish their ability to "manufacture" drama? Do the efforts to make an interface "disappear" give users enough clues for how to effectively interact? Do first-person immersive interfaces allow for enough distance from dramatic situations so that spectators can become empathetic with characters?

Prior media endeavors only go so far towards answering such questions, typically only exploring one or two of these material properties (narrative structure, immersive displays, and interactivity). My own prior work has typically only focused on one or two of the design aspects described in my theoretical framework. Both the Voices of Oakland outdoor location-based audio tour (Dow et al., 2005) and the Four Angry Men augmented reality jury drama (MacIntyre et al., 2003), are solid endeavors of narrative plus immersive interfaces, but do not push strongly on interactivity. In the tangible installation work I helped create in the Topological Media Lab (TML), the goal was to create beautifully abstract immersive and interactive spaces, without a focus on narrative (Sha et al., 2003). None of my prior work go as far as *AR Façade* to effectively combine all these elements in the same experience. From an empirical standpoint, *AR Façade* also proved to be more suited for observational research. In *AR Façade*, not only are participants encouraged to physically interact—unlike Four Angry Men, where interaction is unnecessary—there was also a narrative logic for the things they were doing. Similarly, the TML projects did not incorporate narrative, although they did encourage bodily interaction.

The motivation for creating *AR Façade* has been to explore the user experience in the center of the ENE framework, and to attempt to find empirical answers to questions I have posed in this chapter, particularly with respect to first-person immersive displays. Based on the theoretical foundations and my analysis of the player experience in *AR Façade*, I believe that immersive displays do increase the sense of presence, but that presence alone does not lead to a strong sense of embodied narrative engagement. Efforts to make an interface "disappear" can take away from the equally important sense of agency. Similarly, striving for "immediacy" may not provide the distance desired for dramatic involvement. The overarching goal of embodied narrative engagement is best achieved by providing some mediation, even if it comes at the expense of the sense of presence.

CHAPTER 3

TECHNICAL DESCRIPTION OF AR FAÇADE

For a wicked problem such as game design, exploring design space consists of navigating the complex relationships and constraints among individual design features, while at the same time discovering or inventing new features and approaches that expand the design space. — Andrew Stern and Michael Mateas, Build It to Understand It Ludology Meets Narratology in Game Design Space (2005)

One way to understand a phenomenon is to isolate it through the manipulation of the material world. *AR Façade* is a media experience that attempts to simultaneously create a sense of presence, agency and dramatic involvement. The experience is novel and technically innovative, and allowed us to conduct extensive user studies to understand the effect of immersive interfaces. In this chapter, I will provide a technical description of the experience, including the video-see though AR interface, the mixed physical/virtual reality stage, a short description of the underlying AI-engine and graphics programmed by Mateas and Stern for the original *Façade* (2003), and the Wizard-of-Oz methods for enabling speech and gestural interaction[39].

I will describe our two installations of *AR Façade*—the Atlanta lab and the Beall Center gallery—and the changes we made between the two to improve the overall

39 The technical content in this chapter is also summarized in Dow et al., 2006.

experience for players and group audiences. For the eleven-week show at the Beall, I will describe the infrastructure put in place to support nine undergraduate art students running the show—and performing as wizards—without direct technical supervision from the researchers. Interviews with these wizard docents revealed a range of emergent strategies they adopted for maintaining and performing the experience. I also report on the group/ audience experience, setting the stage for a discussion of the solo player's experience in Chapter 4.

3.1 Making Façade "Embodied"

In this section, I introduce *Façade*, and highlight the system features that helped and hindered the conversion to augmented reality interaction. Physically, the desktop *Façade* experience occurs in a relatively small, fixed setting (Trip and Grace's apartment), so building a matching physical layout was feasible. The design of the apartment is purposefully minimal to focus the player's attention on the characters rather than the apartment (the minimalism also speaks to the characters' "artsy" personalities). The room's objects, the post-modern décor, the character's appearance and utterances are all meant to create a certain mood and social backdrop in *Façade*.

All the objects have symbolic connections to the backstory; manipulating an object evokes conversational references to the associated backstory topics. However, the decision-making engine only monitors whether an object is being manipulated (e.g., picked up, looked at, pointed at), not details of how it moves through space. Since *AR Façade* does not need to track the myriad physical objects precisely, simple approaches can be used to monitor them. These constraints, originally designed to focus the experience on character and story interaction and aid the AI storytelling engine, work to our advantage in implementing the embodied version.

3.1.1 A Short Primer on Façade

My collaborators and I build on the impressive work by Mateas and Stern, who created *Façade*—to much critical acclaim—as a fully-realized "experiment" in interactive drama (2003). *Façade* is the first fully produced, real-time, interactive drama, combining autonomous characters, artificial intelligence (AI)-based story management, and natural language processing to place the player in a dramatic world (Mateas and Stern, 2003). Through conversation, movement and emotive gestures, the player interacts with the characters Trip and Grace, and quickly finds herself entangled in the dynamics of their troubled marriage.

Figure 3.1: Screenshot of *Façade,* with the characters Trip and Grace (courtesy of Mateas and Stern)

As a friend invited over for drinks at a make-or-break moment in the collapsing marriage of the protagonists Grace and Trip, the player unwittingly becomes an antagonist of sorts, forced by Grace and Trip into playing psychological "head games" with them. The player, potentially playing with her own name and gender, may react to the experience with hilarity or anger, or play a number of roles from councilor to devil's advocate. The experience is different each time the player plays it, and unlike most games, the players do not have a clear goal; the player invents goals for herself as the interaction with the characters unfolds. Although there are occasional breakdowns (Mehta et al., 2007), the experience maintains a fluid interaction because the characters constantly respond to the player's unconstrained statements and movements with AI-generated speech and expressions.

Additionally, the story-level choices in *Façade* are intended to not feel like obvious branch points. The designers of *Façade* intentionally avoid stopping the action to present the user a decision menu with a limited number of options. Instead, the story progression changes subtly in response to many small actions performed by the player throughout the experience.

Game players move through a 3D game-like space with the arrow keys, interact with virtual items (to pick up glasses, statues, etc., or to hug/kiss/comfort the virtual characters) using the mouse, and speak to Trip and Grace by typing statements on a keyboard (see Figure 3.2). The interface between the player and the AI engine consists of a graphics engine, keyboard text input, and mouse interaction with objects and characters in the space.

Figure 3.2: Original *Façade*'s keyboard and mouse interaction

Briefly, the AI engine consists of three major components (see Figure 3.3):

- *the autonomous characters*, implemented in the custom reactive planning language ABL (A Behavior Language); ABL supports the dynamic mixing of multiple, simultaneous behaviors, joint intentionality for multiple, cooperating agents, and meta-behaviors that can modify the runtime state of other running behaviors (Mateas and Stern, 2004b),
- *the drama manager*, which dynamically sequences dramatic beats as a function of the player's interaction history; the selected beat modulates the autonomous characters' goals and behaviors (Mateas and Stern, 2000),
- *the natural language process system*, consisting of a semantic parser that parses surface text typed by the player into the underlying discourse acts recognized by the system, and the discourse manager that keeps track of the current conversational context(s) and decides on conversational responses to recognized discourse acts as a function of the active contexts (Mateas and Stern, 2004a).

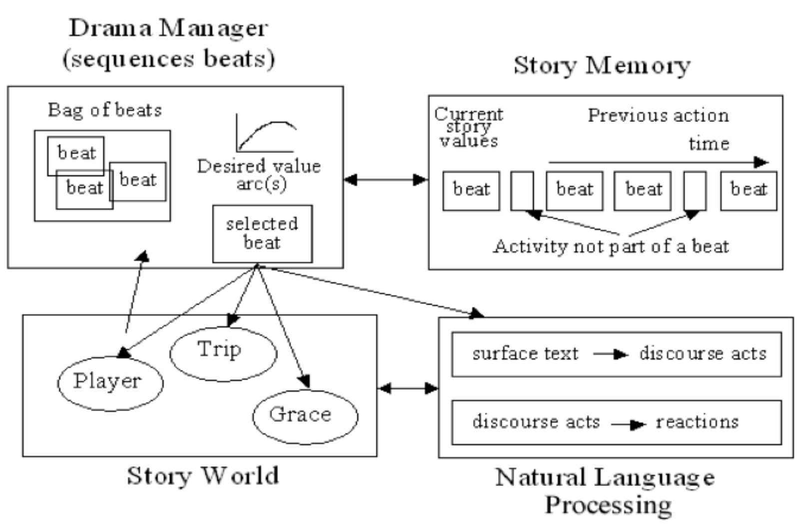


Figure 3.3: Conceptual architecture diagram for desktop *Façade* (from Mateas and Stern,2003)

3.1.2 Putting the AR into AR Façade

The goal for an embodied version of *Façade* is achieved through video see-through augmented reality and a physical stage modeled on the 3D world of *Façade*. In *AR Façade*, a modified graphics engine, physical interaction with objects, and speech handling (Figure 3.4) replace the interface between the AI engine and the player. Several things about *Façade*'s architecture made it very easy to adapt to augmented reality. The 3D environment navigated by arrow keys is nearly analogous to players freely walking around within a physical space. The space itself is a two-room apartment, not a fanciful or particularly massive game world.   The AI engine in *Façade* only responds to the reference to objects manipulated by the player (objects pointed at, picked up, or looked at), rather than detailed object motion, thus simplifying the handling of physical interactions in *AR Façade*. In order to convert the existing AI game engine, only the detailed position of the player must be tracked and all other inputs can be narrowed to a Boolean on/off observation.

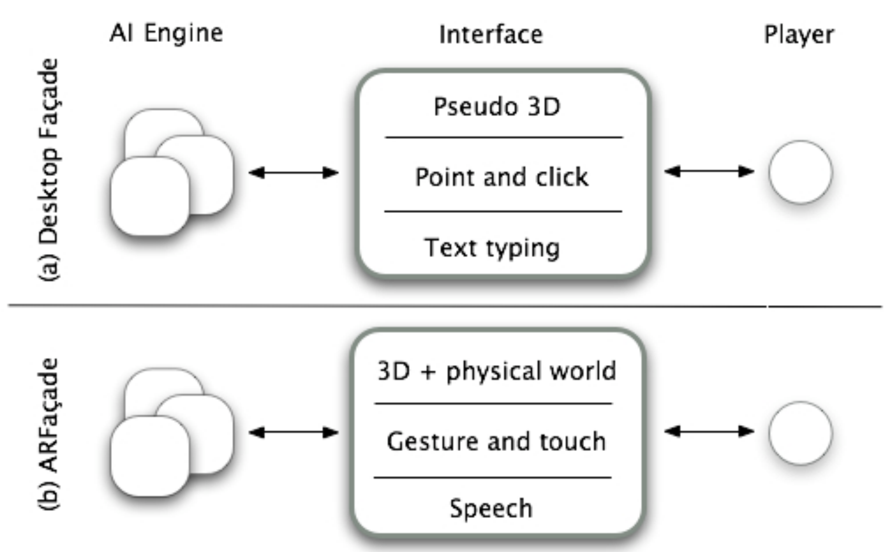


Figure 3.4: The interface for (a) desktop *Façade* (b) *AR Façade*

AR Façade runs on a Windows XP laptop mounted on a lightweight external hiking frame with a small pouch to hold wires and a battery for the camera. The player wears an eMagin Z800 3DVisor head-mounted display, mounted on a medical headband. The display has two bright, high-contrast OLED displays and a 40 degree diagonal field of view, and is integrated with an extended-head Point Grey DragonFly camera (pointing forward) and an Intersense IS-1200 Vistracker (pointing upwards) (see Figure 3.5).

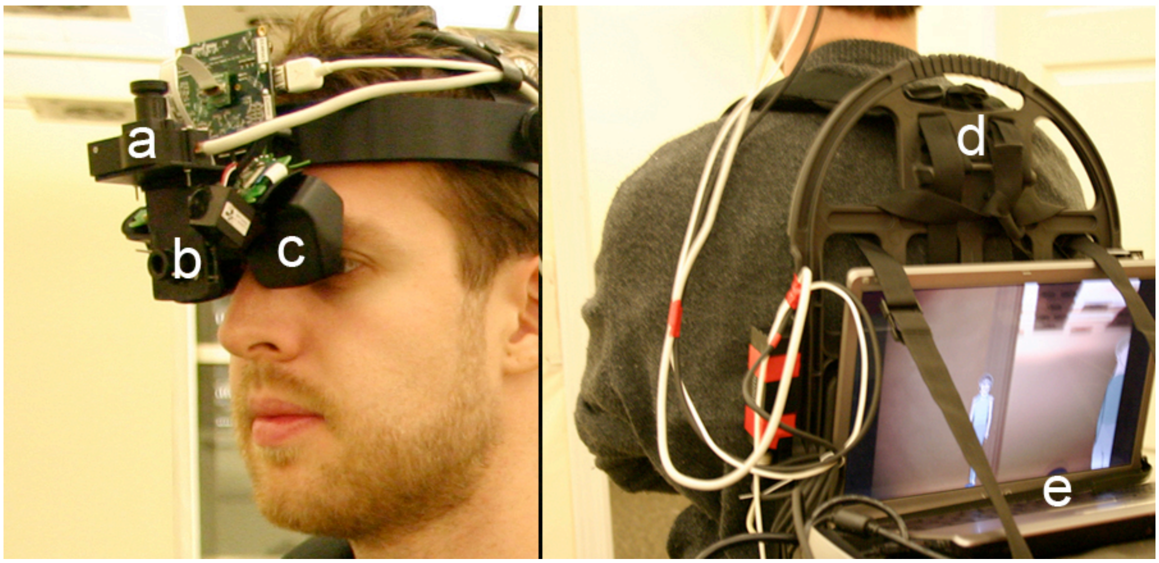


Figure 3.5: Hardware for the Atlanta *AR Façade* installation (a) Intersense IS-1200 VisTracker location tracker, (b) Point Grey DragonFly camera (c) eMagin head-mounted display (d) Intersense Inertiacube3 orientation sensor (e) Laptop computer running Windows XP

For the first Atlanta prototype of *AR Façade*, we constructed the physical space to match Trip and Grace's apartment as closely as possible, within the constraints of the lab space (e.g., see Figure 3.6). The walls are made of wood and off-white, slightly stiff, linen fabric. The paintings are either blown-up versions of original *Façade* images or new images made to look similar to the original *Façade* images, printed on large poster printer paper and framed. To finish the physical design, we gathered other furniture items and knick-knacks to match the stylized post-modern décor in *Façade*.

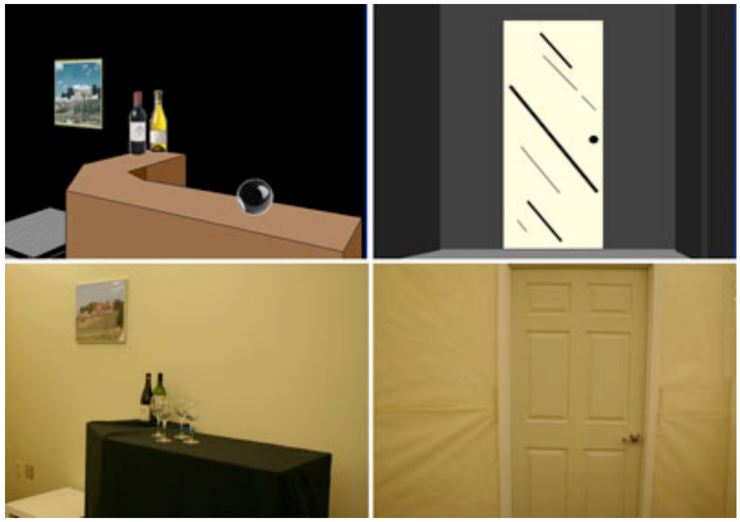


Figure 3.6: Creating a physical replica of the virtual apartment of Facade—the top shows screenshots of the bar and door from desktop *Façade*, the bottom shows the corresponding props in *AR Façade*.

Several changes had to be made to the existing code base. The player's screen consists of video from the physical environment in his field of view overlaid with the virtual characters drawn at their correct location. The video feed is rendered into a background texture and most of the existing virtual objects are not drawn, but rendered into the Z-buffer. Through careful alignment of real-world objects and their virtual counterparts, objects can occlude the characters. For example, Trip appears to go behind the physical bar when he is making drinks. Furthermore, the existing virtual map had to be changed to match the physical setup. All of the hardcoded locations for objects like the couch, tables, bar, etc. had to be modified in the graphics code. Likewise, the AI code contains hard coded values for character staging and path planning, and responses to player locations–these values had to be modified accordingly.

A Wizard of Oz (WOz) interface runs on a second computer outside the apartment, and lets a wizard handle speech input and references to objects in the space. The WOz interface has a series of reference buttons, used by the wizard to signify a player's verbal or gestural reference to physical things, and a text field to type the player's statements (see Figure 3.7). The WOz interface communicates wirelessly with the wearable machine via TCP/IP. We added cameras and microphones to the space to make the players' actions more visible to the wizard. For the Beall Center installation, we made a number of changes to the wizard infrastructure, including two conceptually different methods of wizarding the experience as discussed in Section 3.3.2.

Figure 3.7: Wizard-of-Oz interface in the Atlanta installation of *AR Façade*

Despite many of the advantages *Façade* provided in terms of its conversion into an AR experience, there were a number of important design and technology challenges we had to overcome. In the next section, I discuss these general challenges, our solutions to a number of them, and some remaining open questions. In the subsequent sections, I discuss additional practical challenges that came up when implementing *AR Façade* for an eleven week gallery installation at the Beall Center.

3.2 Design Challenges for Mixing the Physical and Virtual

In this section, I discuss the primary challenges faced when creating the initial prototype of *AR Façade* in our Atlanta lab. Many of the design questions encountered while creating *AR Façade* will be common issues in designing any mixed reality drama:

choosing how to render content, handling dialogue, interactions between physical and virtual objects, and facilitating movement in the space.

While discussing these issues, I also highlight important aesthetic and technical qualities of desktop *Façade* and how they impacted the embodied version. Solutions are proposed where possible and linked back to larger questions of embodied interaction. For more details on the interactive story design in *Façade* see (Mateas and Stern, 2000).

3.2.1 Character Rendering

Although desktop *Façade* does allow a player to navigate 3D space, a pseudo-perspective rendering system is used rather than standard 3D perspective projection. The *Façade* 2D animation engine generates the cartoon-style image layers that make up Trip and Grace's expressive characters on the fly. As a player moves and looks at the characters and as the autonomous characters move their bodies and change their facial expressions, the orthographically projected images are updated to provide the illusion of correct perspective. While dynamically generated 2D cartoon characters have expressive advantages over clunky 3D models or disjointed video content, it was unclear what affect that would have on an AR experience. In particular, when the characters are pulled out of an environment with a similar cartoon appearance and overlaid on a "real" scene, will they still be believable?

As I describe in our study results in later chapters, players found the characters compelling when integrated with a video backdrop, creating an effect similar to *Who Framed Roger Rabbit?*, although many players did comment on their cartoonish effect. While it might be interesting to explore the potential of realistic 3D models or video based characters in an interactive AR experience, much of the control and generative expressiveness would be lost. In *AR Façade*, the complex mental and emotional state maintained by the autonomous character AI engine can actually be visually realized through the procedural animation system. More importantly, it is unclear if more realistic content is actually desired; the theory of the Uncanny Valley would imply that more realistic artificial characters in augmented reality may end up being less appealing after a point (Mori, 2005). An alternative would be to use live video processing to produce a cartoon-style rendering of the real world, creating a consistent style for the mixed virtual/ real space. Mixed reality experiences provide an opportunity for studying the effects of content consistency and researchers are currently creating the mixing technology that will enable user studies in this area (Fischer and Bartz, 2005; Micheal Haller et al., 2005).

3.2.2 Conversation with Characters

In desktop *Façade*, the player starts typing a statement and letters appear on the screen. When the player is comfortable with the words on the screen, she hits the enter key to “say” the statement. A natural language processor (NLP) (Mateas and Stern, 2004) and an AI story engine (Mateas and Stern, 2000) process the utterance and cause the characters to react appropriately. Unfortunately, even in desktop *Façade* there is typically a half second delay between the player hitting enter and seeing any effect, occasionally detracting from the experience.

Moreover, in our early observations of desktop *Façade*, players adapt to the slight delay by strategically using the text buffer. Statements are often typed out and later retracted (the player backspaces over the text before hitting return), especially if Trip and Grace start talking about a new topic. Players quickly learn the limitations of the text buffer size (only 35 letters can be interpreted in one entry by the NLP) because they see letters fill the width of the screen and they hear a beep when the text buffer is full. For the AI engine’s NLP, this effectively constrains the complexity of the sentence structure that must be processed; for the player it provides a temporary buffer and a chance to reflect on the appropriateness of typed statements.

In the AR version (and in the speech-based version[40]) where players can speak continuously with no constraints, these no practical way of enforcing the buffer limitation imposed by the NLP. We imagined having the system beep when the player spoke the maximum number of characters, but we decided to simply show their words onscreen. Just like desktop *Façade*, if they fill the width of the screen, new words no longer appear. The problem is that typing is much more asynchronous compared to speech. Speech interaction in real-life is synchronous and so players expect an immediate reaction.

For *AR Façade*, speech recognition software would clearly introduce additional inherent latency and errors (due to the large corpus of the English language and the open-ended context). The Wizard of Oz method (WOz)—where a hidden operator handles speech to text translation—presents an alternative to actual speech recognition. Utilizing the WOz method in this circumstance turns out to be just as challenging, as discussed by Maulsby in their emulation of a speech-based intelligent agent (Maulsby, 2003).

The problems with both speech recognition and WOz raise some hard questions for speech interaction experiences. Beyond just rapid, high-accuracy input, how do we provide equivalent affordances for players in *AR Façade* as those available to players in *Façade*? Do players need a chance to reflect on and revise verbalized statements? What feedback should the player receive about the system's interpretation of their statements? How is the player made aware of system limitations, such as the maximum buffer size? Will an increase in latency diminish the conversational nature of the experience?

The initial prototype utilized the WOz method with a hidden operator typing in user statements as quickly as possible, but for our second implementation discussed in Section 3.3.2, we created an alternative wizard interface based on matching an

40 In the interface comparison study conducted at Georgia Tech, we included a speech-based desktop version to help us tease out the difference between talking vs. typing and walking vs. navigating with the arrow keys.

interpretation of the player utterance to the underlying constructs in the AI engine. I also discuss the wizards' experience of these two alternative WOz methods in Section 3.4.

3.2.3 Player Movement

Several challenges must be addressed with respect to player movement. In desktop *Façade*, the player controls their position and a single orientation representing both their head and body (which are assumed coupled). In *AR Façade*, we needed to de-couple head and body movement as is done in many first person games (such as in Quake where players can move independently of where they look/aim). Because Trip and Grace pay attention to their location relative to the player (e.g., they try to stay in front of the player when they want to talk), the character engine uses the player's body orientation rather than the player's head orientation (which is used for rendering). The AI engine also watches the frequency of player movement to decide if they are acting "nervous," and has Trip ask them to leave if it decides they are moving too much. The stable, filtered movements of the body, not the rapid movements of the player's head, are used in both cases.

Two trackers were used to decouple the body and head: an overhead hybrid inertial-vision tracker (IS-1200 Vistracker) to track the player's 6DOF head position and rotation, and an inertial orientation sensor (inertiaCube3) to get the relative rotation of the body. The head tracker is accurate to within a few millimeters–good enough to be used for the user's viewpoint in the graphics engine.

More challenging are the times in desktop *Façade* where the AI engine moves the player viewpoint. Most automatic player movement (adjustments when sitting on the sofa, when too close to the walls, etc.) can simply be disabled in *AR Façade*. However in one possible ending, Trip throws the player out of the apartment. Since Trip cannot manhandle a physical player, how do we recreate this ending in *AR Façade*? One idea is to have the entire scene fade into a virtual space around the player, just long enough for the player to see herself get thrown out. We ended up making the screen simply fade to black.

In desktop *Façade*, our team had observed a common trend of players to rapidly explore the virtual apartment before settling into interaction with Trip and Grace. However, since player movement provides interaction cues for Trip and Grace, this game-like exploration is contrary to the social setting of the experience. One of the questions that drove my research in *AR Façade* was if the physical nature of the space would encourage more "appropriate" behavior, or if players will continue this socially inappropriate initial exploration of the space. Furthermore, will the AR gear (head-worn display, backpack with computer and sensors) reduce movement and exploration? Will the video mediated view of the world make it difficult to interact with physical objects, perhaps due to the parallax offset between the display and the periphery caused by the camera not being collocated with the players eyes?

3.2.4 Physical/Virtual Interaction

In desktop *Façade*, the player can interact with the characters and objects, such as hugging, kissing and comforting Grace and Trip, or picking up drinks, trinkets, statues, phone, etc. The AI engine and characters adapt and react to such actions, sometimes apprehensively depending on the context, but always striving to create an engaging situation for the player. *Façade* loosely indexes conversation about objects so that any interaction (explicit touches, staring at an object, standing near an object, etc.) could cause the characters to converse or act on these objects. For example, if a player simply stares at the Italy photo on the wall, Trip will likely start talking about their recent holiday. Interacting with the characters and touching objects is an enjoyable part of the desktop game that we did not want to leave out in *AR Façade*.

For *AR Façade*, conversation takes place around physical replicas of *Façade* objects (see Figure 3.6). This interplay between physical and virtual items can make or break the experience. I discuss which items are easy to deal with and why, and conversely which items present a significant challenge and how we tried to overcome those challenges.

3.2.4.1 Easy Conversion to *AR Façade*

Some aspects of desktop *Façade* were particularly well-suited for AR. The physical objects in the space are only referenced generally. While this may seem overly limiting considering the potential for fine-grained interaction, it actually allows for much smoother and more contextualized conversation, and greatly simplified the infrastructure required for *AR Façade*. A wizard operator only needs to indicate if the player touches, looks at, or even comes close to a physical object of interest (e.g. art on the walls, statues, wine glasses, and the telephone). RFID detection could conceivably be substituted for the wizard interface for physical object manipulation, but the wizard would still be required for recognizing pointing and staring at objects. Some interactions are automatically handled by *AR Façade* based on player movement (e.g., standing near, staring at); all others are managed in a simple WOz interface with buttons to signal object references. We considered using additional WOz operators if emulating speech and monitoring the interactions became too daunting. I describe the wizard experience from both installations of *AR Façade* in Section 3.4.

To increase the sense of presence within a physical space, the audio for the apartment's answering machine and phone are played through physical props connected to additional computers hidden in the space. We chose to physically localize the sound of the phone ringing and the voice leaving a message, rather than using 3D spatialized audio through the headphones. Likewise, if a player were to pick up the phone the voice would speak through the phone. The choice of audio is handled by the AI engine, based on the player's interaction with the phone. In *AR Façade*, we simply forward the commands to play particular audio files across the network to the embedded machines.

3.2.4.2 Difficult Conversion to *AR Façade*

Despite the control afforded by the WOz method, some interactions are very difficult to emulate in *AR Façade*, particularly those involving objects that can be touched by both the players and the virtual characters. It is possible to imagine workable, although inelegant, solutions allowing the player to physically interact with Trip and Grace. Whether detected with sensors or communicated through a WOz operator, an interaction protocol or gesture language could be designed for players to hug, kiss and comfort the characters. Before players entered *AR Façade*, we typically demonstrated how to perform exaggerated gestures so that they were clearly distinguishable for the human wizard.

A number of items in *Façade*, such as the drinks, magic 8-ball, and front door present a challenge because both players and characters can manipulate them. In *AR Façade*, the virtual characters cannot pick up physical objects, but we opted to allow players to perform some manipulations on virtual objects. A player can grasp for the virtual glass placed on the bar by Trip, and if the wizard triggers the "pick up drink" button in the interface the player will see the glass appear attached to the HMD screen. The player can then physically pretend to take a sip; the virtual glass will tip towards the screen and show declining amount of liquid. A similar interaction was created for the 8-ball, where players could grasp for it and then shake an invisible ball to see a new message from on 8-ball screen (e.g. "Outlook not so good").

In our user studies we had questions about how players would approach virtual/physical interaction. Would players be able to figure out these simplistic interactions with virtual objects? What would it be like to "hug" a virtual character? Would the lack of haptic feedback take away from the experience sense of presence or agency?

3.3 An Eleven-Week Gallery Deployment of AR Façade

While many of the general design issues for virtual/physical experiences were addressed in our prototype of *AR Façade*, we faced a new set of challenges when given the opportunity to deploy *AR Façade* at the Beall Center for Art and Technology[41] in Irvine, CA for an 11-week Grand Text Auto[42] exhibition (see Figure 3.8).

Figure 3.8: *GrandTextAuto*'s Beall Center exhibit (Clockwise from the top left) A participant playing *AR Façade*, Mateas' *Tableau Machine*, Wardrip-Fruin's *Screen*, and Flanagan's *giantJoystick*.

---

41 http://beallcenter.uci.edu/ (accessed 9/22/08)

42 http://grandtextauto.org/ (accessed 9/22/08)

In this section, I discuss the challenges of deploying first-person immersive augmented reality experiences to a gallery setting and the design choices and enhancements made to help prepare for these challenges. Of particular interest are the physical layout of the exhibit and modifications to the system architecture to provide a more flexible wizard interface for the novice museum docents who would operate the show for the eleven week period. In the next section I present results from a study of the wizards docents and briefly discuss reactions from audience members who would not be able to participate as first-person players.

3.3.1 Challenges of "Real-World" Gallery Deployment

There were a number of challenges when preparing *AR Façade* for a long-term deployment at the Beall Center for Art and Technology at the University of California Irvine. The gallery installation would be held to higher standards than the initial lab prototype and it would need to entertain audiences of people, not just the individuals who would participate in the immersive experience. More importantly, the show would have to be easy to operate by non-technical, undergraduate, gallery docents who were not part of the research team, nor strongly motivated to ensure the experience operated "correctly".

3.3.1.1 Meet Art Gallery Standards

In the Atlanta lab setting, our *AR Façade* demo was somewhat fragile. It was not unusual to restart the computer or do code debugging on the spot. A work-in-progress laboratory demo would not be appropriate at the Beall gallery where we faced much higher expectations from both the visitors, and from the curators who wanted to maintain a certain aesthetic for the show. I coordinated heavily with the museum curator to insure we could provide a visually spectacular appearance, to deal with audio/visual issues, and to integrate with the gallery layout of nine other installations in the Grand Text Auto exhibit (see Figure 3.9). I sought to overcome many of our prototype's rough edges—such as loss of tracking data and AR registration issues— and to achieve a professional look and feel for all aspects of the experience design, from the HMD to the physical apartment stage.

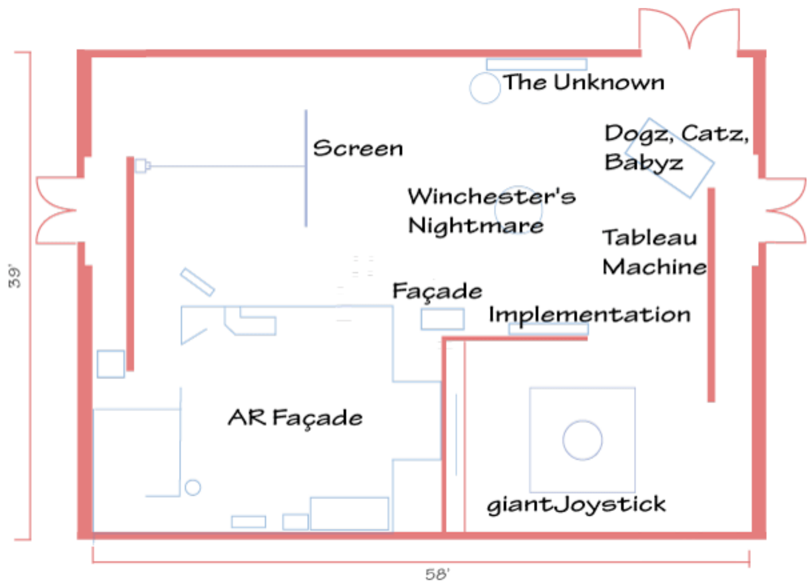


Figure 3.9: Beall Center layout during the *GrandTextAuto* exhibit

<u>3.3.1.2 Be Easy to Operate by a Team of Wizard Docents</u>

Not only would I not be on-site for most of the three-month installation, there would be no real technical support at the gallery. The museum hired a staff of ten modestly-paid undergraduate docents (mostly art students) who would put visitors into the immersive display, explain the basic interactions, and serve as the wizard operator. Although two docents would be present in the gallery, one would have to sit at the front door and greet visitors while another could operate *AR Façade*.

I trained each docent before the show opened, but since we could not debug problems ourselves, the software and the hardware had to be robust and simple to startup and operate. The instructions that would be delivered to players by the docents would have to be short and intuitive. The system would have to accommodate for the wizards' possible lack of motivation and for unforeseen events in the museum. Also, as I describe in Section 3.4, I wanted to support a degree of flexibility so that wizards could develop their own methods for pulling off the entertainment experience for visitors.

3.3.1.3 Entertain a Wide Range of Individuals and Groups

During *AR Façade*'s deployment we anticipated hundreds of visitors, from small groups to individuals who stop in randomly to large pre-planned groups (such as high school field trips). We would have to accommodate a diverse range of physical sizes (both height and width), ages, educational backgrounds, and gender; although, since the Beall Center sits on the UC Irvine campus, we expected the majority of visitors would be students in their early 20s who stop into the gallery when walking between classes.

Larger groups of visitors present a particular challenge for individual, first-person, immersive installations because of low throughput. In *AR Façade*, a single individual can occupy the experience for up to 30-40 minutes between gearing up, hearing instructions, and engaging a story of non-determined length. For Disney and other entertainment producers this is a matter of financial viability, but here we were mostly concerned with giving as many people as possible a satisfying experience.

Visibility would be essential, not only for those people who might not get a chance to experience *AR Façade* as a player, but for those visitors who simply wanted a more hands-off experience. Our prior experience with *AR Façade* in Atlanta taught us that some visitors are very intimidated by the HMD technology and the overall setup. Other visitors would prefer to watch someone else, rather than participate themselves. We sought to support a strong audience presentation so that these visitors could still have a good understanding of what the player experiences.

I attempted to design for universal access, although certain populations—visitors in wheel chairs, blind persons, small children— would unfortunately not be able to participate as first-person players. I also asked children below the age of 13 not to participate even if they were physically big enough because of the objectionable language and adult themes that play throughout the narrative.

3.3.2 Design Decisions for Beall Center Installation of AR Façade

In this section, I describe the design choices made to prepare *AR Façade* for the three-month gallery deployment at the Beall Center, including the physical setup and changes to the system infrastructure to support more flexibility for wizards. While many of the enhancements to the original *AR Façade* are not particularly technical (such as using large posters instead of projected light for the window display), I believe our process of upgrading the prototype experience into a real deployment can be useful for producers and designers of immersive experiences with an eye towards visitors, operators, and audience members.

3.3.2.1 Physical Setup

AR Façade would be part of a free-to-the-public art gallery for 11 weeks, so we devoted a fair amount of time to the construction of the head-mounted display (HMD) (see Figure 3.10, left) and the physical layout of the space. Our research team constructed the head-mounted display from third-party components and custom-created mounts to be robust enough to withstand mishandling. We tracked the HMD using the IS-1200 Vistracker and large visual markers mounted on the 16' tall ceiling (see Figure 3.10, right). Florescent lights lined the perimeter of the apartment at 10' high and provided an even distribution of light across the ceiling and enough illumination for the stage. I did not encounter any noticeable tracking problems—such as jitter and temporary loss of tracking data as we did with our initial prototype—despite a few markers falling from the ceiling during the 11-week period.

Figure 3.10: *AR Façade* at the Beall Center hardware and stage (Left) Head-mounted display; (Right) the *AR Façade* stage during construction

The team focused on the physical layout of the space, because we wanted the experience to be visually appealing and enjoyable to groups as well as individuals. We used black scrim for the long wall of the apartment behind the bar (see Figure 3.11). Since it was light inside the apartment and dark outside, audiences could stand outside the wall and easily see the player's activity, while the player would not be able to see outside though the wall (especially when viewing through the HMD). Moreover, the black walls and the black-painted bar helped maintain the occlusion effect when Trip walks behind the bar; players know the bar is there, but the top of the bar is difficult to distinguish from the wall.

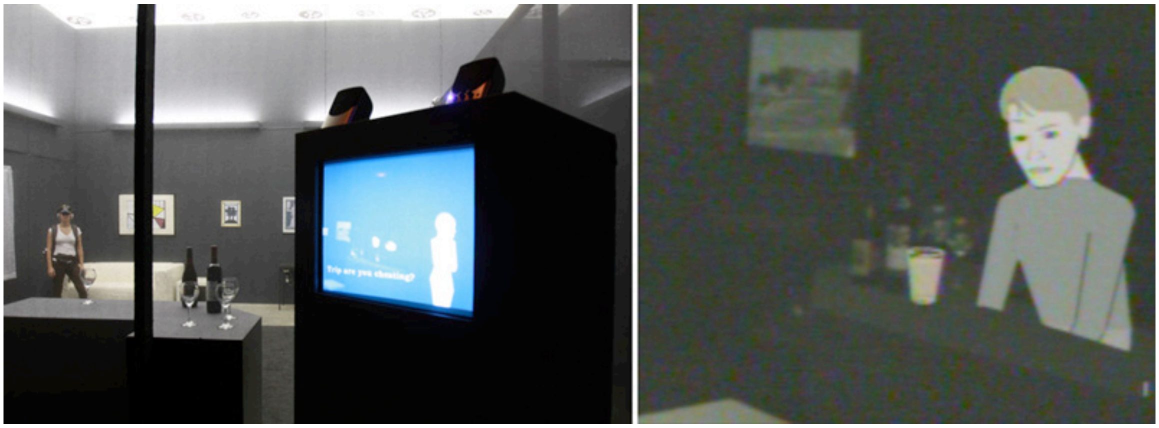


Figure 3.11: Physical setup at the Beall Center exhibit of *AR Façade* (Left) Audience view of the player through the black scrim wall. (Right) Player's HMD view shows the Trip character standing behind the physical bar. Note the black scrim blocks the view of objects outside the wall.

We used a wireless video transmitter to transmit the player's HMD view to a TV monitor sitting to the right of the stage. The wizard—sitting just outside the door of the apartment—could also see the same HMD view, along with a second monitor with video from a ceiling mounted camera. The audio turned out to be one of the biggest challenges. We were competing with several other loud exhibits in the space, so we wanted the audience (and wizard) to be able to hear the dialogue between the player and the characters. To make this work, we used two audio transmitters (one to amplify the player's voice and one to transmit the backpack's computer audio) and a mixing board to mix the two sources and disseminate it to speakers for the audience and headphones for the wizard. The audience could also optionally plug in a pair of headphones into the speakers if they could not hear over the noise from other exhibits. In Section 3.4.1, I reflect on the impact of these physical setup decisions on player and audiences.

3.3.2.2 Wizard Infrastructure

I made a number of changes aimed at supporting the undergraduate docents who would be in charge of operating and wizarding the experience. Most of the enhancements were practical (providing a start-up script, building in network checks, etc.), but I also modified the underlying infrastructure of *Façade* to experiment with some of the issues raised in Section 3.2.2. Dissecting the system architecture of the original *Façade*, I thought it would be interesting to tear out the natural language parser (NLP) and build a wizard interface that could be used to directly trigger higher-level constructs called discourse acts (Mateas and Stern, 2004) (see Figure 3.12). This approach was plausible because *Façade*'s AI engine does not explicitly map user inputs to specific statements and actions by the two virtual characters. Instead, it models the characters' emotional states and attempts to choose lines of dialogue based on local and global contexts. There are about 30 possible discourses (e.g. flirt, agree with, etc.) with optional parameters that can be expressed as local context by the player.

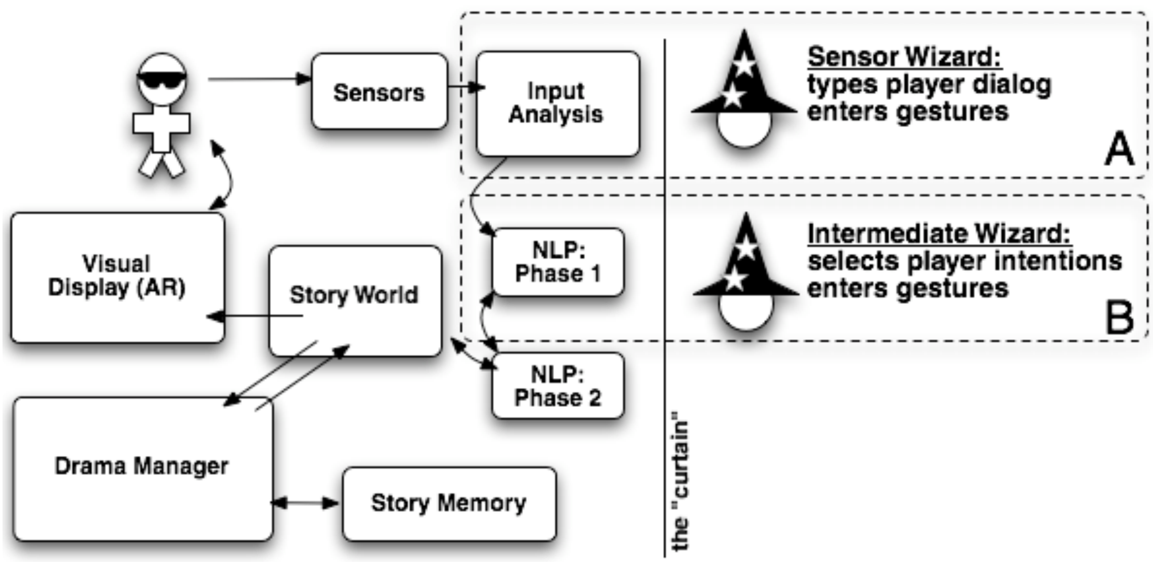


Figure 3.12: ***AR Façade*** **conceptual architecture with two wizard variations**
(A) Sensor Wizard, where the wizard enters text to match the player speech and presses buttons for corresponding gestures; (B) Intermediate Wizard, where the wizard selects from a list of player intentions (discourse acts)

The new wizard interface built in Java had two tabbed panes—"Dialogue" and "Discourses"—both available to the wizard at any time. The Dialogue pane included a text entry field for typing in what players said and large buttons for indicating specific player actions (see Figure 3.13, left). This mirrored our original wizard interface, including the visual and audio feedback to wizards indicating the maximum number of chars that could be entered at one time.

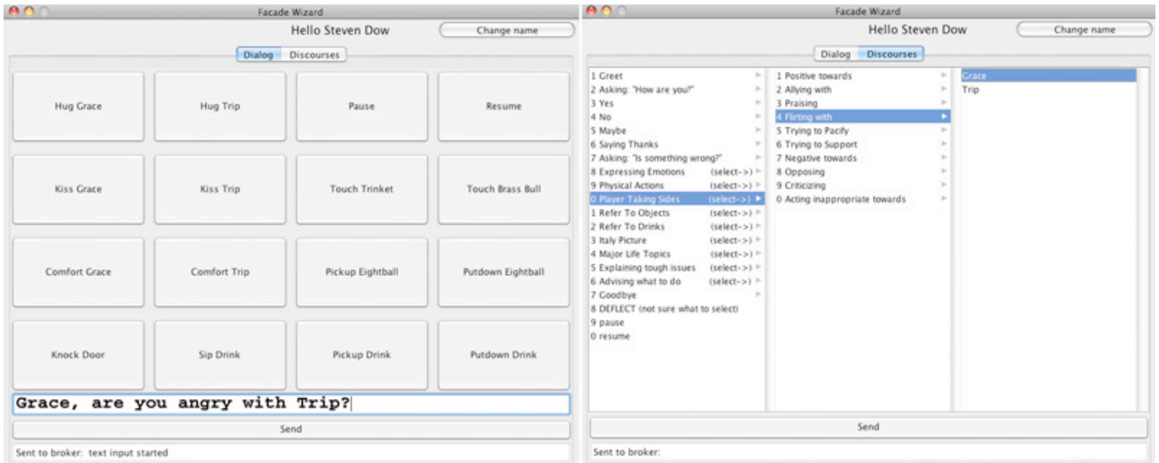


Figure 3.13: Two different wizard interfaces used at the Beall exhibit (Left) Dialogue wizard interface provides a text entry field and buttons for indicating player specific gestures; (Right) Discourse wizard interface provides a hierarchy of discourses or player intentions, such as "Player Taking Sides > Flirting with > Grace".

The Discourses interface provided a hierarchical listing of the higher-level discourses (Figure 3.13, right); rather than typing out what the player says, the wizard selects something that matches what they think the player means (e.g. "Player takes sides by flirting with Grace"). The Discourses interface went though a number of iterations to improve its usability, including incorporating feedback from the docent wizards after the first two weeks of the show. The current interface represents the discourse hierarchy across three columns that can be navigated through any combination of number keys, arrow keys, and the mouse. The hierarchy itself went though changes as we learned better ways to phrase and position the potential discourses. For example, we moved "yes" and "no" to the top-level, rather than as a sub-category of "player taking sides" since they were accessed more frequently. I also provided a "Deflect" category at the top-level, a discourse the wizard could choose it they did not know what other category would be appropriate for a particular player utterance.

From the player's perspective the two interfaces differ with regard to the onscreen feedback. In the dialogue method, the player sees their words appear after the obligatory time for the wizard to type in each word. In the discourse method, I added a subtle revolving line in the lower right corner to provide feedback that the system heard an utterance (it actually starts revolving as the wizard moves through the Discourse hierarchy).

The difference between the Discourse and Dialogue interface raises a number of questions. How would wizards appropriate the use of both methods? Would wizards use the Discourse method or would they rely on the straightforward method of typing the Dialogue? Would the Discourse method be as fast, slower or faster than the Dialogue method? Does it potentially improve on the important time delay problem discussed in Section 3.2.2? How long would it take to learn the Discourses? Finally, how does it effect player experience, if at all? Does the change in onscreen feedback impact the experience? In Section 3.4, I present the results my investigation of wizards including how they adopted the two interfaces and how it effected the experience for players and audiences.

3.3.2.3 The "Pause" Feature

A third design change included in the Beall installation of *AR Façade* was a "pause" feature, where players could pause and un-pause the action at any time. Players could pause simply by saying "pause" and the wizard would either type pause in the dialogue interface or select the pause category in the discourse interface. When pause was activated, Trip and Grace would freeze in place. The player would still view the live video background and the characters graphically aligned in the space, but they were silent and still mannequins. In pause mode, nothing the player says or does impacts the narrative structure. If the player said "resume", the characters jump back to life as if nothing had happened.

I decided to add this feature for a number of reasons. Practically, it would provide players a clean method of exiting the experience; wizards could also use the pause feature if they wanted a player to finish. Hypothetically, I envisioned players appropriating the pause/resume feature in interesting ways to exert more control over their interaction in the drama. I discuss the player adoption of the pause feature in Section 7.5.4.

3.4 Experiences of Non-Player Participants

The majority of my dissertation—Chapters 4, 5, 6, and 7—is devoted to examining the player experience. In this section, I report of the experience of non-player participants, specifically the external audiences and the wizard docent who operate the experience for players. The experience for non-player participants should not be framed in terms of embodied narrative engagement, but should be considered equally important, especially for the ecology of public media exhibits. I will draw on evidence from both the Atlanta and Beall Center installations of *AR Façade* to report on both the audience experience and the wizard strategies. In Chapter 4, when I consider various influences on the player experience I re-examine both the audience and wizard as external factors.

3.4.1 Audience Reactions

In some sense the audience experience is as important as the player's. Often people who "see" first-person immersive exhibits do not get a chance to play due to time and throughput constraints. As I described in the previous section, many of the changes our team made to *AR Façade* for the Beall Center were related to the physical layout (black scrim wall, TV monitor, audio enhancements, etc.) to make the experience more visible for audiences. In the Atlanta version, one of the walls was only built to chestheight so audiences could look in at the player, but inevitably audience members would enter into the space with the player and stand behind her to peer at the laptop computer which displayed the player's view (see Figure 3.14, left). Even after we setup a TV monitor in Atlanta, audiences tended to wander into the apartment, likely due to our more informal approach. At the Beall Center, we explicitly wanted audiences to stay out of the apartment so that players could have the fullest experience (see Figure 3.14, right). Our design changes succeeded in this regard, as I did not witness one non-player enter the apartment during at least a half-dozen large groups of visitors (e.g. high school field trips, etc.).

Figure 3.14: The audience experience in Atlanta versus the Beall Center (Left) Audiences gathering behind the participant at the Atlanta installation; (right) audiences enjoying the experience from outside the apartment at the Beall Center exhibit.

Quite a few audience members became vocal and would shout things through the scrim wall to the player to try to provoke them into saying something outlandish to the characters. If they were not shouting at the player, audiences often generate boisterous laughter in response to exchanges between the characters and player.

Another anecdote worth mentioning is that audiences were much more intent on watching the screen—both the laptop screen in Atlanta and the TV monitor at the Beall—where they could see the characters Trip and Grace. If the player did something notable, the audience could always look to their left to see the player, but the action was clearly happening on screen. Despite the fact that audiences would occasionally miss the player's physical actions altogether, I considered our physical setup a success. If anything it was important to be able to preview the environment for audiences and players. Once the action started the audiences wanted to see the facial reactions of the characters, but it may have been useful to include a second monitor showing a third-person camera view of the player even if it did not show the characters.

3.4.2 Investigations of Wizards

Interactive entertainment experiences—especially in amusement parks or gallery settings—often employ full-time operators to usher visitors in and out of the experiences (Pausch et al., 1996; Koleva et al., 2001). More recently, interactive media have looked to trained actors performing behind the scenes in real-time as animated characters or robots[43], very similar to the hidden wizard in the movie *The Wizard of Oz*. Other experiences use actors more overtly, such as *Can You See Me Now?* where online players try to avoid physical actors who are moving around a city (Benford et al., 2006). In *AR Façade*, our "wizard" docents hold responsibilities somewhere between an usher and an actor.

In this section, I describe my investigations of wizards across both installations. For the Atlanta installation of *AR Façade* where we used the Dialogue interface exclusively, I discovered the challenges our single dedicated wizard faced, her common errors, and the strategies she adopted to be as efficient as possible. For example, to deal with the AI engine's buffer limit, our wizard would paraphrase player statements and/or split statements as two entries. Our wizard experiences with the first installation led to ideas for the Discourse method of wizarding.

---

43 Quasi the robot: http://www.etc.cmu.edu/projects/ibi/platform_hardware.htm; Disney Turtle Talk: http://en.wikipedia.org/wiki/Turtle_Talk_with_Crush;

For the Beall Center installation, I conducted a formal investigation of the wizard docents to find out whether and how docents would fulfill their role given an opportunity to choose between the two conceptually different wizard interfaces, Discourse and Dialogue. I present insights supported by the usage patterns and qualitative sentiments expressed by the docent wizards during a series of open-ended interviews. I include data about how the nine docents learned to use and adopt the two interfaces over the course of the show. I found that the Discourse interface required more cognitive processing, but actually encouraged deeper involvement in the story and to the emotional tone of the players and audience members. Later, in Section 4.5.6, I provide data about the effect of wizarding on the player experience.

3.4.2.1 Wizard Strategies during the Atlanta Installation

During the Atlanta installation of *AR Façade*, our wizard used the Dialogue interface method exclusively. I did not create alternative interfaces because our research team was conducting an comparative study (as I describe in Chapter 4) between different versions of *Façade*, and we wanted to keep the three conditions as similar as possible for the purposes of comparison. We compared three different versions of *Façade*: original desktop 3D (KB), desktop 3D using speech instead of typed text (SB), and the fully immersive augmented reality (AR) version. For consistency, we used the same wizard—a graduate student at Georgia Tech, although not a formal member of our research team—for both the SB and AR version for all players in the study. In this section, I describe the challenges she faced and strategies she adopted to accomplish the task.

As I discussed in Section 3.2.2, the task of the wizard was to attempt to emulate speech and gesture recognition so that it could be compared to the keyboard-based version of *Façade*. Our wizard reported feeling pressured to perform as quickly as possible so that she could match the desktop interface as closely as possible. This was made more difficult in the AR version where the wizard had to multi-task between listening and typing words and watching for gestures and hitting the corresponding action buttons.

Although the tasks were not particularly cognitively demanding, they demanded full attention by our wizard and as a result, she made a number of errors. Despite the fact that the wizard could erase her text as she typed it, she still a made number spelling errors. Erasing text is not unusual at all in KB desktop *Façade* as is seen in the table below; players use the text buffer as a temporary store, and if the moment passes they could erase the entire buffer. In contrast, the wizard simply had to type as fast as possible.

The table also shows a difference between AR, where the wizard had to emulate speech and gesture, and SB, where the wizard only had to emulate speech. The latter wizard task was less demanding so she ended up erasing characters only half as much. The less time spent typing and erasing unneeded letters, the better the wizard does at minimizing the time delay between the player speaking and the characters reacting.

Table 3.1: Total number of text characters erased across three versions (N=12)

Type of interface used in Atlanta study	Augmented reality (AR)	Speech-based (SB) interaction	Keyboard-based (KB) interaction
Total text characters erased (N=12)	161	72	1179

Potentially even more detrimental, the wizard would type in something that could be misinterpreted by the NLP. For example, she once typed “hell” instead of “hello.” Our wizard later reported that she was nervous that the system would incorrectly interpret her mistyping, so she quickly tried to type it in again. In her haste, she typed in “hell” again, but luckily the AI engine did not interpret the statement harshly enough to tell Trip to kick out the player.

A second challenge faced by the wizard is that the user often spoke longer than the NLP-imposed buffer limit, because there are no effective constraints on the player's speech. To deal with this challenge the wizard developed a couple of noteworthy coping strategies. First, she would often find a way to paraphrase the player's statement on the fly without distorting its meaning. For example, one player said "do you have issues with your parents, Trip?" and anticipating that this would be too long for the buffer limit, the wizard typed "you don't like your parents?" If she could not think of a quick way to paraphrase what the player said, the wizard would split the statement into two entries. For example, she might type "but you've been together ten" and hit the end of the buffer. So she would enter that and then add the final word "years" as a separate entry.

As I discuss in Section 4.5.6, the wizard performance does have some influence on the player experience, but it's unclear how or if these specific strategies directly impact the player. Clearly if the wizard misspells a word or paraphrases, the player will see that in the text feedback shown onscreen. Likewise, when the wizard splits up statements, the NLP interprets them independently and so the character responses may not be as accurate.

This analysis of the first wizard experiences with the dialogue interface led to ideas for how to improve the system. One idea is to simply add more wizards and divide the task—perhaps allowing one to watch for gestures and the other to type user statements—so that it is less demanding and results in fewer errors. The buffer limit problem led to the idea for the Discourse wizard interface, where the wizard bypasses the imperfect NLP and selects the higher-order AI engine constructs directly, as I detailed in Section 3.3.2.2 above. Next, I compare these two methods by looking at data from the Beall Center installation where both interfaces were available to the wizard docents.

3.4.2.2 Wizard Interface Usage Patterns during the Beall Center Installation

My investigation of docent wizards during the 11-week installation at the Beall Center was more formal than my observations of wizards in Atlanta. I recorded wizard activity and conducted several open-ended interviews with all nine docents, who were non-technical female art students between the age of 19 and 22. After a short preliminary interview, I trained each wizard to use both versions of the wizard interface and instructed them to create the player experience however they saw fit. During the first 9 weeks of the show—the period where I was not onsite—I allowed the docents to use either the Discourse or the Dialog wizard interface. There were 106 full episodes—some repeat players—distributed across the nine wizards. During this time period, the wizards choose to use the Discourse interface only 15.7% of the time on average, although W8 used the Discourse interface nearly half the time (47.8%) (see Figure 3.15).

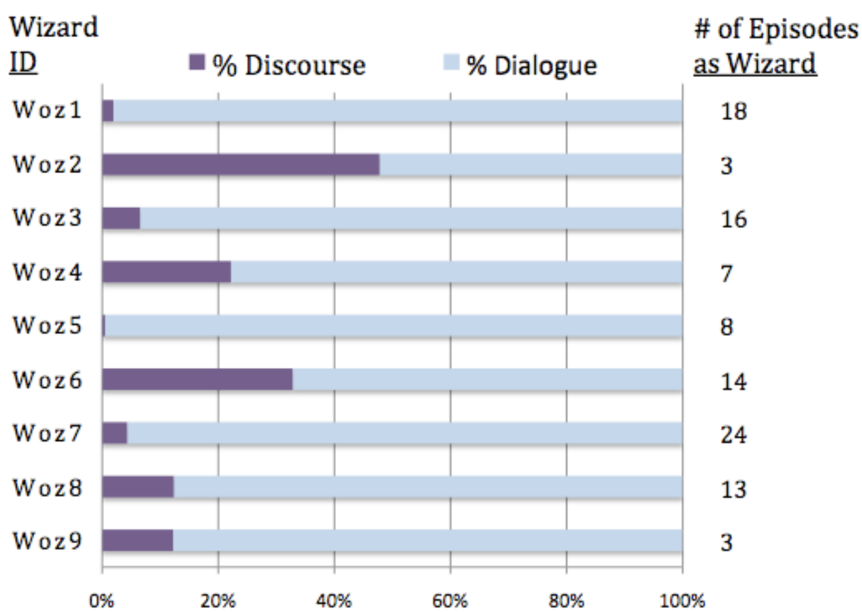


Figure 3.15: Wizard activity during the eleven-week installation, across nine wizard docents (Woz1-Woz9).

During the final 2 weeks, after I noticed that most wizards relied almost exclusively on the Dialogue interface, I forced the docents to learn and use the Discourse interface. I conducted more open-ended interviews during the final 2-week period. The qualitative evidence presented in the next two sections explains differences between the two wizard interfaces and the interactions with players and audiences enabled by the Discourse method.

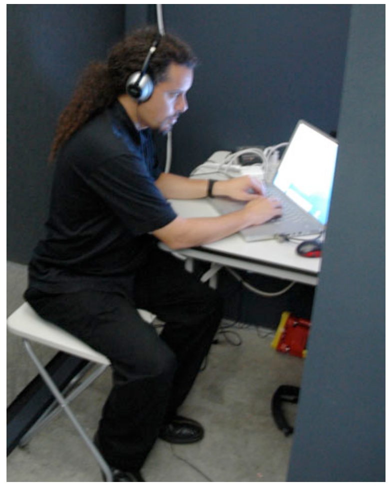


Figure 3.16: The wizard station occluded from the audience view

3.4.2.3 Wizard Insights on the Two Interfaces during the Beall Installation

From the wizard interviews, it is clear why the wizards used the Dialogue more than the Discourse interface. There was a consensus among wizards that the Dialogue interface was simply easier; with the Discourse interface the process of "picking stuff out requires more thinking" (W3). According to W4, "it forces your mind to kind of think in a different way, of not just directly translating specifically what they're saying but kind of attributing it to a larger category of emotion or actions. It depends a little more on your interpretation." And W6 explains, "I feel like I have to interpret more what the players are saying. So it's like more involved, like – you have to pay more attention." W8 pointed to one positive reason to use the Discourse interface: if there was an audio problem, she didn't have to hear the exact phrasing to pick up on the player's intentions. Although the Beall wizards never mentioned it, the players also received more abstract feedback under the Discourse method (spinning wheel vs. spoken words), so they did not have to worry as much about missing words or misspelling words.

For the wizards, the Discourse selection is perceived to require more cognitive load, but it is not immediately clear if it reduces the time delay or helps the conversation flow better. According to W1, “typing it out takes a little bit longer than searching and clicking,” while W4 says “it takes a little longer to sort through them and find the right one then it does to just immediately translate what they’re saying into text.” As I discuss in Section 3.4.3, this question could be answered with a more detailed conversation analysis, but it is not essential to my primary thesis questions. Anecdotally, I suspect that Discourse selection is faster for more verbose statements, but slower for shorter statements (I discuss this further in Chapter 8).

On the other hand, some wizards reported that the Discourse interface gave them more control over the characters and the course of events. W9 claimed “you could definitely shape the player’s experience”. W1 even went further, almost describing the characters as puppets: “I’ll make Trip talk about the picture again and hopefully he will kind of guide [the player] over.” W6 said it was her role “to translate what the people are saying so that Trip and Grace can understand it”, anthropomorphizing the virtual characters.

3.4.2.4 Wizard Interactions with Players and Audiences during the Beall Installation

Wizard interviews not only revealed the affordances of each WOz interface, they revealed patterns of appropriation, especially after I forced the wizards to use the discourse interface. While the Dialogue interface required constant typing and emulation of player speech—especially since the words appear on the player’s HMD—the Discourse interface afforded a degree of experimentation. W8 said if there is a “lull in the conversation... I will select something like Therapy, just to offer a little variation...

because some people would be a little passive in their interactions." For the most part, the wizards were trying to help, but sometimes they got bored and clicked discourses just to spice things up: "I clicked 'have sex' or something because I was hoping that some big explosive thing would happen... I thought that would be fun to see, because Grace seems like kind of an intense chick" (W9).

There were some consequences for this experimentation, as W4 found when she tried using the 'oppose Trip' discourse: "I guess that was too strong of an emotion, because [Trip] kicked him out. haha!" (W4) Wizards reported that the system did not always respond as expected, that they wanted more nuance, and that sometimes they felt their actions did not have any affect.

In their role as wizards, the docents learned to become very perceptive of the state of the player. "I can tell when someone feels awkward or when people are getting really annoyed by just like the tone of their voice" (W3). Some wizards were so attuned to how players felt, the emotions were palpable, "I could hear [the player's] breathing pattern and ... I could almost feel her getting uncomfortable and it's kind of bizarre because you can hear the audio and then you can see what they're seeing and I almost started to get uncomfortable!" (W9).

The wizards had different opinions on what would be most fun for players and audiences. W7 said she would tell players ahead time "the more you mess with [Trip and Grace] the more entertaining it will be." W1 said she also encourages players to play with it "because a lot of the times people tend to like stay in the middle of the room or they don't get near the objects. ...that way there's a little bit more for me to do in the back too." W3 would flatly reveal her role, even invite players and audiences to see behind the curtain. "I would tell people that I'm gonna be back there.... so that they can get the reaction that they want." (W3)

Some wizards felt it was important to preserve the illusion, as W8 states "I would say the illusion is necessary because it frees up this whole possibility of what could happen if they like slapped Trip or something." According to W6, if players saw the wizards they "seemed kind of disappointed. It's like magic tricks. When you find out it's not fun."

3.5 Chapter Discussion

In this chapter, I described the object of study for my dissertation, an immersive and interactive story called *AR Façade*. I outlined the technical and practical issues of converting the conversation-based desktop game *Façade* into the fully-embodied version. While the AR version uses the same character graphics and AI-engine as the desktop version, it employs a mixed physical/virtual reality stage, a video-see-though AR display, and Wizard-of-Oz methods for enabling speech and gestural interaction. I discussed two installations of *AR Façade*—Atlanta and the Beall Center—and the changes made to improve the overall experience for players and group audiences.

Many of the improvements made for the show at the Beall were put in place to support nine undergraduate art students to run the show—and perform as wizards—without direct technical supervision from the researchers. By building a solid infrastructure, supporting different flexible modes of interaction for wizard operators, and exposing the first-person HMD experience to audience members, our research team managed to extend the *AR Façade* experience beyond merely the first-person player participant. The physical setup supported an enjoyable experience for audiences because the player and the physical apartment were clearly visible through the scrim wall, and audiences could get a good feel for the first-person experience through the TV monitor. The team had considered bringing the audience into the experience more explicitly, perhaps allowing the audience to select from discourses in the wizard interface or asking the audience to enter funny statements that the player would have to say, kind of a "karaoke" version of the experience where the player would be stripped of her sense of agency. I decided to leave these ideas to future work and to focus on creating a player experience faithful to the original *Façade*.

The question of whether the Discourse wizard interface outperforms the Dialogue interface with respect to time delays and accuracy remains to be answered. Anecdotally, it appears that the Dialogue interface is better for short utterances and Discourse method works better for longer or more complex statements. As expected, of the two alternative interfaces, the Dialogue typing interface did not require as much mental attention, so the wizards initially preferred it. Once I forced them to use the Discourse interface and they starting experimenting and understanding how it worked, the wizards appeared to become more attuned to the course of the story and the emotional level of the players and audience members. While the Dialogue interface provided a good design from one standpoint (easier to learn and simpler to use), the Discourse interface engaged the wizards as performers of sorts, enabling them to guide the direction of the experience as they saw fit.

Although it was more difficult to master, the Discourse method gave wizards more control over the course of events. My experimentation with the role of wizards suggests potential for other types of wizard approaches. For example, could wizards be in charge of adjusting "tone" variables in an otherwise complete interactive system? Could wizards directly respond to the emotional state of the participants? Rather than discourse constructs like "flirt", perhaps the wizard adjusts system variables representing the player's emotional state and goals, and the system can be programmed to respond directly or subtly to the participant's current state.

This chapter was devoted to the technology, the group/audience experience, and the wizard experience; the next four chapters focus on the player experience. In Chapter 8, I discuss open research questions for immersive and interactive stories, and outline possible future research related to mixing physical and virtual content and using Wizard-of-Oz methods for facilitating a design process.

CHAPTER 4

METHODOLOGY AND OVERALL IMPRESSIONS

Evaluation is the worst form of HCI research except all those other forms that have been tried. —Shumin Zhai, Evaluation Democracy

Our research team created *AR Façade* for two reasons. To create a breakthrough in interactive entertainment and to apply an ethnographic-style approach to empirically investigate the theoretical concepts discussed in Chapter 2. I believe HCI theory and methods have much to offer to experience design and game design, but there have been no standard methods developed to date. The goals of empirical research include observing how phenomena occur in the real-world and analyzing how it impacts future design. Not only are games, interactive dramas, and immersive experiences *ludic*, as opposed to task-oriented, players engage in play in very different ways—and so the evaluation methods should address this diversity. As I set out to study *Façade* and *AR Façade*, I developed a "mixed method" combination of quantitative, qualitative, online, lab, and real-world investigations. My research methodology provides an initial strategy for evaluating user experience in immersive and interactive stories, and speaks to the wider discussion of applying empirical methods to endeavors in media studies and humanities.

In this chapter, I briefly review other methodologies for studying users in entertainment experiences. I describe results from early lab studies of desktop *Façade,* as well as the online survey I administered while we created *AR Façade*. Then I describe the series of studies conducted across two installations of *AR Façade* and my methodology for collecting and analyzing the qualitative and quantitative data. Finally, I discuss some high-level influences on the player experience and the overall impressions of *AR Façade*.

4.1 Background on User Study Methodologies for Experience Design

Methods for evaluating entertainment experiences include many of the traditional methods in HCI, such as self-report questionnaires, interviews, focus groups, and video analysis. At Microsoft's game studios—one of the largest groups studying user experience in games—a process has been designed to gather primarily quantitative evidence of how players play along with some interview data to understand patterns that arise in the data. Their analysis of 3000+ hours of Halo 3 gameplay by 600 gamers led to significant improvements to the game, from simple bug fixes, to subtle new design features to prevent players from running out of ammo, to adjustments in enemy power to maintain the right level of challenge (Thompson, 2007). The usability testing for Halo 3 identified problems by recording as much in-game data as possible so that oddities could be visualized and understood. Unfortunately, play testing of this sort inevitably comes at late stages of development, obviating the possibility for major design shifts.

Recently, Pinelle et al. have suggested a list of ten heuristics to help identify usability problems in games (2008), based loosely on Nielson's heuristics for traditional interface design (1994). Pinelle et al. acknowledge that their methods are appropriate for usability, not issues of user engagement or content (2008). While these classic sorts of game testing are necessary and valuable for addressing surface-level issues, they do not address the user's experience in terms of how they feel and why they feel that way.

In contrast, Lazzaro and her team have been using video analysis techniques and interview methods on some of the more popular game titles towards understanding the range of player emotions and reasons that people play games (2004). Others in the HCI community have invented new methodologies for understanding emotion in user experiences. For example, Isbister et al. developed a novel evaluation tool called the Sensual Evaluation Instrument where users of an interactive system select from small, abstract, hand-held object that best represent their current emotional state (2006). They argue their method is intuitive, transcends language and cultural barriers, works across different types of interactive experiences, and can be more fun for the users (Isbister et al., 2006).

Others have explored the use of physiological sensors towards understanding users' emotional responses to interactive experiences. For example Mandryk et al. presented a method of modeling user emotional state based on a user's physiology measured by galvanic skin response, electrocardiography (electrical activity in the heart), electromyography (muscle activity), and heart rate (2006). They transform these physiological signals into an "arousal/valence" space used by Lang (1995) and Russell et al. (1989) to classify emotions on a 2D grid. They argue the benefit of modeling emotions based on physiology is that emotion can be represented quantitatively, objectively, and continuously over a session (Russell et al., 1989). These quantitative methods are similar to those used in the presence community to measure the sense of "being there" (Meehan et al., 2002), but rather focus on modeling emotion, which can be valuable for non-immersive media experiences.

In a broader context, it is not clear how HCI evaluation can be applied to experience design and art. Zhai argues for evaluation because of the practical consequences of most HCI design contexts[44], stating "a user interface is not a purely personal artistic expression that the audience can either take or leave." Interactive art

---

44 http://www.almaden.ibm.com/u/zhai/papers/EvaluationDemocracy.htm (accessed 9/22/08)

experiences are typically improved by informed critique, not rigorous HCI evaluation. As Paulos boldly stated recently “HCI cannot be used to evaluate art” (Paulos, 2007). While I am sympathetic to this notion, I believe some evaluation techniques can be useful for understanding and refining, even in the most avant-garde experiences.

Moreover, games and interactive experiences—especially those that push on narrative and social relations—are situated experiences that can be observed and analyzed using ethnographic-style methods to explain users’ interpretations and behaviors. These observational methods are intended to be open-ended and grounded in each context. Generalizable evaluation frameworks, on the other hand, can potentially stymie creativity and perpetuate standard game conventions. To summarize, I believe immersive and interactive stories can benefit from traditional usability testing to identify and fix interaction problems, but they should also be studied with qualitative observation (perhaps supported by quantitative data) to understand issues of user engagement.

4.2 User Studies of Desktop-based Façade

Our mission to understand user experience in immersive and interactive stories began with investigations of the original *Façade*. In the months before *Façade* was released to the Internet, the game’s developers recruited a number of game testers to help smooth over some of the interaction issues. The testing strategy involved systematically and creatively trying to break the game, although it was untypical in that the “rough patches” had more to do with awkward communication rather than navigation or the manipulation of complex controls. This was not an attempt to evaluate the user experience as much as it was usability testing the game itself. In this section, I provide a synopsis of several summative evaluations of desktop *Façade*, including investigations prior to my involvement and the online survey I deployed during the development of *AR Façade*.

4.2.1 Previous Investigations of Façade

In 2005, Knickmeyer and Mateas performed a preliminary evaluation of *Façade* where user retrospective protocols to reveal interaction strategies and failures over time (Knickmeyer and Mateas, 2005). Through a specially designed coding method they identified game patterns where players would leverage interaction failures into new opportunistic goals, thus maintaining positive interest. In Chapter 5, I provide additional evidence to support this finding and go further to describe the styles of play that emerge. Moreover, the Knickmeyer study demonstrated the viability of retrospective interviews which I utilized in the interface comparison study (see Section 4.3.1).

Façade quickly became a phenomenon garnishing over 300,000 downloads within the first year[45], winning the grand prize award at the 2006 Slamdance Indie Game Festival, and receiving outstanding media attention ("the future of video games" according to the NY Times, etc.). This growing user base provided an opportunity to understand how *Façade* users experienced the original game in naturalistic settings. For example, there were ample first-hand reports of player impressions of *Façade* from online blogs. One blog community started trading "screen plays" generated from desktop *Façade* game play (e.g., one player pretended to be a zombie, saying nothing but "brains brains... brains...", and posted the resulting script).

Another interesting reflection on *Façade*'s game play came from an undergraduate student's blog which detailed two distinct ways of playing. He reports showing the game to his fraternity brothers and typing things to be disruptive and to elicit character reactions, such as flirting with Grace and using profanity. The blogger said he would also play *Façade* by himself and adopt a more reserved approach because he wanted to see how far he could get in the drama to discover other endings. My analysis of

45 As of Mar 2008, there have been over 525,000 downloads of *Façade* (courtesy of Andrew Stern)

player behavior and interpretations of *AR Façade* described in Chapter 5 validates and demonstrates this range of behavior across a broader set of players, and also shows how having an audience can impact play strategies.

4.2.2 Online Survey of Façade Players

One year after *Façade* was released to the general public, I posted an online survey and advertised it on game blogs and fan sites for *Façade*. The survey created on SurveyMonkey[46] included basic demographic questions (age, occupation, gender, ethnicity, education level, typing ability, game playing and TV watching habits) and a combination of fill-in-the-blank, multiple choice, and open-ended questions about their *Façade* gameplay (see Appendix A).

Among other questions, I asked how many times the respondent played *Façade* and if they would want to play again. I asked questions about perceived agency, common strategies and goals, perceived reasons for communication breakdowns, overall enjoyment, and the next type of game they would like to see produced. Respondents could skip some questions and they could drop out at any point. Out of 129 people who started the survey, 105 kept going past the first couple questions. Although a self-selected set of 105 respondents are not statistically representative of 300,000 players, the survey provided an initial glimpse at the *Façade* player base and helped inform our player recruitment and study designs moving forward.

Most of the results of the online survey are summarized and visualized in Appendix B. I will point out a few of the more relevant statements written in by respondents. Some of the comments speak to one of the issues I address in my thesis, that “naturalness” in the interface decreases the sense of agency because it sets expectations too high:

---

46 http://www.surveymonkey.com/ (accessed 9/22/08)

> *"Providing a natural language parser as a method of input, Façade gives the impression of a large vocabulary."*

> *"I feel that natural language as an interface for communication raises user expectations too high compared to what it can actually deliver. I think you should make more subtle use of other communication channels and try to create a less realistic but more persuasive experience. With all due respect to Aristotle..."*

I will come back to this point in Chapter 6 when I discuss the effect of loosely constrained immersive interfaces on a user's sense of agency.

4.3 Data Gathering for AR Façade

My reason for investigating *AR Façade* was to confirm user patterns discovered in earlier studies of desktop-based *Façade*, to capture player behavior through their actions and statements (clearly more interesting than mouse-clicks and typing), and most significantly, to understand the effect of the immersive interface. In this section, I describe three Phases of data collection for *AR Façade*: an interface comparison study conducted in the Atlanta lab setting, an eleven-week in-situ deployment of *AR Façade* in a gallery space (nine weeks of episode log collection), and a followup player study carried out at the final two weeks of the gallery deployment (see Table 4.1). In total, I gathered in-depth interview and gameplay data from 45 players (N=12 plus N=33) and additional episode data from 106 episodes of *AR Façade* (as well as 126 episodes of desktop *Façade*).

Table 4.1: Summary of research studies of *AR Façade*

Investigation Name	Location	Dates	Type of Data Collection	Number of Participants
Interface Comparison Study (KB vs. SB vs. AR)	GVU Lab, Atlanta, GA	August 2006	Episode video and audio, episode logs, retrospective interviews on video, questionnaires, demographics	N=12
Nine Weeks of Episode Log Collection	Beall Center, Irvine, CA	Oct to mid-Dec 2007	Episode logs (AR and KB), surveys (not linked to episodes)	N=126 (KB), N=106 (AR), 40 surveys
Two Weeks of Player Interviews	Beall Center Irvine, CA	First two weeks of Dec 2007	Episode video and audio, episode logs, interviews on video, questionnaires, demographics	N=33

4.3.1 Interface Comparison Study (GVU Lab, Atlanta, GA)

With the lab prototype implementation of *AR Façade*, my fellow researchers and I compared three different versions of *Façade*: original desktop 3D (KB), desktop 3D using speech instead of typed text (SB), and the fully immersive augmented reality (AR) version. Even with this relatively small number of participants (N=12), we were able to draw useful contrasts between the interface versions (Dow et al., 2007).

We designed our study to gather qualitative data about player experience, exposing participants to the three different variations of *Façade* (in a counterbalanced order) to facilitate a subjective contrast. We used the Wizard of Oz methodology to achieve speech interaction in the SB and AR version and for gesture recognition in the AR version. We explicitly did not tell the participants how the speech was being recognized, nor could they see the Wizard station.

4.3.1.1 Recruitment

My collaborators and I recruited twelve participants through Craigslist.org and other local game forums in the Atlanta area. We screened participants so we could enroll a range of genders (balanced 50/50), races, education levels, and ages (from 18 to 33 with an average age of 25.8, SD 4.0). We also selected players with a large range of professions and prior experiences with computers, games and movies. In the end, none of the demographics appeared to factor into player opinion. Even if we suspected such an effect, it would require a much larger sample size to reach any significant conclusions.

4.3.1.2 Study Setup

Our study took place in a large, dedicated room with infrastructure for the three interface variations of *Façade*. One desktop machine ran the KB and SB version, while the AR version ran on a laptop computer in a backpack carried by the participants. The study lasted about three hours and participants were paid $10 per hour.

After signing a consent form, players listened to a brief explanation of *Façade*. Before each round of play they were instructed on the specific interface, including a short demonstration of possible gestures in AR (e.g. holding arms out in a hug motion) (see Appendix C). Each participant played *Façade* three times, once for each variation (making six possible orders, balanced out to account for learning effects). In addition to allowing participants to contrast the interfaces, asking the participant to play three times enabled us to observe how the player's strategy would change or adapt over time.

4.3.1.3 Data Collection

In open-ended interviews between each episode, I usually started with a general question like "so tell me about that experience," and then asked for additional details as the interview proceeded. Throughout each episode I logged notable events, such as unusual player reactions, apparent conversation breakdowns, and visual anomalies. I dedicated part of each interview to reviewing these moments on a video monitor so that participants could reflect on their experience. After playing all three variations I administered a short questionnaire (Appendix D), which asked participants to compare the interfaces and helped to guide a final interview.

I also collected quantitative data (player and character dialogue, body/head position and rotation, and AI processing logs). During the game episodes and interviews, we recorded video of what appeared on screen (or on the HMD in AR) along with video from multiple third-person perspectives to capture player emotions and physical actions. We analyzed the gameplay video together with the video from the "two week player investigation" and these results are presented in Chapter 5.

4.3.2 Nine Weeks of Episode Log Collection (Beall Center, Irvine, CA)

During the eleven-week deployment of *AR Façade* at the Beall Center, we collected the positioning and conversation logs for the first nine weeks from a large number of game episodes, for both *Façade* (N=126) and *AR Façade* (N=106). Visitors came into the free public gallery and could try any of the exhibits. Visitors could stay as long as they wanted and they could return at any point. It is safe to assume that many of these episodes are repeat plays and that many individuals played both the AR and KB version. At the end of the Beall installation, after filtering out the episodes that did not have any verbal interaction or player movement, we were left with 232 logs between the two versions.

My analysis of the Beall episode log data uncovered several significant differences between the AR and KB worth mentioning here. The AR episodes were shorter on average, 8.8 minutes in AR versus 9.8 minutes in KB. This difference is not statistically significant (t=1.472; p=.142), but both average play times are shorter than the 14.4 minutes average episode time for the 45 in-depth players episodes. Players tended to stand closer in proximity to both Trip and Grace in the AR version than in KB. Players also communicated more, both in terms of text characters used and length of statements in the AR version. I will reference these results again later when I discuss issues of affordances in Chapter 6 and distancing in Chapter 7, but the full analysis of the gallery logs can be found in Appendix H.

AR Façade players were given the option to fill out a one-page survey (see Appendix G). I did not attempt to connect the survey responses to specific episode logs. I gathered 40 survey responses from the Beall Center with the following demographics: 75% of respondents were students, 17 men and 23 women, with an average age of 23.6 years. The overall Likert-scale rating of the experience was 5.42 out of 7; slightly higher for men (5.65) compared to women (5.26). Most of the 40 respondents related the experience to a video game (16), followed by TV/Movies (10) and then Improv (8). The rest of the Beall survey results are summarized and visualized in Appendix H.

4.3.3 Two Week Player Investigation (Beall Center, Irvine, CA)

For the second in-depth study of *AR Façade* conducted during the final two weeks of the Beall Center exhibit, no interface comparisons were drawn, because we assumed most visitors to the gallery would only have time to play either *Façade* or *AR Façade*, and it was unclear to us if repeating the interface comparison from the Atlanta study would yield additional insight. I made three notable theory-directed design changes that impacted the user experience, as I described in Section 3.3.2: I altered the role of the wizard, changed the on-screen feedback for the player, and added a "pause" feature that would allow players to freeze the action at any point.

4.3.3.1 Recruitment

Since the study took place in the Beall gallery during its normal open hours, I did not formally recruit and schedule participants. I did put an announcement on Craigslist.org and hung fliers near the the Beall Center. During slow times I would go to the coffee shop next door to the gallery and ask patrons if they would like to participate.

The majority of the 33 participants for this stage of my investigation entered the gallery with no prior knowledge of my study. I did not screen participants, but managed to get a fair range of people (16 men and 17 women, an average age of 23.6, and 85% students). There were an additional 17 episodes in the final two weeks where the player declined to participate in the interview, but I did collect those game logs and included those in the 106 episodes discussed above.

4.3.3.2 Study Setup

Each session lasted approximately 45 minutes and participants received $5 per half hour rounded up to the nearest half hour. Participants were asked to play *AR Façade* one time and then take part in an open-ended interview. After signing a consent form, players listened to the wizard docents give a brief explanation of the *Façade* story and instructions on what to do in the AR interface. The docents would show the player a few of the canned gestures (hug, kiss, comfort), but mostly just told them they could say and act out whatever they wanted. This was not a controlled laboratory study, so I attempted to interfere with gallery setting as little as possible. I did not control anything on the player side, but did ask the wizards to use the Discourse version of the interface (as discussed in Section 3.4) to find out if the experience could be controlled with that method.

4.3.3.3 Data Collection

After the player tried *AR Façade*, I conducted an open-ended interview, usually starting with a general question like "tell me about that experience," and then probing for additional details as the interview proceeded (see Appendix I). I was not equipped to show players video of their episode to do retrospective interviews—as I had done in previous studies—but I did ask players to recall moments that really stuck out for them.

After the interview, players filled out a short questionnaire which captured demographics and other impressions (see Appendix J and K).

In my thesis proposal I put forth a list of potential measures of engagement and I tried to collect as many of those as possible. I did not attempt to collect physiological measures, partly because of the variability of the sensors, but largely because it would not have been practical to ask gallery participants to put sensors on their body. I did ask players to judge how many minutes they played as an attempt to measure the feeling of passage of time, a key indicator of flow. In retrospect, I would have collected that information differently since most people are not good at estimating time. I could have asked players to say whether the episode felt longer or shorter than ten minutes long. Likewise, the questionnaire at the end had a short quiz to test player's ability to recall story plots from the narrative. This measure was also problematic since players were not guaranteed to hear all of those plot lines due to the variable nature of *Façade*'s narrative structure. The open-ended interviews revealed when players remembered the stories and specific lines of dialog, but it would be difficult to draw quantitative comparisons between players on this matter. I will address the issue of measuring engagement again in Section 5.4.

4.4 Analysis Methods for AR Façade

To analyze this vast quantity of interview and play data from all the studies, I used a combination of qualitative and quantitative analyses. Between the two in-depth studies of *AR Façade* at the Atlanta lab and the Beall Center, I collected approximately 2000 minutes of player interviews and 800 minutes of player episodes from the 45 participants. This section describes the four primary analyses that I conducted (see Figure 4.1): Qualitative interview analysis, player profile analysis, episode analysis, and log analysis.

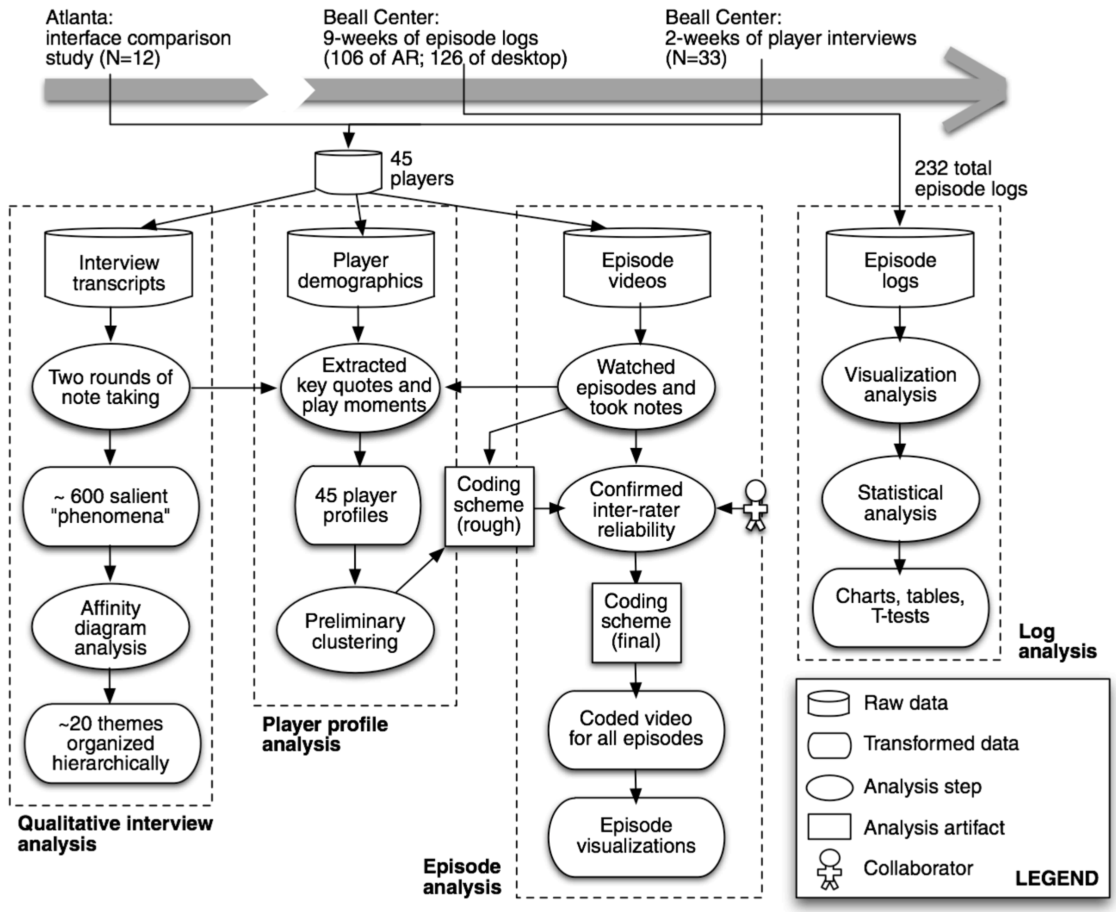


Figure 4.1: Flowchart of data gathering and analysis process for *AR Façade*

4.4.1 Qualitative Interview Analysis

I transcribed the interviews for the interface comparison study and we outsourced the transcription for the two-week player investigation. I then followed a grounded approach for understanding the situation (Strauss and Corbin, 1990). My process was to go through the 2000 minutes of transcripts twice, first to take open-ended notes (see Figure 4.2) and second to highlight more salient "phenomena" or ideas expressed by players.

Abridged transcripts	Researcher notes
I want something to talk about besides the objects he sees at the house	Player talks about limits of the conversation
Trip shows painting but I see my reflection instead	Technical issue
The arguments are realistic; I have to complement that… the girl with her career and Trip trying to be overprotective. G feels that T is preventing her from advancing her career	Player describes the fight and the plot
Made me feel like leaving… I'm not their counselor	Towards perceived role
Said to G that she sounded stressed and I guess the computer took it as "depressed"	Interpretation of system
Great… Grace was already mad at me. And that elevated the situation vs. mellowing it	Perceived feelings of Grace towards player

Figure 4.2: An excerpt from the interview transcripts with researcher notes

In total, I had about six-hundred phenomena codes, such as "Player talks about role in the story" and "Player feels limited by typing buffer," listed into an Excel files with multiple player quotes for each phenomena. I kept a pointer to the media file and the specific time each statement was spoken so I could listen to the exact wording of quote in later analysis. The phenomena codes were printed out and cut onto individual pieces of paper, along with numbers corresponding to the row of each code in the Excel file. On a large table surface, I conducted an affinity diagram analysis where related phenomena codes were physically grouped together. For example, all codes that had something to do with "game strategies" were initially grouped together and physically labelled with a post-it note (see Figure 4.3, left).

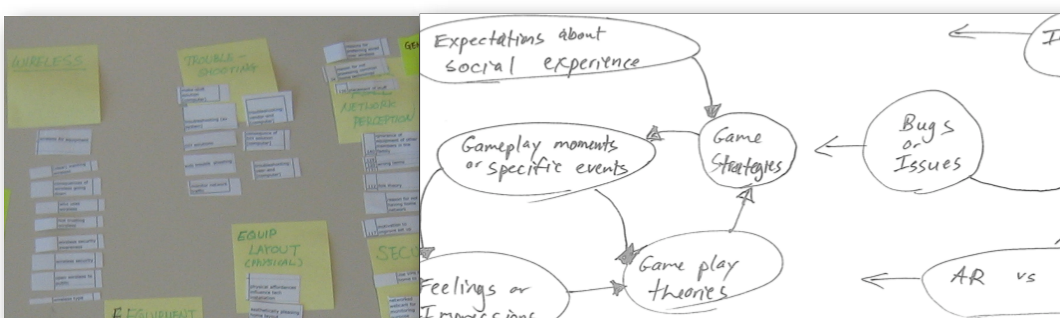


Figure 4.3: Sketches of the affinity diagram analysis method (left) Small pieces of paper with researcher notes or "phenomena codes" (right) high-level groupings of themes and how they conceptually relate to each other.

With the ~600 phenomena codes and ~20 initial themes laid out on the table, I went through an iterative process of describing each theme and conceptually linking them to each other (see Figure 4.3, right). This was a sketchy high-level thinking process that happened on paper and white boards, and often involved rearranging the physical layout of the themes and phenomena codes. I formed a single hierarchy of the themes and then went back into Excel and digitally organized the same themes with a paragraph to describe each theme along with the top three to five quotes for each theme. This hierarchical organization of themes served as an outline for writing.

4.4.2 Player Profile Analysis

For each player (N=45), I created a profile page with key demographics, quotes, and time-stamped moments from the episode where the player did something interesting. I wrote a paragraph or so describing each players' strategies and their interpretation of the experience. This analysis also included a high-level indication of engagement. In my dissertation, I reference a metric of "Overall Rating," a combination of five Likert-scale questions on the questionnaire from both the Atlanta and Beall studies. These five questions are on a scale between 1 and 7, and the sum ranges between 5 and 35:

- Content rating (Hated it vs. loved it)
- Curiosity about the outcome (Not interested vs. very interested)
- Character believability rating (Not believable vs. believable)
- Physical presence (Not consistent with real space vs. felt like being in a real space)
- Influence of interaction (No impact vs. big impact)

Throughout this dissertation, I might look at only one of the five metrics (e.g. just the Influence rating when talking about agency), but I will mostly refer to the overall rating as a subjective indicator of overall engagement. For my analysis of each player, I realized that many players could be conceptually linked in terms of their styles of play. I formed rough clusters of player profiles, again by physically organizing paper versions of the profiles. The clusters of players led me to conduct a more rigorous coding analysis of the episodes, and formed a preliminary basis for the "styles of play" concepts presented in Chapter 5.

4.4.3 Episode Analysis

Through an iterative process, I developed a method to code gameplay behavior. This task was particularly challenging because there are few prior examples of video coding interactive entertainment experiences. Furthermore, many coding schemes are very specific to the domain. Lazzaro conducted an extensive video analysis for players in traditional video games, coding emotional reactions, such as surprise, fear, and fiero (personal triumph), by interpreting facial expressions (2004). This type of video coding was not possible because we did not have a clear view of the player's face, since players were free to move anywhere and the HMD obstructed their face. Behavioral social scientists, such as those working in Applied Behavior Analysis[47], use video coding methods in specialized domains such as evidence-based care of autistic children (Hayes et al., 2008). My video coding method is similar to behavioral analysis, but designed specifically for *AR Façade*. The method could potentially be adopted for other Immersive and Interactive Stories, but it would be less relevant for experiences that do not involve speaking and physical gesture (such as most traditional video games).

The video coding scheme evolved from the qualitative note taking and analysis of the player episodes. After seeing the episodes live, I watched the recorded video for each episode at least twice. My initial coding scheme was based on my knowledge of the

47 http://rsaffran.tripod.com/whatisaba.html (accessed 9/22/08)

observable features in the video collection and through my initial clustering of player profiles described in the previous section. Since I recorded audio and video from at least two cameras (the player's HMD view and either one or two external camera views), I could observe occurrences of both player and character speech, gestures, and technical obstructions such as loss of 3D tracking, poor virtual/physical registration issues, or AI logic errors. I faced the challenge of creating a coding scheme that would not be so simplistic that it does not reveal anything useful, nor could it be as complex as conversation analysis. The goal was to create a coding scheme solid enough to allow another researcher to repeat the coding on the *AR Façade* data set within acceptable inter-rater reliability rates, and for it to be potentially generalizable for other immersive and interactive story experiences.

The coding scheme classifies player speech and gestures, determines if characters fail to respond to direct requests, and marks technical breakdowns (see Appendix L for the coding scheme reference sheet). I chose this interval method for coding; so I marked whether if occurrences happened within 15-second intervals in the video. This method was useful, not only because it would be less complex than marking the exact time and length of occurrence, but because it was necessary for performing inter-rater reliability comparisons. With an average episode length of 14.4 minutes, this resulted in about 60 discrete measurements of events per episode. This level of granularity was appropriate for Façade's conversational structure—short enough to provide a detailed overview of the episode and long enough to capture approximately one player statement and character response. If a short occurrence happened right at the end of one interval and the beginning of another, it would be coded as part of the latter interval.

The coding scheme is based as much as possible on observable events (speech, gestures, technical errors). To make it more descriptive, I introduce a categorical differentiation of player actions that is plainly open to interpretation. I identify each player speech or gesture into one of three categories: normative, divergent, and meta.

- *Normative speech and gesture*: The player's action is 'normal' for the social situation represented in the narrative. The player is following the characters' leads, usually only speaking and acting in reaction to social cues. (For example, the player might say "It has been a long time, how are you guys?")
- *Divergent speech and gesture*: The player's action is 'abnormal' for the social situation represented in the narrative. The player is intentionally diverging from the story role, provoking the characters or acting overly dramatic. (For example, the player might say "Trip! I love you. I love you. I love you…")
- *Meta speech and gesture*: The player's action is not part of the social situation represented in the narrative. The player is experimenting with the interface interaction or referencing action outside the experience. (For example, the player might say "Wow, this is like Grey's Anatomy.")

This aspect of the coding scheme is open to interpretation, particularly since what is considered normal or abnormal in a social situation is subjective. There are player actions that do not clearly fit into one category or another, and some actions that could be simultaneously classified into two types at once. I felt it was important to distinguish between these types of actions, because it is relevant to the player's level of emotional engagement and to their suspension of disbelief.

I sought external validity for the coding scheme by conducting inter-rater reliability with a research partner. We both coded 2% of the video data under my initial scheme and then I ran Kappa's statistic for inter-coder reliability (Cohen, 1960). We came to a more succinct agreement of the codes and then I created a coding scheme reference sheet (see Appendix L). My research colleague and I each coded another 5% of the data. I again ran the Kappa statistic for inter-coder reliability. While subjective data coding is always open to interpretation, my colleague and I achieved concurrence above 80 percent using Cohen's Kappa statistic on five percent of the total video data (details in Appendix M). As the primary coder, I coded the remaining 95% of the video.

The raw data from the episode video coding is a collection of integers representing each code for each 15-second interval. I used Excel to tally the codes and calculate inter-rater reliability statistics. There are some free, yet buggy, software tools for video coding, but most of the good software tools for coding video were out of our price range. I also used Excel to create visualizations for each episode, as I describe in Chapter 5, along with the key results from the visual analysis. Appendix N provides the video coding visualizations for all episodes.

4.4.4 Log Analysis

In addition to creating the episode visualizations, I wrote Java programs to read the log files and extract or calculate key metrics for each episode: length of play, # of statements by Trip and Grace, # of player statements, average length of statements, # of gestures, type of ending, choose name, total movement, and interpersonal distances between the player and Trip/Grace (IPD-T and IPD-G). The quantitative log metrics along with answers to the questionnaire data were dumped into several Excel spreadsheets: (1) online survey of desktop *Façade*, (2) optional Beall survey data, (3) nine weeks of Beall episode log data, and (4) in-depth player data (N=45) for both the Atlanta interface comparison study and the 2-week player investigation at the Beall. I imported the data into Tableau[48] and used standard information visualization techniques to discover anomalies and trends. I take advantage of both Tableau and Excel for graphing data throughout this dissertation. The primary utility of information

48 http://www.tableausoftware.com/ (accessed 9/22/08)

visualization tools was to identify trends in the player demographics and episode data that could be later checked with statistical analyses. The 232 total episode logs from the first nine-weeks of the Beall installation provided the best opportunity for drawing out statistical differences between the two versions of Façade (keyboard-based and augmented reality), although I did calculate statistical correlations on the 45 players where I have demographic information as well as gameplay data.

4.5 Influences and Overall Impressions of AR Façade

In this section, I will present internal and external influences beyond the immersiveness of the display that impact how players engage *AR Façade* along with their overall impressions of the experience. Generally speaking, when players entered into either our Atlanta lab space or the Beall gallery to take part in *AR Façade*, they did not know what to expect. Many players had reservations since they were visible to people outside the space, yet the experience also tapped into player excitement about taking part in something new and novel. After players entered the room and adjusted to the interface, the experience elicited an interesting set of behaviors and reactions. Some players struggled to understand the technology, often relating the experience to other media or to encounters in real-life. Overwhelmingly, players did not know how the system worked or that we used human wizards to enable the speech and gesture interaction. Other players fixated on our style choices, such as the cartoon rendering of the characters. Beyond all of those influences, the player experience was obviously influenced by the story content, and by personal characteristics that might impact how they relate to the story themes. My thesis claims that the immersive interface played a role in how players engaged in *AR Façade*, as I describe in Chapter 6 and 7. In this section I describe the other factors germane to engagement. Based on my qualitative analysis, I believe the following influences must be considered while analyzing player behavior:

- Setting (lab vs. gallery)
- Novelty of the medium
- Technology distractions
- Knowledge of technology
- Prior media
- Wizard influence
- Style choices (cartoon rendering, music, expressions)
- Story content (plot, characters, genre)
- Other personal influences

As I describe these other external and internal influences on player behavior, I refer to players by number: P1-P12 played *AR Façade* in the Atlanta lab version and P13-P45 played it at the Beall Center. In subsequent chapters I discuss player behavior during the experience and provide evidence of my main thesis. The main question I consider with each of these factors is how (if at all) it influences player behavior?

4.5.1 Setting (Lab vs. Gallery)

Many players commented on the setting for this intervention—both the lab setting where they were under the examination of researchers (and an array of cameras) and in the gallery where any visitor could play witness. As player 13 expressed, "I was self-conscious about the fact that I needed to interact with them ... to just try to keep the game going" (P13). For player 1, "I felt like it wasn't helping you do research if I didn't try to interact" (P1). There was a notion of appeasing the research team—that falls under the concept of the Hawthorne effect (Parsons, 1974).

Several player were self-conscious for having to speak out loud, as these Atlanta players revealed: "It's weird because you are talking to a computer and you (the researcher) are a complete stranger too" (P8) and "it's a certain amount of self-consciousness just sitting in a room talking to yourself" (P4).

This self-consciousness did carry over from the lab setting to the Beall where the players dealt with the influence of an (often unknown) audience. "This is more of like psychology sort of thing ... since there were people watching what I was saying and what I was doing, I was like more conscious of like my actions" (P30). Audiences did indeed pay attention and often scrutinized the player during and after the episode, as P29's friend commented to her as she was removing the HMD: "it was funny how you stand like with your hands in your pockets."

The presence of audiences and researchers did influence player behavior. I acknowledge that we do potentially sacrifice some of the more natural styles of play, that may only happen when players are at home alone or "fooling around" with their friends, as we saw with the blog anecdotes. However, as I discuss in Chapter 5, players exhibited a diverse range of behavior despite having people and cameras looking in at them. My video coding analysis indicates similar patterns of play styles across the two settings. The setting did not appear to impact players overall rating (24.5 in Atlanta, 24.1 at the Beall, with an overall rating of 24.2 across both locations).

4.5.2 Novelty of the Medium

Some of the first sentiments expressed by many players are similar to these: "I've never done anything like that before, so… I didn't know what to expect..." (P19) and "I didn't really have any expectations" (P27). Players did not know what to expect and so the experience was "pretty cool…" (P6), "very modern" (P28), and "kinda weird" (P10). Player 26 said the experience is "pretty cool" because "it exactly mirrored what I was seeing through the headpiece ... like the actual set" and because it enabled interaction with the characters "when you got closer to them they obviously got bigger" (P26).

While Player 18 also pointed to the novelty of having the virtual characters in a physical space: "I thought that was really neat how [Trip and Grace] were interacting with the room" and how "they were looking at me and they're able to watch me" (P18). Player 22 was impressed with the technology, but felt he needed more time to get acquainted, both to the medium and the characters:

> *"Yeah, that was pretty cool! Just the mix of the virtual space with the real space. It's pretty impressive put it together in one.... It took some digesting ... it moved too fast for me to just sort of settle into the environment... You need to have an initial, more introductory kind of phase and get more comfortable with the characters before you interface." (P22)*

This sentiment was shared by P12 who stated "it took a few minutes to get used to. Once I oriented myself, it was much more fun and engaging." (P12). While some players described the phenomena of the medium as fun: "just cause it's a little more fun to be in the action. You can walk around, sit down, look at the things…move and get closer to objects" (P8). Other players— usually in reference to the intense dramatic content— back off the word fun: "It was a really good experience, but it's not the typical feeling of fun [laughing]" (P32). In Chapter 7, I describe the range of emotions expressed by players that go beyond the preliminary novelty effect.

4.5.3 Technology Distractions

Part of becoming acquainted with the immersive medium was recognizing and dealing with a slew of technology bugs. The head-mounted display and backpack worn by players for an average of 14.4 minutes did not became a major point of conversation.

P18 said, “I didn’t notice the backpack at all... it wasn’t too heavy or anything....” Some players were hesitant due to the large backpack, such as P23, “I was afraid to sit on the couch because of the equipment I was wearing, and didn’t know how far off my back it stuck off for where I would sit.” Player 31 said “I wanted to sit... on the couch, but then I was like, ‘Oh, I'm afraid I'm gonna break something’, so I didn’t” (P31).

Due to the narrow field of view of the display and the fact that the characters could move around, players often lost their orientation to the characters, as P43 stated, “what was hard for me was to get the sense of where they were. It’s hard to find them to see what they’re looking at.” This meant having to “keep turning my head” (P18) and “look around the whole room for her” (P29). Player 17 said, “it wasn’t like I could use my peripheral vision to know where people were.... I would have liked to not have as much periphery vision”, suggesting that the periphery view directly into the space may have caused more of a distraction. Most participants made the necessary adjustments to the AR experience and when asked specifically said something along the lines of “the HMD didn’t really bother me” (P8).

Another persistent technology issue was the registration issues of virtual content with the physical world, leading one player to say “part of the reality was taken away when I noticed the glass was floating or he was floating.” (P31) For some players this technology anomaly was interpreted within the context of the story and then later reconciled as an actual bug. Player 2 said “at one point I thought Trip was standing on a table…. I was thinking he was fixing a picture… then I messed with the glasses and he came back down” (see Figure 4.3).

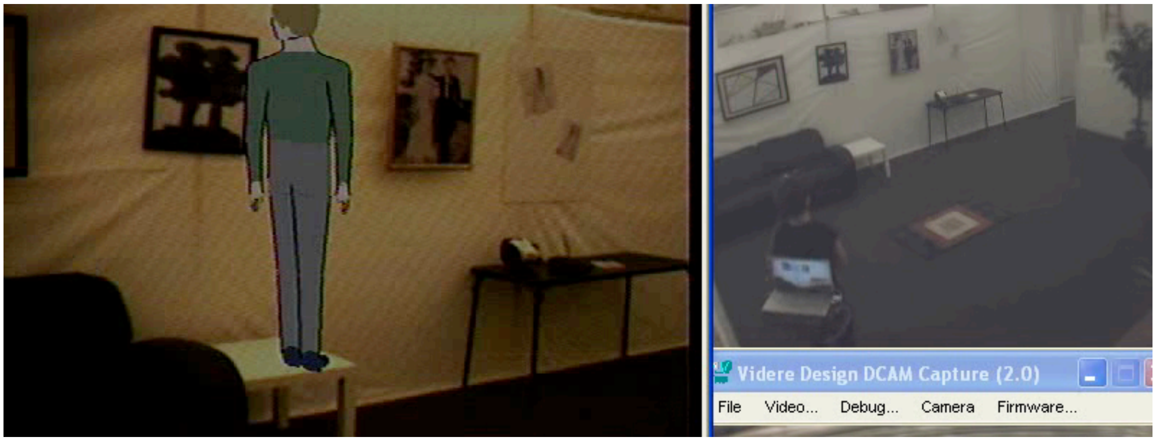


Figure 4.4: Player 2's experience of the registration error (Left) Player 2's HMD view where Trip appears to be standing on the table; (Right) a third-person view of the player in the apartment.

The interviews revealed a tendency by players to assume technology anomalies were part of the drama. When Player 10 saw Trip stuck in a momentary path-planning loop he thought that it related to the social situation: "at first I thought maybe he wasn't sure if he wanted a drink or not" (P10). On several occasions, the AI engine crashed while the graphics engine kept on rendering the characters in the HMD. Players would continue to see the video-mediated space and the characters along with their surface expressions. A couple of those players continued to probe Trip and Grace for quite some time before I had to inform them that the experience had ended.

For the most part, technology bugs in the AR version, such as imperfect graphics registration, tracker errors, and path planning glitches, were ascribed to the research nature of the experience. Players seemed to accept the unrefined technology, saying "It didn't really affect me at all, negatively… I didn't think about it. I was really paying more attention to the story line" (P3) or "it didn't really take away from it" (P7). It appears that players preferred or disliked the AR interface despite the limits of the technology, rather than because of them.

Technological anomalies were peripheral to larger issues of poor interface affordance and emotional distance from the drama, as discussed in Chapters 6 and 7 respectively. The momentary "breaks" in presence may only have a minimal effect on the experience, although they seem to be very useful for gauging player engagement and for understanding what a player is currently thinking about the experience. In Chapter 5, I consider how these moments of technology breakdown can be indications of how the player was currently engaged.

4.5.4 Knowledge of the Technology

How knowledgeable were players of technology? How did they think this experience works? Does it matter? Is it actually better to be blissfully ignorant of the technology? Looking at the self-report statistics, the players appear to have solid prior experience with computers, video games, and TV (see Table 4.2).

Table 4.2: *AR Façade* players (N=45) self-report use of technology, averages with standard deviations

Average Computer Experience (where 1=never and 7=expert)	5.5 (0.9)
Video Game Experience (Average hours per week)	4.5 (6.5)
Television/Movies Experience (Average hours per week)	7.3 (8.4)

When asked to describe the technology behind the immersive and interactive story experience, most players struggled: "You asked the worst person for this, like technology-wise. I have no idea." (P26) and "I don't know how that works really." (P28) Player 18 offered the following explanation of how the characters moved around the space: "I want to say like a track or something that they were going on. You know what I mean? I really have no idea how that works. [Laughter]" (P18)

While most players did not venture an explanation of the technology, others provided laymen's descriptions based on their first-hand experimentation.

"I tried to touch the characters to see where my hand fit in. I guess the characters are almost like a slide, but the slide would be clear... I don't know if that makes sense." (P19)

"I don't know what they are, graphic images, but I guess the room is set and then there are set parameters like where they can go in the room and that bit. Like, when he went behind the bar you couldn't see his legs." (P25)

Some of the theories put forth by players actually extend the abilities of *AR Façade*. The system does monitor the location of the player, but several folks imagined the *AR Façade* reading into their actions much more than it actually does.

"I'm gonna guess it checks the range of your voice, like the pitch of how you end the question. You kind of raise your voice so they interpret that as a question maybe." (P38)

"I thought that I should start using their names to get their attention and it really helped with eye contact. Not so much necessarily with communication, but when I said 'Grace' she turned to me." (P21)

"Trip was saying, 'the focus is on us, not just one person.' ...So I think he wanted me to look at him. ... in that situation, I had to like maneuver around the room to try and get both of them in the same perspective, and then they stopped saying it to me." (P23)

These explanations can provide insight on the strategies players use to interact with *AR Façade*, a point I will talk about in great detail in Chapter 6 when I talk about the affordances of the environment. Many players suspected that the system performed word matching.

> *"It goes to like a certain word, and then they try to match that word to like a set of questions. I don't know. I'm not much of a computer person." (P29)*

> *"I think there were a few trigger words. If I would say something when they would say something, I would have to say one of the few words to trigger another emotion or response from them." (P27)*

It seems that the player's actual knowledge of the technology matters less than the strategies players develop to make the interaction work. As player 4 explained, he just needed to learn the skills.

> *"There was a learning curve… if I did AR 5 times… I would feel more comfortable with it… I have the basic skill set, just not the VR skills… I was raised with a desktop computer…" (P4)*

Players enjoyment of the experience does not hinge on having a fundamental understanding of how it works. Considering how much player read into the abilities of the system, players likely benefited from naive understanding of the technology, supporting anecdotal evidence of Coleridge's notion of the "suspension of disbelief".

4.5.5 Expectations Set by Prior Media Conventions

Some media theorists have claimed that users experience of new media technology is highly dependent on their experience of prior media (Bolter and Grusin, 1999; Voida, 2008). This was apparent during the interviews as players used the language of prior media to speak about their experience.

“Almost like being in a movie in a way except I was just kind of like an extra. (Laughter)” (P17)

“Like a video game ‘cause .... it’s like a different world, kind of thing” (P13)

“It definitely felt like I was put into an Improv situation, like in a drama class or something” (P37)

“It's like one of those books where there's like a bunch of alternate endings.” (P31)

“a disneyland-esque experience where it's kind of like surreal where it's almost like you're part of a ride...” (P30)

“...like an interactive movie...” (P15)

“...like role playing games...” (P24)

On the questionnaire given to players after the interviews, one question asked the player to relate *AR Façade* to other forms of media, such as TV, video games, and improv theatre. The question asked players to select a medium that most resembles the *AR Façade* experience. Most players compared it to video games or improv, and a fair number compared it to movies and TV. Interestingly, the players who compared it to video games generally rated the experience less favorably (22.5 out of 35) compared to players who related *AR Façade* to TV or movies (25.7). Although it is not statistically significant, it might be an indication that players who related *AR Façade* to video games had more conventional game expectations for the experience; thus, some players were not satisfied by unconventional nature of *Façade*.

For some players, the novelty and physicality of the AR experience elicited a set of expectations that seemed to build less on their prior experience with video games and other media, and more on their everyday experiences in real-life, as described by Player 4:

> *"When you are standing in the real world with a headset on and you are interacting with them… it didn't feel like a video game as much as it felt like real life." (P4)*

The fact that players pulled from their experiences in everyday life indicates that an hermeneutic analysis, as suggested by Voida (2008), would be useful for contextualizing immersive story experiences. I reference these findings again in Chapter 6 when I discuss the effect of player expectations on embodied narrative engagement.

4.5.6 Wizard Influence

Players were not told that wizards were used to achieve speech interaction[49], and most players did not notice a human element in the system, as revealed in statements like, "I said to Grace that she sounded stressed and I guess the computer took it as 'depressed'...is that a problem with voice recognition?" (P1). Even when players were stunned how well the system performed, they often did not suspect a hidden operator behind the scenes:

> *"The technology works pretty well for me I suppose... I didn't know if there was a mic or anything around the backpack, but I was surprised that it could hear and decipher what I was saying even in a low voice." (P37)*

49 Our Wizards reported that they informed players about their role during the 11-week Beall Center deployment, but during the final two weeks during the in-depth player interviews, players were not informed about the presence of wizards.

Only a few players commented on the presence of human operators, like Player 25 who said “I feel like someone was controlling it, ‘cause [the characters] would change what they were saying halfway through” (P25) and Player 30 who asked “is somebody like typing what I'm saying...?” (P30). One player with experience as a theatre technician “assumed there was someone typing” because he knows that “adjusting microphones is a huge pain” (P4).

As I describe in Section 3.4, the docent wizards at the Beall Center exerted their influence on the experience, particularly as they adopted the use of the Discourse interface. As the wizards learned the Discourse interface and sought to make the experience more engaging, they would drop in unwarranted discourses and choose the wrong discourse, occasionally resulting in the player getting kicked out. Players did not explicitly attribute these errors to a human operator—since most players did not notice that a wizard was part of the system—although the wizard performance likely did have an influence on the player experience.

Most wizards managed to get player ratings above 23 on the 35 point scale, but W4 and W8 both averaged overall ratings below 20 from their six participants. Those two wizards invoked the two lowest averages for discourse selections per min (1.4 and 2.1, respectively), but any variances in discourse selections per minute were not statistically different and so I cannot draw strong conclusions about the low ratings for those two wizards’ players. It might be possible to get further insight into the effect of wizard performance by conducting a detailed conversation analysis of each player, to figure out exactly what each player spoke, and then see how it matches up with what the wizard entered.

4.5.7 Style Choices

Some players pointed out the surface features of *Façade*, the design choices such as the music, facial expressions, and character rendering techniques. The music really set the backdrop, as Player 17 stated "the music really made me feel like something was wrong between them" (P17), similar to player 31 who said "the music was like really haunting" (P31).

The characters' facial expressions and body language were very important for communicating the emotional state to players, as expressed by these participants:

> *"That arched eyebrow thing she was doing… the facial expressions in this aren't half bad… They convey a lot…" (P4)*

> *"The facial expressions were pretty decent… like rolling her eyes at him and snarling… they also use some body language which helps give you a sense…" (P8)*

Interestingly, not many participants mentioned the fact that the characters are cartoon renderings against a real backdrop, only expressing sentiments like, "they seemed more animated than human" (P1) or that the AR interface is like "dealing with partial bits of reality" (P25). The cartoon appearance of Trip and Grace appeared to be a good fit and kept players out of the "uncanny valley", a theory introduced by Mori to explain why humans reaction negatively towards robots, as robots are made more human-like—but not quite exact—in appearance and motion (2005). However, for some players, the cartoonish renderings had a negative effect:

> *"I couldn't take it seriously 'cause the characters themselves were just like cartoons basically." (P17)*

> *"The graphics seemed kinda two dimensional... 'Cause you feel like you're in a 3D dimension and then this 2D figure pops up in front of you, – I lost my concentration." (P38)*

So on one hand, our tight manipulation of facial characteristics and Trip and Grace's non-photorealistic appearance may have strengthened the presence of Trip and Grace, rather than decreasing it (just as one would predict from Mori, 2005). On the other hand, there did seem to be a desire for more realistic renderings, perhaps driven by the 'three-dimensionality' of the AR experience and would suggest that the immersive interface raised player expectations.

4.5.8 Story Content

Beyond the technology issues and the surface-level design choices, players were also influenced by the story content and character design. The dramatic situation was one of the first things players mentioned, as expressed by these players: "the topic was a charged one" (P33), "it felt like I walked into my friend's house, where her and her husband are about to kill each other! (laughing)" (P29), "it was really a soap opera"(P21), and "it totally reminded me of Days of Our Lives or something like that" (P21).

The *Façade* story is indeed a soap-opera. When I recruited participants for both the lab and the gallery studies, I tried to minimize the outright rejection of *AR Façade* due to content preferences, although this was impossible to avoid completely. Before players volunteered for my study, I would let them know about the content without giving away the plot. I succeeded in getting players who would not be turned off by the drama genre, as indicated in the demographic questionnaires that asked players about their favorite genres of video games and movies. The top two preferred video games were "adventure" and "role-playing", while "drama" was players' favorite genre of movies.

Players had different reactions to the charged drama in *AR Façade*. Player 21 "couldn't help but laugh the whole time" (P21). Player 13 was curious and "wanted to know what happened" (P13). Player 11 wanted to get "out of the situation and let them work it out" (P11). Players' emotional reactions to the situation influenced how they behaved, as I describe in Chapter 7.

Not only does the *Façade* content represent strong human emotions, the characters and the situational premise were grounded in reality.

> *"I thought that [Trip and Grace] were really amusing and tragic… I thought that their personalities were really believable… they remind me of real people." (P3)*

> *"The premise for the story line is very realistic... to me seemed very real because I've seen and met people that are like that. ... it could have happened." (P14)*

The characters and story were very believable because they build on common experiences and cultural understandings of relationships, marriage, and gender differences. Player impressions would likely be different if *Façade* had a different story or a happier tone. The intense, uncomfortable setting of witnessing a marriage in shambles likely helped to accentuate the finding that players need 'distance' from the medium, but as I discuss in Chapter 7, the overall influence of the immersive interface appears to be independent of the content.

4.5.9 Personal Influences

It would be impossible to understand all of the personal influences that impact how people engage an experience. Individual moods, personalities, cognitive capabilities, everyday circumstances, and previous encounters come to bear on media experiences.

The effect of personal influences has been researched in the context of reading by Green, who found that a person's ability to engage is highly individualistic (Green, 2004). Readers with prior experience relevant to themes in the story are more likely to feel "transported" or to be able to "identify" with a particular character (Green, 2004).

For the Beall portion of the study, I attempted to 'measure' personality using an abridged version of the Myers-Briggs[50] personality test (see the study instruments in Appendix J). I wanted to know, for example, if extroverted personalities favored the *AR Façade* experience over introverts. With only 33 participants contributing to that study, I was not able to draw any conclusive arguments, but for a longer discussion of personality types and its influence on the game experience, see Appendix P.

Some of the player demographics appear to have a relationship with players' behavior and enjoyment of the story themes, as I discuss further in Appendix O. Looking at gender differences, for example, there was no difference in overall rating, but men and women did react differently to the experience, as I discuss further in Chapter 7.

Perhaps the most salient personal influence was how players could relate the situation to their own lives, as some players expressed:

> *"If I had to relate it to anything – it would be the scene that would have played out in the lives of people I know who have gotten divorced or who have broken up and it was just kind of irrational arguments – not trying to connect at all..." (P21)*

> *"I've faced them before with like dealing with my own friends and stuff." (P24)*

---

50 http://www.myersbriggs.org/ (accessed 9/22/08)

Although this scene is familiar to many players, not everyone was able to relate:

> *"I've never had to work with friends who are breaking up... This was a very foreign situation to me." (P36)*

More importantly, the social situation represented in the simulation became a springboard for players to reflect on their own lives and on how they could better deal with that particular situation:

> *"I understand those moments since I'm married as well. (Laughter) Sometimes it gets like that... now I kinda see if me and wife are arguing in front of people, you can kinda feel that way." (P38)*

> *"I could have done a better job of facilitating the argument...and then help them come to an agreement." (P13)*

While the intervention itself persuaded players to deal with a situation, the post-interviews extracted deeper self-reflections. The opportunity for self-reflections did not stop with the players, as I discussed in the previous chapter. Wizards not only followed along to perform the task, but also reached insights about the players, saying stuff like "I can definitely tell a lot about a person... If they're outgoing... if they're shy... if they're creative or not... if they're smart" (W7). Section 3.4 provides a discussion of the insights revealed by our team of undergraduate docent wizards.

External viewing audiences also used the intervention for reflection. I talked to some audience members informally, usually just friends of the player who came along and watched. Player 32's friend said of his friend, "she reacted much like she does in real life... it like drew real emotions and .. reactions out of her... even her body language was very similar." Player 29's friend said: "It was like watching a movie. ...the scenario was pretty believable ... and then seeing how she dealt with it, 'cause I know who she is. It was like I was like, ‘Ah, she would definitely do that.’ ” Insights from players themselves—as well as wizards and audience members—on how players deal with the social situation speak to the potential of the medium.

4.6 Summary

In this chapter, I described the empirical studies we conducted to study *AR Façade* across two installations: an Atlanta interface comparison study and the in-depth player investigations at the Beall Center. I outlined our data collection procedures and analysis methods. Our methodology for empirically investigating *AR Façade* is itself a contribution, as it informs the evaluation of future immersive and interactive stories. In the last half of the chapter, I presented initial findings about the influence of setting (lab vs. gallery), novelty of the medium, technology distractions, existing knowledge of technology, prior media, wizard effects, style choices (cartoon rendering, music, expressions), story content (plot, characters, genre), and other personal influences. While all of these have some impact on the overall user experience, my dissertation focuses on the effect of the immersive interface on the overall sense of embodied narrative engagement.

CHAPTER 5

PLAYER DEMOGRAPHICS AND STYLES OF PLAY

Direct observation reveals details about player emotion.

—Nicole Lazzaro, Why We Play Games (p7)

Anthropology demands the open-mindedness with which one must look and listen, record in astonishment and wonder that which one would not have been able to guess. —Margaret Mead

In the previous chapter I introduced my methodology and discussed some of the external and internal influences that could impact player behavior and their interpretation of the media experience. In this chapter, I take a closer look at the players from both installations of *AR Façade* and share an in-depth, qualitative account of player behavior during the game episodes. The reason for conducting in-depth investigations of game episodes is to more deeply understand questions about *embodied narrative engagement*: How do players engage in immersive and interactive stories? What influences engagement during an episode? Are their different varieties of engagement?

With each new formulation of computational media, researchers have sought to make sense of how participants behave and engage with the media artifacts. Classification labels, such as hardcore gamer versus casual gamer, are initially useful for understanding player habits (Bartle, 1996), but can often fall short of fully explaining the dynamic relationship between participants and the medium. Such distinctions, however, do communicate something useful about the diverse personalities and approaches of the participants. Such analyses hopefully lead to more diverse content production and provide opportunities for applying adaptive intelligence techniques.

In this chapter, I will present player demographics and general game statistics. I illustrate my method for visualizing episodes of *AR Façade* and use the visualizations to outline player behavioral patterns and how they are effected by events in the experience. I present a qualitative analysis of five “styles of play” for immersive and interactive stories, as evidenced by players’ different goals, interpretations, and appropriations of *AR Façade*. I illustrate the five styles of play—engager, performer, partaker, tinkerer, and observer—through five case studies consisting of episode visualizations, excerpts, images, and player quotes. I roughly classify players into one of the groups and investigate the in-game quantitative differences. Finally, I discuss player modeling research and the possibility for run-time detection of play styles towards more adaptive immersive and interactive story experiences.

In this chapter, I explore the notion of player types for immersive and interactive stories and present three main ideas:

1) Players do not exhibit just one form of engagement in immersive and interactive stories, but a variety of equally valid styles of engagement (my analysis explains five: engager, performer, partaker, tinkerer, and observer), as evidenced by their different goals, interpretations, and appropriations of *AR Façade*.
2) Players’ style of play may change throughout an immersive and interactive story and is influenced by story-related and technology-related events.
3) Players’ style of play may be correlated with certain game statistics and player demographics.

5.1 Player Demographics and Episode Statistics

Across two study settings—Atlanta lab and Beall Center gallery—I conducted in-depth interviews and episode analysis of 45 players. My interview methodology is explained in Chapter 4. This section presents player demographics, in-game data, and some commonly observable phenomena from the game episodes.

5.1.1 Player Demographics

Overall I recruited a diverse group of individuals. Occupations varied from physicist to homemaker, although the majority of participants in the Beall Center portion of the study were students. The age of participants was fairly young, ranging from 13-45 with a mean of 24 years old. We balanced gender (23 females to 22 males). About 10% of the participants had already played *Façade* before trying *AR Façade*. Figure 5.1 shows demographics for each player and provides an initial glimpse of the episode length and game ending. The first twelve players are from the Atlanta study; 13-45 are the thirty-three additional participants at the Beall Center.

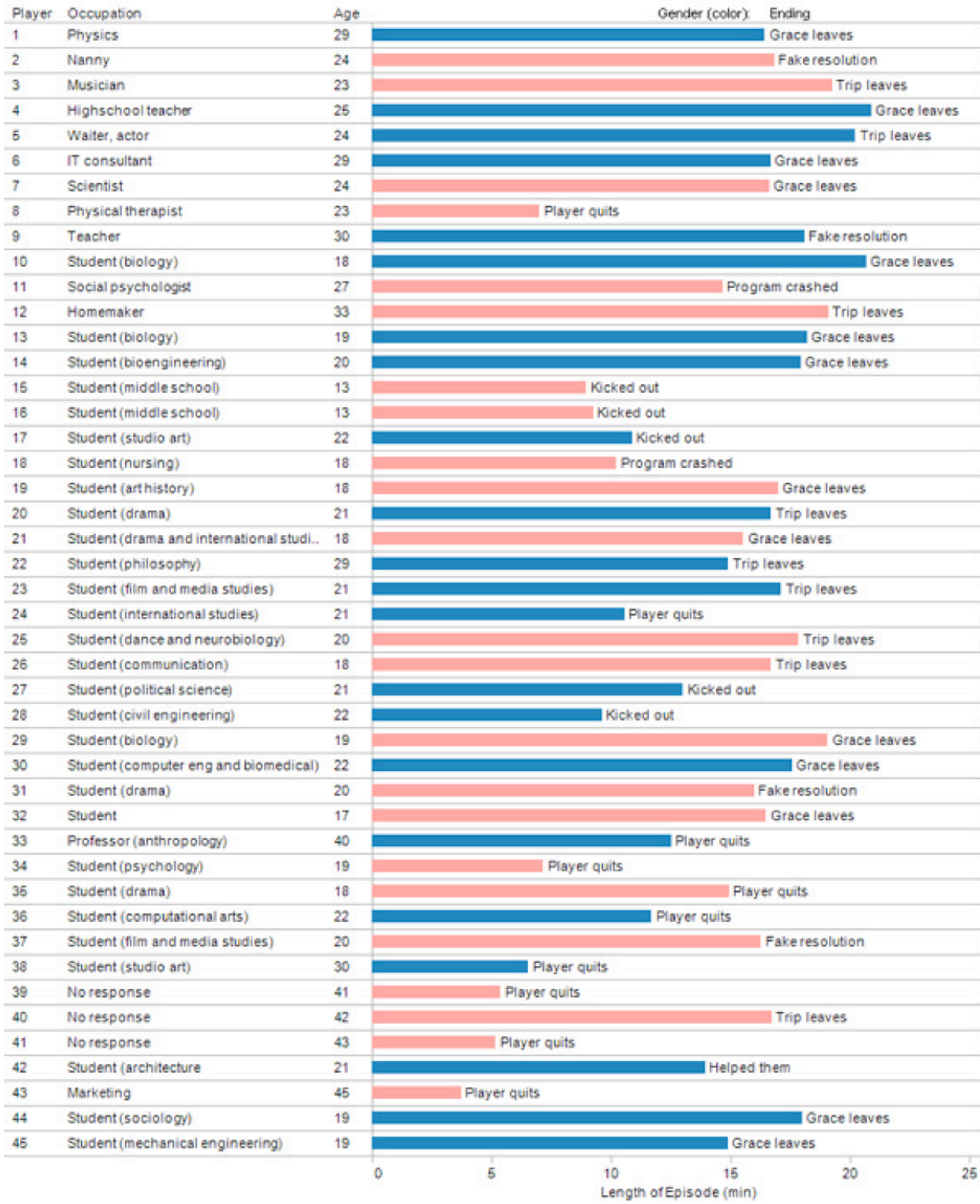


Figure 5.1: Key player demographics and episode data (N=45), from left to right: Player ID, occupation, age, gender (color of bar), episode length (length of bar), and episode ending.

5.1.2 Episode Statistics

In Figure 5.1, the bar shows the length of the episode with the particular episode ending to the right of the bar. The average length of play was 14.4 minutes, but it was different between sites (17.2 minutes on average in Atlanta and 13.3 minutes at the Beall). This difference can likely be attributed to the gallery setting where people visited in-between classes and felt pressure perhaps to let other visitors have a try. Across both installation, 22% of players quit the experience before a scripted game ending occurred, but 9 out of 10 of those came during the Beall installation. Overall players encountered all the endings, but only one player actually achieved the “winning” ending of helping both Trip and Grace (see Figure 5.2).

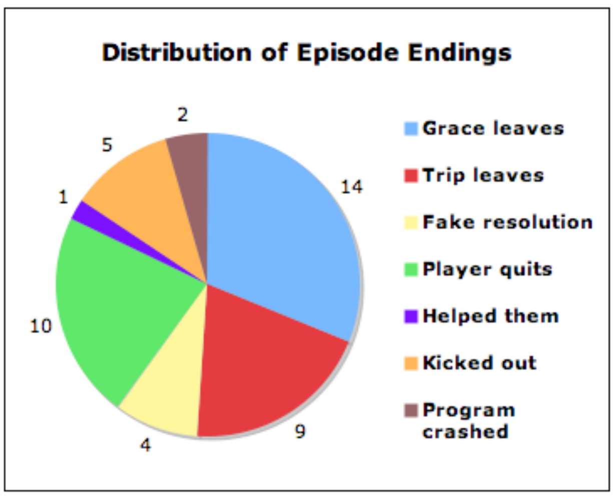


Figure 5.2: Distribution of episode endings (N=45 players) for *AR Façade*

The game logs provided a few more statistics about player actions. Atlanta players used an average of 3.8 lines of dialog per minute, while the Beall players registered an average of 2.3 discourses per minute (see Table 5.1). The discrepancy between dialog in Atlanta and discourses at the Beall center does not necessarily imply that players spoke much more in Atlanta versus the Beall. It likely has more to do the strategies developed by our wizards, as discussed in Section 3.4. In the dialogue interface used by wizards in Atlanta, longer statements were often typed as two or three separate entries. Also, the

Beall wizards would deviate from what players said and occasionally did not enter anything.

Table 5.1: Player averages for key episode data across both installations (N=45)

Length of Episode (minutes)	14.4
Number of lines of Dialog (Atlanta only) per minute	3.8
Number of registered Discourses (Beall only) per minute	2.3
Number of "special" game gestures per episode (hug, kiss, comfort)	3.0

On average, players used about 3 canned gestures per episode. Again, this does not accurately represent what players actually did with their bodies. Wizards may have missed game gestures (since the players were viewed through camera monitors and their primary attention is devoted towards the speech entry task). Plus many of the gestures players used are not part of the game. The video coding described below will give an overall glimpse of the actual frequency of speech and gesture.

Another game statistic listed in Table 5.1 is the average interpersonal distance (IPD) between the player and each character, measured by averaging the game distance between the parties. There was no difference in IPD between the player with Trip vs. Grace (both 2.11 meters), nor was there a significant difference between the Atlanta and the Beall settings (for Grace: 2.13 and 2.11, respectively; and for Trip: 2.06 and 2.13, respectively across the two locations). However, IPD becomes part of the discussion when I talk about comparing different interface versions and 'media distance' in Chapter 7. Additional demographics and episode correlations are summarized in Appendix O.

5.1.3 Reoccurring Episode Behavior

Beyond the game statistics, I observed certain phenomena happening time and again: interacting with physical objects, moving near story objects in the apartment, adjusting their position with respect to the characters, and reacting to social cues delivered by the animated characters. In this section, I highlight some of the recurrent behavior that motivated me to conduct a more formal episode video analysis.

Many players felt at ease interacting with the physical objects, such as sitting on the couch, closing the physical door, and picking up the wine glasses, trinkets, or phone. Most players moved around the space at the request of the characters, especially when Trip asks the player to look at the "Italy photo" (see Figure 5.3). If players can identify which photo Trip is referring to, they generally move over to that part of the room to get a closer look. This same 'compliance' behavior happens when conversation turns to the couch, the bar, the trinkets on the shelf, the window view, the wedding photo, and the artwork above the couch.

Figure 5.3: Player 38 directed to look at the Italy photo

Only two players actually followed a character back into the kitchen space after Trip and Grace have their first fight. Player 1 commented on that moment where Grace left the room, saying his "instinct was to look for her (Grace)", but that he did not explore the back room because "it was totally inappropriate during a heated argument." Player 10 posited a more practical reason for not exploring the back room when comparing the AR to the KB version: "I felt easier walking into the kitchen in (the KB version), because I knew I wasn't going to get hit with anything."

Generally players looked to the characters and their subtle expressions for cues on what to do and when to speak. Player 10 postulated, “if they wanted me to say something they would look at me…“ (P10). Player 23 claimed that Trip was saying his name to get his attention. “I think he wanted me to look at him... I had to like maneuver around the room to try and get both of them in the same perspective” (P23). Players spent a fair amount of time just readjusting their perspective in the space to have a better view of the characters, especially in the later stages of the fight where the player is more prone to simply observe. In Chapter 6 and 7, I provide many more examples of how players behaved socially and used their physical bodies in the flow of the dramatic action.

5.2 Episode Analysis

Towards forming a deeper understanding of each episode of *AR Façade* and how players reacted to the game situation, I conducted a video analysis of the episodes as described in Section 4.4.2. In this section, I briefly describe the narrative structure of Façade in terms of its three distinct “phases”: affinity game, therapy game, and end game. I introduce visualizations I created from the episode video codings and provide some general observations about player behavior based on the visualizations. I include the episode visualizations and game play images to help illustrate my points, only occasionally referencing the full results of the visual analysis (Appendix N). Finally, I will discuss the effect of *Façade*’s phased narrative structure and some of the more severe technical errors, both of which appear to influence how players behaved over time.

5.2.1 Three Phases of Façade

In an attempt to provide a loose dramatic arc, Mateas and Stern designed *Façade*’s narrative structure with three phases, often referred to as the affinity game, therapy game, and end game (Mateas and Stern, 2000). In the affinity game, Trip and

Grace try to pull the player onto their side; in the therapy game, players' actions contribute to each characters' degree of self-realization; and in the end game, characters usually reveal something about their past and then the experience reaches one of the possible endings. There is also a "hot-button" game that occurs in parallel with the affinity game where the player can trigger satellite topics to learn more about the characters' back story. The player can leave or get kicked out at any time.

In each episode visualization, I mark three points: the transition from Phase 1 to 2, the transition to Phase 2 to 3, and the end of episode. These transition points occur at different times for each player episode because of the generative narrative structure of *Façade*—the character dialogue unfolds based on player interaction and is different in every episode. I identify the two transition points by two common exchanges that happen in every episode. The first transition occurs when Trip or Grace have their first fight and one of the two of them storms into the kitchen space. The second transition occurs when either Trip or Grace declares "I've been paying close attention to what you've been saying tonight" and then delivers offers statements about what the player supposedly said throughout.

5.2.2 Visualizing the Episodes

Each player episode is represented by a bar with different colored triangles protruding up or down from the bar marking the event occurrences. The full bar length represents the maximum length of episode (21 minutes). Each episode bar includes up to two blue triangle markers for the transitions between phases and a black marker at the end of the episode. Triangles pointing up represent player gestures (brighter) and speech (more opaque), while triangles pointing down represent potentially distracting events (technical errors and failed character responses). If events occur in subsequent intervals they appear as a solid trapezoid, not a triangle. Figure 5.4 provides a legend for the visualizations, which are included in their entirety for all 45 players in Appendix N.

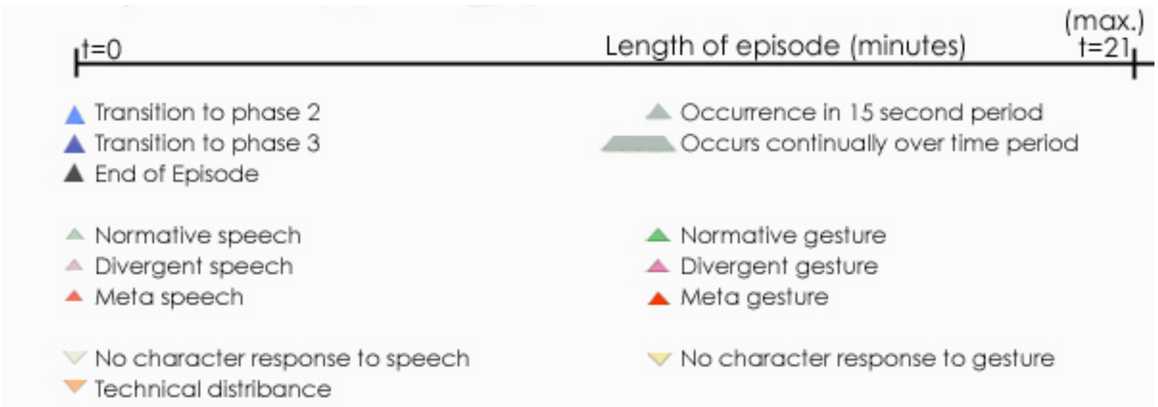


Figure 5.4: Legend for game episode visualization (triangles represent occurrences in 15-second intervals, colors represent different categories of actions, brightness differentiates between speech and gesture.)

Similar to my coding scheme, I developed the visualization method to suit my needs using the tools I had available. I first attempted to use established visualization tools (Tableau[51] and InfoScope[52]), but I ended up using Excel so I could control things like the length of the bar and the shape of the occurrence. The tools for visualizing time-series coded events are somewhat sparse, pointing toward a need for future work on tool support for qualitative analysis and visualization of activities within a physical space.

5.2.3 General Observations about Player Behavior

Looking at the episode visualizations as a whole (Appendix N), several general observations emerge. Most of the episode visualizations are overwhelmed by green (normative) speech and gesture. Some players like 3, 14, 26 and 35 interact consistently and 'normatively' throughout, even in the late stages (see Figure 5.5). Every player had at least some occurrence of green normative behavior.

51 http://www.tableausoftware.com/ (accessed 9/22/08)

52 http://www.macrofocus.com/public/products/infoscope/ (accessed 9/22/08)

Figure 5.5: Player 26 shows normative behavior throughout the episode

Other players tend to dabble with the other categories of action, seen in the visualizations as occurrences of pink and red triangles peppered in among the green. Other patterns also stand out. Many players—for example, 1, 2, 5, 9, 13, 17, 18, 32, 37, 39, 42, and 44—exhibited red, meta behavior within the first minute or so of play; when players first entered in the AR environment they often "felt out the edges" to try to figure out what they can and cannot do. When players get into the third Phase of the *Façade* narrative (after the dark blue triangle) they often interact less than in the first two Phases. For example, the frequency of conversation for players 5, 12, 14, 19, 20, 25, 40, 42 , and 44 drops significantly from their first two stages. I attribute some of this to narrative structure during the third Phase, but some to continual conversational breakdowns.

Figure 5.6: Player 42 exemplifies transitioning between different styles of play

Player 42's episode is representative of how some players changed behavior during the experience. He starts out enacting primarily meta-speech, then continually interacts with the story during the second Phase, and then doesn't interact much at all during the third Phase. Notice in Figure 5.6, Player 42 encounters a technical disturbance within the first minute and then throughout Phase two the characters fail to effectively respond to him. Like many of the players, P42 transitions how he behaves during the course of his episode based on specific events he experiences. Next, I take a closer look at the effect of the three Phases of the *Façade* narrative structure by investigating specific examples.

5.2.4 Behavior During the Affinity Game (greetings, drinks, and decorating)

The first Phase of *AR Façade* involves greeting the characters, talking about the decorations in their apartment, Trip fixing drinks behind the bar, and other exchanges as Trip and Grace subtly (or not so subtly) try to get the player to form an affinity with them. The player's actions are revealing during this phase. When players first enter the apartment, how do they greet the character? How do players adjust to the technology? For example, some players follow Trip's invitation to "come on in!" and feel comfortable speaking and interacting with the characters right from the start, like Player 40 and Player 16 who greet Grace with a big hug (see Figure 5.7).

Figure 5.7: Players greeting the characters at the beginning (left) Player 40 and (right) Player 16 giving Grace a hug upon entering the space.

Other players are much more hesitant with the technology, like Player 13 who expressed his uncertainty during the interview:

> *"At first it was like weird, because I wasn't sure if I should speak or if I was speaking loud enough or or if I was speaking too loud... they kept asking me questions and I'd say something and I wasn't sure if they heard me." (P13)*

Player 17 actually took quite a long time to adjust to the interface as seen from the red indicators in his game episode visualization (see Figure 5.8).

Figure 5.8: Player 17 had interaction problems throughout and quit early

Player 17's exchange of greetings with Trip at the very beginning of the episode reveals that he speaks more towards the researcher outside the experience, rather than towards Trip.

Table 5.2: Player 17's opening exchange with Trip

Time	Player and character statements (with notes in parentheses)
0:40	Trip: Peter! Hey, I thought I heard someone out here.
0:43	P17: Whoa!
0:44	Trip: Great to see you. It's been a while. How's it going?
0:47	P17: Ahh
0:49	Trip: Yeah, yeah... Come on in.
0:51	P17: What happens if I don't say anything? (towards researcher outside)
0:52	Trip: Let me go get Grace
0:53	P17: Ok, hold on.
0:53	Trip: Oh, ah... well....
0:54	P17: Come in?
0:56	Trip: Just stay here.
0:58	P17: He said to stay here... (towards researcher who instructs him to go on in)

The episode visualization in Figure 5.8 also shows that P17 experienced quite a few technical errors early on, which is evident when looking at the HMD video data. After Trip asked the player to "come on in," he did a strange loop through wall near the door (see Figure 5.9), which caught the player off guard:

> *"[Trip] would walk in through the door, you know, and then would just disappear. You know, Trip walked through the door and through the wall and miraculously came out some where." (P17)*

In Player 17's case, he never did fully recover from the rough start and so he ended up quitting shortly after Phase 2 started.

Figure 5.9: A technical disturbance where Trip appears to walk through the wall

The narrative around the decorations and the drinks also exposes the players willingness to follow along with the social situation. Some players would try to change the subject. Some players would start to ignore the characters. Some players played along like Trip and Grace were old friends.

5.2.5 Behavior During the Therapy Game (first fight, one-on-one, counseling)

The second Phase of *Façade*'s narrative starts with the first fight between Trip and Grace, where one of them ends up storming out of the room, leaving the player with the other character. A moment later the other character returns and the player finds themselves in the middle of an ongoing argument between the two. This Phase of the experience demonstrated the greatest diversity of reactions from players. For example, Player 8's episode visualization shows that she treated the situation normally before she quit after Trip and Grace had their first fight (see Figure 5.10).

Figure 5.10: Player 8 acted socially appropriately, but quit the experience

Looking at Player 8's detailed episode interaction, it became evident that she tried to politely escape the social setting for well over a minute (Table 5.3).

Table 5.3: Player 8 exchange near the end of her episode

Time	Player and character statements (with notes in parentheses)
5:34	P8: Ah, think it'd be ok if I come back at another time? (walks towards door)
5:42	P8: Maybe I can come back another night.
6:34	P8: Maybe I'll come back another time.
6:42	P8: OK, I'm going to leave now.
6:43	P8: Bye... (backs up towards door, hand on door handle)
6:51	P8: Bye! (turns and leaves)

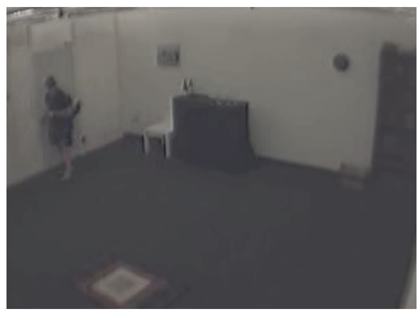


Figure 5.11: Player 8 by the door trying to politely excuse herself

P8 acted as if Trip and Grace really were old friends, but she did not want to mediate their fight. She thought if she stood by the door (see Figure 5.11) and threatened to leave, these actions might change the subject. When her attempts did not have any impact, she decided to give up and leave, as she explains in the interview:

> *"They kept calling me in to fix it as I am trying to leave… I didn't feel like I could help them because everything I said they would jump back to what they were going to fight about anyway. Nothing I would have said would have influenced the outcome…" (P8)*

Player 34 had a very similar exchange during her episode, hanging around the door before finally leaving, whereas Player 43 left much more abruptly as soon as she sensed a hint of tension. The therapy game also unveiled other player reactions, such as

P26 (see Figure 5.5) who continues to speak to the characters and act as the therapist throughout Phase 2. For player 7, when the fighting started her behavior changed (see Figure 5.12). She went from acting appropriately in the situation to behaving disruptively. She tells Trip to "calm down, calm down.. here have a drink." and then tries to force Trip to drink from the glass she got from the bar.

Figure 5.12: Player 7 style of play changed when the fighting started

When some players got to Phase two—like Player 9—their exchanges with the characters happened much less frequently, usually only responding to the yes/no questions posed by Trip and Grace (see Figure 5.13). Direct questions from the characters often revealed what the players were thinking and whether they were paying attention and taking it seriously. Some players would ignore direct questions (P31); some would answer the questions cynically (P17); Other would answer them deliberately and clearly as if they felt this was the only means to get the characters to respond (P6).

Figure 5.13: Player 9 spoke less as the experience carried on (usually only responding to yes/no questions from the characters)

5.2.6 Behavior During the End Game (character revelations and resolution)

Assuming players stick around long enough, the final Phase of the experience is where either Trip or Grace, or both, reveal secrets about themselves (Trip had an affair, Grace secretly paints, etc.). Soon after the characters share their revelations the experience ends, and as Figure 5.2 illustrates the most common endings—occurring over 50 percent of the time—are when either Grace or Trip declare the relationship over and then leave the apartment. Many players reacted to this moment, again revealing how the players were engaging with the experience. Do players try to stop the characters? Do they just stand and watch them leave? For example, there is an interesting contrast between

Player 3 who tries to stop Trip to convince him to work things out with Grace (see Table 5.4 and Figure 5.14), and Player 29 who tries to run out with Grace because she doesn't want to be left alone with Trip.

Table 5.4: Player 3 exchange near the end of her episode

Time	Player and character statements (with notes in parentheses)
18:03	Trip: I think it's over.
18:06	Grace: Trip, what?
18:08	P3: Trip wait! (moves in front of Trip and extends her hand to stop him)
18:10	Trip: It's over Grace, don't you see that?

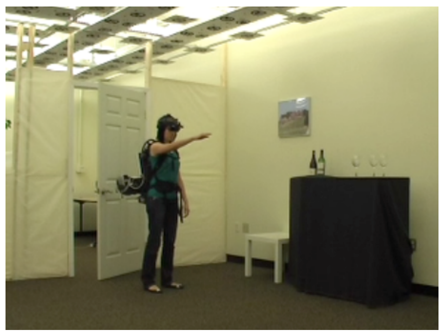


Figure 5.14: Player 3 quickly moving to get in front of Trip before he leaves

During the retrospective interview, the player explained her actions at the end, showing how she was involved in the dramatic moment and thought she might be able to effect the outcome:

> *"I was kind of using desperate measures… I didn't know what actions would have an effect so I thought I would try anything… I wanted to keep him there because I thought they could still talk about things."* *(P3)*

Player 29's exchange during the same narrative sequence is much different. In contrast to Player 3's urgency to mend the relationship, Player 29 acts as if Trip will do her harm and runs out the door towards Grace (see Table 5.5, Figure 5.15, and Figure 5.16). Her behavior demonstrates her desire to perform for her friends who were outside laughing at her actions.

Table 5.5: Player 29 exchange near the end of her episode

Time	Player and character statements (with notes in parentheses)
18:00	P29: Wait, Grace I'll go with you! (walks towards and reaches for Grace)
18:02	Grace: Thank you Jane, you really helped me.
18:04	P29: Ahhhh, don't leave me here with Trip! (running away from Trip)
18:08	P29: He's going to kill me!! He's mad! (Pointing at Trip)

Figure 5.15: Player 29 running towards Grace and shying away from Trip

Figure 5.16: P29 performing divergently towards the end

5.2.7 Behavior During the Severe Technical Errors

Technical disturbances are the other significant observable occurrence that influenced player behavior and revealed the level of engagement of the player. When major technical disturbances occur, do players look past them? Do they distract the player and cause them to wonder how they should be interacting, as was the case with Player 17 (see Figure 5.8)? Do players assume the errors are part of the experience, like the example in Section 4.5.3, where Player 2 interprets a registration error problem as Trip standing on the table to adjust a painting?

One of the most disturbing technical errors occurs when the AI-engine crashes, but the display system containing the character animation subsystem continues to run. The characters are still rendered in the video-see-through display, and continue blinking and looking at the player. However, they no longer react, move around, or talk. The experience is effectively over, but it is not immediately obvious to the player. On the few occasions where this error occurred, the player would not assume the action had ended and would carry on trying to get the characters to talk (see Table 5.6 and Figure 5.7). The following excerpt is from Player 18's exchange with Grace, starting just before the AI-engine crashes.

Table 5.6: Player 18 exchange after encountering an AI crash

Time	Player and character statements (with notes in parentheses)
9:00	Grace: We're not normally like this… (last statement from Grace before crash)
9:06	P18: It's Ok, I know. Every couple has their problems
9:12	P18: It's Ok... (comforts Grace)
9:20	P18: Everything's going to be Ok.... Trip's just in one of those moods...
9:30	P18: You know how he can get... he's always been like that....
9:42	P18: Ok, do you need a hug? (hugs Grace)
9:51	P18: Is everything ok?
9:56	P18:(Yelling loudly towards the back room) Trip?
10:16	P18: Please talk... (smiling)

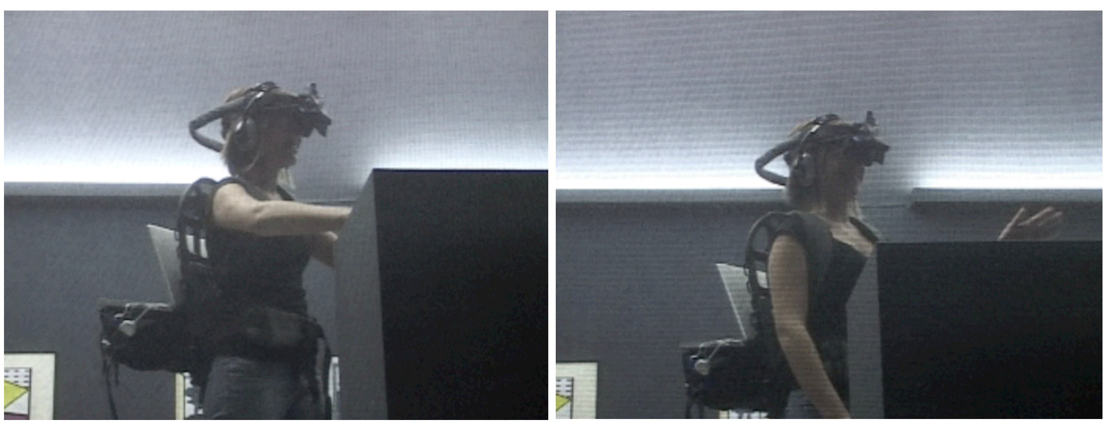


Figure 5.17: Player 18 hugging Grace after the AI engine crashes

It is clear from this exchange that the player was really into drama. She stayed in the apartment for nearly 90 seconds after the last statement from Grace. She physically interacts with Grace and tells her it will be okay. She calls for Trip and even tells Grace to "please talk." Later during the interview, Player 18 reflected on this moment saying: "...when it stopped at the very end... I was like 'oh – I can't do anything to bring them back'. But I didn't know if he was gonna listen to me" (P18). The AI engine also crashed on Player 3 during her first AR episode (before restarting and trying again) and she had a very similar exchange where the silence really drew her in.

While such technical breakdowns can be viewed as errors, they can also be seen as "breaching experiments" (Garfinkel, 1967). They reveal how players currently feel. The errors encourage recovery behaviors, as players struggle to figure out how to make the interface work. Their reflections on those moments afterward expose their knowledge, or lack of knowledge, about how the system works. Such "breaks" have actually been explored as a research method by Slater et al. for measuring presence (2003). The situation is not unrelated to Garfinkel's ethnomethodological practice of violating social situations to shed light on how people construct social reality (1967). In that sense, the awkward social scenario simulated in *AR Façade*'s is also a breach that reveals how people make sense of, and deal with, such situations.

5.3 Styles of Play (Case Studies)

In the previous section, I introduced the range of player behaviors exhibited in *AR Façade* across both installations, and I sought to visually code and analyze the episodes. In this section, I identify the most salient "styles" of play. Based on my analysis of player actions and their interpretations offered during the interviews, I define five styles of play and exemplify each through a player case study. The visualizations also helped identify patterns and confirmed for me that these should be referred to as play styles, not player types, because any individual player can exhibit one style at one moment, and then switch to another style the next moment. Therefore, the five cases studies are only meant to serve as examples of the emergent styles of play. In an attempt to leave gender biases out of the discussion at this point, all five case studies feature female players. In each case study, I present the episode visualization along with images, excerpts, and post-reflections from the player. An understanding of the play styles not only underscores the fundamental differences in people, it points to opportunities for future design.

5.3.1 Engager

The first play style that emerges in immersive and interactive stories is where the player fully engages with the experience physically, socially, and emotionally. It is when players exhibit this style of play that they come closest to the idealized notion of embodied narrative engagement defined in Chapter 2. The "engaged" player accepts the illusion of being in an apartment with fictitious people and able to influence the course of events through their actions. Players are also emotionally involved in the drama, invested in the characters, and curious about the plot. When the players engages, they take on the role that has been scripted for them in the plot without cynicism or mockery.

In *AR Façade*, when the "engager" play style surfaced, players exhibited emotions on their face, as well as through their actions and words. To illustrate this style of play, I present a case study of Player 32, who epitomizes the emotions that can arise in such a dramatic situation. Table 5.7 provides a sample of her interactions through transcripts, images, the episode visualization, and her comments during the interview.

Table 5.7: Player 32 exchanges exemplifying the engager play style

Time	Player and character statements (with notes in parentheses)
5:15	Trip: You're never happy, you're never ever satisfied!!!
5:20	P32: Maybe I should go? (bites lip, looks uncomfortable)
5:36	Grace: I'm stifled Trip!.... artistically… (P32 puts her hand to mouth in disbelief)
...	
7:14	P32: Maybe you should go talk to her... (points with both hands)
7:23	Trip: Chris, you are saying I am not communicative? (P32 takes a step back, and bites her lip)
...	
10:15	Grace: Chris, you blame me for this don't you?
10:18	P32: Nooo, I don't blame you for this at all!! (jaw drops, reacts emotionally)

Time	Player and character statements (with notes in parentheses)
...	
13:10	Grace: I've been paying close attention to what you've been saying tonight.
13:15	Grace: You've really been pushing me.
13:18	P32: (jaw drops) I'm sorry, I didn't mean to push you....
...	
15:45	Grace: Chris, I'm sorry. Thank you though. You really helped us...
15:47	Trip: Grace! (P32 walks quickly after Grace as if she will try to stop her...)
15:50	P32: I'm sure that's not all you can say... (towards the back of Grace)
15:58	Trip: Jesus Christ.... I should have told her...I could have told her.
16:08	P32: It's OK. You could still tell her. She's just outside. (pointing towards the door)

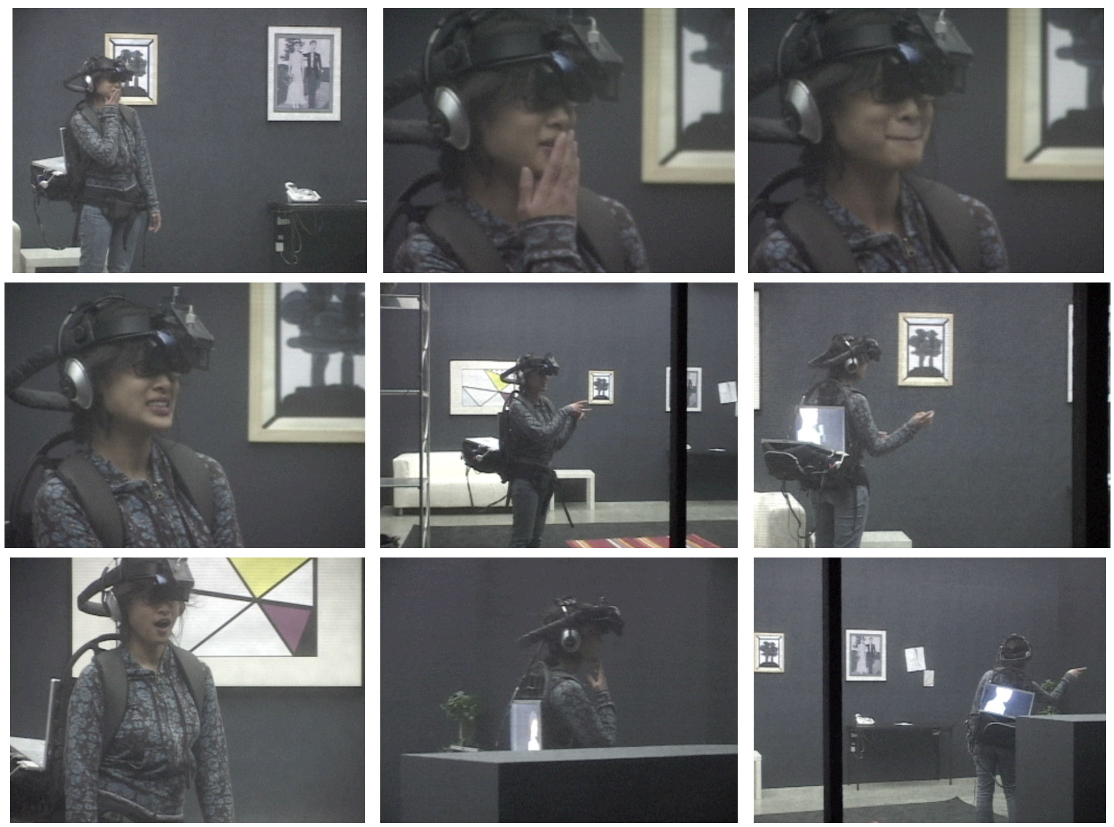


Figure 5.18: Images of Player 32 exhibiting a "engager" style of play

Figure 5.19: Visualization for engager Player 32

During Player 32's episode, she first tries to leave before getting pulled in emotionally. She does try to help the characters throughout, even running after Grace as she leaves (see Figure 5.18). Player 32 was paying close attention to the characters and responding as if she was actually stuck in an uncomfortable setting:

> *"Especially when Grace said like, 'You've been saying this and that and this and that, and is there any meaning behind that?' I don't know, it just kinda put me on the spot, very much so.... I kinda felt caught in the middle (chuckle) between both of them." (P32)*

Even more telling is her tendency to take comments seriously from Trip and Grace. She appears to feel guilt at times, even apologizing when Grace says she's been pushing her. During the post interview Player 32 said this:

> *"I didn't think they'd actually be like really paying attention to what I said. Like it wasn't just I was saying something and they were reacting to it; it was like they actually sort of had emotional reactions to what I said." (P32)*

Not only did she feel a sense of her own agency, she felt like the characters were emotionally reacting to her. She fully accepted the illusion of an encounter with old friends, although as she expressed later, she did not feel Trip and Grace were necessarily acting appropriately:

> *"I hadn't seen them in a long time, right? Right, so if I hadn't seen them in a long time and all of sudden they're fighting and I don't really know what to say." (P32)*

The whole situation caught her off-guard. Her reactions to the fighting seemed authentic; her boyfriend watching from outside even commented that she acted exactly how she would in that situation. She was not trying to be someone else. She was not reacting to the novelty of the medium, as much as directly reacting with the simulated social situation. Besides Player 32, the players in Table 5.8 also exhibited the "engager" style of play.

Table 5.8: Summary of other players who exhibited an "engager" style of play

Player ID	Highlights from gameplay	Key statements from interviews
P1	Says "come on guys, I came here to catch up, not to fight"; Chases Grace towards the door as she leaves	"I felt like it was so real. It was unexpected, yet logical..."
P3	Tries to stop Trip from leaving at the end by standing between him and the door. (see Section 5.2.4)	"I liked the story, it was believable. I was emotionally caught up, just like I would if I was experiencing it in real life."
P8	Tries to politely excuse herself for a couple minutes before leaving. (see Section 5.2.4)	"I felt like if they need to talk about their problems, they should do that after I leave… I don't have to be here....so I was upset."
P12	Strong reaction to Trip's misunderstanding of her statement "No, no... you do [love her]"; moved quickly to avoid Grace stomping out of the room	"I was trying to get out of her way. I was like whooaa, she's pissed! That felt really real... like she was going to brush my shoulder."
P14	Effective exchange with Trip and Grace during the therapy game. "No, I really think therapy will help. Really."	"Honestly it seemed as if both of them were real... How could you like honestly like – even if I'm a friend – how can you do that?"
P18	Comforts Grace near the end: "Trip is just in one of those moods..."; also continues to interact with Grace for several minutes after an AI bug. (see Section 5.2.5)	"Uncomfortable. It made me feel more that I wanted to help them – I couldn't fix their marriage – but to make things pleasant for both of them."
P25	Gives a long explanation of her idea of marriage to the characters. Seems to be getting upset when Trip and Grace do not listen.	"When you go in there your role or whatever is as a mediator ... the ideal thing would be to solve the problem."
P26	Picks up on the marital problems right away. Longwindedly, tells Trip and Grace what she thinks are their problems. Lots of emotional responses (e.g. hand over heart).	"Well the whole time I didn't want to abandon them, but it really didn't feel like it was any of my business either, because ... you do pick sides and you hurt feelings."
P28	Gets defensive at times: "Why are you getting mad at me? I didn't say anything about you!"	"I kinda felt like that in-between guy. It felt like they were just getting mad at me."
P35	Offers to leave a few times. Tries to back out politely. Turns back when Trip uses her name.	"I kept trying to get out, but I didn't wanna be rude... They kept like dragging me in..."

5.3.2 Performer

A second style of play happened when participants play off of the absurdity of the situation represented in the simulation. When the players perform, they "riff" on the characters and come up with actions and dialogue meant to mock the scenario. Unlike the engager style of play, performers do not take the situation seriously and seek to disrupt and make fun of it through situational humor. When a player performs, they are usually very physical and vocal, but they do not necessarily believe their actions will impact the situation. They want the attention to be on them, not the characters or the plot. The experience is treated like a stage performance.

In *AR Façade*, the performer style of play is gratified by funny reactions from characters, such as when Trip becomes bashful and uppity if the player talks about sex. To illustrate the performer style of play, I present a case study of Player 16, who epitomizes the absurd, divergent things players tried, especially to entertain their friends watching from outside. Table 5.9 provides a sample of her interactions through transcripts, images, the episode visualization, and her comments during the interview.

Table 5.9: Player 16 exchanges exemplifying the performer play style

Time	Player and character statements (with notes in parentheses)
0:50	Grace: It's been so long since we've seen you.
0:52	P16: I knowww! (Goes in for a big hug)
…	
4:20	P16: Can you hurry up and make my drink?
4:22	Trip: Trying to get onto Grace's good side tonight.
4:24	(Player starts throwing punches at Trip)
…	
5:56	P16: I'm on Grace's side (Player does the "finger snap" towards Trip)
…	
6:55	Grace: You keeping trying to tell me you love me Trip, that you're so romantic.
6:57	Trip: Oh God!
6:58	P16: You guys sound like my parents! Shut up!! (big laugh from her friends)
…	
8:35	P16: Trip I love you! I love you!!! (gets down on one knee with hands out)

Figure 5.20: Images of Player 16 exhibiting a "performer" style of play

Figure 5.21: Visualization for performer P16: showing she tried to act appropriately for a few minutes before completely devolving into divergent play

Player 16 came right in and gave Grace a big hug (recall Figure 5.7) and then after a minute or two of playing along to see where the story would go, nearly all of her actions were meant to be goofy and ridiculous. Player 16 used the situation to make jokes for her friends and, as one of her friends explained afterwards, her performance seemed to have the desired effect:

> *"It was cool – like you saw them hugging and they're like hugging air and it's like, yeah. (Laughter) And then like [player's name] was kissing him like crazy. (Laughter) ... like, you are like crazy! Yeah, it was funny." (P16)*

The performer style of play probably emerged the least often, and usually only when there were people there to watch. I witnessed a number of groups tour the gallery—students during class trips, conference attendees, etc.—and usually the bravest of the group would enter the experience with all eyes on her. The presence of an audience definitely impacted how players acted, but it did not necessarily always extract this behavior (for example, P32 had her boyfriend watching). The performer style of play did not happen very often if no audience was present, although as Player 4 justifies: "some of the things I did simply for my own amusement" (P4). Besides Player 16, the players in Table 5.10 also exhibited the "performer" style of play.

Table 5.10: Summary of other players who exhibited a "performer" style of play

Player ID	Highlights from Gameplay	Key statements from Interviews
P4	Tries to force each character to the middle of the room to talk it out face to face. He says "Can I just drag you?" while laughing.	"Some of the things I did simply for my amusement… I didn't think they would catch, but I didn't let that stop me… I was trying to be difficult. They refused to let me cause that much trouble…"
P15	Uses overly dramatic gestures. Runs over to Trip, extends her arms, and then overly kisses Trip (ends up getting kicked out due to her antics.)	"It was really fun.'Cause they were annoying us" (Laughter from group) They have some problems!"
P29	When Trip admits his affair, she yells and points at Trip "It is your fault!" then at the very end she runs away from Trip screaming "he's going to kill me!" (see figure 5.16)	"It felt like I walked into my friend's house, where her and her husband are about to kill each other... it was kind of funny."
P39	After Trip says something about therapy the player says "Everyone I know has a therapist!" and her friends outside start cracking up. She seemed to enjoy the fight, laughing at the dramatic moments.	"I don't need to be in this situation (laughter)... they're not my close friends that I haven't seen in a while."
P45	Tells Trip to "Sit down!" and then tries to grab him. At the end, when Trip says it's over he says "It's not over. It's not over!", while his friends laughed.	"I like how you can ... (reach out to signify touching the characters)... but I tried not to hug her 'cause I always figured that was going to be a problem."

5.3.3 Tinkerer

A third style of play occurred when participants spend time figuring out the system limitations, fidgeting with the technology, and doing things that are not necessarily related to the characters or the story. Like the performer, a tinkerer pays less attention to the story, but rather fixates on the medium itself. The tinkerer style of player emerges because a player is curious about what they can do and what they might need to do in order to "win" or master the system. They enjoy the physicality of the experience, but they do not feel their interactions will actually matter. They tend to remain outside of the drama and experiment with the interaction mechanisms, taking nothing for granted.

In *AR Façade*, the tinkerer treated the interaction less like an actual social situation and more like a game that they should be able to manipulate. The characters seemed like cardboard toys that they can poke and manipulate. They tried to figure out the "keywords" so they can "see what kind of reactions" (P6) they can get. Player often enjoyed messing with the graphics more than the listening to the fight between Trip and Grace. To illustrate the tinkerer style of play, I present a case study for Player 21. Table 5.11 provides a sample of her interactions through transcripts, images, the episode visualization, and her comments during the interview.

Table 5.11: Player 21 exchanges exemplifying the tinkerer play style

Time	Player and character statements (with notes in parentheses)
0:47	P21: How are you doing?
0:49	Grace: No, we're fine. We're fine. Everything's fine.
0:53	P21: How are YOU doing? (Very clearly, with emphasis on YOU)
...	
4:30	(Player laughs when she sees her hand on the screen. Plays with the graphics.)
...	
5:30	P21: (Plays with graphics again...) This is like a soap opera. (laughs)
5:38	Grace: I don't want to look at you. Dammit Trip! (Player laughs)
5:40	P21: This is like Grey's Anatomy actually.

Time	Player and character statements (with notes in parentheses)
...	
8:45	P21: Maybe I should just poke you. (Starts poking at Trip)
8:53	Grace: Trip, you and my parents are always... (P21 still poking in the air)
8:54	P21: Poke, poke... (haha)
...	
10:15	Grace: Therapy? You think we need therapy?
10:20	P21: Ohh! Yes! Cool, you understood me. Yes....you.. do.. need.. therapy...

Figure 5.22: Images of Player 21 exhibiting a “tinkerer” style of play

P21


Figure 5.23: Visualization for tinkerer P21

Player 21 enjoyed jabbing the characters (see Figure 5.22) and spoke to Trip and Grace very deliberately and slowly. As she explained afterward, she wanted to reach the winning resolution, so her strategy was to talk to Trip and Grace, not like people, but like some kind of robots.

> *“I did want there to be a resolution. And I was trying to find ways to do that, but I wasn’t having much success. I tried to calm down, relax, things like simple words.” (P21)*

For Player 21, her sense of agency was only affirmed when the interaction mechanism is very clear, like yes/no questions.

"I really hoped that they would ask me more interactive questions, like yes or no .....like a video game... Whatever you say will direct the rest of the game." (P21)

Player 21 did most of her tinkering in Phase 2 (see the episode visualization in Figure 5.23); she may have been getting bored, or perhaps she felt deflated from the technical errors and the characters' failure to respond. Also during that period—as the transcript above illustrates—she tended to make a lot of side-comments to herself because she had no faith that her meta-talk would be understood by the characters:

"I had asked them some pretty simple questions and they couldn't understand me. I knew that something as complicated as 'this is like Days of Our Lives' or 'this is like 'Gray's Anatomy' – that would go way over their head." (P21)

The "tinkerer" style of play often arose when players did not get the interaction feedback they expected. Their experimentation not only fulfilled their curiosity about the novel medium, it served as a process of seeking out more concrete interaction mechanisms. Besides Player 21, the players in Table 5.12 also exhibited the "tinkerer" style of play.

Table 5.12: Summary of other players who exhibited a "tinkerer" style of play

Player ID	Highlights from Gameplay	Key Statements from Interviews
P6	Talks very deliberately and slowly. Repeating statements several times. Says "How are you?" 5 times in a row.	"I just felt like I was getting bored talking to myself… I was just throwing curve-balls to see if I could get anything different. I am just trying different things to see what kind of reactions I can get them to say and do..."
P11	Spends a lot of time trying to pick it up the eight-ball,...then played with the floating 8-ball, making it intersect with Trip and Grace.	"I didn't know how to go about grabbing...."; "I realized I had to be clear and concise. I couldn't talk to them like real people."
P13	Finds enjoyment in comforting... devolves into obsessively playing with the graphics.	"At first...I wasn't sure if I was doing it right. But then... I was kind of getting into the game. I could have done a better job."
P17	Suffers through several awkward exchanges, including the first five minutes or so as he tries to feel out the environment. Uses the pause mode to poke the characters. Says "The picture. You can talk about the picture."	"You try to assimilate what's going on, but at the same time there isn't that instant reaction you get from people."; "It just felt like the computer just knew my name, it would just squeeze into the script..."
P30	Picks up the phone. Plays with a trinket. Laughs while comforting Grace. Plays with hand in front of the HMD.	"I guess I didn't really have a big influence ... I couldn't like social engineer them into like thinking certain things...."
P33	Frequent meta-talk: "Can I talk to them? Can I do this?" Takes out his phone at one point to check the time. Never gets his bearings.	"I have enough problems in my life, ...I mean...it's just like sort of, you know, 'Why do I care?' ...this kind of imaginary figure."
P37	Asks Grace if she is able to hug. Lots of edge feeling throughout. Talks to herself.	"So what should I say and what should I do? What would the character say and what would the character do...?

5.3.4 Observer

A fourth style of play is evident in what the player does not do, rather than what she does do. The observer style of play emerges when a player stops interacting (or speaks and uses gestures infrequently), but continues to watch the story unfold. When a player observes they are involved in the story lines and characters, but they do not necessarily feel like an active participant. Like the tinkerer style of play, the player does not know what to do, so instead decides to do nothing. They remove themselves from the social situation and passively absorb the scene, much like a film.

In *AR Façade*, observers paid close attention to the story and became emotionally invested in the characters, but they did not converse much with the characters. When players observed, they seemed to stand away from the characters and just let the drama unfold. Observers were different from engagers who might choose "listening" as a strategy for interaction. Many of the observers would start out trying to interact, but as they failed to get the responses they expected and as the fighting intensified between Trip and Grace, they would become hesitant and perhaps only respond to yes/no questions, if at all. To illustrate the observer style of play, I present a case study for Player 19 (see Table 5.13).

Table 5.13: Player 19 exchanges exemplifying the observer play style

Time	Player and character statements (with notes in parentheses)
6:10	Grace: Emily, yes or no.
6:13	Grace: Do you think its wrong for one person in a relationship to listen too much to the other?
6:20	Trip: What?
6:21	P19: No.
6:22	Grace: To trust your husband or wife too much to— what?
6:28	Grace: Oh alright. Goddamn it...how can I be happy when you act this way Trip? ...

Figure 5.24: Images of Player 19 exhibiting an "observer" style of play

Figure 5.25: Visualization of observer P19 shows she interacted less as the experience went on, and not at all during Phase 3

Player 19 went from talking quite a bit with the characters (see Figure 5.25), but after about four minutes, the more she tried to interact, the more she decided to take a more passive approach:

> *"I felt like I should have wanted to step back and watch... when they were just conversing between themselves—especially on separate sides of the room—I just wanted to watch the conversation." (P19)*

Player 19 may have started out as an engager, but her illusion of agency did not last very long. As Player 19 explains, she was "overseeing" the experience, more like a book or film. She was somehow removed from the conversation, but dramatically involved and curious where it would go.

> *"I knew it was a drama piece and everything, but at certain points I just wanted to see where the conversation went. Like I wasn't assuming they were interested in where I fit in the conversation, but I wanted to know where their conversation would lead to... it was almost like a book – like I was in a situation and I was kind of overseeing, even though I was supposed to be interacting with them." (P19)*

She was aware that she was supposed to be interacting, but did not feel compelled to do so. Her style of play was to observe and listen and absorb the story like a book. In fact, Player 19 paid close attention to the story lines and wanted to know what would happen with Grace:

> *"I wanted to see where that would go... most of it was focused on her and how she wanted to become an artist and how she was forced into*

advertising by Trip... I wanted to see what, in her view, was important to their relationship..." (P19)

For players who displayed the "observer" style of play, they became generally interested in the story, but did not actively participate in the player-character role. In addition to Player 19, the players in Table 5.14 also exhibited the "observer" style of play.

Table 5.14: Summary of other players who exhibited a "observer" style of play

Player ID	Highlights from Gameplay	Key Statements from Interviews
P9	Only responds to yes/no questions.	"I took a little more of an observer role in this one.. I guess it's different when you are actually expected to speak versus typing"
P20	Most of the time just looking between the two, not saying anything. One word dialogue.	"I had thought —'it's a computer, it's trying to read what I say'— and I didn't know how much, so I was hesitant. It provoked a lot of emotion in me..."
P31	Really just moves around and does not say anything, even to direct questioning. No response to yes/no questions. No reaction when Trip says "Lisa, are you ready to hear it? Say something!"	"I don't know why, like a part of me just was like kinda paralyzed with fear. The music really kinda - I really thought at one point like something really scary was gonna like happen... I just didn't want to talk. I kinda wanted to see like what they were doing and their reactions. I think I just felt weird talking"
P36	Spends most of his time just watching. He does say goodbye then waits at the door for a few minutes before just leaving without a final goodbye.	" I already know that there's no consequence. Like any attempt I was trying to do to detract from it, like they didn't even notice that I was trying to do that. Like they didn't even acknowledge that I was uncomfortable or that I was trying to change the subject"
P42	Transitions from tinkering with the medium to just hanging back and watching the characters fight.	"I wasn't sure like exactly when I was supposed to respond or how I was supposed to respond... they kind of had the argument without me having to interfere too much."

5.3.5 Partaker

The final style of play is a more nuanced version of the engager style. The partaker figures out the interaction mechanisms and follows along with the social situation to some extent, but they do not get as dramatically involved as the engager style.

They will speak and gesture throughout the experience (and appear visually similar to engagers in the information visualizations), except they keep an emotional distance that is apparent in their reactions. Partakers do not necessarily resonate with the content being represented, which is not to say they could not engage in a situation that is more interesting to them. Like all of these styles of play, the partaker style is open to interpretation and players might be interested one moment only to become disinterested the next.

For *AR Façade*, the difference between engagers and partakers was apparent in the player's reactions. If the player was smiling and laughing throughout, even if they were saying all the right things, then they exhibited more of a partaker style of play. If the player appear to be disturbed and emotionally "in-tune" with the characters, then I identify that as the "engager" style. To illustrate the partaker style of play, I present a case study for Player 43 (see Table 5.15).

Table 5.15: Player 43 exchanges exemplifying the partaker play style

Time	Player and character statements (with notes in parentheses)
1:43	Trip: Julie, remember it was exactly 10 years ago tonight you introduced us?
1:47	(Player puts hands out, big smile on her face)
1:48	Trip: Senior year of college....
1:50	(Player starts laughing loudly) P43: Ha, ha... Fabulous! Something to celebrate.
1:58	Trip: We really want to thank you for years and years of...
2:02	Grace: Pain...
2:04	Trip: Ah... eh... agony...
2:07	P43: Pain can be good (said with a smirk on her face and then she laughs).
2:16	Trip: Yeah... ah love... yeah.
3:25	(Player turns and walks to the door and gives a "timeout" signal, still smiling)

Figure 5.26: Images of Player 43 exhibiting a "partaker" style of play

Figure 5.27: Visualization of partaker P43 shows how she played for just over 3 minutes and then bailed

Player 43 saw the *AR Façade* experience as something novel and out of the ordinary, so she wanted to give it a try.

> *"I was kinda like this is a cool experience, let's check it out. It was like oh, how does this work. So I was curious..." (P43)*

She was open-minded, and didn't really know what to expect. She played along for a few minutes and then quickly picked up on the tenuous tone of the characters (see Figure 5.27). It was not something she wanted to take part in for very long.

> *"That's a disturbing situation (laughing) to come into, you know... and it's clearly a dysfunctional relationship. I found it hard to relate to." (P43)*

> *"You know you're walking into what is not gonna be a pleasant situation. You just wonder should I just turn and maybe like say 'I got lost',.. give 'em a call on the cell phone and say, 'I'm so sorry, I came down with something' (laughing)" (P43)*

As the excerpt above relays, Player 43 said things that you might say to old friends in that social situation, but she kept laughing at everything that happened. During the interview she explained her tendency to laugh:

> *"Well, it's – you know what, it's almost that nervous laughter because you're trying to figure out.... you know sometimes couples play off of it and it's more fun" (P43)*

Once Player 43 realized the Trip and Grace were not actually joking around, she left the experience by signaling for a "timeout" (see Figure 5.26). Not all of the players who exhibited the "partaker" style would leave quite so early, but many of them did not enjoy the experience because of the story situation. I am not including a table of example for partakers since there are not specific representative behaviors or reactions. The partaker style of play is the default mentality that occurs when players are not emotionally engaging the simulation, tinkering around with the interaction, performing for their friends, or observing the action from the sidelines. Partakers are generally less interested in the experience as indicated my Player 40's sentiment: "It's aggravating listening to people bicker unless you're one of the bickerers... it's like listening to a baby cry. If you're not the mother, you wanna strangle that baby" (P40).

5.4 Towards a Measure of Play Style

Thus far I have offered an assessment of play styles for immersive and interactive stories grounded in observational evidence from one media experience, *AR Façade*. I analyzed only observable, "face value" behaviors of participants without looking at the underlying episode statistics, story decisions, or other measurable features (such as physiological measures). The play styles are my interpretation of engagement through both installations of this particular experience. My primary reason for doing this qualitative assessment of play styles is to be able to make sense of the data I present in Chapter 6 and 7 on the effects of immersive interfaces. Players' behavior and statements about the immersive interface will be better understood within the context of how the player plays. For example, performers seem to thoroughly enjoy physically acting out, indicating a sense of presence, but their actions do not necessarily mean they felt a strong sense of agency.

The five play styles operate at a level above specific content related strategies and can potentially provide a framework for evaluating other immersive and interactive stories. Moreover, if the styles of play can be identified through in-game features, they have the potential to contribute to better player models for adaptive interactive narrative. If a system could determine these styles of play as they happen, it opens up possibilities for interactive stories to more actively play off the emotional state of players. In this section, I present some previous work that seeks to identify player types in other gaming environments and to model player behaviors towards a goal of more adaptive narrative. Then I discuss several trends that exist between the in-game episode statistics, player demographics, and the play styles.

5.4.1 Previous Work on Descriptive Models of Play Behavior and Player Modeling

There are two threads of research germane to my work: empirical studies of player behavior (either qualitative or quantitative) aimed at creating descriptive player models, and player modeling research, which builds on the tradition of computational user modeling in HCI, but leverages the player type research to specialize user models to game situations.

5.4.1.1 Descriptive player models

Many researchers investigating games and interactive experiences have communicated that players can have diverse, but equally valid approaches for engagement. In the gaming world a salient distinction is drawn between hardcore and casual gamers, although Juul and others have called for re-examining this narrow differentiation (2008). In Bartle's discussion of multi-user dungeons (MUDs) he identifies four types of players: killers, achievers, socializers, and explorers (1996). Laws performed a very similar analysis of role-playing games (RPGs) differentiating between six types: the power gamer, the butt-kicker, the tactician, the specialist, the method actor, and the storyteller (2001). Bateman and Boon performed a cluster analysis of gamers where they administered surveys, collected Myers-Briggs personality types[53], and labeled four primary types of gamers: conquerer, manager, wanderer, and participant (2005). Lazzaro also conducted in-depth players studies of contemporary video games, pointing out individual player differences (2004).

Descriptive models of players typically only make sense within a particular gaming context and have a lot to do with the interaction mechanisms available (e.g. the "killer" archetype may not be possible in a game that does not permit killing). We are also cautious about type-casting players into one category or another since most people have dynamic personalities. This chapter presents empirical research and a descriptive model of play styles for an immersive and interactive story. Although our description of play styles is not "operational", it may help designers to explicitly design interactive stories to provide satisfying interactions for the different player types. (Bartle's work is

53 http://www.myersbriggs.org/ (accessed 9/22/08)

extensively used in this regard by MUD and MMO designers, who work to make sure that MMOs offer game mechanics that satisfy Bartle's different player types). Moreover, our analysis could serve as the basis for creating computational player models for interactive storytelling environments.

5.4.1.2 Player modeling for games

In other gaming contexts, descriptive models of players have provided a starting point for adaptive storytelling. In Magerko's adaptive drama, Haunt 2, he models player behavior on Bartle's player types, and continually updates it based on game actions by the player (2006). Likewise, the PaSSAGE system created by Thue et al. attempts to model participant's style of play using Robin Laws' rules as the basis for the model (2008). Their system looks at key plot-points and the path players take through the story to determine weights for each player type. This automatically maintained player model is then used by the system to dynamically select story events. Their preliminary evaluation of the system showed that players felt the adaptive version of the story was more fun and provided more sense of agency than the non-adaptive version.

Other approaches to player modeling and adaptive storytelling have sought to model a player's emotional state without using prescribed "primitive" player types. Sharma et al. created a drama manager with an internally referencing player model based on four features of player interaction (2007). The primary finding of this investigation was that the average time spent by the player to perform game actions discriminates between gamers and non-gamers (Sharma et al., 2007). Other approaches include modeling users based on performance theory (el-Nasr, 2007) and using physiological sensor data to modify the behavior of virtual agents (Prendinger et al., 2005).

While *Façade*'s story architecture does not create a predictive player model, it does model patterns of player activity (e.g. whether the player systematically sides with one character over another). Existing player models developed for interactive stories and RPGs are not appropriate for *Façade*, where the player makes decisions at multiple levels of abstraction, from detailed social interaction, through the various social games (affinity game, therapy game, hot button game), up to major choice points (yes/no questions posed by the characters). Many player models assume a single level of player decision making, and a single progression (rather than the multiple simultaneous progressions active in *Façade*). In my discussion of future work in Chapter 8, I consider how a descriptive account of play styles could add nuance to *Façade*'s story architecture.

5.4.2 Understanding Play Styles with Quantitative Data

Looking at the game statistics from *AR Façade*, in light of the play styles I have identified, I can conduct a course analysis by clustering players together based on those five styles of play. As I have already acknowledged, these play styles are a subjective interpretation of observable behavior and reactions. They also do not accurately capture the whole of the player, since players change their focus of engagement over time. However, even this gross classification can be useful. In particular, classifying players into one of the five types, some interesting supportive quantitative patterns emerge (see Table 5.16).

Table 5.16: Clustering study participants (N=45) into one of the five play styles

Play Style	Players who most exude this style	Count
Engager	P1, P3, P8, P12, P14, P18, P25, P26, P28, P32, P35	11
Tinkerer	P6, P11, P13, P17, P21, P30, P33, P37	8
Performer	P4, P15, P16, P29, P39, P45	6
Observer	P9, P19, P20, P31, P36, P42	6
Partaker	P2, P5, P7, P10, P22, P23, P24, P27, P34, P38, P40, P41, P43, P44	14

In Section 5.3, for each style of play I included a table that lists similar play behaviors in the game episodes and comments from players to help explain each play style. Table 5.16 lists those players that most exude the various play styles. Participants are classified most often as partakers (14 players), followed by engagers (11 players).

Looking at patterns across those groups for average episode length and average amount of conversation per minute (see Table 5.17), some interesting points emerge. For player conversational patterns, the first important point is that the relative quantity of conversation between the play styles is consistent across the two installations, despite the two different wizard tasks described in Chapter 4. Second, performers and engagers were the most verbose players, while the observers were the least verbose.

Table 5.17: Play styles and episode statistics (averages/standard deviations) — discourses per minute (Beall), dialog per minute (Atlanta), and episode length (across all)

Play Style	Ave. discourses per min	Ave. dialog per min	Ave. episode length
Engager	2.52 (0.84)	4.66 (2.1)	15.0 (4.2)
Tinkerer	2.25 (0.81)	3.39 (2.0)	15.3 (2.5)
Performer	2.57 (0.46)	5.27 (only)	13.1 (6.2)
Observer	1.27 (0.49)	1.88 (only)	15.6 (2.3)
Partaker	2.44 (0.92)	3.21 (2.2)	13.4 (5.7)
All types	2.27 (0.85)	3.78 (2.0)	14.4 (4.5)

Despite talking very little, observers had the longest episodes (15.6 minutes, compared to the average of 14.4 minutes). The two statistics taken together provide evidence that observers are genuinely curious about the outcome of the experience even if they had no sense of personal agency.

The episode ending also served as a rough indicator of certain play styles (see Figure 5.28). Unsurprisingly, performers are the most likely to be kicked out by Trip and

Grace due to inappropriate behavior. Partakers are more likely to quit early— probably due to their disinterest in content—backing up the fact that they also played the shortest episodes. Interestingly, the players who perhaps did the least to guide the direction of the story—the observers—were the ones who managed to pull off the two "neutral" endings (the fake resolution and genuinely helping the characters).

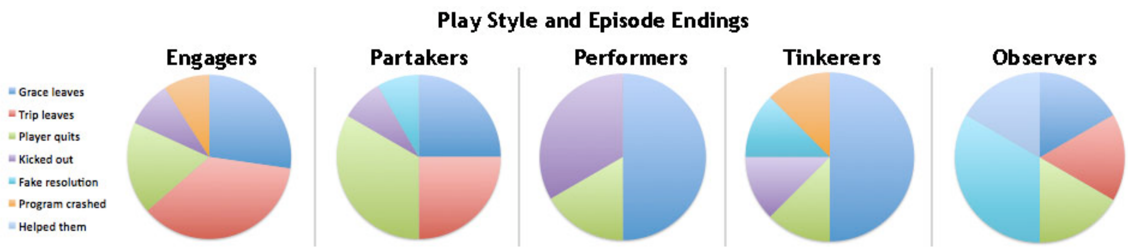


Figure 5.29: Endings across the five play styles for all participants (N=45)

Partakers conversed relatively little and quit early on, illustrating a distaste for the *AR Façade* content. As the partaker Player 24 stated:

> *"Initially I kind of wanted to see if I could help them try to fix it, but then after... just like back and forth ....I was thinking, 'Well, this isn't gonna go anywhere', so I'm just gonna end the game." (P24)*

Comparing the play styles to the player demographics for *AR Façade*, I discovered the partakers were the oldest participants (along with tinkerers) at an average age of 24.6 years, while the youngest participants turned into performers (who averaged 17.8 years old). Also, more men were partakers (8 men to 6 women) and more women were engagers (8 women to 3 men). In Chapter 7, I come back to this point and argue that the emotional tenor of *AR Façade* resonates more with women than men. Although these are only rough trends, they suggest that immersive and interactive stories can be appealing to certain demographics. In Chapter 8, based on my familiarity of the data collected for *AR Façade*, I reflect on potentially measurable features that could contribute to a continuous and automatically-updated player model for immersive and interactive stories.

5.5 Chapter Discussion

In this chapter, I presented a qualitative analysis of play styles based on interviews and episode data collected from participants across two installations of *AR Façade*. My mixed-method analysis included episode video coding and visualization, interviews, and quantitative data about players and their episodes. I suggested five prevalent styles of play illustrated through case-studies and supported through additional quantitive data. The five play styles operate at a level above specific content related strategies and can potentially provide a framework for evaluating other immersive and interactive stories. Moreover, if the styles of play can be identified through in-game features, they have the potential to contribute to better player models for adaptive interactive narrative. If a system could determine these styles of play as they happen, it opens up possibilities for interactive stories to more actively play off the emotional state of players.

To develop a more structured HCI evaluation method for immersive and interactive stories, one strategy would be to measure the component parts of embodied narrative engagement: presence, dramatic involvement, and agency. There are a number of issues with this, however. The “objective and quantitative” measurements of presence—the physiological sensors—do not exclusively measure the sense of being “in” a medium. The peaks and valleys of human physiological response have as much to do with “world” events—such as, authored content, social interaction, and situational phenomenon—than the medium in which these events are transmitted. Survey questionnaires might also attempt to measure embodied narrative engagement, but this also has problems. People may not be able to segment and rate their sense of the component parts—even if they understood what was meant by presence, agency, and involvement. An overall rating would be useful, but would obviously be tied closely to the participants enjoyment of the content.

In my analysis, embodied narrative engagement has different flavors, and thus an observer's experience be as valid and satisfying as an engagers. Since a stimulated mind can trigger involuntary responses in our bodies, an observer who is very dramatically involved in an immersive and interactive story may have physiological readings similar to an engager. Two players with different styles of play could report high overall ratings, but for different reasons. A more tractable question might be: can the style of play be measured or detected? Based on my observation of the data for the forty-five players in *AR Façade*, I believe some of these play style differentiations may be possible in run-time. For example, as I illustrate in Table 5.17, it appears that monitoring the conversational activity identifies the observer style of play in participants who do not converse much. The more talkative players are likely either engagers or performers. In Chapter 8, I discuss the open research questions and future research related to descriptive and operational models of play styles.

CHAPTER 6

THE EFFECTS OF IMMERSIVE INTERFACES ON AGENCY

At the seashore, between the land of atoms and the sea of bits, we are now facing the challenge of reconciling our dual citizenship in the physical and digital worlds. —Hiroshi Ishii

In this chapter, I return to the theoretical framework for embodied narrative engagement outlined in Chapter 2. I present evidence from two studies of *AR Façade*—the Beall Center exhibit and the Atlanta lab setting—to illustrate the effect of immersive interfaces on the psychological concepts of presence and agency (I address dramatic involvement in Chapter 7). Specifically, I argue that the less mediated AR interface induced a sense of presence, but failed to meet players' expectations for interactivity. Players felt a stronger sense of agency with the less immersive desktop interaction, because the interface presented appropriate constraints.

The evidence presented in this chapter shows that players felt perceptually immersed and connected to the story environment in *AR Façade*, reinforced by the physicality of their own body and the space. I demonstrate that most players experienced a sense of freedom and "naturalness" with their interactions in AR. Players could speak and move without having to think, but they were not satisfied with the interaction. Many players were expecting verbal communication to be on par with everyday conversation. I present contrasting examples of interaction strategies to illustrate how some players achieved a fluid conversation, while others ultimately failed to feel agency over the course of events. Finally, I show that players actually experienced greater agency in the desktop version of *Façade*, because the interaction better matched their expectations. While the desktop environment did not invoke feelings of presence, it did provide clearer affordances for what objects can be manipulated and how to verbally communicate.

6.1 Immersed and Unconstrained

In this section, I look at examples from both installations of *AR Façade* and the data collected from player interviews to demonstrate that players felt like they were perceptually immersed in the story environment. They felt free to do and say whatever came to mind. Players had many reasons to take action and there were few impediments to carrying out those actions. Participants acted "naturally" in the AR interface, even attempting gestures that fell outside of the prescribed set of actions (hug, comfort, kiss). The survey results from the comparative study with Players 1-12 showed that participants felt the AR interface was most like reality and at the same time the most challenging when compared to desktop interactions.

6.1.1 Perceptually Immersed

As I described in Chapter 3, the head-mounted display interface for *AR Façade* only covers a portion of the visual field, but players see the same environment unmediated in their periphery. There is no "interface"—the participant can only see the characters through the display and interact through verbal communication. Player 3 describes the phenomena of the immersive augmented reality interface:

> *"When you are sitting at the desktop you are conscious of that, but when you are walking around (in AR) your senses are taken over… and because*

> *you are immersed in the environment, there is really nothing to distract you on the outside. That's what makes it real." (P3)*

The participants did not perceive anything outside of the story environment. As I discussed in Sections 4.5.3 and 5.2.5, there were some technology distractions that appeared to bother some players, but most players saw past the imperfections. Compared to desktop interaction, the peripheral information in the visual field of the AR experience is relevant, as Player 36 says:

> *"On the desktop it's all framed. You have a very small screen and, you know, it's very easy to get distracted by background music, someone walking down the hall that isn't supposed to be in the setup. Whereas like in (AR), you try to fill my vision up. ... it seems more realistic. You know, it feels more like it's actually happening as opposed to like on the computer, I always feel like it's being relayed to me..." (P36)*

The environment presented in the video-see-through display was simply an extension of their actual environment, although as Player 17 pointed out "it wasn't like I could use my peripheral vision to know where the people were." Players not only talked about visual perception, but the physicality of the space and the ability to move around and physically connect with the stage. For Player 35:

> *"It adds to the feeling that you're actually in the area because, you know, you're feeling your body move and you're actually changing position. And there's actual furniture that you can touch and sit down on." (P35)*

The fact that players could see the physical apartment stage before entering the experience also reinforced their connectedness to the environment, as Player 27 describes:

"I obviously saw the room before I went in there, and it looked exactly like that... The backpack wasn't too heavy, so it just felt kind of natural moving around." (P27)

The effect of being perceptually immersed, of being in a physical space with the characters, allowed people to feel more a "part of the action" or in the game. As Player 6 related, "you feel part of things because you're walking around the room instead of just looking at the screen…It felt like you were in the action instead of just observing" (P6). Player 12 felt "less like an eye in the sky" (P12) contrasting her experience in AR with the view on the desktop.

Players felt a strong sense of presence in the story environment, although they did not necessarily describe the feeling as predicted by the virtual reality presence community. Reflecting back on Chapter 2 and the discussion on the theoretical differences between physical presence in VR versus AR, participants in *AR Façade* were not talking about being "there"; they were already in a real place. They would talk about feeling "conscious of my body" (P1, P7) or believing the experience more "because you are actually doing the actions and actually speaking…" (P9), supporting Biocca's findings of self presence in AR (1997). Several players talked about the experience of feeling "in it" (P12), more "connected" (P21, P30) or in a "realistic" environment (P4, P16). The language used by the players to describe the experience provides evidence of their sense of physical presence and self presence, but it is also evident from the players' actions.

6.1.2 Free of Constraints

Not only was the interaction "human" (P33), players had motivation for performing actions. As I describe in more detail in Chapter 7, the *AR Façade* experience provided a story environment with strong clues for players to decode their role. Most players could see the "writing on the wall" about where the scene was going and this provided enough impetus for what the player character might want to do. As Player 25 indicated, "when you go in there your role is as like a mediator" (P25).

The previous chapter demonstrated that not all players immediately fall into this counselor role, but most players had sufficient formal cause for action, a requisite for agency according to Mateas (2001). In fact, if players did not take action, the characters would provoke the players into action, as Player 37 related, "the absence of words would probably provoke them in a certain way." (P37) Players had motivation to perform actions and did not feel constrained to act how they wanted.

In *AR Façade*, players felt like they were part of the environment and were free to do what they wanted, as expressed by Player 21 "I tried to immerse myself in it and enjoy the freedom I had in it" (P21). Players did not need to make cognitive adjustments to understand how to move, they just moved their body:

> *"I didn't have to think about where I was going to be… If you want to touch your nose with your finger, you can just do it, you don't have to think about it… it was more like me doing the action… like me stepping in front of him was just natural.. I didn't have to think about it at all… the space felt like it was to scale and everything." (P3)*

Player 3 said she "didn't have to think" (P3) when she wanted to stand between Trip an the door as he was about to leave (see Figure 5.14). Player 1 describes the movement afforded by AR as "much more continuous" and "much better than just moving the mouse" (P1). Player 12 enjoyed the possibility of "touching things…" (P12). Player 8 also thought the freedom of interaction was fun:

"...it's a little more fun to be in the action. You can walk around at sit down, look at the things… It was just a little more real I guess… having the physical… able to walk and move and get closer to objects… and to do what you want to do." (P8)

The players felt free to do what they wanted to do and they didn't have to think about it, they just did it. A couple of case studies will help to demonstrate players' feeling of freedom to interact physically and verbally, including Player 4's attempts to physically manipulate the characters, and Player 25's long verbal rants.

In the first example, Player 4 got frustrated with Trip and Grace during a fight sequence. The player said "Can I just drag you?" and tried to put his hand on Trip to pull him to the center of the room, and then he walked behind Grace and attempted to push her (see Figure 6.1). Later, during the retrospective interview Player 4 recalled that he thought "pushing them together would allow them to talk to each other instead of over the shoulder to each other" (P4). The lack of physical constraints supported Player 4's "performer" style of play. Player 4 made a comment about how he felt immersed at times, but that things would happen that reminded him that it wasn't real:

"In desktop it was consistently a 5 on the scale for realism where as in AR it was jumping between 0 and 10… so in desktop you were immersed, but it was the shallow end…" (P4)

His sense of presence perhaps impelled him to reach out to the characters, but the lack of feedback—neither physical haptic response from the virtual content or conversational reactions—hindered his enjoyment.

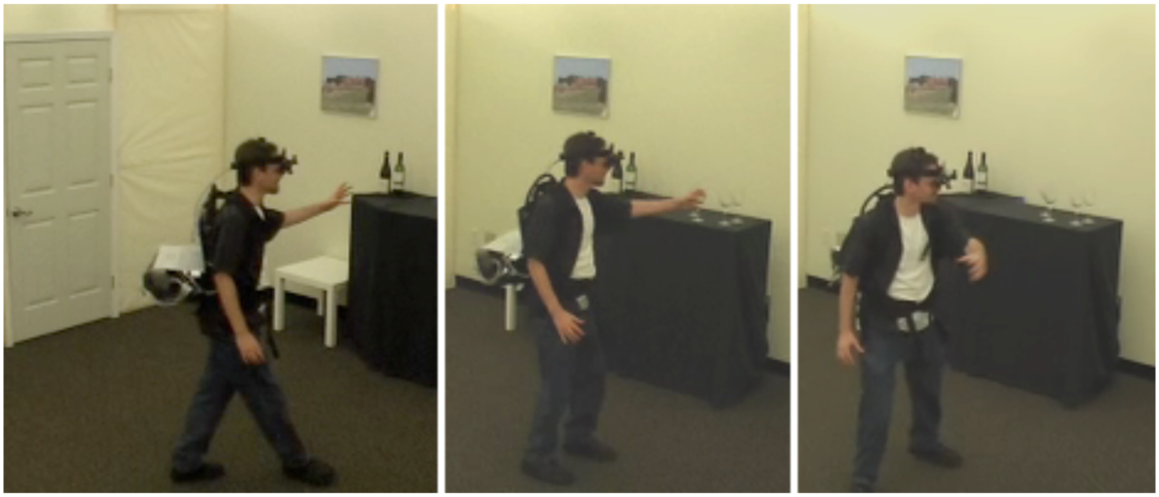


Figure 6.1: Player 4 trying to drag the characters together

In terms of verbal communication many players enjoyed being free of constraints, like Player 9 who said he felt “more freedom with what I said” (P9). Likewise, Player 6 who also had the opportunity to contrast the different interfaces stated, “(typing) isn’t my natural form of communication… speech is, so there is nothing to constrain your conversation” (P6). Another demonstration of players’ feeling of freedom to interact comes from Player 25 who did not recognize any constraints on her verbal input as she went on long diatribes about marriage on several occasions. In the following excerpt from her episode, she talks continuously for about fifteen seconds:

Table 6.1: Player 25 speaking a continuous 187-character-long utterance

Time	Player and character statements (with notes in parentheses)
6:35	Grace: Can you trust your husband or wife too much? To rely on them too much? Is that wrong? Yes or no?
6:40	P25: I think that it’s wrong to become overly dependent on someone. But, I think marriage is about two people being about to be together and being able to trust each other and help each other.
6:56	Grace: Oh that’s fine. That’s fine

Rather than taking the obvious route of answering the yes/no question, Player 25 offered her view of marriage, illustrating her “engager” style of play. She felt free to speak as long as she wanted within the conversational cues left by Trip and Grace, even speaking over the top of the characters at times. As I discuss further in Section 6.4.2, players did not face any hard constraints for what they could say, particularly at the Beall Center where players did not even see their own words appear. Player 25's statement was 187 text characters and 15 seconds long; in the traditional text-based dialog interface under *Façade*'s 35 character buffer limit, it would have required 5-6 long text entries (and much more time) to express her sentiment.

6.1.3 Realistic, but Challenging

Players felt immersed and unconstrained in the augmented reality version *Façade*, but how did this compare to the desktop versions of *Façade*? The survey results from the comparative study with twelve players provides a glimpse of how players compared the different interfaces described in Chapter 3: augmented reality (AR), speech-based interaction (SB), and keyboard-based interaction (KB). While 8 out of 12 players thought the AR interface was more like reality, an equal number believed it was the most challenging to learn (see Figure 6.2). Players ranked the KB interface easiest to learn. Players were virtually split on their opinion of which one was more engaging with six choosing AR and five choosing KB. The internal details of the survey indicate that players' preferred interface lined up with their choice on one of two other questions: players who were most engaged in AR thought AR was most realistic and players who were most engaged in KB thought KB was easiest to learn (see Appendix F). This shows that players probably based their decision for engagement on different criteria; some players enjoyed the novel interaction, while others enjoyed ease of use.

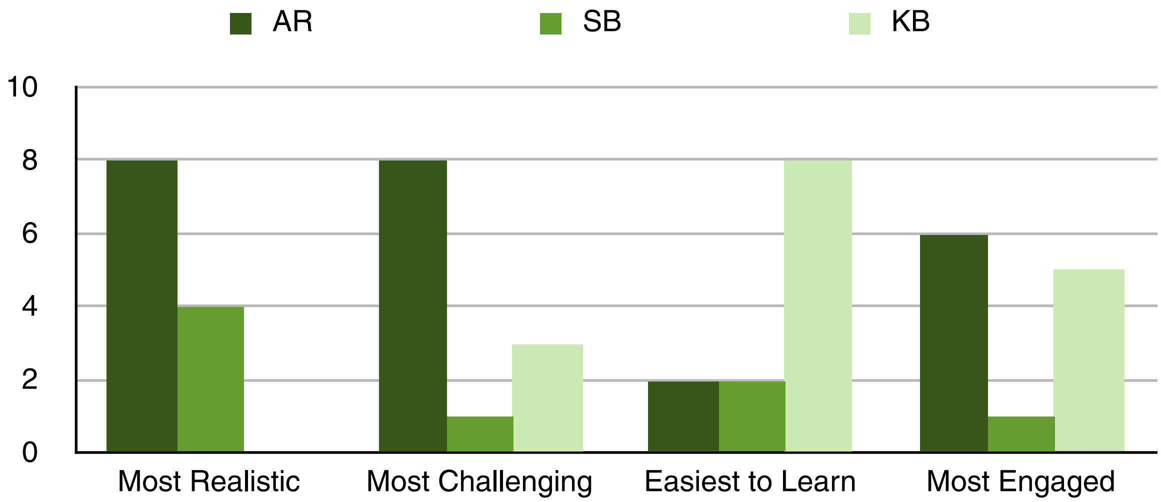


Figure 6.2: Survey results for comparative study (N=12)

The question I will look at next is why players view the AR interaction as both realistic and challenging at the same time. The immersive and unconstrained interface supports a sense of presence, but it appears to come at the cost of agency. Take Player 21's comment about connecting visually, but not verbally:

> *"In terms of the visuals – I felt connected to the world I was in completely.... But in terms of communication, that's where the barrier happened. I could imagine that they were there... but, I really couldn't communicate with them…" (P21)*

In the next section, I investigate the reasons why communication broke down and player agency suffered despite the "natural" and "realistic" interaction of the AR interface.

6.2 Diminished Agency

While the AR interface helped players feel part of the action, their feelings of agency appear to have suffered. Some players felt unlimited freedom to act, but still did not know where to start. "I felt sometimes that I didn't know what to do, but I felt that I could do a lot. I just didn't know what..." (P20). He felt limitless in what he could do, but was confused in understanding what he should do. Unlike Player 14 mentioned above, he suffers from lack of formal cause, of knowing what his player character should be doing. Most players, however, read the social cues and understood what they should do; they were just not successful in exerting the influence they wanted. In this section, I present examples of verbal and physical interaction problems with *AR Façade*.

6.2.1 Verbal Communication Problems

Evidence from both installations of *AR Façade* indicates that most players did not think the conversation with Trip and Grace felt like natural communication. Human spoken language is not without problems, considering the miscues that occur when communicating with a non-native speaker. Researchers who study linguistics and analyze conversation argue that people generally manage to co-construct mutual understanding despite language breakdowns (Schegloff, 1991). In *AR Façade*, players had little faith they would be able to overcome communication breakdowns. As Player 43 aptly stated, "you don't really necessarily feel like there's a flow" (P43). There were a lot of reasons flow did not occur.

The time delay issue discussed in Chapter 3 was one of the more consistent interaction failures, and participants such at Player 28 would make comments afterward, "a few times when they'd ask a question and I'd respond, their response to me would be a little wait" (P28). Player 23 noticed the delay on his physical gestures as well:

> *"There was some kind of delay between like when I would bring up something and when that reaction would happen. So, for example, if I patted them or hugged them on the back, there would be like a delay between when some reaction might happen from there." (P23)*

Players eventually picked up on the fact that there was a small delay before they received any response from characters, based both on their verbal and gestural inputs. The delay was a problem because it became difficult for players to know when they should talk. As Player 25 pointed out, “I was trying to respond to what they were saying, ... but it’s like you couldn’t get a word in.” It made players start to feel like their interactions did not matter:

> *“I would say something and then they would have to finish what they were saying. So I wouldn’t be able to interrupt them and it just seemed like they were on their own kind of agenda. So it was kinda pointless for me to be there in a way.” (P17)*

Players based their verbal interactions on normal human conversation and so many would try to wait for the characters to finish talking. The problem was that the breaks did not always happen. Player 3 was waiting for the characters to “open up some window where I could respond…, but (laughing..) that window never came…” (P3). Player 2 picked up on a subtle design feature in *Façade*:

> *“When I tried to talk when they are talking, it just disrupts the flow… it’s a little bit easier to wait for a break. Obviously I can say anything at any time… but it could stop them if I don’t wait.” (P2)*

When the player speaks, particularly in the first Phase of *Façade*’s architecture, the characters are programmed to stop talking and look at the players intently. Many players interpreted this as if they were interrupting the characters. As Player 4 explains, he did not want to talk over the top of the characters:

"At times it was kind of frustrating… especially the delay. Conversation is so much about timing… I would say something and then I would be interrupting… and we would end up talking over each other…" (P4)

Beyond time delay issues, communication breakdowns occurred in the conversational context. Sometimes the players expected a reaction from the characters, but got nothing. Player 34 said she "kept like saying things and it was like I didn't say anything at all" (P34). According to Player 17, it was as if the characters did not have ears:

"I didn't really feel ... like they had ears. It was just like... it went in through one ear and it went out the other, and they just had their own goals in mind." (P17)

Many of the players described the characters as being "on their own tangent" (P2) or "following a script" (P24). Player 3 felt as if Trip and Grace "would just kind of just continue with whatever they were saying." For Player 4, the repeated non-response to his actions and the characters' tendency to continue speaking led to a moment of realization:

"That's when it clicked… I am here to watch them fight… I am not getting out of this…I sort of saw my place in the script… the fight is going to occur no matter what.. you felt railroaded at that point…" (P4)

While some players theorized that the characters followed a script, others could not make sense of the characters reactions: "there were points where it didn't make sense..." (P25). Other players simply felt as if Trip and Grace were misinterpreting what they said. Player 21, for example, explained that the characters would make stuff up:

"It felt like they were making more assumptions about what I was saying..., they would allow that to make the scene go somewhere else, but I

didn't intend it to at all. Or they would make up something I had asked." (P21)

These mischaracterizations of player statements can be attributed to some extent to assumptions built into *Façade*'s architecture. My colleague and I published an analysis of player reactions against the AI responses and found that more violations occurred when the system made gross assumptions of player intentionality (Mehta et al., 2007). From many players' perspective, the characters would read something said by the player and then draw overly strong conclusions. Player 32 explained one of these situations:

"I had told Trip, you know, 'maybe you should go talk to her' ...'cause they're on like opposite sides of the room, and they thought I was saying they're not being communicative (Chuckle). So, and that's not like what I meant." (P32)

Sometimes the characters would understand the player's statement, but miss the nuance of what they were trying to say. As Player 22 related:

"I said 'no' and he just – it's sort of a broad yes or no but the situation was more complicated so it seemed like he took what he wanted to take from that no. It wasn't necessarily exactly what I meant." (P22)

What is perhaps most interesting is that players expected the characters to be able to communicate with the same nuance they do in everyday life. As Player 24 said, "I was trying to talk ... how I would just talk. I think this added to the toughness of trying to communicate with them." Players would look for subtle things in character's vocal patterns to match their expectations for interpersonal communication:

"You can hear beats in people's voices...like the way I'm talking with you. You don't talk right in the middle of when I'm saying something, unless you're intentionally interrupting." (P26)

The move to immerse players in speech-based interaction is a daunting challenge because of the subtleties of actual communication. Our first attempt to enable "natural" verbal communication was not without problems, but as I demonstrate in Section 6.3, players went to great extents to see past the problems and in a few cases players believed this illusion and felt as if they were actually having a conversation.

6.2.2 Physical Interaction Problems

In addition to the miscues for verbal communication, players suffered from the same sorts of uncertainty when performing physical actions. Player 27 described the difficulty of interacting with the virtual drinks:

"...I couldn't get the drink. I didn't really know how to do that, and then also the physical interactions. I wasn't exactly sure when I did it right or when I didn't." (P27)

As I discussed in Section 6.1, most players felt like they were in the room with the characters—some players even moved to avoid the characters as they came towards them. Player 4 felt as if he should be able to touch the characters, as he explained what it was like to hug: "If there is no body there, it doesn't feel right… like I should have felt them when I reached out my hand" (P4). Player 8 reflected on a moment when Grace crossed her path:

"I think I walked right through her… she just kind of disappeared… you would expect in real life for them to back up or still be there, but I walked

> *right through her… ha ha! Then, it was like 'where'd she go?' (scanning head back and forth)… 'is she behind me?…in the bedroom?' " (P8)*

Intersecting with characters was an unsettling experience for players, and again this can be attributed to high expectations established by perceptual immersion and naturalistic interaction. As Player 30 stated, he wanted the experience to take advantage of his physicality, he was "expecting a little bit more interaction with real world objects." Player 17 wanted the characters to be as real as the real space he was in:

> *"The actual space was so real with like the real wine glasses and the real sofas and the actual like shelves and artworks that was hung. I felt like the people could have been real too." (P17)*

The fact that players were paying attention to conversational beats and to their physical interaction points to the challenge of achieving the sense of presence and agency. Feeling a strong sense of presence has traditionally meant striving for a "transparent" and non-mediated interface, but the lack of clues about how to effectively interact caused problems for players in *AR Façade*.

6.3 Players Strategies for Achieving Agency

Many players struggled to maintain an illusion of agency despite feeling a strong sense of presence. In this section, I present evidence of the strategies developed by players to deal with the problems encountered with verbal communication and physical gesturing in the loosely-constrained immersive interface. I then present two case studies to illustrate the extreme differences in the sense of agency players felt. Player 6—whom I categorized as a tinkerer in Chapter 5—spent much of his time looking for the right "commands". In contrast, Player 14—whom I categorized as an engager—had long stretches of fluid conversation with Trip and Grace and ended the experience feeling like he impacted the story.

6.3.1 Interaction Strategies

As I revealed in Section 6.2, players encountered a number of interaction issues; this section shows how players tried to deal with some of those issues. I already mentioned one player strategy for dealing with the time delay (e.g., Player 3's attempt to wait for a "window"). Player 23 offered his timing strategy for contributing to the conversation:

> *"I was gonna ask them how long they were married or things, but ...I didn't know how in-depth I could get, and when I did talk, when I thought there would be a break in their conversation... So I kind of waited for like a really, really downtime, when they were both just moving, and then that's when I would try and speak, but I didn't find another one of those before it ended to talk to them more." (P23)*

Many players found the "waiting" strategy to be somewhat disappointing since the characters would just keep on talking. Player 9 said he would "really try to gauge" (P9) when the characters were about to finish talking so he could jump in with a statement. Player 28 talked about his strategy for anticipating the characters' reactions after a few seconds:

> *"Once I kind of got the drift of it, then I know when the response is coming. Well when he or she would say something to me and then I would respond, they'd continue going on about something else, and then after they went on about that then they'd go back to the response that I had." (P28)*

Some players resorted to overly simplifying their verbal communication. Many players tried speaking just Trip's or Grace's names as if that would help get their attention. Player 9 said he consciously used short and simple utterances, because he thought it would give him more agency, saying:

> *"I was training my self to be a little more simple with my speaking… less long sentences… or one or two word responses… I was a little more conscious of that." (P9)*

Player 19 also used simplified phrases like "I agree" and "Thank you", but he later said "I don't know if they understand this .... they don't know what I'm saying." (P19) Players wanted the characters to understand them. For Player 23, "The only thing I really got across that I thought they understood was, 'Are you okay?'" Player 31's comments were also informative considering he did not say much during the episode:

> *"...Should I choose the words that are most easy for a computer to detect or something like that? I didn't know what things I could say, like what things the computer would actually read. I knew of two things that I could do – the comfort and the hug – but I didn't know of anything else I could do." (P31)*

Player 31 talked about changing her style of communication to match the experience, but as an observer, she ended up saying very little. Some players even developed strategies around performing the gestures, like Player 30 who said he overly emphasized his gestures because he wasn't always sure if it worked:

> *"I had to like kind of ... exaggerate my movements. I was trying to comfort them and it wasn't working, I guess, so I'd try to like exaggerate it a little*

bit more and felt kinda silly doing that, (Chuckle) in my head there's like that, 'Oh, I have to make huge gestures.' But maybe that was just like what I thought I needed to do... I don't know." (P30)

Most players resigned to a nebulous sense of agency and hoped to only nudge the scenario, rather than have a fluid conversation. Player 30 revealed, "I could only direct it in a certain way and hope that they would kind of make revelations on their own" (P30). Similarly Player 20 said he tried "indicating to [Trip and Grace] to sit down together... but that didn't seem to work" (P20).

Generally players would start out speaking naturally and then fall back on to more succinct, specific verbal utterances in hopes that they could feel a sense of agency. As Player 2 aptly commented, "You don't talk that way in real life" (P2), but it was how players adjusted to the conversation patterns of *AR Façade*. The following case studies show how some players struggled to make these strategies did work, while other player experiences seemed to flow more smoothly.

6.3.1.1 Case Study #1: "Keyword" Communication

Player 6's game episode provides an interesting example of the sort of unnatural communication patterns that players invented to try to achieve a sense of agency. Table 6.2 presents several of the exchanges for Player 6 where he looks to exploit key phrases.

Table 6.2: Player 6 demonstrating his "keyword" strategy for communication

Time	Player and character statements (with notes in parentheses)
3:55	P6: How are you?
4:00	Grace: What? No, I'm fine, we're fine. Everything's fine.
4:06	P6: How are you?
4:12	Grace: What? Hey, no, there's nothing.. .you keep... everything's ok alright.
4:19	P6: How are you?
4:26	Grace: Bob! Uh (disgusted)
...	
12:40	P6: Are you OK? (towards Grace)
12:49	...Grace: Bob, look, I know you are trying to help, but right now we just need to get this all out in the open.
12:53	P6: Are you OK? (towards Grace)
12:49	Trip: Bob, stop it with that stuff!
13:10	P6: Are you OK? (walks right up into Trip's face)
13:20	Grace: Bob, I know what you are trying to hint at about me, about our marriage.

Player 6's game episode demonstrates how far some players went towards distorting conversation into a game-like interaction. As Player 6 stated afterwards, he tried to "keep it simple to see if I get more information" and so he tried a number of things:

> *"Maybe you have to have one-word statements for them to understand… Maybe if you use a question voice inflection you have to use a question word with it… I am just toying with the game at this point." (P6)*

He used voice inflection, one-word statements, repeated statements, and character name referencing. This particular episode excerpt does not do justice to the "tinkerer" style of play strategies used by Player 6 to try to get some sort of payoff. Player 6 felt no sense of agency, lamenting "I'm trying to desperately see, are they listening to anything" (P6). He tried everything he could think of and none of it had the effect he expected. According to the player "it seems like the more naturally I speak, the less they understand" (P6). Interestingly, as much as Player 6 deformed how he would normally speak, he said he still preferred speech-based interaction over typing:

> *"I am not crazy about typing. It isn't my natural form of communication… speech is. Anytime you have something in the middle it's going to constrain your conversation. With speech it's just air." (P6)*

Player 6 delighted in the fact that the speech interface allowed him to say whatever he wanted without constraints, but his behavior would indicate that he needed constraints and better feedback on what the story engine actually understood. Even though he had very little sense of agency, he said he was "most engaged in the AR version" because:

> *"Instead of just looking at the screen you can walk around… I would love to have that with Madden ™ (football game). If I had that helmet and I was the coach walking around in the living room looking out on the football game, that would be great." (P6)*

Player 6 is a hardcore gamer, as his comment about playing Madden™ football illustrates, and he really wanted to figure out how to win. Player 6's enthusiasm towards the AR interface can be attributed to the novelty of speaking and walking around, not because it gave him more control over the game scenario.

6.3.1.2 Case Study #2: Fluid Conversation

Player 14 provides a great example of how some players managed to have an ongoing conversation with the characters while getting wrapped up in the dramatic situation. Table 6.3 presents a number of exchanges from Player 14's eighteen minute long episode.

Table 6.3: Player 14 demonstrates "fluid" conversation

Time	Player and character statements (with notes in parentheses)
14:25	P14: Trip, relax... (hand out to comfort)
14:27	Trip: No, Mark, we just need to get this under control, that's what we need.
14:33	P14: Have you thought about therapy?
14:36	Trip: No, we need to talk about us both not just one.... ah ha!
14:40	Grace: Mark, you think we should see a therapist?.
14:44	P14: Maybe...
14:46	Trip: Grace, maybe a therapist could help us figure out why we always fight about visiting my parents.
14:52	P14: It's worth a try.
14:55	Trip: Ah, thank you. That helps us.
14:57	Grace: I don't know. I mean, therapy.
15:00	P14: It's not a bad thing.
...	(Grace and Trip start bickering back and forth…)
15:10	P14: It might be good to talk it out with a therapist. (player using hands forcefully)
...	(silence…)
15:19	P14: Really.
...	(Grace and Trip bicker back and forth again; player seems to be getting frustrated)
16:00	P14: Maybe you two should face each other when you are talking! (with hands out)
...	
16:15	P14: (loud outward sigh) Trip!
16:18	Trip: Mark, give me some advice here
16:20	P14: Relax...
16:24	Trip: Mark, I know what you are trying to hint at...
16:32	P14: What is that?
16:34	Trip: Do you really want the truth? Huh?
16:36	P14: Sure, tell me the truth.

Figure 6.3: Player 14 having a fluid conversation with Trip and Grace

Even though Player 14 does not always get the characters to understand him or do what he suggested, he carried on communicating with them as if he was really in that situation (see Figure 6.3). His actions indicate that he felt a physical sense of presence and afterward he also said “the reason this feels more real would be because you’re starting to introduce (physical) interactions ... you know, patting, hugging” (P14). Player 14’s actions also show that he was dramatically involved in the characters and their marital problems. He said in the interview that the simulation was “kinda like real life. I’ve seen people like this before in my life” (P14). He fits the “engager” profile I described in Chapter 5.

While Player 14’s conversation with Trip and Grace seemed to flow fairly well, one of the first things he talked about when he sat down for the interview was the known time-delay problem:

> *“One of the things immediately when I was there is there’s a lag while it’s trying to compute what you’re saying ... If you could minimize that lag, this would be like really –* ***would really feel real.****” (p14)*

I pressed Player 14 to explain how he reconciled the time delay and he offered an explanation that shows he learned how to anticipate what the characters would talk about. It was not only about timing his statements as much as getting into the head of the characters:

> *"The way I adjusted was – I mean, you sort of have to predict what the new thread's gonna be. It becomes you're trying to predict ... to try to figure out where it's going, which is what you do in real life, but you have to think a little bit faster than what you'd have to do in real life in order to make it work." (P14)*

Like many of the engagers, Player 14 was able to "make it work" because he was intimately involved in story lines and the drama surrounding these characters. He could relate to the issues they were facing, he had heard these arguments from people in his real life. All of this helped Player 14 "predict" what the characters would say and so when the characters actually said those things, he felt a strong sense of agency even if the ending did not turn out as he hoped.

6.3.2 The Effect of Style of Play on Agency and Overall Enjoyment

Reflecting on the play styles from Chapter 5, the rough play-style groups showed differences in the overall rating of *AR Façade*. Tinkerers like Player 6—who spent much of their time feeling for edges and struggling to achieve some sense of agency—gave lower ratings than both the performers and engagers. Tinkerers had the lowest overall rating (22.3 out of 35) and the second lowest interaction rating (3.1 out of 7), second only to observers who did not say much at all (see Figure 6.4).

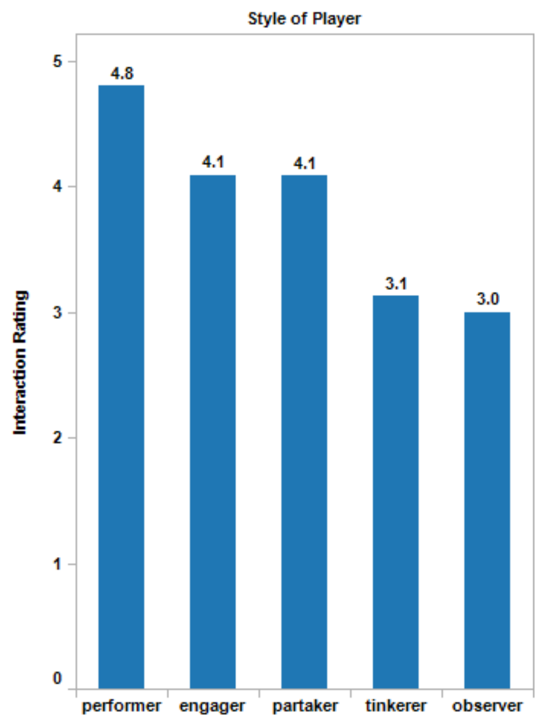


Figure 6.4: Interaction ratings across play styles (N=45) The difference between performers and observers was significant (T=2.49, p=.034), although the difference between performers and tinkerers was not (T=1.88, p=.087)

Like Player 14, most of the "engagers" genuinely felt a sense of agency, believing that their verbal and physical actions were being understood to some extent. Player 32 said "it was like they actually sort of had emotional reactions to what I said." Engagers tended to accept responsibility for interface gaffes, saying things like "I was probably asking too general a question" (P18). They were more forgiving of the characters, often attributing violations of agency to reasonable human error (e.g., "I wasn't sure if they heard what I was saying" (P3)) or to the fact that Trip and Grace were too wrapped up in their fight to pay attention to the player (e.g., "either she doesn't want to talk about it… or she's ignoring me…I don't know" (p8)).

Performers seemed to have a strong sense of agency and gave the experience a strong overall rating. Performers may have held lower expectations for agency because they were more focused on their own performances, rather than the characters' cues. The "partakers" felt a little bit of agency, but typically did not really care to get involved, as Player 2 points out "this is definitely about them and not about me" (P2). Observers often seemed too subdued to try things and thus felt a diminished sense of agency. Partakers and observers both gave the experience overall mediocre ratings, but not nearly as bad as the tinkerers. It seems that *AR Façade* failed to adequately support this style of play, perhaps because tinkerers like Player 6 were expecting more game-like interaction.

6.4 Trading Presence for Agency

The question becomes, how many of the problems with agency can be attributed to the immersive first-person AR interface to *Façade*? Many of the reasons players suffered a diminished sense of agency were also part of the traditional desktop *Façade*: script non-sequiturs, and characters failing to respond or misinterpreting the player. For players who had a strong sense of agency, as is the case with the engagers, it can be partly attributed to that individual's ability to believe the illusion. As Figure 6.2 indicates, players felt the keyboard-based (KB) desktop interaction was easiest to learn, but does this correlate to a stronger sense of agency using KB interaction? In this section, I use evidence from the Atlanta interface comparison study to help isolate the effect of the immersive interface. While the KB version of *Façade* suffers from its own usability issues, in contrast to the AR, it provided clearer affordances for interaction. Hence, it gave players a stronger sense of agency.

6.4.1 Clearer Affordances in KB

This section explores the results from the Atlanta installation where I conducted the interface comparison study with the first twelve participants. As I explained in Chapter 4, the speech-based desktop interaction was included to help us tease out the effect of speech alone (SB) vs. speech plus embodiment (AR). First, there were some relevant quantitative differences between the three versions.

Table 6.4: Player activity across three versions in the interface comparison study (N=12).[54]

	AR	SB	KB
Text input (average chars per min)	52.0	45.7	37.4
Average statement length (chars)	13.9	14.0	14.9
Text erased (average chars per min)	0.57	0.32	4.97
Hug, kiss or comforts (average per min)	0.16	0.35	0.24
Average episode length	17.4	18.3	20.6

According to my paired-samples T-test (summarized in Appendix F), there were significant differences between AR and KB for text input (T=2.328, p=0.04), between AR and KB for text erased (T=4.075, p=0.002), between SB and KB for text erased (T=4.511, p=0.001), between AR and SB for gestures used (T=2.216, p=0.051), and finally KB episodes were significantly longer than both the AR and SB episodes (T=3.56, p=0.004 and T=2.681, p=0.21).

Examining these statistics in terms of the conversational input, players entered much more text in the AR and SB versions than in the KB version (see Table 6.4). For the AR and SB versions, the wizard operator is typing the input through the remote interface. As I explained in Chapter 4, the wizard frequently missed words or changed what the player actually said to be able to keep up with player statements and to stay within the text buffer limitation. So the actual average characters entered per minute were likely higher in SB, and especially in AR where the wizard's attention was divided between two

54 Normalized values based on the episode length from each version of *Façade* during the interface comparison study (n=12). For AR and SB, the text input, text erased, and gestures input to the system are performed by a single wizard-of-oz operator (see Chapter 4 for a description of study setup).

tasks: speech and gesture recognition. The average statement length across all three versions was nearly equivalent since the wizards also had to abide by the text buffer limits, and thus would enter long utterances as two system entries if necessary.

The text erased statistic also reveals an important usage strategy for desktop-based typing dialog. Players erased 4.97 text characters per min on average (about 13% of their eventual text input). It was commonplace for players to type something and then erase it a moment later before pressing enter to commit their statement. For AR and SB, since the wizard was trying to match the players' statements as quickly and closely as possible, the text erased values amount to mistyping and error correcting. More that twice as many text characters were erased in AR than SB, again likely due to the more demanding wizard task.

In the AR version, the designated gestures (hug, kiss, and comfort) towards Trip and Grace were also entered by the wizard who probably missed some of the players actions due to poor visibility, etc. Looking at the difference in number of gestures entered by players in SB versus KB, it is not surprising more gestures were used in SB because players did not have to worry about typing as they navigated the space and used the mouse to interact with Trip and Grace.

Despite the limitations of typing, such as poor typing ability, spelling errors and the buffer limitation (criticized by some players, such as Player 9 who said "you couldn't type long... and it kept beeping at me, so I had to keep rewording my statements into something simpler" (P9)), players ranked the KB interface as the easiest to use (see Figure 6.2). Player 9's statement also partly explains why there were far more text characters erased in KB (see Table 6.4).

The speech-based interaction was particularly challenging as players started to recognize the delay between speaking and their words appearing on the HMD screen. As

Player 12 explained, "It was like my words had not caught up yet…like the computer had not yet generated what I said" (P12). Player 7 preferred to type because it gave her a more immediate response:

> *"It seemed more immediate if I typed something in that they would stop talking and…actually think about what I was saying." (P7)*

Moreover, typing statements out provided an opportunity to reflect and visually process statements before they were entered, as expressed by Player 2: "I am typing my words and I can see them, so it seems more concrete for some reason" (P2). The speech interaction required a greater degree of commitment, since it could not be undone. Some players found it hard to listen while speaking, "I was concentrating on listening to them while speaking, while typing I could still listen while I was typing" (P8). On the other hand, the speech interaction freed up players' ability to simultaneously move and verbally interact, summarized best by Player 11, "When I type I can ensure accuracy, but I cannot walk and talk at the same time" (P11). That also helps explain why players talked more in AR and SB, where they did not have to type. Other players were simply uncomfortable speaking to the computer: "I usually don't play games and talk to my computer… it's awkward... I'm comfortable with typing." (P8).

The desktop interaction also provided clearer affordances for physical interaction, especially for the designated gestures towards characters. As Player 4 says, the desktop interface was "easier to interact with because it was a little more traditional....you could tell when you are having an effect. You're not just waving your hands at someone..." (P4). For Player 11, AR was challenging, as she said, "I didn't have the mouse so I couldn't see the comfort/hug icons." (P11) Player 7 stated:

> *"Desktop is not as neat (as AR), but it was easier because if I click the mouse on something, like the hug.. that it would actually hug, where as in AR I would not know if it got picked up or not." (P7)*

The affordances in AR were not adequate. The strict "transparency" of the interface did not clearly communicate how the system operates. In AR, players encountered a litany of issues related to poor feedback for both speech and gesture. Players felt more agency in the KB version than they did in the AR version, because they felt that typing provided "more control" (P7) and that "(the characters) respond much better" (P1). Whether players preferred typing or speech, their rationale points to their perception of what is more natural for interaction with the story. As the survey results from the comparative study indicate (see Figure 6.2), some players were partial to the "freedom" (P9) of speech, while other players favored the dependability afforded by typing.

6.4.2 Feedback Changes in Speech-Based Interaction

Unlike the first twelve participants who took part in the Atlanta interface comparison study, the players at the Beall did not see their text appear on the HMD screen, as I describe in Chapter 3. The Atlanta players would see their verbal statements appear as text as the wizard typed. At the Beall Center, the only feedback seen by players was a small spinning wheel in the lower left-hand corner. The spinning wheel would appear and spin only while the wizard was searching for the appropriate discourse to select.

During the Atlanta installation, the text feedback had several effects: (1) players could see when their speech utterance was picked up by the system, (2) players could see a system limit in the number of words that could be interpreted at one time, and (3) players could see if the wizard mis-entered their statements. The first effect was actually useful for informing players about the time-delay issues, leading to some of the strategies explained in Section 6.3. In the second case, players began to notice the buffer limitation for how many words could be spoken. Player 8, for example, observed, "they don't understand really long sentences, so it is hard just to say like five words…" (P8). Thirdly, the wizards' mis-typed entries had a minimal effect on the player experience since the wizard did a decent job of typing player statements accurately. However, occasionally the wizard mistyped or entered something slightly different and the player would notice. For example, during the interviews Player 11 said:

> *"[Trip and Grace] interpreted it as if I suggested they needed help. Also the 'do' was missing ... I said 'do you need any help?' but the statement appearing on the screen was 'you need help'." (P11)*

The exact effect of having the text appear on screen is not clear since players were never able to compare the methods directly, but it did have some influence judging from player behavior in the Beall installation. Most players did not even notice the spinning wheel icon in the corner, but for those that did, it provided a subtle clue about the time-delay problem. For Player 14, commented on the feedback saying, "it helps out a little bit, ... it reminds you that there is a lag in the system." Anticipating the delay was important part of Player 14's conversational strategy, as I described in Section 6.3.1.2.

The lack of text appears to have removed the notion that their speech had to be limited (recall the example of Player 25's one statement of 187 text chars). These sort of unconstrained and long verbal interactions were possible since the wizard only had to select a corresponding discourse category rather than type in each spoken word (and operate under the 35 text char buffer limit). Under the new feedback design the only cue for how the system interprets their statements comes from how the character reacts. Many of the same agency problems occurred, but players were more likely to attribute the problems to their ambiguous notion of the system as a whole. When the player sees their text, they know that the system picked up exactly what they said. If the system still responds inappropriately, it becomes harder to blame it on speech recognition, for example.

I believe the "upgrade" we provided for the Beall installation of *AR Façade* had an overall positive effect on agency, but it still did not do enough to smooth over the problems with speech interaction (and perhaps even further highlighted the fact that the experience is not a real conversation). A number of other interface design ideas could be considered in future research including providing the list of discourses to the player directly, and allowing the player to enter or clear their verbal statements before pushing them into the system. As I discuss in Chapter 8, both of these ideas involve interface mediation and could take away from the "illusion of non-mediation" towards a sense of presence.

6.5 Chapter Discussion

In this chapter, I presented qualitative and quantitative evidence from two installations of *AR Façade* with a focus on the immersive interface, the interaction mechanisms, and the effect they had the play experience. I demonstrated that players had a strong sense of presence in the immersive interface, but that most players' sense of agency suffered, despite attempts to mentally adjust to the interaction. The desktop interaction provided a less immersive environment, but the affordances for dialogue and gestures towards the characters was much clearer. The data presented in this chapter reveals evidence of a tradeoff between unconstrained immersive interfaces that strive for presence and carefully-constrained interaction mechanisms that emphasize agency.

The novelty of AR as a medium may also be factoring into these results. While players in desktop *Façade* are likely to approach the interactive drama as something between video games and film, our players seem to relate *AR Façade* more to everyday life, perhaps setting overly high expectations for interactivity and player agency. If these expectations are not met, the player may feel disengaged. Current references to AR in popular media tend to imagine a seamless, undetectable integration of virtual content in the physical world, reinforcing this connection between AR and reality. Future work could explore the role of expectations on the user experience, and find ways to explicitly manage expectations for better game play.

Interestingly, the occurrences of strongest agency generally happened when *Façade*'s narrative structure spurred moments of clearly afforded interactivity. Yes-no questions are an example of a more obvious prompting for player interaction, as Player 20 reflected:

> *"I appreciated the times where I was led on to like yes or no or like here's my choice of answers. ...So that was cool and I appreciated those times where I was led in to a way I could have a big effect. But the times where it was just open and there was just talk, I didn't know how much it would [understand] so I was hesitant." (P20)*

While yes-no questions provide blatant hooks for when and how players could respond, and thus strengthened the sense of agency, they often did not provide enough nuance. As Player 3 mentioned "I didn't want to say yes or no… I wanted to have a whole conversation about it… I felt my answers had to be more concise." If the narrative structures are too rigid they risk making the player feel like the script and the story is too predetermined.

Other narrative design techniques in *Façade* were more subtle, like late in the experience after Trip and Grace have been fighting intensely, and one of them poses the question "Help us out here. What should we do?" Although the question is more open-ended than a yes-no query, the situational constraints would suggest a fairly limited number of responses (e.g. therapy, divorce, etc.), many of which are anticipated in the *Façade* script. As Player 13 related excitedly, "I mentioned therapy and they responded to that. That was pretty cool!" (P13). In general, players enjoyed the hard constraints of desktop interaction and narrative prompts (such as the yes-no questions), savoring moments of empowerment over the situation.

Achieving "natural" verbal interaction with virtual characters remains an open question for future research. While the concept of direct manipulation—essentially balancing "natural" movements with sensible constraints—makes sense for tangible interaction, it is not clear how the concept applies to speech interaction. In Chapter 8, I discuss some of the open questions posed in this chapter: For dialogue-based immersive and interactive stories, would "unnatural" interface mechanisms take away from a players' sense of presence? What interface designs are appropriate for designers striving for strong presence and strong agency? How do explicitly mediated constraints for verbal communication with virtual characters impact players' overall sense of embodied narrative engagement?

CHAPTER 7

THE EFFECTS OF IMMERSIVE INTERFACES ON DRAMATIC INVOLVEMENT

The birth of a new communication medium is both exhilarating and frightening. —Janet Murray, Hamlet on the Holodeck (p1)

Pretending that the action is real affords us the trill of fear; knowing that the action is pretend saves us from the pain of fear. —Brenda Laurel, Computers as Theatre (p113)

In this chapter, I provide evidence from two studies of *AR Façade*—at the Beall Center exhibit and in the Atlanta lab setting—to understand the effect of immersive interfaces on the psychological concepts of presence and dramatic involvement (agency is covered in Chapter 6). I argue that the combination of the first-person immersive AR interface, the second-person narrative voice, and the true-to-life social scenario simulated in *AR Façade* led some players to seek distance from the medium. The data shows that the desktop version of *Façade,* as well as other more mediated interfaces dramatic content, better supported players' desire for distance. Thus, the strategy of minimizing mediation to create the sense of presence did not maximize the overall sense of embodied narrative engagement, because it did not provide sufficient means for users to manage their distance from dramatic content.

My analysis will show that players generally understood and committed to the character role provided for them in the script. I provide evidence that the immersive interface (combined with *Façade*'s narrative structure) reinforced a transformation into a player-character. Many players treated the characters as believable social partners, acted within the dramatic moment, and exhibited raw emotions as the scenario played out. I present strategies sought out by players to give themselves more psychological distance from the simulation. Some players expressed preference for less immersive forms of media (e.g., the desktop version of *Façade* or TV/movies), specifically because it would better support their desire for managing their emotional distance from content.

7.1 Second-Person Narrative Voice and the Player-Character

Both the immersive interface and the second-person narrative voice in the script reinforce a player's transformation into a player-character role within *AR Façade*. The characters, Trip and Grace, look at and speak to the player in second-person narrative voice, referring to the player using their chosen name. The story is not about dragons or butterflies. Players are free to be themselves as they visit old friends from college. The script only makes a few assumptions about how the player supposedly knows the characters. My intention is to not only highlight a few of *Façade*'s successful narrative hooks, but also to illustrate the proactive quality of second-person narrative voice. This section provides evidence that players understood their role the script, and that many players played into the story and made up backstories to fill in some of the missing history.

7.1.1 Becoming Aware of the Player-Character Role

Unlike many fairy tales and futuristic science fictions, the *Façade* story builds on a real-life scenario that many people have encountered, or can imagine encountering, in their everyday life. As Player 43 related "it's not like you're coming in and you're a butterfly and then you have to try and figure 'what does a butterfly do—how do they move?'" The players know how to move, and to some extent, they know what to say because they have seen this scene before. *Façade*'s story was designed to be open to interpretation and appropriated in different ways, but most players shared a similar view of the scenario they were thrust into. Player 37, for example, said "you're kind of placed into the role of... a marriage counselor." Players could see that they were there to "help them" (P24), to "listen to both sides of the story" (P3), and to "advise them" (P1), and to "play the counselor" (P8). Player 21 breaks down what she thinks she can do in that situation:

> *"When someone's around to witness an argument you always want to get them on your side. So they were both trying to get me on their side. So that's the part I played – I played the witness that was trying to get pulled on either side. .... I played the innocent bystander. And usually the friends who witness those kinds of arguments don't get involved that much because they're in their own world and you don't want to make it worse somehow, although you do try and help in the way you can which is just calm them down." (P21)*

Player 21 understood her place as a "witness" or an "innocent bystander" which underscores a subtle intent of *Façade*. *Façade*'s narrative structure provides players an interactive role, but masks the technology limitations by exploiting a social scenario that often leaves people feeling powerless. This design technique is similar to the Eliza project which uses the guise of a psychiatric interview to hide the limits of its natural language parser (Weizenbaum, 1966).

Second-person narrative voice is used sparingly in novels, rarely if ever in films, and in limited ways in video games—although this is changing as more contemporary titles seek to infuse story into games (Ellison, 2008). In *Façade*, Trip and Grace look directly at the player and say things like "you think this is all my fault, don't you?" They also refer to the player by the name the player choses. Anecdotally, during our studies of *AR Façade* players first attempted to find their own name from the list. In terms of choosing a male or female name, only one out of forty-five players switched to the opposite gender. In contrast, desktop *Façade* players tend to experiment a bit more with names and genders.

Although players did not have to transform significantly to become a character in the situation, players did exercise their imaginations, as Player 35 stated "it's just the whole idea of being something other than yourself for a little while" (P35) and as Player 37 said "you have to actually decipher what's going on in the scene." Although the story has familiar themes, the social situation is contrived. Players tended to play along, but only to an extent.

7.1.2 Pretending to be a Character

Before a player enters Trip's and Grace's apartment to speak with them, she is told[55] that she introduced the couple back in college and that they are meeting to catch up. In my investigation of forty-five players, I witnessed numerous examples of players playing along with that scenario, particularly within the first few minutes of greetings. Player 15 said she just "went along with what they were saying and acted like they were

---

55 The player is supposed to hear a voice mail message from Trip inviting her over for the evening, but in the *AR Façade* installations the player often missed the intro while putting on and adjusting the gear and so she would hear the short backstory from an attending docent.

regular people." Similarly, Player 17 talked about his initial strategy of catching up with the characters:

> *"When I went in there, I really just wanted to get to know them and talk to them and catch up. You know, the whole 10-year thing." (P17)*

After the initial round of greetings, Trip often reminds the player that "it was almost exactly 10 years ago tonight that you introduced us... senior year of college." Many players reacted similarly to Player 1 at this moment in the episode; he told Trip: "yeah, I remember that, I hope you are still doing great." In the interview, Player 1 said "I thought Trip and I were college buddies and that I was good friends with him first. I guess I extended what you told me." He assumed his relationship to the couple started with Trip and used his assumption to play into a number of other narrative hooks. For example, while Grace is in the kitchen Trip asks Player 1 about a women named Veronica at work; when Grace returns abruptly Player 1 tries to cover for Trip. Later he talked about how he enjoyed that:

> *"That was the best part that I experienced so far. I felt like it was so real, … boys talking about this sort of thing and the wife thinking there was something fishy going on. I never expected Grace to come out…" (P1)*

Moments like this in the narrative structure worked to transform the players into the old friend of Trip and Grace, particularly in the early stages of the game. Once the characters started to fight, players reacted differently, as I detailed in Chapter 5. As an engager, Player 1 reacted to the fighting similar to the other moments, telling Trip and Grace he was "there to catch up, not to fight." Although engagers like Player 1 were taking cues from real-life social situations, their actions revealed a degree of make-believe. Player 18, for example, said she felt like she was "playing along", especially when she confidentially told Grace "Trip is just in one of those moods" (P18) after Trip stormed back into the kitchen. Player 18 later described her strategy:

> *"...To remind them of good times we had together. Or just to get them not to fight and remind them that there are good things that they liked about each other to begin with, because clearly I'm the one who introduced them." (P18)*

She assumed that Trip and Grace had good times together at one point and that she could somehow uncover the good aspects of their relationship. This is not a far-fetched strategy for dealing with this situation in a real-setting. Engagers provided interesting examples of interaction because they treated the situation similarly to real-life situations, as Player 25 explains:

> *"If you're pretending that they're good friends and they're people that you care about, and they're married, then I feel like that's what you would be doing anyway." (P25)*

A "good friend" of the couple would listen intently and try to help them—precisely the strategy explained by Player 3:

> *"I was just trying to think about what it is she wants to hear, what would warm her up, to make her receptive… I wanted to get her to open up." (P3)*

While most of the players played along through the initial greetings with Trip and Grace, not all players reacted to the fighting like the engagers. As I describe in Chapter 5, performers reacted to the situation with absurdity; observers tended to go into a shell; tinkerers took more interest in the interaction mechanisms; and many of the partakers checked out early.

7.1.3 Filling in Backstory Gaps

For many players, their ability to pretend and make-believe in this particular scenario was hampered by a lack of history with the characters. As Player 20 explains, the characters might claim that she knows them for ten years, but "....in my experience I just met them, so I didn't really know what to say." (P20) Player 28 also pointed out that Trip and Grace "were going on about stuff that I didn't know in the past, but like once I like started to understand the story a little bit it was good." (P28) However, when Trip and Grace hit their fighting sequences some players wish they knew the characters better. Player 22 explains the feeling of being abruptly thrust into the scenario:

> *"It seemed sort of abrupt. It's like there wasn't initially time to get to know the characters first. I didn't have a chance to get accustomed to the whole environment and to the idea that I can relate to these people." (P22)*

Player 22 did not have enough time to establish a relationship with the characters. Similarly, Player 27 struggled with imaging a friendship that actually doesn't exist:

> *"You know, like if I had a lot of background information on them, you know, like a 10-year friendship, then I might have more things to say... It seemed like a very hostile environment... especially with two people you're supposed to know, but you kind of don't. So it makes you wonder like what you're actually to do." (P27)*

To adjust to the missing backstory players would necessarily make assumptions about their supposed relationship with the characters. Similar to Player 1's assumption of a longer relationship with Trip, Player 35 said she "got the feeling that I knew Grace better, like we were better friends than me and Trip." (P35) Interestingly, even though many players assumed they knew the character of the same gender for a longer time, most male players preferred Grace, while more female players preferred Trip (see Figure 7.5). Player 38's exchange with Trip and Grace provides a good example of how players would "fill in" for backstory gaps (see Table 7.1).

Table 7.1: Player 38 filling in backstory for the player-character role

Time	Player and character statements (with notes in parentheses)
2:10	Grace: Rich, seeing you again makes me remember the wonderful times we used to have.
2:13	P38: Yeah, it's been a while, what, five years?
2:16	Trip: Haha, yeah.
2:18	P38: What are you up to Trip?
2:21	Trip: Yeah, tonight is special. Rich, remember it was almost exactly 10 years ago tonight that you introduced us... senior year of college. Remember that?
2:32	P38: I do. What was it? Disneyland?
...	(Grace and Trip do their "pain, agony, love" exchange)
2:54	P38: So, where are the children?

First Player 38 tries to guess the length of their relationship ("five years"), then he offers an event to explain how he introduced the couple ("Disneyland"), and then he suggests that Trip and Grace already have kids that he knows. According to Player 38 he was there to "have a couple drinks to say hello — a 'long-time-no-see' visit..." (P38) His casual approach hit a snag when the characters started fighting and so he quit the experience shortly after. For many players, getting to know the characters was a matter of piecing together their history. According to Player 36:

> *"You don't have the background to really even give an informed decision, so you just have to kinda make it up... it was like I had to construct the opposite of what they were now from the experience of them now." (P36)*

Ultimately, players wanted "a little bit more background history" (P24), but they used their imaginations to play along with Trip and Grace and fill in the ten years of missing history. While this section focused on *Façade*'s interactive narrative structure and the effect of second-person narrative voice, the next section looks at the role of the first-person immersive interface in shaping players' transformation into the player-character role.

7.2 First-Person Interfaces and the Player-Character

As I noted in Chapter 6, the first-person immersive interface used in *AR Façade* supported a strong sense of physical presence within the narrative world. In this section, I use player comments and episode examples from *AR Façade* to illustrate that the players' sense of physical presence reinforced their connection to the characters and the dramatic moment. Players treated the characters like they would treat human conversational partners and reacted to the situation dramatically. The immersive interface helped participants complete a transformation into the player-character—essentially *becoming* rather than *portraying* the role.

7.2.1 Becoming versus Portraying the Player-Character

The immersive interface helped participants connect to the story world and to complete a transformation into the player-character. The immersiveness allowed players to treat the characters like "life-size" (P11) social actors and to physically enact the role established for them. The qualitative descriptions from the player interviews help to explain how AR feels different than less immersive forms of media (players 1-12 could directly contrast desktop and *AR Façade* where as the other players would contrast the experience with TV or video games).

Unlike many contemporary video games where players perform actions in the world through an avatar, in *AR Façade*, players are inserted into the world and masquerade as a character role using their physical self. As Player 29 stated:

"(In KB..) It's like there's a middleman. Here (in AR), there isn't. It's just you and then the other characters." (P29)

Desktop *Façade* does not use an avatar either, but there are important differences as some players point out. Player 36 said "on the desktop it's all framed… I always feel like it's being relayed to me" (P36), supporting Player 29's notion that desktop interaction asserts an in-between layer. Player 9 expressed a difference between "portraying" a character and having to "be" a character:

"You feel like the person in the game vs. portraying someone in the game (in desktop interaction). You are supposed to be the person, but I think you believe it more when you are in the physical space, because you are actually doing the actions and actually speaking." (P9)

As Player 9 explained, interactors are scripted to take on a player-character role in desktop *Façade*, but transformation into that role is much more believable when every action and every word become part of the interaction. Player 9 also explained desktop interaction, saying "even though you are looking at it in first person (on the PC), you are still thinking about it in third person" (P9). Player 28 describes his sense of physical and social presence when interacting in *AR Façade* by contrasting the experience with watching a film:

"Like it feels like I'm physically there, like I'm not on a computer. This makes me feel like I'm actually there, and I'm actually with people and interacting with those people. It makes it seem more real, I guess. I don't know if it makes it more entertaining or if it makes it ... I don't know... It has a different effect than like going to see a movie or something. It kind of

makes you feel like you're a part of the situation and everything that was going on, like you had a say in everything that was happening." (P28)

Player 28 said it made him feel like he was "actually there", like he was "part of the situation" and "interacting with those people." He said the immersive interface seemed "more real" and had a "different effect" than seeing a movie, but he hesitated to say it made it more entertaining. Other players talked about how the experience required a greater degree of commitment. Player 25 mentioned that she felt more "attached" to things she spoke:

"I feel like [typing] is very different than from actually saying it, because when you say it, you attach yourself to it." (P25)

Once you say something or do something in an unmediated AR interface, it becomes part of the fabric of the social situation. There is no chance to reflect on the interactions. There is no undo. As Player 9 found out, there is no room for inner thoughts:

"When I first walked in I couldn't find Grace and when I finally saw her I just blurted out… 'oh, there you are!' but I didn't want to SAY that …I guess this means that I don't get any of my own dialog… [laughs]… I really just meant to just think it." (P9)

Unlike the typing interface, the speech interface offers no opportunity to reverse actions. That is not to say an immersive experience could not be designed to mediate speech interaction in a way. One reasonable design would be to allow players to speak something, show them what they said, and then allow them to discard or enter their statement. As it stands, the speaking interface works like it would in everyday conversation. Once a statement leaves the lips, it is in the air and subject to potential misinterpretation. Like real-world arguments, players are forced to consider what they say before they say it. As Player 37 described:

> *"It's less reactionary through a computer screen because you can kind of think about it... the thought process that it takes to develop what you're going to say, it translates through your hands is a little bit slower. Where if you're in the middle of things and if you are looking towards a character or another, … they ask you a question, you immediately say: 'oh, you're looking at me?... you're asking me a question? what do I say?' So it's much more an on-the-spot reaction inside than it is on the screen…" (P37)*

Player 37 thinks a typing interface would be "less reactionary" because it takes time to develop the thought and then translate it into words. Speech is more immediate. Moreover, Player 37 talked about having to provide "on the spot" reactions to the characters who are probing her for some response. The narrative structure of the experience combined with the "natural" and unconstrained mode of speaking to the characters reinforced many of the players sense of connection to the drama.

7.2.2 Feeling the Social Presence of Characters

The fact the players could physically move around within a tangible apartment not only helped players feel as if they were "in" a shared environment with the fictitious characters, it bolstered their relationship to the characters. Player 19, for example, related that when Trip "asked me to go see the picture – I felt that I was a friend and he had just asked me to see something." The physical act of moving in reaction to the characters strengthened players' sense of social presence, as Player 37 explains:

> *"You are in the same space as the virtual characters and so having to turn around and actually have to orient yourself towards the different characters, you are kind of very literally in the middle of things." (P37)*

Players would often use their arms and hands in unconscious non-verbal communication, for example Player 14 holding his arms out, palms down and telling Trip and Grace to "just relax" (P14) or Player 7 stepping back from the characters and holding her hands up in front of her body, as if trying to shield herself from the awkwardness of the situation (P7). As the case study for Player 25 illustrates below, she unconsciously used a number of physical gestures common in everyday conversation to help convey her thoughts.

A number of players deliberately moved in concert with the characters, not only following "staging commands" (e.g. the Italy photo, the couch, etc.), but trying to steer clear of the characters. On several occasions, I witnessed the player quickly step out of a character's way. Several players commented on the moment when one of the characters storms out of the room during a fight. Player 12 said "I was like whooaa, she's pissed! (laughing)…so I was just trying to get out of her way" (P12) as if Grace would do harm to her. Player 10 also eluded to Grace's emotional state reflecting that "…I thought I made them a little bit angry (smirking) and that Grace might make a run at me" (P10).

For players 1-12 in the interface comparison study, the contrast between desktop interaction and AR interaction made it easier to highlight the effect of a having "life-size" characters (P11). Player 4 said the AR version "felt more social" (P4) and player 12 felt "more connected to them" (P12). Player 2 reflected that she felt "in the room with them," (P2) especially as Grace was sharing her feelings. Players were more conscious of the social conventions in the situation and often commented on whether they met or violated expectations. Player 2 for example revealed her standards for social protocols through her comments on greetings and farewells ("normally people say some sort of salutation") and drink pouring protocol ("he was holding two drinks, but I wasn't just going to reach over the bar and take it because it might seem rude (laughing)" (P2). Player 3 felt it was rude for players to turn their backs to her, saying "they're supposed to be my friends and they had their backs to me a lot. That was a little upsetting."

Some players expected the characters to be emotionally deeper and more conversational in the AR version, saying "they had more weight as characters" (P4) and "you feel like they should be even more humanistic… you feel like you are one of those characters and you should be able to interact even deeper" (P9). The immersive physicality of AR helped to strengthen players' sense of connection to the other "social" beings and to become involved in the dramatic moment.

7.2.3 Reacting to the Dramatic Moment

A number of players commented generally on the dramatic nature of the simulated scenario, such as "you weren't kidding about drama…" (P9). Their actions were even more revealing, since many players intentionally used their bodies within the dramatic moment. While some of these physical actions can be viewed as physical manifestations of social presence—when players move out of the way of characters, for example—other actions were the result of a greater sense of involvement in the drama. I have already presented a number of examples of this in previous chapters; recall Player 3 who wanted to keep Trip from leaving at the end and actually moved between Trip and the door and held her hand out defiantly (see Figure 5.14). Similarly, Player 32 wore her emotions on her face and tried to point Grace towards Trip near the end of the scene (see Figure 5.18), and Player 4 tried to drag the characters to the middle of the room to talk (see Figure 6.1). These actions arose in reaction to specific story situations.

In Chapter 6, I referenced Player 25's long diatribes to illustrate how the unconstrained nature of the immersive interface supported her desire to expound on her views towards Trip and Grace. Her episode also elucidates how some players employed their bodies and their physical presence within the dramatic moment. Exemplifying the engager style of play from Chapter 5, Player 25 becomes involved in the story, believes she can impact the course of events, and demonstrates her sense of dramatic presence through her physical enactments and emotional reactions (see Table 7.2 and Figure 7.1).

Table 7.2: Player 25 reacting to the dramatic moment

Time	Player and character statements (with notes in parentheses)
11:27	Grace: Jane, all this time I can't stop thinking that I should have been painting.
11:32	Trip: Why don't you just do it?
11:34	P25: Well, why don't you paint? (offers hands out to Grace)
11:36	P25: If it's something you really want to do
11:37	(talking over player) Trip: Be a goddamn artist!
11:39	P25: Trip! Shhh (holds hand out at Trip to silence him) We need to at least let her try.
...	
12:33	(gesturing her hands towards both characters) P25: I think all of us just need to cool down and look at this in a sensible manner.
12:36	Grace: Jane, are you saying this is my fault?
12:38	P25: I don't think so. I think a marriage is when two people come together. It's not just one person's job...
12:43	Trip: Jane...(interrupting the player)
12:44	P25: or the other person's, but both people (talking with hands...)
12:47	Trip: Jane, I thought we talked about that already.
12:50	P25: Alright, so... what is your problem with Grace? (towards Trip)

Figure 7.1: Images of Player 25 showing her dramatic involvement

Player 25's behavior demonstrates the intentional (and unconscious) occurrences of physical gestures discharged during the rising tension of the drama. Her conversational actions are not extraordinary, but they only make sense in the context of mediating an arguing couple. In the interview Player 25 reflected about how weird it was to be provoked by "characters in a computer":

> *"I'm feeling awkward and these people aren't even real that I'm interacting with, which I guess it's wild to think that I can get that sense from people that aren't there. They're just characters in a computer." (P25)*

Player 4 described the increased sense of dramatic presence drawing on his background in theatre and acting, saying that in AR "you would commit to the scene and to your character" (P4). Whether emotional reactions or improvisations, these physical acts do more than simply illustrate the unconstrained nature of immersive AR; they exemplify physical involvement in the dramatic moment, leading players to perform actions (such as Player 3's attempt to stop Trip from leaving) that would not have happened unless they were closely following a dramatic story and submerged in the physical context, "on stage" to extend the theatre metaphor. In other words, I believe that if *Façade* did not have a story arc, players would not have performed many of the actions they did because there would have been no reason to do so.

7.3 Emotional Involvement

In this section, I demonstrate that *AR Façade* succeeded in eliciting raw, genuine emotions, especially as the tensions rose between Trip and Grace. I will present two case studies and numerous comments from participants to illustrate the level of emotional investment from some players. When players make comments like "I felt the emotion... I felt the tension" (P19), it shows that the *AR Façade* enjoyed a certain degree of artistic success. Mateas and Stern's expressed intent of *Façade* was to "offer a satisfying dramatic experience for players" (2005), and to communicate an overarching theme: "to be happy you must be true to yourself" (Mateas, 2001). While many players echoed this theme during post-interviews, I focus on the authors' goal of creating a satisfying dramatic experience. The emotional responses appear to actually reinforce many players sense of physical presence, as Player 39 indicates:

> *"Well, I was definitely a part of it in the sense that, you know, my emotions ran, I was anxious being in the room, I felt a party to what was happening. I felt a party to the drama that was unfolding." (P39)*

The tense nature of *AR Façade*'s narrative combined with the immersive interface appeared to fortify the illusion of situational presence and, as the examples below illustrate, elicited players to display authentic emotions.

7.3.1 Case Studies of Emotional Display

In this subsection, I present two case studies of players (both engagers) and their outward displays of emotion. I start with an excerpt from Player 26, who intermingles feigned dramatic actions (like holding her hand to her heart) with authentic, unconscious emotional reactions (see Table 7.3 and Figure 7.2).

Table 7.3: Player 26 reacting emotionally

Time	Player and character statements (with notes in parentheses)
7:27	Grace: Can you trust your husband or wife too much? To rely on them too much?
7:34	P26: Relying and trusting are two different things though. (hands up at sides)
7:37	P26: So I'd say yeah, you shouldn't trust anybody too much.
7:40	Trip: What?!
7:45	P26: Well you don't trust her too much do you? (looking at Trip now)
7:47	Trip: Fine. That's fine. Arrrh, I keep trying to help... (player cringes up and looks frightened and guilty.)

Figure 7.2: Images of Player 26 showing her emotional involvement

Player 26 exemplifies the "engager" style play (from Chapter 5) and illustrates how some players readily take on the role of a mediator. She also had several long exchanges near the end where she appears to take offense at some of the characters' statements, and she yells at them "you both need to turn around!" (P26). During the interview Player 26 describes the situation as "uncomfortable," but expressed that she felt somewhat obligated:

> *"Well the whole time I didn't want to abandon them, but it really didn't feel like it was any of my business either, because ... you do pick sides and you hurt feelings." (P26)*

Player 26 got wrapped up in the story and took the situation seriously enough to worry about hurting the characters feelings. Throughout her episode Player 26 tries to excuse herself, saying things like "I could just leave" (P26), but she ended up sticking around until Trip leaves. Player 35 on the other hand starts out quite enthusiastic and responsive to Trip and Grace, but then gets extremely uncomfortable to the point she has to leave (see Table 7.4 and Figure 7.3).

Table 7.4: Player 35 reacting emotionally

Time	Player and character statements (with notes in parentheses)
7:50	(The player's frustrations are audible. She sighs and huffs and squirms.)
7:57	Grace: Oh and you... I know what you're going to say the way you keep talking about Trip's goddamn drinks.
8:04	P35: What? (angry and defensive)

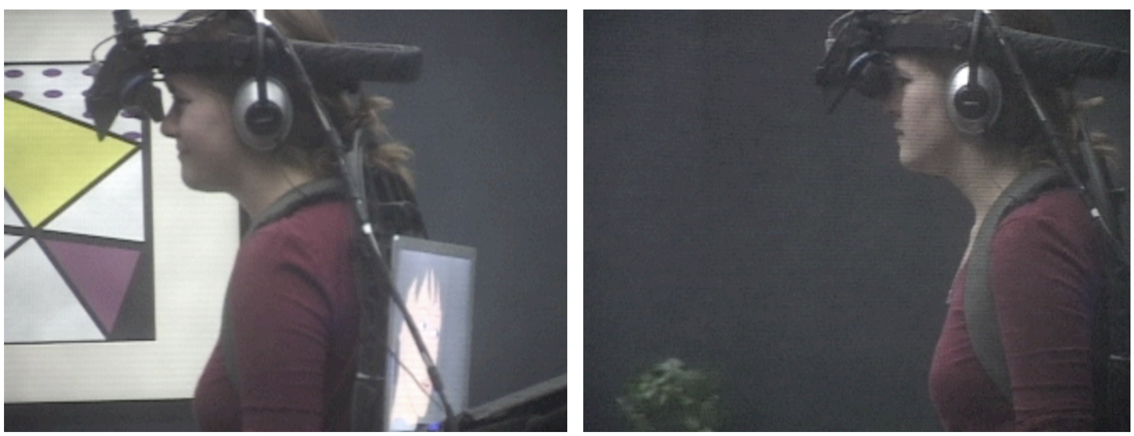


Figure 7.3: Images of Player 35 showing her emotions change

Player 35 also tries to politely excuse herself for about five minutes before finally just leaving. In her interview she said "didn't want to be rude" but she really wanted to "leave them alone to fight." She said that Trip and Grace "kept dragging" (P35) her in, but that she really got pulled in when they used her real name:

"You're aware that it's not real, but I think you definitely get into it. And I think having my own name might've gotten me in more because, you know,

> *they keep calling me Tanya and so you're like, "Oh, okay, that's my name, that's my name, that's my name, that's my name..." (P35)*

The personalization of name dropping led Player 35 to get deeper into the social dynamic, to the point where she “took it personal” and was affected by many of Trip and Grace’s comments. Her frustration was evident in her unconscious reactions and in her audible sighs and huffs.

7.3.2 Testaments of Emotional Effect

In addition to innumerable examples of visually observable emotional reactions by players, they also expressed the emotional effects during the interviews. I heard testaments from players cataloging wide range of emotions, including feeling “helpless” (P3), “embarrassed” (P14), “alienated” (P17), “creeped out” (P25), “scared” (P31), and “blamed” (P29), “uncomfortable” (P37) and “awkward” (P42) to name a few. The following interview segments explore the emotional effects induced by the experience and point to some of the subtle features of *AR Façade* that strengthen players’ emotional involvement.

AR Façade simulates an occurrence that people can relate to, but one that does not happen often. When it does happen, most people do not want to be a party to it. The simulation is real enough that it elicits the same feelings they would have if they were in that situation in real-life, as Player 42 explained:

> *It gave like the same like awkward feeling that I guess I would probably have if I was in that situation myself. (P42)*

Player 37 related the experience to listening to her parents or friends argue, a scenario that feels “wrong” to even watch:

> *"But other times when it's like your parents arguing or when it's close friends that are arguing, you're seeing it and it feels really uncomfortable to watch. And you just, you feel like, 'Really, I shouldn't be watching this'... because even though it is two people that you're supposedly friends with, it's just kind of wrong to watch." (P37)*

Several players talked about how uncanny and strange it was that the simulation actually made them feel emotions. Player 25 said "It's creepy. It's like being watched and watching... and the fact that they aren't real makes it really strange." (P25) It was strange for her that a mere simulation would incite emotions generally reserved for real social scenarios. As Player 38 explains, his genuinely felt emotions were an indication to him that his actions were also sincere:

> *"At the end when they get really upset, I was not pretending. I was even thinking to myself like that's really weird that I'm feeling like awkward. It's just like, if I was pretending, I wouldn't be feeling that. But this is like their argument and me being there and interacting with them, it made me feel awkward." (P38)*

AR Façade extracted real emotions and, in many cases, genuine behaviors that would have been applied to the scenario in real-life. The experience triggered such an effect for a number of reasons. As I mentioned above, the immersive interface made the characters seem more "human" and gave players the freedom needed to be emotionally expressive. Also, the player has a role within the script; their character can effect the course of events and is subject to critiques leveled from Trip and Grace. As the above example of Player 35 illustrates, one clever design feature of *Façade* is having the characters use the player's chosen name. For Player 28, the angry use of his name by Trip and Grace made him feel like the characters were blaming him for their problems:

> *"They started like blaming me for the problem. They kept saying my name kind of in an angry voice... I was kind of like the in-between-guy ... when they were both kind of taking their anger out on me instead of talking to each other about it. It felt like they were just getting mad at me." (P28)*

Player 28 felt like Trip and Grace were getting mad at him and blaming him for their problems. Several of the players eluded to *AR Façade*'s story "hook" about introducing the couple during the senior year of college, which seemed to evoke a notion of culpability. Players would make comments like, "I guess I just introduced them..." (P28), and then go on to describe why that made them feel guilty. Player 37 said "it felt almost as if they were holding me responsible for something." It led Player 5 to exclaim "that's ridiculous, it wasn't my fault they are fighting!" Similarly, Player 29 said:

> *"In a way I felt like they kind of blamed me for introducing them... because they were both looking at me really angry, and I was like, 'Ah, not my fault!' ...It felt like I walked into my friend's house, where her and her husband are about to kill each other. There was so much tension and it was really awkward, yeah." (P29)*

In Chapter 5, I presented Player 29 as a "performer" because of her funny outbursts towards Trip (see Figure 5.16). She felt emotions similar to an engager like Player 28, but she reacted differently to the situation. The awkwardness led Player 29 to react with absurdity and silliness so that she did not have to deal with the emotions directly. Player 30, an "observer" according to my analysis, said "I actually like felt the emotions of it. Like I was too scared to talk..." (P30), indicating that her silence throughout the episode could be partially attributed to the emotions brought out of her. On the other hand “engagers”, like Player 3, seemed unafraid to deal directly with the emotional crisis. The following was Player 3’s sentiment just after her episode where she tried to stop the character (recall Figure 5.14):

> *“I feel really bad… (repeated) oh man… I felt like I could have helped a little more… like I could have stopped her from leaving or something. I wanted to help... but I felt helpless. Nothing that I was saying was really helping and I was probably doing more harm than good by the things I was saying… I thought that eventually I would get a chance to help if I listened long enough.” (P3)*

Player 3 felt genuinely distraught and guilty over not being about to help them more. Her style of play indicated that she dealt with that situation as she would in real-life. She really would have listened to her friends’ problems and tried to help. Despite the different “styles” and approaches to the situation, many of the players felt some emotional response. In the next section, I investigate how players dealt with the emotionally-charged scenario and argue that many players developed tactics to maintain distance from the situation.

7.4 Player Tactics for Maintaining Distance

The evidence to this point suggests that many participants readily eased into the player-character role aided by the second-person narrative voice and first-person immersive interface. Not only did players take on the actions and thoughts of the player-character, they reacted with authentic emotions that would be appropriate for the context represented by the simulation. If *Façade*’s story was about something else, if it represented a happier theme for example, players obviously would have exhibited a different emotional response. In *AR Façade*, players wanted some psychological distance from the contentious nature of the content.

In this section I explore the tactics employed by players to manage their distance from the drama. First, some players would remind themselves it was merely a simulation as a means to deal with its emotional intensity. Second, players exhibited styles of play that signify a desire to escape the situation. Third, players managed their physical interpersonal distance with the characters.

7.4.1 Emotional Distancing through Mental Modulation

Several players commented on how the confrontation between the characters and the emotions it conjured caused them to mentally take stock during the episode. One tactic employed by some players for dealing with the intensity was to mentally modulate the experience, to deliberately not engage their emotions. For Player 23, when the drama got intense, it occurred to him that the simulation was not real.

> *"At that point it hit me that I wasn't actually in a room with the characters, and that made me smile because it caught my attention enough to where I was actually paying attention to what they were saying and caring about what they had to tell me or they were trying to convey..." (P23)*

Player 23 was wrapped up in the characters and sufficiently suspended in the illusion that he momentarily forgot the characters were fake. As a partaker, Player 23 reminded himself that it was fake and continued to play along with it, but then later reflected "I talked even less in this than I might have if it was a real scenario." (P23)

In the back of many players' minds, they never let go of the fact that they were in a simulated environment. Player 35 for example, said "you were aware that ..it wasn't real, but it was like you were still uncomfortable watching them fight.” Many players were “caught off-guard by the fighting” (P1) and were put “on the spot” (P36) by the situation. Some players dealt with that emotionally intense situation by reminding themselves of the simulation. Player 34 said:

> *“It was like, ‘Oh okay, this isn't real.’ I can just stand here and watch it and not say anything. And I don't have to have any guard up or anything because it's just not real.” (P34)*

According to Player 34, she gets hit by the intense drama and has to tell herself “this isn’t real” (P34). Then she decides that she doesn’t really have to behave like she would in a real situation, she does not have to really put a “guard up” (P34). She then admitted that “in a normal situation, I probably wouldn't have just been standing there watching” (P34). Player 34 rallied behind the fact that it wasn’t real so that she didn’t really have to engage those emotions, a strategy common to partakers.

A number of players talked about what they would have done differently if they actually encountered the situation in real-life. Player 24 outlined his real strategy if he really came into this fight:

> *“I’d probably just go to the bathroom or get a drink.... I probably would have cracked a joke or two or something to ease the situation or, you know, try to get them away from that, going into combat with each other.” (P24)*

Similarly, Player 31 hints that her real-world strategy would be to get out of there or to change the subject:

"I'd probably find a way to get outta there sooner. I'd probably like take out my cell phone and give someone a call, (Chuckle) you know, something along those lines." (P31)

As the statements from Player 24 and 31 indicate, most people would distance themselves from the setting if it was actually happening. Indeed, many players carried out these kinds of strategies within *AR Façade*, cracking jokes and trying to change the subject were common approaches, especially by performers and partakers. Many players were able to tolerate and participate in the dramatic content only because they reminded themselves it was not real. As Player 29 explained "you're not as connected to these people as one of your best friends" (P29), while Player 13 said she could have "gauged their emotions a lot better if it was like real people." (P13) Trip and Grace were not actual friends. The drama was not actually happening. Within a computer simulation, players had an opportunity to act things out differently than they would in a real setting, which helps to explain the range of play styles that emerged.

7.4.2 Distancing Through Play Style

Referring back to play style arguments from Chapter 5, the different styles of play can also be analyzed as different tactics for maintaining emotional distance from a clearly uncomfortable setting. Performers goofed off—perhaps to lighten the moment—but largely because they wanted to showboat for people watching. Many observers behaved like they were not there and did not even answer questions directed at them from the characters. Tinkerers spent a lot of time ignoring the fight, poking at the AR graphics, and making out loud reflections to themselves. Player 21 explained her meta-commentary as something she felt she could do because she was not completely connected:

> *"Yeah, I was just making lots of commentary to myself... I felt because I didn't feel completely connected to the world I was in... these weren't people I knew, I could just make the commentary out loud." (P21)*

Performers, observers, and tinkerers not only illuminate how players can become fascinated with ancillary aspects of the experience (e.g. the audience, the story, the medium), they indicate that many players did not want to deal with the social situation in *AR Façade* directly. This is not surprising, as most people would avoid this real-life social scenario if they could.

The play style distinction between engagers and partakers provides a chance to reflect on players' level of emotional engagement. In my analysis, players who exhibited the engager style of play are more emotionally and socially involved than partakers who seem to take the scenario less seriously. Where the engagers went the farthest to engage the experience emotionally (as the case studies for Player 26 and 35 captured), partakers did not engage their emotions even though they said and did the socially "correct" things. The only "inappropriate" action by partakers was to laugh during the middle of a dramatic moment. Player 22—whom I would classify as a partaker—explains that he was engaged, but not emotionally connected:

> *"I felt engaged in it, but sometimes I did feel like I was external. They were just going off—sort of the two of them—and I somehow didn't have that much of a hook into their situation. Like I did still feel a little bit like an outsider. I mean, I guess I was supposed to know them, but it still had a sort of distant feel to it for me." (P22)*

As I pointed out in Chapter 5 there were some quantitative differences between partakers and engagers during their episodes. Engagers' episodes were slightly longer on average—15.0 minutes (SD 4.2) to 13.4 minutes (SD 5.7), and they spoke a bit more per minute on average than partakers. Partakers gave the experience a lower, although not significantly lower, overall rating: 23.3 out of 35 compared to 25.9 for engagers (see Table 7.5). The internal details of that overall rating show some statistical differences. While the two groups give nearly identical average ratings for content, interaction, and presence, a difference stems from their contrasting curiosity about the outcome and believability about the characters (see Table 7.5).

Table 7.5: Rating differences between engagers and partakers (the first five ratings are on a 7-point Likert scale, while the overall rating is the sum of the others; standard deviation in parentheses).

	Engager	**Partaker**
Content rating	5.0 (1.3)	4.9 (1.3)
Curiosity about the outcome	5.6 (1.5)	4.2 (1.5)
Character believability	5.6 (0.9)	4.9 (0.9)
Physical Presence	5.5 (0.9)	5.3 (1.3)
Interaction	4.1 (1.5)	4.1 (1.3)
Overall Rating	25.9 (4.2)	23.3 (4.0)

Partakers were also twice as likely to quit the experience, showing their lack of interest in the outcome and their significantly lower rating of curiosity about the outcome (T=2.32, p=0.03). As Player 41 explained:

> *"I felt that tension of being in a room with people arguing that you're like ok, I don't need to be in this situation (Laughter)." (P41)*

Partakers were also less pleased with the characters (rating the believability at 4.9 versus 5.6 from engagers) (the difference is nearly significant with T=1.89, p=0.07). On the questionnaire, players were asked to select a preferred character with the following options: “Trip”, “Grace”, “Liked them both equally”, and “Hated them both equally.” While none of the engagers chose the last option, five of partakers said they hated both the characters (see Figure 7.4).

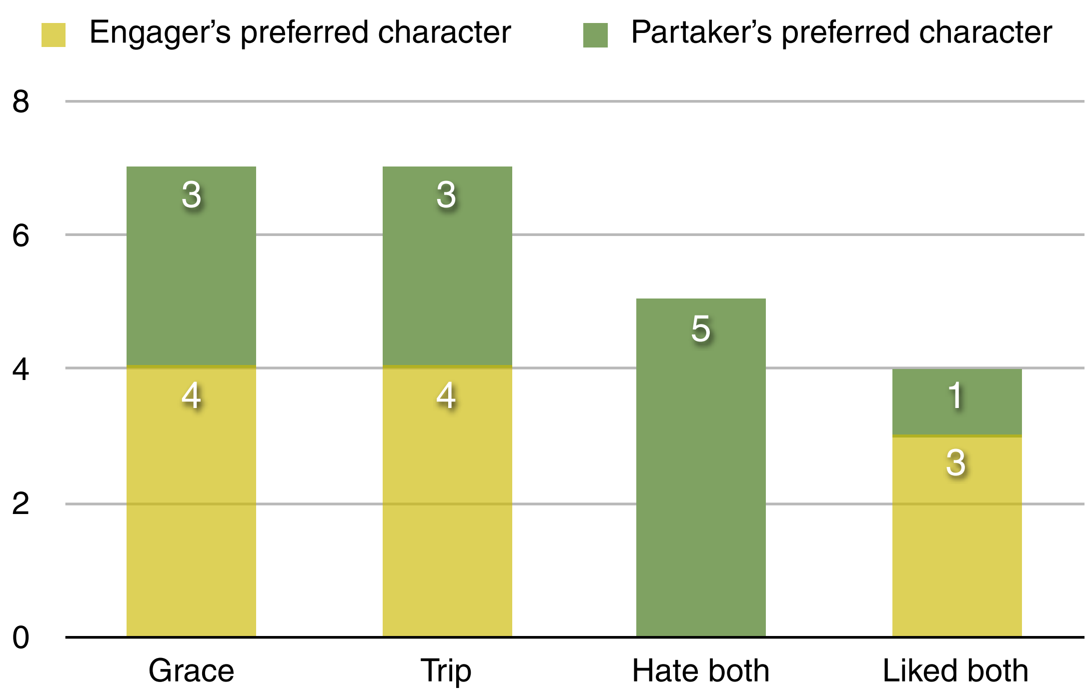


Figure 7.4: Difference between engagers and partakers for preferred character

This same question about preferred character also yielded differences among the players’ genders. Male players were more likely to say they “hated them both equally”, eight men answered that way compared to only three women (see Figure 7.5). The same question also revealed that players generally preferred the character of the opposite gender, especially for men where twice as many (6 to 3) liked Grace more than Trip.

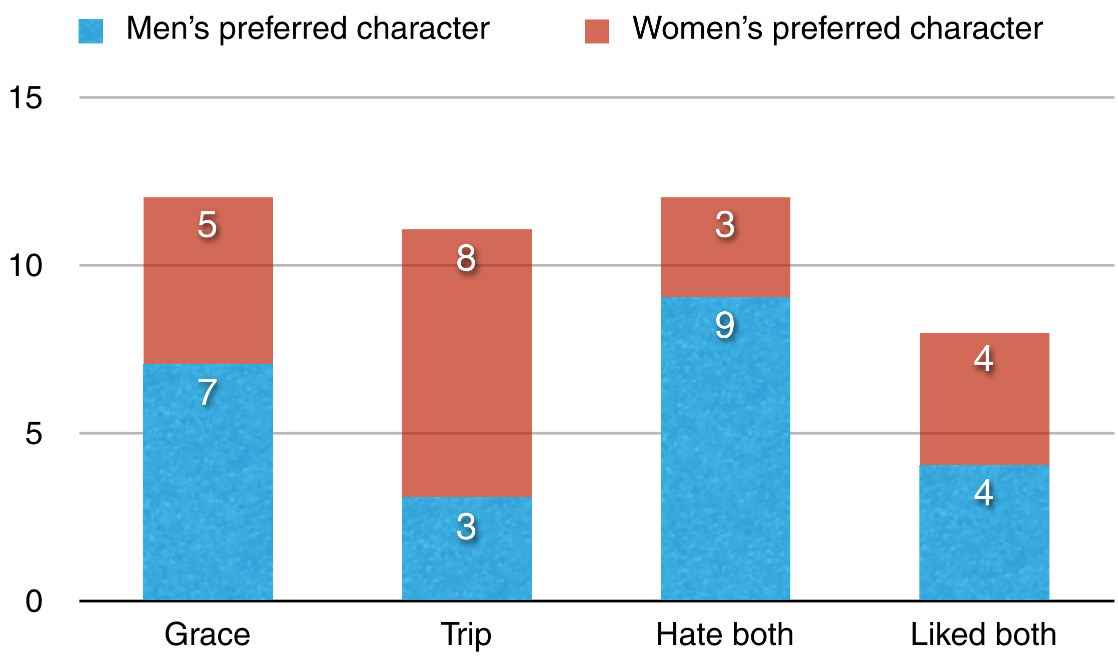


Figure 7.5: Difference between players' genders for preferred character (N=45)

While there was no difference between the genders on the overall rating of the experience (24.0 for men versus 24.5 for women) across all play styles, the anecdotal evidence indicates that the female players were more likely to engage the content emotionally and socially. My qualitative breakdown of player types indicates that males were more often categorized as partakers (eight men versus six women), and females as engagers (eight women vs. three men) (see Figure 7.6).

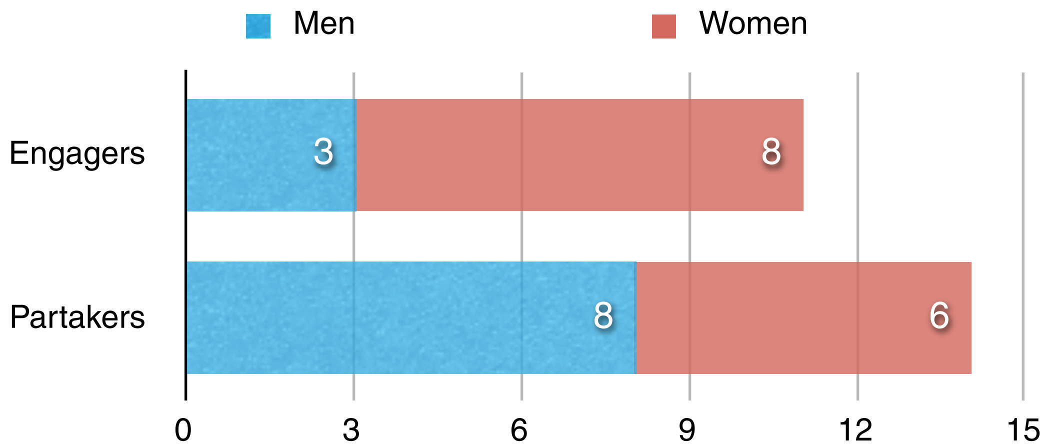


Figure 7.6: Number of females and males classified as engagers and partakers

My analysis that more women adopt the more emotional "engager" style of play during the *AR Façade* studies corroborates previous research on gender differences. Some social physiologists researchers have hypothesized that men are socialized to devalue and restrict emotional expression, leading to greater levels of alexithymia[56] (e.g., Brannon, 1976; Levant, 1992; O'Neil, 1981). Stokes found that men are more likely to avoid emotional disclosure with intimate acquaintances than women (Stokes, 1980). Similarly, Liebler and Sandefur report that women are more likely characterized as emotional support exchangers versus men who are more likely to exhibit low exchanger patterns (2002).

---

56 Alexithymia is the inability to express feelings with words (http://en.wikipedia.org/wiki/Alexithymia) (accessed 9/22/08)

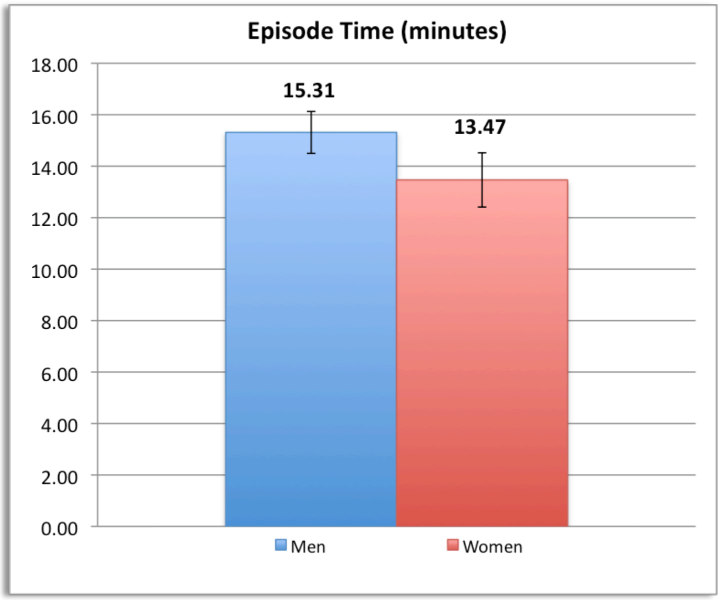


Figure 7.7: Average episode time for men and women (N=45)

As Figure 7.7 illustrates, men also had longer overall episodes across all forty-five players, 15.31 compared to 13.47 minutes (although not quite a significant difference, t=1.379, p=.175). I hypothesize that the gender difference in episode lengths is due to the fact that female players took the situation more seriously than men and thus decided to abandon their episodes earlier to avoid the confrontation. I believe further analysis of the *AR Façade* data will contribute to the discussion on gender differences and emotional intimacy.

7.4.3 Maintaining Interpersonal Distance

Interpersonal distance (IPD) is the distance between two social actors, and the study of IPD is known as proxemics (Hall, 1966). In everyday social life humans maintain an unconscious and unspoken physical distance from other people. In this section, I present results related to IPD between the player and characters in *AR Façade*. In my studies I was able to effectively calculate the IPD between the player and each character because I logged the position of each social actor at each second throughout the episode[57]. I argue that players who treat the experience more like an actual social experience maintained greater IPD. The evidence also indicates that some players used physical distance as an avoidance tactic—similar to their enacted styles of play—distancing themselves emotionally by building in more interpersonal distance from the characters.

Social science researchers often look at IPD as one dependent variable in studies of human intimacy. As one example of such research, Argyle presented an "Equilibrium Theory" to explain why people avoid eye-to-eye contact with decreases in IPD (Argyle, 1988). The theory essentially states that people avert their gaze[58] or maintain physical distance from other social actors from whom they want avoid high amounts of intimacy. If someone's desire for personal space is encroached, they may widen their physical distance as a countermeasure. As an element of social politeness, the acceptable amount of IPD and the various forms of compensatory adjustments are different across cultures (Baxter, 1970). Hall's observations and analysis indicate that 4 feet (1.22 meters) is too close for conversation with strangers; IPD less than four feet are generally reserved for personal friendships (Hall, 1966). The mean IPDs in *AR Façade* are larger than this minimal barrier for intimacy, but there are some significant differences, particularly between genders.

---

57 The IPD calculation remains accurate on the average despite small tracker variation. The tracker did occasionally completely loose its barring and mistakingly place the user's position somewhere outside the physical room, although this only happened once or twice per episode for the first twelve participants in the Atlanta installation when the tracker was not as consistent. I filtered out the outliers by looking for sudden changes in the position of the player and then manually matching the same timeframe in the video. I replaced each outlying timeframe with an average interpolated position from before and after the timeframe. While small variations in a players static position did not effect the IPD calculation, they did effect my calculation of the overall movement of the player, likely inflating the value from all the incremental changes. Again, I could have created a filter to bring the overall movement down to a more realistic value, but it has not been an essential part of my argument as this point.

58 We are not able to capture eye gaze without a specially equipped HMD.

In Reeves' and Nass' extensive research on their "media equation" theory, they use IPD as one way to show that human behavior extends into media environments (1996). In one study, they found that participants deemed faces of people in photographs as more intense when the faces were closer (Reeves and Nass, 1996). Yee et al. also studied IPD in the context of the shared virtual environment Second Life and found that mutual gaze was inversely correlated with IPD, confirming their prediction of participant behavior according to the equilibrium theory (Yee et al., 2007). Bailenson et al. conducted a similar series of experiments in immersive virtual environments and found participants gave more personal space to virtual agents who engaged them in mutual gaze (2001). Participants also maintained greater distance from virtual humans when approaching their fronts compared to their backs (Bailenson et al., 2001). In a followup study they revealed that participants maintained larger personal space bubbles for more familiar tutoring agents than for less familiar stranger agents (Bailenson et al., 2003), and offered an explanation that participants demonstrated more politeness towards the more familiar tutors.

From the previous work, one could theorize that participants in a media environment with virtual social actors maintain greater IPD when the virtual agents seem more "human". This theory appears to be consistent with the data from *AR Façade*. The players that I contend are more emotionally involved and treat the characters like believable social actors—the engagers—give the characters a larger space bubble (on average ~0.1 meters larger than partakers). While the result is not statistically significant, partly due to the imperfections of lumping players into one style of play or another (as I indicated in Chapter 5), it does show a trending difference between players who otherwise act and speak appropriately for the situation.

There were no consistent IPD patterns among the other play styles in how they interacted with the characters. Tinkerers on average stood close to Grace (1.92 meters) but far away from Trip (2.23 meters), while Performers stood close to Trip (1.91 meters) and far away from Grace (2.21). It makes sense for *AR Façade* that when an IPD to one character goes down, the IPD to the other goes up, but it does not explain why Tinkerers stood close to Grace and Performers stood close to Trip.

Another way to examine the interpersonal distance data is to look at gender differences. As I argue in the previous section, females participants in the *AR Façade* studies appear to be more emotionally and socially engaged than males. On average over four-forty participants (N=22 males, N=23 females) I found that female IPD was significantly larger than male IPD toward Grace (t=2.133, p=0.039), and nearly significantly larger towards Trip (t=1.661, p=0.104) (see Figure 7.8). The details of an independent-samples T-test is provided in Appendix Q.

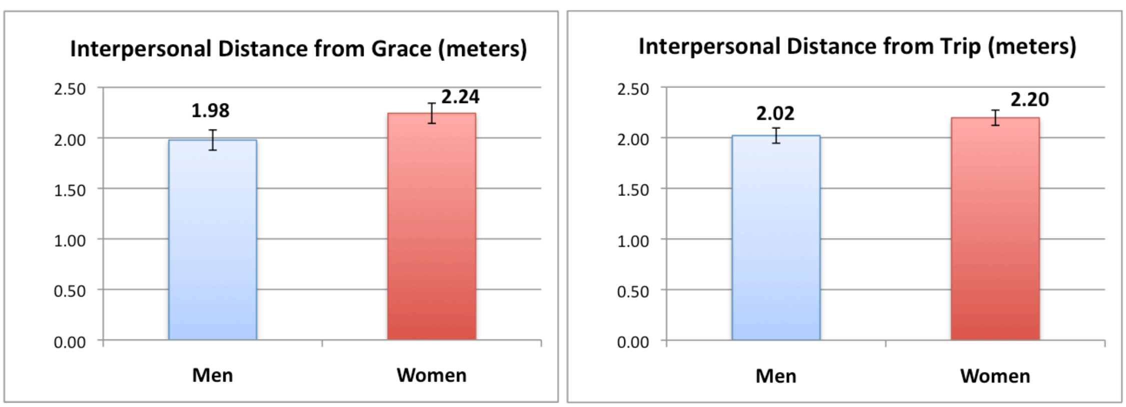


Figure 7.8: Interpersonal distance across genders (N=45) (Left) IPD with Grace in meters. (Right) IPD with Trip in meters.

On average, men stood closer to both characters, at 2.0 meters, compared to women who stood at 2.2 meters on average. This did not seem significant until I created a 2-D plot that shows each players' average position towards Trip and Grace (see Figure 7.9).

Figure 7.9: Interpersonal distance in meters among men and women players (N=45) and characters, Grace (X-axis) and Trip (Y-axis).

Visually there appear to be at least five females—Players 8, 18, 21, 25, and 29—who are on average farther away than any male (three of those five are also engagers). Likewise the four closest players to Trip and Grace are all men, Players 1, 6, 38, and 44 (two of which are partakers). Referring to the work by Yee et al. on non-verbal social norms in Second Life, one of their findings is that male-male dyads have larger IPDs than female-female dyads (so females stand closer to other females than male pairs) (2007). Our data does not support their finding (in fact male players stood closer to Trip than female players to Grace), but I believe the contentious nature of the story content confounds the effects of typical same-sex discomforts.

One explanation of larger IPDs in female players could follow from social psychology research that says males are typically socialized to be more dominant and willing to physically step into a confrontation to resolve conflicts. Another explanation is that males were less effected overall by the situation and so they could tolerate standing closer to both characters, while females would step back to give themselves more of the requisite emotional space. The latter explanation makes sense considering prior media-based social psychology research that says participants who regard virtual agents as more "human" maintain larger space bubbles (Bailenson et al., 2003). In my analysis, females players in *AR Façade* were more likely to treat and regard the characters as humanistic; males were more likely to de-humanize the characters. Both the play-style and gender-based quantitative data serve as indications that the players who yearn for and accept intimate and social relations with the *AR Façade* characters are also more likely to have a larger IPD with the characters.

Interview data further qualifies the importance of interpersonal distance. The physicality of the experience imposes itself on some players, particularly during those tense moments. For some players, the feeling of connectedness in AR heightened as the tension rose between Trip and Grace, leaving one player feeling "cornered" (P4) and another feeling "trapped between them" (P5). Moreover, some players appear to expand their personal space bubbles, building in physical distance from the characters to provide emotional space, and maintaining their degree of intimacy through IPD. Player 19, for example, describes her tendency to stand back from the characters.

> *"I watched them play out their situation and their conversation. So I stood – even my positioning in the room with them – I even took a step back and I didn't want to get too close." (P19)*

As an observer, she "took a step back" to avoid getting "too close" and then just watched them play out their conversation. Player 29 also employed a tactic of physical distancing and later noted that it made the experience feel less personal:

> *"I felt like I was in the room with them, but sometimes I didn't, 'cause it seemed like I was just watching them at times. ...I had to like stand back and then standing back just feels less personal." (P29)*

In general, players deal with the intense emotional situation through a range of adjustments, from mentally reminding themselves of the simulation, to adapting their style of play, to maintaining physical interpersonal distance in accordance with real-world social interaction.

7.5 Interface Effects on Emotional Distance

As the previous section indicates, players constructed various emotional barriers to deal with the intense situation simulated in *AR Façade*. Clearly the nature of the content influenced how people felt, but what role did the immersive interface play? In this section, I present evidence that directly contrasts the immersive AR interface with less-immersive experiences, such as the desktop version of *Façade* and filmic forms. Player 1-12 were part of an interface comparison study and so they could draw from their first-hand contrasts of the two experiences. Other players could speculate on the differences based on their experiences with other media.

First I revisit the interpersonal distance (IPD) arguments laid out in the previous section and analyze data that shows players were physically closer to the characters in the AR version than the desktop version. The desktop version supports greater IPDs with the characters, and in general allows for more distance between the player and the drama. I argue that the AR interface encapsulates the experience in a way that impinges on some player's ability to engage comfortably. Players commented how the desktop version of *Façade* and other less interactive media allow a certain distance that the AR interface did not afford.

7.5.1 Comparing Interpersonal Distance Across Different Interfaces

Looking at quantitative data from the Atlanta interface comparison study and the 11-week Beall exhibit, I found significant differences between the AR and desktop (KB) versions of *Façade* with respect to interpersonal distance between the players and characters (see Figure 7.10 and Table 7.6). In both instances where I could compare the relative position of social actors across interface versions, players stood consistently closer to the characters in the AR version. In this section, I will discuss reasons why this happened, and then consider the effects on players perception of each version of the experience.

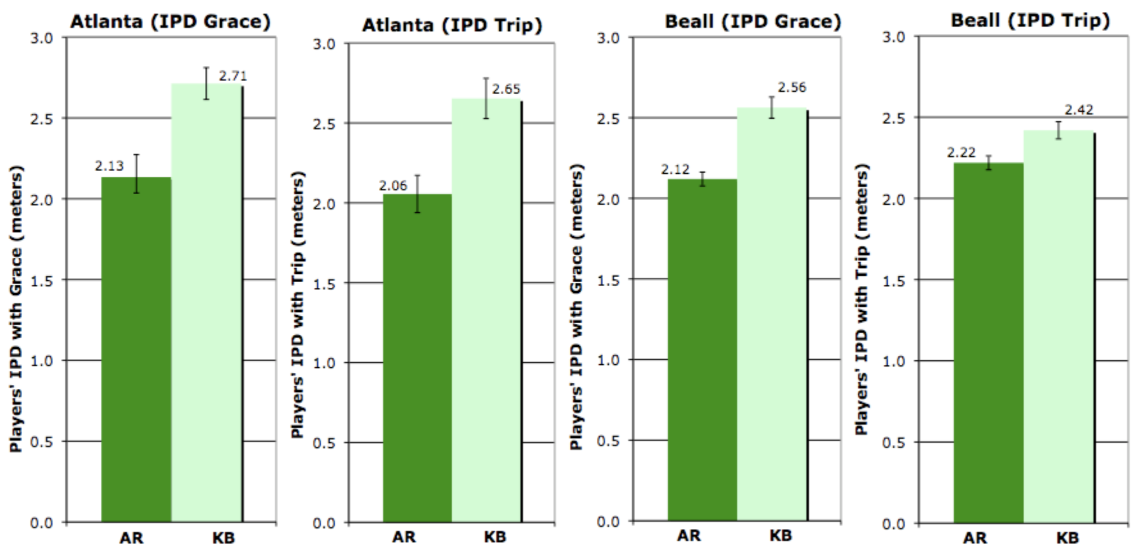


Figure 7.10: Interpersonal distance across two installations (in meters with std. error bars). (from left to right) IPD with Grace during Atlanta interface comparison study (N=12); IPD with Trip during Atlanta interface comparison study (N=12); IPD with Grace during 11-week Beall installation (N=126 for KB, N=106 for AR); IPD with Trip during 11-week Beall installation (N=126 for KB, N=106 for AR)

Table 7.6: Interpersonal distance across two installations (in meters with standard error). All four measures resulted in statistically significant p-values. For Atlanta: a *paired-samples* T-test showed that IPD in AR was significantly smaller than KB (For IPD-Grace, t=3.490, p=0.005; For IPD-Trip, t=4.131, p=0.002). For the Beall: an *independent-samples* T-test showed that IPD in AR was also significantly smaller than KB (For IPD-Grace, t=5.563, p=0.000; For IPD-Trip, t=2.854, p=0.002).

	AR	KB	T-value	p-value
Atl: Player IPD with Grace (meters)	2.13 (.09)	2.71 (.14)	3.490	0.005
Atl: Player IPD with Trip (meters)	2.06 (.11)	2.65 (.13)	4.131	0.002
Beall: Player IPD with Grace (meters)	2.12 (.04)	2.56 (.07)	5.563	0.000
Beall: Player IPD with Trip (meters)	2.22 (.04)	2.42 (.05)	2.854	0.005

There were a few differences between the Atlanta study and the Beall installation that should be kept in mind. In Atlanta, our twelve players tried both versions (AR and KB) as well as a third version, speech-based (SB) desktop interaction where the player speaks rather than types. At the Beall Center, the episode logs for AR (N=106) and KB (N=126) were collected as independent groups, although it is possible some participants tried both versions.

Moreover, we modified the field of view (FOV) for the virtual camera, going from narrower in the Atlanta installation to wider at the Beall Center[59]. Looking just at the KB versions from both installations, there appeared to be an effect of changing the FOV (see Table 7.7). The IPD value for both Grace and Trip reduced when going to the wider FOV (although the results were not statistically significant; see Appendix Q). A narrower field of view caused the characters to appear closer to the screen, as was pointed

59 The FOV for the Atlanta installation was 30 degrees, to match the FOV of the physical camera in *AR Façade*. I am not able to report the exact FOV for the Beall installation (original desktop setting), because that version used an orthographic projection as opposed to 3D perspective. By eye, the two versions were obviously configured with different FOV values, but it will require future work to determine how IPD mathematically relates to FOV.

out by Player 8 who said "sometimes on the (desktop) computer you felt like they were right up in your face."

Table 7.7: Interpersonal distance in the keyboard version across two FOV values (N=12 for Atlanta; N=126 for Beall). An *independent-samples* T-test showed that there is no significant difference between the IPD Trip and IPD Grace with the different fields of view.

	Narrow FOV meters (std error)	**Wide FOV meters (std error)**	**T-value**	**p-value**
Player IPD with Grace	2.71 (.14)	2.56 (.07)	0.977	0.343
Player IPD with Trip	2.65 (.13)	2.42 (.05)	1.719	0.106

Turning back to the difference between AR and KB, the values of interpersonal distance across all instances of KB interaction were greater than the IPD values for AR interaction (see Appendix Q). There are several logical explanations for why players' IPD with characters was smaller in AR than in KB interaction. Player 8 pointed out a perceptual difference between the two versions:

> *"They didn't get as close (in AR)… they seemed to be a little more distant… in terms of physical space." (P8)*

Although the characters were actually abiding by the same path-planning algorithms and taking up the same portion of the screen in both versions, players initially perceived the characters to be closer in desktop interaction. So as a countermeasure, players would back up, as Player 8 reported, she "would have to back away" in the desktop version in order to get a better perspective of the space and the characters. In AR, players had a chance to preview the apartment before entering and they could use their peripheral vision to see the space during the episode. So players moved back to get "more perspective" as Player 7 explained:

> *"(In KB) one thing I did was kind of back away from Trip, even though in real life I would have stayed close to him… because that gave me a little more perspective." (P7)*

Player 7 backed away from the characters in the desktop interaction because she wanted to be able to see more of the apartment and to understand the game context. Unlike the AR version where the player is surrounded by the game world, the desktop screen provides the only view of the game world. A more practical explanation of the wider IPD in KB interaction can be attributed to the AR gear. Perhaps players wanted to be able to back up more in the AR version but felt prohibited. This argument could also be supported by this statement by Player 8 talking about the AR version:

> *"If had been standing in the corner of the room it would have been a lot easier, but since I was standing in the middle I had to keep (spins finger around in a circle) scanning… with the others (other versions) I just backed up into the corner." (P8)*

Player 8 was complaining about having to constantly turn towards both characters. She says her strategy in desktop interaction is to simply back into a corner so she doesn't have to look back and forth. She also implies that it was not possible to back into a corner in the AR version, which is true since players carried a computer on a backpack. The desktop environment provides the affordance of sliding along the walls and finding a resting place for the point of view.

In the SB version from the Atlanta study, players' average IPDs with Grace and Trip were 2.56 and 2.46, respectively. These values were both significantly larger than the AR values (IPD-Grace, T=3.348, p=0.007; IPD-Trip, T=3.951, p=0.003), and smaller (although not significantly smaller) than the KB values (see Appendix Q). I believe the slightly smaller IPDs for SB interaction can be attributed to the fact that in KB players are forced to type statements as well as maneuver the space. This provides more evidence that players in KB would simply plant their perspective and focus on communication—the primary means of effecting the game.

Perhaps more important than what caused the IPD differences is the effect it had on players' sense of engagement with the characters. Player 4 talked about physically reacting to the characters' positions and how standing between Trip and Grace took on a distinct feeling in the AR space compared to the desktop environment.

> *"I reacted physically to the characters, like moving out of their way ... when they cornered me. When they are standing on opposite corners you can look between them. (On the keyboard-based desktop version) you have do the keyboard… you don't feel as trapped between them as you do in AR." (P4)*

The AR interface takes on a distinctly different feel than desktop interaction. In AR, players stand in the middle of action with close views of the characters' faces. Players must use their own body to look back and forth, giving them a sense at times that the virtual characters were behind them. With the desktop interface, players could employ the tactic of backing up into a corner. Not only did that provide players necessary perspective on the space, it gave them a larger interpersonal space bubble and implicitly more social/emotional distance. Players felt less embodied using desktop interaction, because they only needed to move their fingers to connect with the space.

The fact that several players commented on interpersonal distance as a difference between the interfaces demonstrates their disparity. Despite the fact that we did nothing to alter the motion-planning algorithms between versions, players perceived differences and it points to the future possibility of contrasting specific, measurable social cues while varying specific interface features (virtual field of view, screen size, etc.).

7.5.2 The Safety of Desktop Interaction

In this section, I provide evidence that some players would prefer to portray the character on the desktop screen, rather than literally be in the situation. The qualitative evidence will help to explain the questionnaire results from the first study—where players felt the AR version was most realistic (8 of 12), but many players felt the less immersive version allowed them to engage the way they wanted to. Only 6 of 12 said the AR was more engaging (see Figure 6.2). Some players rationalized the preference as a matter of comfort or familiarity, such as Player 3 who said "desktop feels more comfortable to me…I am more familiar with the desktop." (P3) Similarly, Player 19 welcomed the familiarity of desktop interaction:

> *"I think that the desktop version would be more comfortable. You wouldn't feel the tension nearly as much, and you would be even more outside of the whole experience." (P19)*

The most revealing part of Player 19's statement is her implied desire to be "outside of the whole experience."(P19) She did not talk about preferring desktop interaction because she wanted to sit or she liked the comfort of typing. The desktop interaction was seen as comfortable, because she would rather not endure the immediacy of the immersive interface. Player 24 expressed that the AR interaction would somehow be "easier" if he could operate an avatar rather than be "in the middle of it":

> *"It was like first person ... it felt like me being in the middle of it. I think that's why it was a little more awkward, ... whereas if it was an avatar*

then I think I would have been a lot easier to interact with the characters." (P24)

For a number of players, the less immersive desktop interface resulted in a more enjoyable experience, not only because it provided a safer emotional distance from the dramatic tension, but it also offered more freedom to be someone else:

"Here (in desktop) you feel like you are playing a role in an environment and (in AR) you feel like you are the role. You can say some stuff (in desktop) that you might not say (in AR)… like you are somebody else." (P10)

Similarly, Player 4 explains his preference for desktop interaction, precisely because the environment is less realistic and it provided fodder for his "performer" style of play:

"You get used to playing a character in their world on their level…It's almost because it's not as realistic. You can relax more. Goof around more…" (P4)

As a performer, Player 4 did not want to take the scenario seriously, so the desktop interaction gave him more space to "goof around" (P4). Both Players 4 and 10 preferred desktop interaction because it allows them to escape into a persona or into a fantasy world. Desktop interaction took the pressure off the social situation and allowed players to interact without considering the consequences. Player 7 related her impression of KB interaction, "I felt like there was less of a worry about what I was saying." (P7) Similarly, Player 8 explained, "...it's worth a try… I could type something completely ridiculous just to see how they react…" (P8) Player 24 speculated that he would say more and do more in a traditional desktop environment because it "won't affect you… emotionally or mentally."

> *"It's as if I have this wall that I can hide behind, in a sense. That's why I can say more things if it was an avatar. With an avatar it's not real in that sense.... there's that distance, where like you're just controlling someone that's not you, so you can do whatever you want. In essence, you can say whatever you want, and it won't affect you that much emotionally or mentally." (P24)*

According to Player 24, controlling an avatar is not like doing something yourself. The non-reality of operating through a proxy provides emotional and mental distance, a "wall" to hide behind. The head-mounted augmented reality interface in *AR Façade*, affords unconstrained interaction, removes the need for a proxy, a virtual avatar, and forces a participant to enact their role physically and directly. The less immersive version of *Façade* not only provided distance from the emotionally intense situation, it allowed some players to play more freely and to experiment with the social situation.

7.5.3 The Safety of TV and Movies

For the Beall exhibit where players did not have the benefit of directly comparing *AR Façade* with the desktop version, many players contrasted the experience with the feeling of watching the same drama on TV or in a movie. Player 29 compared the experience to the "normal" feelings elicited by TV:

> *"Being there and being able to talk to these characters, it makes you feel things that you wouldn't feel normally, like if I was just watching it on TV." (P29)*

The feeling of being there and talking to the characters was different than the feeling of watching it on TV. Player 37 implies that screen-based media like television or desktop interaction provides some sort of "gap".

"Looking at the screen... you're never going to be able to close this gap that you can't really get across." (P37)

Player 22 expressed the notion that the distance of television actually allows him to delve deeper into the characters, as if they are easier to understand and connect to their emotions from a more observational stance.

"On some level, even just watching TV there tends to be somehow an easier time just connecting to the characters that you see on a TV show." (P22)

His comment suggests that the emotional empathy he feels towards the characters can actually be deeper when there is a safe distance. According to Player 36 who contrasted the experience with watching a movie, people do not experience the same "mental processes."

"When you watch a movie of someone else doing it, they're the ones who came up with it and you're just observing that they come up with in that situation. You're not really feeling all the same emotions and the same mental processes. As when you're in the situation, you're like, 'Oh, okay, now I'm on the spot I have to pull up something to help address this issue'." (P36)

At times *AR Façade* was so stifling, it left players in a state where they "didn't know what to say." (P18) Many of the participants reported feeling like they were forced to react, because the characters kept on looking to the player for answers. Player 26 said she felt subjected to "force mediation" against her will, so this made it "uncomfortable." Player 43 also talked about what it felt like to be "forced to react":

> *"When you're in there it's different because you're forced to react. I mean it really does feel like you're part of it—that there's some responsibility to try and impact what's happening—versus when you're watching TV and you're like, ok, this is dysfunctional, I'll turn to another channel." (P43)*

Her sentiments capture a common feeling among players that "you're part of it" and obligated to "impact what's happening." (P43) She offers the juxtaposition of television where participants are not immersed or allowed to interact with the characters, and where exiting displeasurable content is as easy as flipping to another channel. Similarly, as Player 19 conveyed, it was not like she could just "block out the scenario that was at hand."

> *"I felt awkward in the situation. And I mean, some of it might have been due to… the simulation with virtual reality, but you can't automatically like block out ... the scenario that was at hand. It had an effect on how I was feeling through the whole thing." (P19)*

The immersive simulation did not offer Player 19 a means to step out of the scenario or to "turn the channel" (43). She felt a desire to "block out" the scenario to give herself more emotional distance. As a single unmediated social experience, *AR Façade* does not provide affordances for channel surfing although there is no reason that cannot be part of a future instantiation of the immersive and interactive stories.

Audience members who watched their friends try the experience on the TV screen also enjoyed having distance from the drama. Player 32's friend talked about his ability to see the big picture because he was not forced to react and really deal with the situation at hand:

> *"As an external observer, you know, like I have this big view on the situation and like I'm not sucked into it. You're still isolated from what's happening ….like you aren't actually in there. Whereas if you're (in the AR space), it just draws a different behavior and reaction out of you. She's caught up in the moment in there... you know, she's on the spot and just things are happening all around her." (P32 friend)*

Outsiders to the experience have perspective on what they might say and what they might do. They can reflect on how they might react in that situation without having to actually choose a direction. Player 25's friend illustrated this notion that audience members do reflect and think about the situation, but they do so from the safety of the sidelines:

> *"I think when you're just watching something, it's a lot easier to make decisions about how you feel about it, whereas if you're dealing with it like around you in a pseudo-reality." (P25 friend)*

As I have argued throughout this chapter, the AR interface combined with the second-person narrative voice conjures up emotions that are appropriate to the player-character. In *AR Façade* in particular, the player-character is forced to take on emotions that many participants were not be prepared for. Without any mediation to the interface—no way to directly turn the channel or block out the content—some players were ill-equipped to handle the situation. It explains why a third of the players either quit the experience (10) or got themselves kicked out (5) with their performative antics. As a whole, the comparisons players made with other forms of media expressed the desire by many players to have critical distance from the experience. Explicit mediation would help players manage their emotional distance from the drama.

7.5.4 The Pause/Resume Experiment

As I described in Chapter 3, I designed in a pause/resume feature into *AR Façade* before the Beall installation to see how and if any players would take advantage of the possibility of pausing the action. The concept derived from comments made in the Atlanta comparison study, such as Player 9's comment: "I guess this means that I don't get any of my own dialog." I was curious if players would use the pause feature, for example, to make comments towards people outside of the experience.

In apparent support of an argument for mediation, several players demonstrated a need for a "safe word" so they could distance themselves from the emotional intensity. Player 24 only used the pause feature when he was ready to quit. Similarly, even though she did not use the word "pause", Player 8 held up a time-out symbol when she wanted to escape the situation.

For the most part, the pause/resume feature was not used often during game play. Player 30, however, used the pause/resume on several occasions in support of his tinkerer style of play. Player 30 expressed his desire for wanting to be a "social engineer" to get Trip and Grace into "thinking certain things." (P30) He said it was not satisfying to merely "guide their conversation on like an overarching scale" (P30), so he imagined using the pause/resume feature to be more specific with his interactions.

> *"...to react to certain situations where you only have a small window of opportunity, I thought you could like PAUSE, like try to plan out, 'Oh, if I*

say this, then this will happen maybe' or 'If I say this, this will happen.'" (P30)

In practice Player 30 did not use the pause feature in response to fast dialogue, as much as to give himself a chance to experiment with the system. Player 30 went on to envision how he would control the experience if he could not only pause, but rewind and replay a section of the drama:

"If you could ...rewind, you know, change the situation and kind of evaluate… the conversation that they were talking about… Like where Grace asked me a question, 'yes or no,' you know and I answered like, 'Yes, it's weird for people to be super dependent on their spouses.' If I could rewind and say like, 'no' or if I could rewind and say like, 'Oh, Grace, stop talking.'" (P30)

Façade's original design intent was to support Murray's notion of replay story (Murray, 2004) by supporting multiple replays of the entire evening with Trip and Grace. Player 30 imagined being able to replay specific social interactions on a micro-scale. He even had ideas for how the interface should work:

"I think like saying [pause] out loud kind of like destroys the fantasy a little bit... maybe if you have a glove with like a button or something like that you could just rewind. ... you could see the characters like moving backwards, then you can say, 'Oh, okay, I wanna stop here.'" (P30)

Interestingly, Player 30 said his "fantasy" for pause/rewind would not be well-served by verbal commands. He wanted the fine control of tangible buttons and the immediate feedback of seeing the characters reverse their positions. Where a button interface may disrupt the illusion of "reality" for most players, Player 30 conceptualized

AR Façade as a simulated game world. His fantasy would be disrupted by interfaces that are too much like reality.

Player 30 had actually played the original version of *Façade* before taking part in the two-week player investigations at the Beall, and so he also had a number of observations about the difference between the two versions of the interactive drama. In particularly, he appreciated being able to replay the desktop version of the game to try out different things.

> *"The main advantage I see with the desktop version over this is you have the ability to replay over and over and over again like without consequence. You're kind of in this uninhibited environment… In the PC version, I was just more like experimenting with different things. And I was just kind of like, 'Oh, this will be fun to say...' I could type whatever I want pretty much. I could tell them to like go F-off or whatever." (P30)*

Player 30 preferred the desktop version because could say whatever he wanted and he didn't have to worry about consequences. As a tinkerer style of player, the desktop environment gave Player 30 a chance to be more exploratory:

> *"(referring to KB...) this game's like great for like, you know, experimenting with social interaction, so, you could like go up and just start kissing Grace over and over again (Chuckle). And, I don't know, that just would seem kinda weird if you were to do it in like this sort of setting (pointing towards AR)." (P30)*

Player 30's preference for desktop interaction was commensurate with his vision for the pause-resume feature. Player 30 was not looking to become emotionally connected to the characters, but to distance himself and to dehumanize the characters. He wanted fine-grained control of the social situation, but the unmediated AR interface failed to give him clear affordances for how to significantly influence the simulated events.

7.6 Chapter Discussion

In this chapter, I provide evidence that both the first-person immersive interface and the second-person narrative voice of the script supported players' transition into the character role in *AR Façade*. Players understood their place in the script, and in the absence of necessary back-story, some players would improvise to "fill-in" a supposed history with the characters. In addition to the physical presence induced by the interface, many players felt socially connected to the characters and would act out physically and verbally within the dramatic moment. The combined effect of the story, interaction, and interface led people to feel "on-the-spot" and uncomfortable with the immediate situation. Many players reacted emotionally and sought tactics to give themselves more distance from the intense drama. Players' desire for emotional distance was evident in their play styles and through quantitative differences in physical interpersonal distance. My analysis leads me to argue that women were more emotionally involved in *AR Façade* than men. However, both genders expressed a desire for more distance and an appreciation for the desktop interaction which appeared to provide that distance.

An interesting finding from this work is that an increased sense of presence in the immersive AR experience did not result in an overall increased sense of embodied narrative engagement for all players. While most players felt a heightened sense of presence in AR, many would still prefer a desktop environment (the first 12 players were split 50/50 on their preferred interface version).

Critically, some of the players preferred the desktop system, not because of technical limitations in the AR system, but rather because they deliberately wanted a sense of distance in order to engage more comfortably with it. Player 3 provides a clear case of this finding. As an engager, she said things such as, "I feel really bad…oh man… I felt like I could have helped a little more" (P3) and later, after all three versions, sharing, "I liked the story… I was emotionally caught up in it, just like I would be if I was experiencing it in real life" (P3). She spoke at length about feeling presence in the AR version; she physically acted out during dramatic moments; and said she wasn't bothered by technology anomalies. Yet Player 3 preferred keyboard-based desktop interaction. She felt stifled in the AR interaction because it seemed too realistic, intensifying the severity of the dramatic moments. Players that represented other play styles also specified a preference for desktop interaction when they felt it better supported their desires to observe, tinker, or so on.

In the context of the notion of a "magic circle" that defines the boundaries of the game experience, these players seemed to be having difficulty creating a safe "circle" to escape to and "goof off" in (Huizinga, 1971; Salen and Zimmerman, 2003). *AR Façade* puts people into a very realistic space and a real-to-life social scenario that may not provide the distance players need to enjoy the simulation. Players talked about having a hard time being someone else in *AR Façade.* Assuming another persona in desktop *Façade* requires typing what they say; the immersive interface in *AR Façade* implies the need to adopt the behavior of the persona as well, which is a very different kind of play experience (i.e., talking vs acting). Some players indicated that the ingredients for dramatic involvement—such empathizing with characters and reflecting on the social scenario—came easier when are they are a step removed from the situation.

In this chapter, I argued that the true-to-life second-person narrative structure and first-person immersive interface collaborated to elicit emotional reactions by players in *AR Façade*. However players really wanted to have more emotional and intellectual distance from the character-role so they could engage the way they wanted. Media designers should provide explicit mediation—even at the expense of presence—to give players the ability to manage their distance from the medium and to support the wider goal of embodied narrative engagement.

The findings from this chapter raise a number of research questions. At a high-level, will players have an inherent desire for emotional distance from all genres of immersive and interactive stories, or is this desire explicit to the intensity of drama? While I could speculate, it would be difficult to answer this without more content exploration of immersive and interactive dramas. On a lower-level there appears to be a relationship between field-of-view, interpersonal distance, and players' subjective opinion of closeness. This is a relationship that could be studied within a controlled lab setting. Moreover, the "safety" of desktop interaction is just one example of mediated distance—what other forms of mediation would also allow for distance and divergent styles of play within a first-person immersive and interactive story? In Chapter 8, I discuss some of the key questions that emerge from this chapter and present ideas for future research.

CHAPTER 8

FUTURE RESEARCH FOR IMMERSIVE AND INTERACTIVE STORIES

I am enough of an artist to draw freely upon my imagination. Imagination is more important than knowledge. Knowledge is limited. Imagination encircles the world. —Albert Einstein

Once you are liberated, you are forced to ask who you are. —Jean Baudrillard

My dissertation explores immersive and interactive stories conceptually, tangibly, and empirically. In Chapter 2, I define the key material properties of the medium and the experiential pleasures that can arise at the integration of immersive interfaces, interactivity, and narrative structure. I pose the theoretical construct of embodied narrative engagement to describe a potential for simultaneously feeling the illusion of *presence* (a feeling of being within an environment), *agency* (a feeling of empowerment over events), and *dramatic involvement* (a feeling of being caught up in the plot and characters of a story). In Chapter 3, I present the collaborative project, *AR Façade*, as a prime exemplar of an immersive and interactive story. I describe the project's genesis (with details about the original *Façade*), the technical details (including mixing virtual/physical content and enabling natural speech and gestural interaction), and the construction of two installations of the AR experience (one of which took place in a free, public gallery over a three-month period). In Chapter 4, I outline my empirical research goals and my strategies for data collection and analysis. I begin an analysis of the player data by commenting on the range of factors that might influence the experience of *AR Façade*. In Chapter 5, I reveal different styles of play that emerged during both installations of the experience and support this analysis with a series of five case studies with episode visualizations, episode excerpts, and player quotes.

Towards the primary thesis of my dissertation, in Chapters 6 and 7, I attempt to isolate and investigate the effect of the physically immersive interface. In Chapter 6, I provide evidence that the immersive AR interface provided human-like affordances for interaction, but largely failed to maintain agency. I argue that to optimize a sense of agency, participants should not only be invited to enact a wide range of actions, the actions must be met with an adequate response. In Chapter 7, I show how players naturally masquerade in the character role only to find themselves with nowhere to escape and looking for emotional distance. Some participants seeking dramatic involvement may benefit from a degree of detachment—distance that would allow them to safely reflect on the story themes and empathize with other characters.

The primary argument of my dissertation is that achieving the theoretical concept of embodied narrative engagement (ENE) requires some mediation. The strategy of perceptually immersing the user and minimizing mediation to create the sense of presence does not maximize the sense of embodied narrative engagement, because complete transparency hides interaction mechanisms that strengthen the sense of agency and does not provide sufficient means for users to manage their distance from dramatic content.

ENE may not be the ultimate goal of all media designers, nor does ENE emerge as a similar experience across all people. I argue that media designers exploring immersive and interactive story experiences should provide mediation, because complete transparency hides interaction mechanisms that strengthen the sense of agency and does not provide sufficient means for users to manage their distance from dramatic content. While a move against immediacy would be seen as counterproductive by many presence researchers, mediating the experience—by explicitly building in constraints, structuring manipulation hooks, and allowing for story experimentation—reinforces the overall engagement of the experience by strengthening agency and providing for emotional and intellectual distance. While some players managed to achieve a strong sense of presence, agency and dramatic involvement—an overall high sense of embodied narrative engagement—most players would have benefited from clearer affordances and the ability to better manage their distance from the drama.

In this chapter, I reflect on the findings and pull in additional player comments to outline a future research agenda around immersive and interactive stories. Based on our experience creating *AR Façade* and the results of my empirical studies I present a series of research questions and possible future work towards answering the questions. Many of these ideas are speculative, long-term lines of inquiry; others build on specific findings and can be investigated with near-term empirical study.

My reflection on research questions and for future research follows five themes: (1) the design process and construction of immersive and interactive stories, (2) descriptive and operational models of play style, (3) methods for constraining loosely-constrained immersive interfaces, (4) considerations for managing distance in immersive interfaces, and (5) limitations and potential for future immersive and interactive stories.

8.1 Constructing Future Immersive and Interactive Stories

My experience designing and creating *AR Façade* revealed a number of possible research vectors to explore. In Chapter 3, I described several key challenges for augmented reality versions of immersive and interactive stories, including character rendering, speech and gesture recognition, and physical/virtual interaction. I also presented findings from the Beall installation of *AR Façade* that revealed interesting strategies employed by our wizard docents to maintain an engaging experience for players and audiences. Our creation process for *AR Façade* raises questions and possibilities for future work on specific implementation details and design methodologies for immersive and interactive stories.

8.1.1 What effect does the rendering style of virtual characters have on the player experience in immersive and interactive stories?

In the *AR Façade* experience, despite the cartoonish renderings and the mismatch between the characters and the live video backdrop, players rated the characters as quite believable (5.2 out of 7 for N=45 players) and appeared to experience a strong sense of social presence. One could examine this question empirically, using a mixed-methods approach. My hypothesis is that more visually realistic characters would not bring about a stronger overall sense of embodied narrative engagement because it would continue to raise expectations for interactivity. As my thesis indicates, if expectations for interactivity are not matched, a sense of agency diminishes and the sense of embodied narrative engagement cannot be maximized. This raises another question at the intersection of presence and agency: what is more important towards an overall sense of embodied narrative engagement and character believability—visual appearance or behavior? My intuition tells me the more important factors are character behavior and the responses that define a character's personality and fortify the player's sense of agency, not how the character looks.

My claim is not unlike other research that indicates users can feel a stronger sense of agent believability with non-anthropomorphic forms (Reeves and Nass, 1998). I believe a comparative study of different character renderings would show us that players are more forgiving of agency gaffes with a cartoonish Trip and Grace, than with 3D avatars or video-based actors. I would also hypothesize that measures of social presence are higher for more "reality-like" renderings, but only when the characters are in passive, non-interactive modes. Of course measured improvements in both the interactive narrative mechanics and visually-realistic characters renderings would likely result in higher ratings for social presence and agency—and thus a stronger sense of embodied narrative engagement. This research could help us to further understand the relationship between agency and presence.

8.1.2 Can Wizard-of-Oz methods be used to guide the design and implementation of immersive and interactive stories?

I believe improving the design setting will encourage more artists and media designers to explore the immersive and interactive story medium. New prototyping tools and methodologies will be advantageous for building "naturalistic" interfaces (such as speech and gesture interaction) and creating better interactive plots and characters. Based on my investigations of "Wizard-of-Oz" methods during our installations of *AR Façade*, I am particularly motivated to look at methodological improvements to the design setting.

My initial research strategy is to deploy and investigate Wizard-of-Oz (WOz) methods in various application contexts (not limiting myself to only IIS experiences) to simultaneously advance technology and design goals. I have previously argued for the use of WOz methods throughout an iterative design process, and demonstrated that a wizard can change their role over the course of a project, gradually shifting responsibility to the underlying system (Dow et al., 2005). I would like to create a mature theory of wizard-guided design and implementation that would explore the use of one or more collaborating "wizards" fulfilling various system tasks, with varying levels of responsibility, throughout a user-centered data gathering process. My agenda will be to quickly construct prototypes that can be evaluated with people while gathering useful design-specific and system-specific data. Wizards can be deployed early in the design process when the system and specific technologies are ill-defined.

Working with uninitiated wizard docents at the Beall Center exhibit of *AR Façade* has fueled my motivation for a more extensive use of wizards throughout a design process. As I discussed in Chapter 3, these non-technical wizards not only successfully performed complex tasks (serving as either a speech recognizer in the "dialogue interface", a natural language parser in the "discourse interface", and a gesture recognizer in both), they provided insights about player behavior on par with members of the research team. Interviewing the wizards turned out to be fruitful for my player investigation, because the wizards played an intimate role in making the application work. The wizards could speak to the overarching goal of engaging the players, as well as to low-level problems that cropped up repeatedly. The unique perspective of repeatedly fulfilling a system function can yield operational insights.

A wizard-guided design process would not only lead to deeper insights about players, it could result in domain-specific wizard-tagged input data. From the system technology perspective, this corpus of "training" data could be analyzed towards the development of sensors, algorithms, and software architectures to optimally satisfy the design goals. Within a full design cycle, the extensive use of wizard methods could uncover eventual roles for human operators. As a "wizard-guided" application evolves, it would become clear if a specific wizard's role can be eventually replaced with technology or if a human will perpetually fulfill some piece of the system function (i.e. an outsourced service role). My goal will be to explore the idea of a wizard-guided prototyping and implementation method within a specific domain while documenting how a team of wizards utilizes the method during a design process.

8.1.3 Can Wizard-of-Oz methods iteratively improve speech and gesture interfaces?

The conversational and corporeal nature of immersive and interactive stories solicits a need for better speech and gesture recognition technology. Ultimately there will need to be technological improvements in sensors, digital signal processing, and software engineering to achieve reliable natural speech and gesture interfaces. Until there are significant improvements in the technology, speech and gesture interfaces will have to be designed and prototyped within the constraints of specific application domains (e.g., a video game, a digital medical chart, an educational tool). In *AR Façade*, I investigated the idea of a wizard-guided design process—specifically for speech—and assumed a constant human presence behind the scene.

There are two valid approaches to designing speech interfaces with the aid of wizard methods. One approach, implemented in Suede where realistic error rates are actually artificially injected into perfect wizard responses (Klemmer, 2000), asks: how will users perform under practical technology constraints? Another approach is to ignore specific technology limitations and ask: how will users behave under ideal conditions and does that negatively effect their experience? In *AR Façade*—as is the case for many immersive entertainment experiences—it is acceptable and necessary to strive for recognition rates that may never be achievable through technology. As I pointed out in Chapter 3, most players did not have a clue how the speech interaction worked, nor did they suspect a hidden human operator. I believe the overall player experience would have suffered from the error rates of actual speech recognition technology. In fact, despite the "perfect" technology, the conversational issues (e.g., time delay) were the most significant problems for players to overcome.

I took advantage of the constant human presence to investigate specific different wizard infrastructures at the Beall installation and to push towards more fluid conversation with virtual characters. As I described in Chapter 3, I developed two wizard interfaces—a speech recognition “dialogue interface” and a user intention “discourse interface”— used to mediate the speech and gesture interaction within the experience. In both interfaces the novice wizards achieved the goal allowing the players to use natural speech and gesture, however without deeper investigation it will not be clear which interface formulation is better.

The first open question would be to calculate whether the time delay is shorter for the dialogue or discourse method of wizarding? Based on my observations and interviews with wizards I would hypothesize that response time is a function of the complexity and length of utterance. I sketched a theoretical “delay curve” in Figure 8.1 to explain this relationship. The Dialogue method takes progressively more time to perform the task, while the Discourse method remains steady since it is simply requires an interpretation of meaning. There would be a point at which the length and complexity of the player utterance would cause the Dialogue WOz method to take longer to perform (marked by the ‘?’ in the diagram).

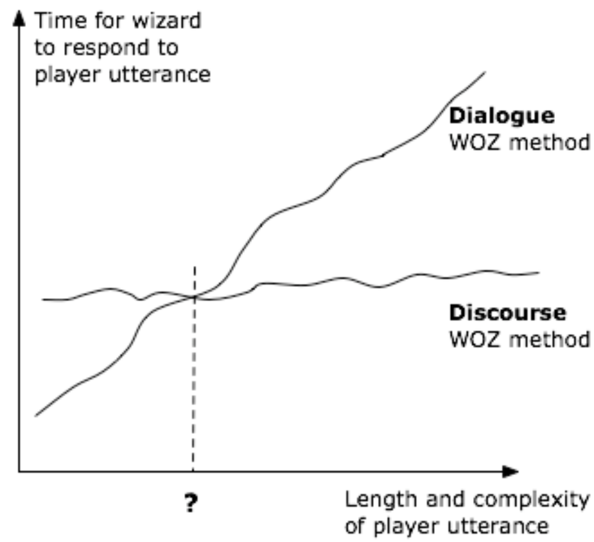


Figure 8.1: Theoretical "delay curves" for the two wizard methods

Other specific questions include: what is the average baseline time delay of the discourse method? what is the time delay curve of the dialogue method? at what point do the two curves intersect? I believe it is possible to answer these questions quantitatively from the data I have collected. It would require an extremely detailed conversation analysis marking times for the start and end of each player utterance as well as the time to enter dialogue or select a discourse. I already attempted to segment the audio from the game episodes to locate the start and end of each speech utterance, but I was not successful using the best free audio analyzer I could find (Annosoft Lypsync).

The specific answers to these questions—such as the location of the '?' along the x-axis of Figure 8.1—are highly dependent on the conversational context and the system constraints. For example, as the number of high-level discourses represented in the system increases, the more difficult it will be for wizards to discern and select between them (it will also be harder for an NLP to accurately process language into those categories). While the exact value of this calculation is not generally useful, the concept that different wizard infrastructures have different affordances can provide tremendous flexibility for designers. Not only does it speak to final performances of immersive and interactive stories, it provides a wider breath of strategies for hoisting a partially complete system. For example, could wizards be in charge of adjusting “tone” variables in an otherwise complete interactive system? Could wizards directly respond to the emotional state of the participants? Rather than discourse constructs like “flirt”, perhaps the wizard adjusts system variables representing the player’s emotional state and goals, and the system can be programmed to respond directly or to distort the participant’s current state. My vision for wizard-guided system design and development would allow developers to experiment with underlying software architectures.

8.1.4 Are there specific software design tools that would aid the design process for immersive and interactive stories?

My ideas for software tools are mostly geared towards the immersive aspects of these experiences, rather than interactive story elements. However, since the release of *Façade* in 2003, there have not been any other full-featured interactive dramas, partly due to a lack of knowledge for how to create one. Mateas argues for the “programmer-artist” model, and has suggested a need for stronger computational education in media programs to bolster the knowledge required to program interactive stories (2005). When Mateas and Stern created *Façade*, they also created ABL (A Behavior Language), a software language for simplifying to development of interactive characters (2004b). While I agree with the need for programatic knowledge, I think more can be done to improve the design environment for developing future interactive dramas.

Staying within my expertise, my ideas for software design tools and potential future research work will focus on the design of immersive environments, particularly those that seek to integrate virtual and physical space.

8.1.4.1 Replay reality tools

Mixed reality and ubiquitous computing applications generally incorporate various sensor data (including cameras) from the physical environment and then provide multi-media output to displays for human users. An effective design environment for these types of applications would continuously capture sensor data so that designers could selectively archive chucks of "reality" as building blocks for testing applications. For example, we integrated the idea of capturing and playing back sensors into DART (the Designer's Augmented Reality Toolkit) (MacIntyre et al., 2004). Under this model, media design can be iterated and perfected "on top" of sensor archives and deployed later to the actual sensor environment. When an application is deployed for user testing, the design tool could continue to archive sensor data which would be based in part on contextualized user actions.

8.1.4.2 Simulating physical and virtual worlds

For applications that span virtual and physical worlds, it would be useful for a design environment to simulate the entire design setting. For augmented reality applications, I imagine a hypothetical software tool with a 3D environment (initially desktop-based) with two layers: the modeled physical world that will be the stage for the application and media content that would be added as "virtual" augmentations of the world. For many AR applications, if would be particularly useful to be able to quickly sense and construct a real-world model to serve as a basis for the "physical reality" layer, such as the work proposed by Lang et al. for exploring AR in Second Life (Lang et al., 2008). If such a "sensor" existed, it would allow designers to build AR applications for a wider-variety of locations and to adapt to different formations. I imagine the software tool supporting a "user-view" to preview the users' view of content as it is overlaid on the physical world (similar to the "stage" in Macromedia Director). This hypothetical tool would also support a "world-view" where all physical models and virtual content could be manipulated independently of how they appear in the user-view. This independent view would be particularly useful for working on the tracker/camera alignment with the physical world. I imagine a timeline (or multiple timelines) for controlling the appearance of virtual content and for placing selectively-archived chunks of "reality" as temporary input for the system.

8.1.4.3 Run-time tools

While creating immersive experiences it would be beneficial to be able to tweak certain features of the design in run-time, as an experience is taking place. In this hypothetical tool, I imagine including support for run-time manipulation of some features to allow for on-site changes beyond the designs constructed in the simulated environment. For example, the position of virtual characters might be initially positioned within the simulation environment, but the designer should be able to tweak and test the character staging while under an actual sensor infrastructure (i.e., 6DOF trackers). Again Lang et al. have explored some of these concepts in Second Life (2008). For AR applications, in addition to being able to fluctuate the positioning of virtual content, I imagine a run-time manipulation tool providing control over a host of properties of the virtual content, cameras, and sensors. We actually built one run-time manipulatable feature into *AR Façade* for the purpose of manually aligned virtual content with the physical set. It was difficult to accurately predict the position of furniture on the *AR Façade* set before we arrived at the Beall Center. To quickly re-position modeled physical objects—such as the model of the bar which is used to occlude Trip if he walks behind it—we built in the ability to transition models between transparent (as they are normally) and opaque (so they can be visually and manually aligned with the actual physical object in the camera).

While I have provided a few rough ideas for a future design tool for AR/VR, the challenge would be to demonstrate an improvement over current design environments, such as DART (MacIntyre et al., 2004). The research agenda would first involve building an elaborate software design tool and then evaluating it to prove it works. Both of these tasks would involve an enormous time investment, but could enable others to develop immersive and interactive stories which could lead to answers to research questions about the player experience.

8.2 Modeling Play Style

In Chapter 5, I proposed five unique styles of play—engager, performer, tinkerer, observer, and partaker—in immersive and interactive stories based on my qualitative analysis of forty-five players across the two installations of *AR Façade*. I arrived at these descriptive categories after a post-hoc analysis of player behavior and documented them through case studies. As I point out, pigeon-holing players into one particular style of play does not capture the nuance and variability of all personalities, but it provides a starting point for developing run-time, quantitative measures of play style. There are two research questions that could provide a path for future work: how can play styles be detected and why does it matter?

8.2.1 Can style of play in immersive and interactive stories be detected in run-time?

As I illustrate in Table 5.17, it appears that monitoring conversational activity may help to identify the observer style of play—it calls out participants who do not converse much. The more talkative players are likely either engagers or performers. Episode time and endings are not useful indicators during run-time, but I assert there are other ways to detect play style. Towards future work on understanding a player's experience of immersive and interactive stories, I suggest two paths forward: surface-level environmental sensing and context-specific language analysis.

8.2.1.1 Environment sensing

Physiological measures could be useful towards a detection of play styles, but they are problematic in emotionally engaging immersive and interactive stories. For example, the physicality of performing wild gestures probably raises the heart rate above a passive observer, even if the observer is highly emotionally engaged. I believe there are less intrusive ways of detecting behavioral patterns that take advantage of sensing the "reality-like" interactions in immersive interfaces. Simple audio level detection could detect the difference between loud performers and soft observers. Similarly, I noticed audible heavy breathing and sighs, particularly from those players who were emotionally engaged in the drama. In terms of physical behavior, exaggerated gestures are more likely part of the performer experience, and again observers would likely have no gestures at all. The use of frequent repetitive gestures is a potential sign of the tinkerer style of play. When these environmental outliers are considered within the story context, they can become even more telling.

8.2.1.2 Language analysis

Studying the details of what participants say can also contribute to a detection of the play styles I have identified. For example, a simple measure such as whether or not players repeat statements might help identify tinkering. Short one word statements are more likely from observers and tinkerers. A more complex language analysis could detect whether players stay or diverge from the story context. For example, if players in *AR Façade* start talking about topics that are not part of the current beat and not within the scripted 'satellite' topics, they are likely to be a performer or tinkerer. Engagers and partakers are more likely to stay within prescribed story lines. As other researchers have investigated, language analysis can also be used to discern content-specific strategies (detecting killers vs. explorers). In *AR Façade*, this might mean discerning between subtle content-related strategies used by engagers (do they tell Trip and Grace to seek help or do they try to get them to open up?)

Table 8.1: Hypothetical run-time indicators of potential play styles

Play styles **Possible indicators**	**engagers**	**performers**	**tinkerers**	**observers**	**partakers**
Audio level (volume of speech)	normal	loud	normal	quiet	normal
Non-language utterances	sighs, heavy breathing	laughing, screams	normal	none	laughing
Physical gesturing	normal	exaggerated	repeated, mechanical	none	normal
Quantity of language use (see Table 5.17)	frequent	frequent	normal	low	normal
Repeated statements	occasional	no	yes	none	no
One-word statements	few	few	yes	yes	few
Use of language in relation to story context	Stays on topic	Diverges from topic	Diverges from topic	Stays on topic	Stays on topic

Table 8.1 hypothesizes how these different styles of play could be differentiated in run-time through a combination of surface-level sensing (e.g. audio levels) and context-dependent analysis (on topic vs. divergent). This is meant to be starting point for a potential map of indicators. Some of these (such as detecting non-language utterances) may not be possible with existing technology. The Wizard-of-Oz method could be utilized as a stop-gap measure to “detect” all of these features in run-time using a small team of hidden operators (e.g. one wizard marks non-language utterances, another keeps track of topic convergence, etc.). As I outlined in Section 8.1, this WOz-guided approach

has the benefit of generating a corpus of "tagged" data that can be later analyzed toward the development of actual detection technologies.

8.2.2 How could an immersive and interactive story adapt to players' style of play?

Assuming progress on the previous research question, the underlying interactive story architecture could be adjusted automatically by "playing off" of these play styles during run-time to create a more satisfactory experience. If the system thinks the player is exhibiting a performer style of play, perhaps it piles on the absurdity and tries to match the crazy behavior of the player. Tinkerers might benefit from highlighting the interaction mechanisms and providing fun rewards for messing around with the interactions. Or perhaps, the experience could sarcastically make fun of itself, calling out the fact that the characters are computer-simulated representations. If the system detects the partaker style of play, an indication of disinterest, perhaps it should change topics. If it detects engager play, perhaps it should continue to push for the intended emotional responses. Observers might need to be called out of their shell (in *Façade* for example, Trip and Grace will ask 'what's wrong?' if the player has not interacted for a while). Or perhaps observers simply need space to stand back and watch the story unfold with having to interact. While this conversation is somewhat speculative, it could provide a way forward to more overall engaging experiences in immersive and interactive stories, because it accounts for the different approaches people take towards the medium.

With progress on adaptive story architectures, methods would have to be developed to understand their effects on player experience. One challenge for evaluations of adaptive story experiences is that there would be no pre-determined classification of players and it would be difficult to understand how the experience appeals to various play styles. A mixed-method approach would be appropriate—quantitatively measuring how players interact within the system and qualitatively gathering accounts of player satisfaction. Success would be marked by a broad exploration of the narrative space across different players and overall high ratings for player satisfaction compared to non-adaptive story forms.

8.3 Constraining Loosely-Constrained Immersive Interfaces

In Chapter 6, I demonstrated that players felt a strong sense of presence in *AR Façade*, but their sense of agency suffered due to a lack of interaction affordances. The desktop version of *Façade* did not perceptually immerse the player, but it did provide clear affordances. Players felt a stronger sense of agency at the desktop computer than they did in AR. My analysis suggests that there is a tradeoff between unconstrained immersive interfaces that strive for presence and carefully-constrained interaction mechanisms that emphasize agency. Based on my preliminary studies, I believe a number of empirical investigations can be derived to help understand this tradeoff.

8.3.1 Do clearly afforded interaction mechanisms increase a sense of agency and do they sacrifice a sense of presence within an environment?

I want to further investigate the premise that a tradeoff exists between agency and presence. This could be investigated empirically through a series of lab experiments where players play two versions of *Façade* (although this could be conducted with other game experiences). I would record how players interact (amount of movement, dialogue inputs, etc.) within different formulations of *Façade*. At the end of each game episode, I would attempt to "measure" players' sense of agency and presence (initially through a questionnaire).

There are two transformations of the interface that I believe could help elucidate a possible tradeoff. First, I would compare the normal typing interface to a menu-based system where the user selects from a list of discourses (i.e., a list of 3-5 things the player could say at any particular moment throughout the conversation, much like current RPGs employ). The menu-based interaction would not rely on a natural-language parser, but would prompt a list of possible statements based on a traversal of currently available character behaviors. The hypothesis is that a menu-based system would provide players a stronger sense that their inputs gave them control over the scenario; the dialogue interface would provide players a stronger sense of expression and social connection to the characters.

A second interface comparison could be established to investigate speech-based interaction. I would compare the unconstrained speech interface employed in the Beall Center exhibit with a speech-based interface that allows players to preview their verbal statements and then either "enter" or "clear" their utterance. The "preview" feature would be partially achieved through the Wizard-of-Oz method where the wizard would type what the player says (as we did in the interface comparison study), but it would be up to the player to decide when to "enter" the statement. A similar hypothesis is behind this research study: the unconstrained speech interface would provide a more expressive and socially connected interface to the conversation with virtual characters, but the "preview" feature would give players a stronger sense of agency. Assuming these studies would help to illuminate a tradeoff between agency and presence in immersive and interactive stories, there would be different design philosophies moving forward—whether to attempt to "optimize" the tradeoff or to use mediation creatively to throw presence and agency out of balance.

8.3.2 If a tradeoff exists, is there a way to optimize the sense of presence and agency?

Assuming the goal is to create an interface with "reality-like" interaction (i.e., natural speech and gesture), how can the system be designed to maintain a strong sense of agency? Some designers may continue to adopt the philosophy of simulating "reality" for immersive and interactive stories. I believe the way to achieve agency within this of experience is to subtly embed interaction affordances as conversational and cultural cues. The yes/no questions in *Façade* are an example of an interaction affordance built into the conversation. Conversational cues can have the ill-effect of making the script seem too planned and limited (e.g. Trip asks the player to describe the Italy photo in just one word), but they can also be subtly integrated into the social context (e.g. the characters ask the player to "help us out here"). The narrative context can also be constrained based on common social cultural protocols. *Façade* succeeds in this regard by playing on players' knowledge of greeting dynamics and drink requests. The experience also includes culturally understood artifacts, such as the wedding photo and the "8-ball" toy. People brought up in western culture generally understand how to interact with these objects (i.e., people know they can shake an 8-ball toy and ask it an odd-ball question.)

This philosophy will lead designers to subtly constrain users and to smooth over technology limitations in future such conversation-based interactive stories. For example, an experience could minimize the effect of the "time delay" problem by simulating a social situation about conversing with foreigners and speaking through a language translator. Or, perhaps the game context forces the player to speak through a "walkie-talkie", where players would be required to press and hold a button while they speak providing the system a definite window for processing natural speech, and it would lower expectations for an expedient response. At a high-level, the key design idea is to lower player expectations for the conversation without making it feel artificial. At a low-level, much more can be done to "tune" an interactive story for speech, such as adding better non-verbal cues for face-to-face conversation. As Player 14 hinted there are subtle communication cues that happen in the flow of conversation: "You and I are talking, right? You're listening to me, you're nodding your head, I'm listening to you." *Façade* actually does a decent job with facial expressions, as many players noted during interviews. However, since *Façade* was not originally designed for speech interaction, it is missing much of the "real-time" feedback present in everyday interaction such as nodding and common utterances like "uh-huh" that communicate to a speaker that the non-speaking party is actually listening.

Moving forward with these research ideas would involve iteratively designing and evaluating how various conversational and cultural cues are received by players. Immersive and interactive story prototypes could be simulated (or partially simulated) using WOz methods (see Section 8.1) and then play-tested by large numbers of players. At the level of scripting the narrative, the goal would be to maximize the narrative coverage (i.e., minimize the number of things that the experience does not respond to). At the level of speech interaction, the goal would be to build in enough of the subtle conversational cues that players believe they are having a fluid conversation. I believe the way to improve speech interaction with virtual characters is to study cognitive science and linguistics (also referred to as "cognitive linguistics") to understand the semantics of language co-construction and generative grammar.

8.3.3 Do "non-real" interaction mechanisms lead to more embodied narrative engagement for players?

An alternative design philosophy to embedding "transparent" interactions is to provide more blatant and obvious mechanisms for exerting control over immersive and interactive stories. Rather than disguising the means for achieving interactivity, the environment could adopt conventions from video games (e.g. step-by-step "tutorials", visual hover cues over manipulable objects, heads-up mapping displays, arrow clues, etc.). These types of design features go beyond "reality-like" interaction (i.e., hovering virtual arrows are not part of our everyday experience), but they attempt to mediate the interaction towards a stronger sense of agency. The “purest” philosophy would be to adopt affordances from reality (such as turning a door handle as one would do in the everyday world), where as the “gaming” philosophy would be to trade-off reality for special powers (such as being able to open a door from a distance).

Conversation-based interactive stories that employ speech interfaces could also design more blatant interaction mechanisms. One idea is to provide thought bubbles above the heads of the virtual characters[60]. While the idea goes beyond what is possible “in reality,” it provides a creative way to give the player feedback on what the characters are thinking and how one might be able to influence them. Another idea would be to show the player their own words emanating into the air and then highlight the words that actually impact the direction of the narrative. These sorts of ideas are creative ways of supporting the expressiveness of speech while also providing some conversational feedback. As a research program, these ideas could be iteratively designed and evaluated with play-testers with an eye towards overall player engagement and satisfaction.

It will be interesting to see how the two philosophies for embedding interaction mechanisms—transparent vs. mediated—influence the development of immersive and interactive stories over time. How will an active design practice for immersive and interactive stories evolve under the forces of “immediacy” and “fragmentation” as suggested by Bolter and Grusin? If compromises between agency and presence are traded off, where do they lead the medium—towards some sort of immersive “human-like” haptic simulation? If immersive interface are not destined to become “transparent” simulations of physical reality, than where will the physical form of these interfaces be in the future? These broad questions may only be approachable as immersive and interactive stories are explored more pervasively.

60 Thanks to Kathryn Isbister for this idea.

8.4 Managing Media Distance

In Chapter 8, I provided evidence that both the first-person immersive AR interface and the second-person narrative voice of the script supported players' transition into the character role in *AR Façade*. Many players felt socially connected to the characters and acted out physically and verbally within the dramatic moment. The combined effect of the story, interaction, and interface in *AR Façade* led many players to react emotionally to the immediate situation. Some players developed tactics to give themselves more distance from the intense drama. Players' desire for emotional distance was evident in their play styles and through quantitative differences in their interpersonal distance with the virtual character. Many players expressed a preference for less immersive forms of the experience, not because of technical problems, but rather because they explicitly wanted more distance from the simulation in order to engage more comfortably with it.

The findings in Chapter 7 raise a number of interesting new research questions. There are the broader questions about immersive and interactive stories that would be difficult to answer without more exploration of the space of possible content: Do players have an inherent desire for psychological distance from all media, or is it contingent on specific kinds of content? The socially uncomfortable situation created by *Façade*'s story context strongly contributes to why some players found it difficult to engage. So, for example, will players seek out this distance from "mellow" or "happy" content?

Even if the desire for emotional distance is unique to dramatic-style content, will players eventually become over-exposed and adapt to the most intense and realistic media experiences? Empirically this would be very difficult to answer. This question may only be approachable from a historical perspective garnered after years of media practice in immersive and interactive stories.

Although I argue that a mediated interface can support more diverse styles of play, are there content circumstances where rigid intensity and a lack of mediation are actually more appropriate? The increased sense of consequence experienced in AR may offer a significant advantage when the application provides a scenario where engaging in meta-play ("goofing around" or other deviant styles of play) detracts from the intent of the experience, such as training situations. Moreover, there may be entertainment experiences where a lack of mediation better supports the goal of the genre, perhaps immersive and interactive stories derived from from the "love story" genre of film.

In general, I believe the goal of presence—the "illusion of non-mediation"—does not broadly support the breadth of embodied narrative engagement. I would argue that for most content situations, most players would like to be able to explicitly manage their distance from the scripted character role. The "safety" of desktop interaction is just one example of mediated distance. The "pause" feature added for the Beall installation of *AR Façade,* and exploited particularly by Player 30, only scratches the surface for possible methods of distancing the player from the immersive action.

8.4.1 What other forms of mediation help players manage their emotional distance within a first-person immersive form?

The formulation of a media experience most in line with "reality" is a first-person immersive camera point-of-view and a second-person narrative voice, but adhering to this strict form may not provide designers with the variety they need to explore story themes. My point here addresses an essential design consideration behind my thesis: start with the premise of "reality constructs" and then deconstruct it along various lines. Deconstructing the classic formation of the *Holodeck* reveals new design ideas, including techniques that will provide the mediation that some players desire.

8.4.1.1 Role-switching

Breaking from the notion of a single designated player-character role, role-switching form would explore the ability to play one character one moment and then immediately transform into another. Designers can explore the expressive possibilities of allowing players to act out the same event from different character perspectives. We already explored this to some extent in Four Angry Men where players can leave the body of one juror and move around the table to see the perspective from each jurors' eyes (MacIntyre and Bolter, 2003), but there would be many ways to explore this concept, especially in interactive narratives where the player-character can effect the direction of the plot. For example in *AR Façade*, with a more flexible AI architecture, we could allow the player to switch roles between the visitor character, Trip's character, and Grace's character.

8.4.1.2 Out of body experiences, dream states, flashbacks, other non-embodied states:

These common filmic techniques do not conform to the strict "reality-like" form of immersive and interactive stories, but they could be adopted to create disembodied sequences to great effect. Designers could make use of non-photo-realistic rendering techniques and other forms of "hyper-mediation" that would classically break a players sense of presence. Along this line, designers can explore more avant garde immersive experiences, intentionally distorting the notion of a "normal" experience.

8.4.1.3 "God-like" overviews:

This common design technique in video games could be adopted to immersive and interactive stories by allowing the player to "rise out" of their character-role to view themselves (perhaps as an avatar) and to see a broader section of the simulated world.

Designers can choose to provide players with super-powers allowing them to do things that are not possible in reality.

Generally, our phenomenological experience of the real world is based on the first-person perceptions enabled through our bodies, but the medium of immersive and interactive stories do not have to follow this strict formation. Even HMD-based augmented reality experiences—where the camera and display is mounted on the participants' head—does not have to abide by a first-person point-of-view. One could imagine shifting to a 3rd person camera viewpoint—such as a ceiling mounted camera pointing down on the player (likely creating a disjointed experience). I believe the craft in designing these experiences lies in thinking about the assumptions built into "reality-like" interactions and then deconstructing and artistically manipulating them to reach a desired effect. Just as the three material properties of immersive interfaces, interactivity, and narrative structure can be balanced towards an overall goal of embodied narrative engagement, they can be purposely thrown out of balance as a design choice. Thus far this discussion has raised largely broad questions that are difficult to answer, but there are several concrete directions for future research to understand the effect of such transformations.

8.4.2 Do players feel less emotionally involved in a first-person immersive experience that does not use second-person narrative voice?

The combination of first-person immersive interfaces and second-person narrative voice most closely resemble a person's experience of the everyday world. If one diverges from that specific formation, does it provide media distance? Do participants feel like the experience is less intense? Designers may want to explore other narrative voices in immersive interfaces so they can depict intense drama without overwhelming the user. Player may feel less "on-the-spot" if the characters do not refer to the player character by name. This could be evaluated in an empirical setting by comparing two experiences with the same interface and the same story, but altering narrative voice as an independent variable. So in one case, characters look at and refer directly to the player as “you”. In the other case, characters never look at the player and refer to another person in an abstract sense. The hypothesis is that the experience using second-person narrative voice will elicit more emotions (perhaps measured through physiological sensors) and higher subjective ratings of intensity.

Do players feel less emotionally involved in an experience that employs second-person narrative voice, but diverges from the first-person immersive interface? This question addresses the other side of the coin: what happens when the immersive interface is not strictly first-person? Do participants feel like the experience is less intense? Designers may employ second-person narrative voice where characters in the story refer to the player by name and by use of the pronoun “you”, and yet explore a third-person viewpoint (even in immersive HMD or CAVE displays). Participants would see a representation of themselves (as an avatar in VR; perhaps a 3rd person camera view of themselves in AR). The question is whether this alternative formation of the interface feels less intense? Do players feel less “on-the-spot” because the characters are looking and talking to a representation of the player, not directly at the player?

Again, this could be evaluated in an empirical setting by comparing two experiences with the same narrative voice (second-person) and the same story, but altering the camera viewpoint as an independent variable. So in one case, characters look directly at the screen and in another characters look at a representation of the player, which the player sees from an outside view. The hypothesis is that the experience through the first-person camera viewpoint will elicit more emotions (perhaps measured through physiological sensors) and higher subjective ratings of intensity.

8.4.3 How does field-of-view effect interpersonal distance and players' subjective ratings of emotional connection?

The data presented in Chapter 7 indicated that as the field-of-view of the virtual camera in desktop-based *Façade* increased the players needed less interpersonal distance between themselves and the characters. The evidence also suggested that players felt a stronger sense of social connection and emotional "closeness" to the characters when standing closer (in AR versus KB), which supports prior social science research on IPD and intimacy. There appears to be a relationship between field-of-view, interpersonal distance, and players' subjective opinion of closeness, but it would require further empirical investigation to understand how they relate. Assuming a constant interpersonal distance, one hypothesis is that players' subjective ratings of the intimacy of virtual characters will increase as the field-of-view narrows because the characters will feel physically closer. Another way to investigate it would be to allow players to set their IPD at a distance they feel comfortable, a "socially appropriate" distance. The hypothesis would follow that players will stand much closer to characters when they experience a wider field-of-view. I envision an experiment much like the Bailenson experiment on IPD, only with field-of-view as an independent variable.

8.4.4 What is the effect of gender on players' subjective ratings of emotional connection in immersive and interactive stories?

My analysis from Chapter 7 led me to argue that women were more emotionally connected than men to the dramatic scenario in *AR Façade*. Women were more likely to fit into my "engager" style of play, while more men were categorized as "partakers." Women seemed to take the situation more seriously, gave the characters more interpersonal space, and departed the experience earlier. Men were more likely to "hate both equally," and seemingly dehumanized Trip and Grace and the overall scenario. This is my subjective analysis of the data collected in the *AR Façade* studies, but I believe further investigation is required before making strong claims about gender differences. As a starting point, I have additional qualitative data to analyze from the *AR Façade* studies, but more media examples would have to be designed to really understand if women are more effected by conversation-based interactive drama. In *AR Façade*, both genders expressed a desire for more distance and an appreciation for the desktop interaction which appeared to provide that distance.

8.5 Exploring the Potential and the Limitations

In this chapter I have discussed practical design considerations for the construction of future immersive and interactive stories, and further research questions that have emerged out of my player investigations of *AR Façade*. In this section, I will consider the future directions for the content of the medium. What other content would be appropriate for immersive and interactive stories? What did the *AR Façade* players feel about the potential of the medium? What sort of experiences are possible, considering the limitations of the medium? How do the limitations serve as useful design constraints? Do the themes discussed throughout this dissertation apply to human-robot simulations?

One way to ponder the question of future content is to think about current genres in film and video games and then project forward. People have a wide variety of interests—not just dramatic content—so for immersive and interactive stories to thrive, media designers will create content that fulfills at least the most popular of genres: comedy, action, love stories, documentaries, pornography, news/politics, science fiction, etc. Imagine how these genres will play out when the participant has a first-person role. Education and training exercises benefit from participants' ability to try/fail/re-try various activities until they understand in body and mind. Flight simulators are already used to teach Air Force pilots—why not create simulations for business leaders to experiment with negotiation scenarios?

Participants in the *AR Façade* studies offered a number of ideas for the potential content directions for this sort of experience. Many players point out an obvious content direction based on the situation they just experienced. As the following quotes illustrate, the medium might serve as a marriage counseling tool because couples will get exposure to, and be able to explore, the dynamics of a marital argument:

> *"You could almost use it as kind of a counseling—a therapy tool. If you kind of retooled it a little bit and put someone in the situation of either Grace or Trip... you're actually doing those life problems. And then you could see all the different sides of the argument, and maybe just figure out what's wrong in your relationship I guess." (P27)*

> *"It's good for like a marriage counseling advice thing. Putting 'em through it or putting someone through it and just saying, 'Hey, watch this.'" (P38)*

Some of the *AR Façade* participants suggest that immersive and interactive stories will serve as formal training for certain social scenarios. Again, the obvious projection is that counselors-in-training can use the simulation to "get their feet wet" (P30) for counseling people with problems:

> *"Clinical psychology type of application where.... people kind of get their feet wet without getting their feet wet so to speak..." (P30)*

> *"Some kind of psychotherapy....maybe even like counseling kind of things.... Like how to train counselors or therapists to just engage with different people...." (P36)*

Other participants viewed the experience as an opportunity to train for potentially stressful activities, such as surgery or military situations:

> *"I think this has actually potential for like surgery or medical simulations ....they can have like this virtual setting without actually like putting lives at stake, you know. That's just a simple example." (P30)*

> *"Military applications..... real life situations like let's say an interrogation or, you know, a stressful environment..." (P32)*

Participants also emphasized the power of this type of experience to elicit real emotions and to expose people to uncomfortable situations as a way to overcome their fears. Player 25 points out that if virtual spiders help people overcome arachnophobia, it could it could also work for the fear of people:

> *"For people with phobias... I know they do it for arachnophobia or they have tried it, I think a little virtual spider. I don't know, dealing with people. Agoraphobia. I see it. ... for kids who have a hard time like figuring out social cues and like how to deal with people, it might work." (P25)*

Player 43 also emphasized the possibility of exposing people to diverse or uncommon social dynamics, to allow people to "try on" different skin colors or different genders:

> *"It's really more about the social interaction... So anything dealing with social dynamics. It could be something with diversity.... see what it's like to be, you know, a black person or...assign a different sex and say 'ok, try this on, and see' ... I mean it's not as simple as that, to do that, but to play around with some of those dynamics." (P43)*

Other participants suggested less "practical" applications of immersive and interactive stories, and rather pointed to entertainment as a sensible content direction:

> *"The entertainment aspect... you could actually have the actual characters of the films and the video games in your living room or even the theater....like characters could walk down the aisles and stuff and that'd be really cool." (P38)*

> *"I would love to that with Madden. If I had that helmet and I was the coach walking around in the living room looking out on the football game, that would be great." (P6)*

At least two different participants suggested that embodied virtual characters, like Trip and Grace, could serve as temporary replacements for actual friends or dates:

> *"If all my friends are gone and I needed a friend, I'd talk to them." (P29)*

> *"Like there are people who might not be able to get a date, but might enjoy going on here and have a date." (P12)*

Player 26 imagined the medium as an extension of current communication tools, so rather than speaking to virtual characters people would be able to talk to remotely located business partners or family:

> *"Maybe like virtual meetings for a company that's 3,000 miles away from another company it does business with. They can be in the space of somebody else and it has more of a personal feeling to it. Maybe ways of communicating with family that's not really close, but I don't think it's a replacement for being actually physically with someone." (P26)*

Not all participants from the *AR Façade* studies held such an optimistic viewpoint for the future of this kind of experience. Player 17 feared for the societal effects of creating such realistic simulations:

> *"I feel like it's gonna be this total move to this mechanized society where you can't really tell the difference between what's virtual and what's real." (P17)*

The participants saw the possibilities for training people for certain jobs, exposing people to difference social scenarios, entertaining people by providing virtual social partners, and connecting remotely located people. Each of these concepts explores the fanciful notion of the *Holodeck*, but without serious consideration of practical limitations. In the near future, immersive and interactive stories are subject to certain hardware and software limitations that will constrain the content possibilities, but also suggest design trajectories.

I will briefly discuss some reasonable schemas suggested by the technology limitations. For example, the space requirements for immersive displays limit the extent of a simulation. Large VR worlds call for some sort of special navigation equipment, like magic carpet in the VR Aladdin ride (Pausch et al., 1996) or the boat in Pirates of the Caribbean ride (Schell and Shochet, 2001). If participants are able to freely walk around (as is the case in *AR Façade*), the story environment then must abide by stricter space constraints (*Façade* takes place in a two-room apartment). Particularly for AR, content will be limited to experiences that can take place in a single room—an apartment, a bar, an underground lair, etc. Multiple scenes and story settings can happen in the same location in an AR application, but then will not take advantage of the physical environment.

Similarly, these experiences cannot take place for a long period of time since players could get exhausted wearing an HMD or other speciality gear. This limitation calls for story experiences that make sense within a short period of time, such as a random encounter or a visit to old friends—the schema explored in *Façade*. The time limitation leads to other constraints. The player does not have much time to form relationships with the characters, which is why some players in *AR Façade* complained that they did not have time to catch up with Trip and Grace before they started fighting. The content for such experiences could be limited to social situations where the player supposedly knows the characters, but not very well—old friends, co-workers, a doctor, a waitress, etc.

Other schemas are suggested by the limitation of natural-language processing in interactive storytelling. Since the system cannot always fully understand what the participant says, content should be built around scenarios where characters might logically ignore or misunderstand the player. The marital argument schema in *Façade* works since couples often do make 3rd parties feel like they are not in the room. The represented social scenario should support the fact that the system does not correctly interpret all player statements. Just as *Façade* plays off the social capital of a “conflict”, future content can play off the communication limitations presented by language and cultural differences (e.g., communicating with a foreigner), education gaps (e.g., talking with a child), or an external distractions (e.g., other people talking loudly, or communicating through a proxy).

Another way to setup the content scenario is to place the player slightly outside the main social interaction. Script the interaction so the player serves a secondary role to a protagonist, so that the story is told mostly through the non-player characters. Actually *Façade* does this to some extent by relaying the story themes primarily through Trip and

Grace. Future content can situate the player as a sidekick or an assistant where they can occasionally communicate to a main character who drives the action and moves the narrative forward.

Every aspect of the experience should be considered and carefully designed to create an effective immersive and interactive story. For example, what is the story behind wearing the immersive HMD? Perhaps the player is putting on a fighter helmet or a costume. Well-designed narrative and interaction mechanisms go a long way to supporting an illusion.

8.6 Summary and Final Thoughts

In this chapter I provided a number of research directions and design considerations that could influence the development of immersive and interactive stories. As Player 37 aptly points out, *AR Façade* is unique in how it creates a feeling of "being provoked by technology... it is definitely a good advancement." (P37) Our collaborative endeavor is an exemplar of the design practice around immersive and interactive stories. The purpose of my dissertation has been to reveal how players encounter such an immersive experience, and to forewarn of the problems that arise if designers focus entirely on a "lack of mediation." I argue for explicit mediation to maximize a sense of embodied narrative engagement—the combination of presence, agency, and dramatic involvement.

In close, I would echo the same challenge raised by Turkle in her discussion of video games: "The cultural pervasiveness of simulation is a challenge to develop a more sophisticated social criticism. This new criticism would take as its goal the development of simulations that actually help players challenge the model's built-in assumptions" (Turkle, 1995, p71). I worry that future media designers will put people in situations to enact violence or perpetuate social stereotypes without building in room for critical reflection. Along with others, I believe some individuals will be vulnerable to addiction and to blurring the line between the simulated world and their own “real” world. Mediation can serve as one means for clearly defining the boundaries for simulated worlds, and creating space for critical reflection.

As one of the great science fiction writers, Isaac Asimov, stated “science fiction writers foresee the inevitable, and although problems and catastrophes may be inevitable, solutions are not.” The immersive and interactive storytelling medium has the potential to become a powerful tool for rhetoric and persuasion, as have all other prior media. Future designers should be challenged to ponder whether the good will outweigh the bad—to think about the ramifications and how to deal with them.

Lastly, I want to acknowledge my awareness of the tremendous privilege this technology currently bestows upon such a small percentage of the population of the world. At this time and for the foreseeable future, immersive and interactive stories will only be available to a small fraction of those living in technologically-advanced “first world” countries. The problems I present in my dissertation pale in comparison to the enormous troubles and injustices facing the world. While my work by no means serves to alleviate human suffering, my hope is that it will enable media creators to explore cultural issues and the depths of human psyche, towards an understanding and resolution of conflicts in society.

APPENDICES

APPENDIX A: ONLINE SURVEY OF FAÇADE

(SURVEYMONKEY.COM)

This survey was deployed using SurveyMonkey.com and advertised on *Façade* fan sites in the summer of 2006.

1. Gender:
a. Male
b. Female

2. Year born (you should be 18 years or older please):
(text field)

3. Race/ethnic background (check all that apply):
   a. American Indian or Alaskan native
   b. Native Hawaiin or other Pacific islander
   c. Asian
   d. Hispanic or Latino/Latina
   e. Black or African American
   f. White
   g. (other)

4. Occupation:
(text field)

5. Highest education level
   a. Middle school (or equivalent)
   b. High school (or equivalent)
   c. Two-year college or technical degree
   d. Undergraduate degree
   e. Masters degree
   f. PhD degree

6. Estimated typing ability:
   a. Beginner (0-15 words per minute)
   b. Intermediate (15-30 words per minute)
   c. Average (30-45 WPM)
   d. Above average (45-60 WPM)
   e. Expert (60 + WPM)

7. Estimated hours playing video games per week:
(text field)

8. Favorite kind of games (check all that apply)

   Action Adventure
   Role-Playing
   First Person Shooters
   Strategy Games Including Real-Time Strategy
   Adventure

Sports
Puzzle
Sim series games (e.g. Simcity, Sims roller-coaster)
Massively Multiplayer
Casual games (web based)
Other ________________________

9. Estimated hours watching TV/movies per week:
(text field)

10. Favorite kind of TV/movies (check all that apply)

Action
Drama
Comedy
Mystery
Detective Stories
Documentaries
Love Stories
Science Fiction
Thriller
Art films
Westerns
Animated
Family
Musical
Reality TV
Soap-Opera
Other ________________________

Game play questions:

11. How many times have you played *Façade*?
   a. Once
   b. 2-5
   c. 6-9
   d. 10 or more

12. Rate your overall impression of the experience: (Likert: 1=poor, 7=excellent)

13. How much did each of the following activities influence the characters and story?
   a. navigating Trip and Grace's apartment
   b. having a conversation with Trip and Grace
   c. interacting with objects in the space
   (Likert for each: 1=my input had no effect, 7=had a big influence)

14. How much did your interaction throughout the experience influence the ending of the story?
(1=no effect, 7=big influence)

15. When do you feel your overall input (navigation, conversation, interaction with objects, etc.) had the MOST influence [of the experience]?
   a. Beginning of the experience
   b. Middle of the experience
   c. End

16. When do you feel your overall input (navigation, conversation, interaction with objects, etc.) had the LEAST influence?

a. Beginning of the experience
b. Middle of the experience
c. End

17. When do you feel you had the most difficulty communicating with Trip and Grace?
    a. Beginning of the experience
    b. Middle of the experience
    c. End

18. Breakdowns in communication with Trip and Grace occurred because (choose all appropriate):
    a. I cannot type fast enough
    b. I cannot think of things to say
    c. I did not want to interrupt what they were saying
    d. Trip and Grace do not understand me
    e. Trip and Grace do not listen to me
    f. The situation was tense and awkward for everyone
    g. The computer program has errors
    h. other [text field]

19. How (if at all) did you adapt your statements to have a successful conversation with Trip and Grace:
    a. I waited for Trip or Grace to finish speaking before I spoke
    b. I interrupted Trip and Grace
    c. I spoke in short sentences
    d. I tried to guess what Trip and Grace are going to say
    e. I erased and changed statements often
    f. I limited my vocabulary
    g. I was reluctant to type statements b/c I felt like they were not have any affect
    h. I typed without punctuation to save time
    i. I paraphrased
    j. I used emoticons (such as ,)
    k. I frequently hit the end of buffer (I typed too much and heard a beeping sound)
    l. Other [text field]

20. Which of the follow specific strategies did you try? (check all that are appropriate)
    a. I tried to reconcile Trip and Grace's differences
    b. I sided with Trip
    c. I sided with Grace
    d. I tried to seduce Trip
    e. I tried to seduce Grace
    f. I was being disruptive to both
    g. I tried to get Trip and Grace to team-up against me
    h. I role-played (I acted like someone I am not)
    i. other [text field]

21. Check all of the facts you learned about Trip and Grace:
    a. Grace and Trip met during their senior year of college
    b. Tonight is their 10-year anniversary
    c. They took a recent holiday to Italy
    d. Grace did not enjoy the trip to Italy
    e. Grace does not like Trip's bull statue
    f. Trip invented a drink called Grace's inner soul
    g. Trip was a bartender in college
    h. Trip had an affair with a client in Barcelona
    i. Grace slept with someone the night before Trip proposed
    j. Trip proposed to Grace in front of her parents on Christmas Eve
    k. Grace felt pressured to marry Trip
    l. Grace's parents are wealthy

m. Grace was coddled and controlled as a kid
n. Trip's family spent 6 months in a shelter when he was a kid
o. Trip is ashamed of his parents
p. Grace wants to be an artist
q. Grace painted the paintings above the couch
r. Trip hangs out at the local bar
s. Trip's mother gave him the 8-ball
t. Grace dislikes her decorating
u. Grace and Trip are in advertising
v. Grace and Trip work at the same company
w. Grace designs magazine ads
x. Grace hates her job
y. Trip pressured Grace to go into advertising
z. Grace is using Trip as an excuse not to be an artist
aa. Grace has refused to go to therapy

22. Which of the following endings have you experienced?
  a. Trip left the apartment
  b. Grace left the apartment
  c. I was thrown out
  d. Trip and Grace are going to work on their differences
  e. I think Trip and Grace reconciled their differences, but I am not sure
  f. Other [text field]

23. Which of the following endings have you tried to make happen?
  a. Make Trip leave
  b. Make Grace leave
  c. Get thrown out
  d. Cause Trip and Grace to quit fighting
  e. Other [text field]

24. Would you like to play again? Why or why not? (text field)

25. The following types of experiences are being considered for the future. Please rate your interest in each of the following: (Rate all that apply)

Mystery (not interested maybe somewhat interested very interested)
Action
Dramatic
Comedy
Erotic

26. Comments on any other aspect of *Façade*? (text field)

27. We would like you to submit the game logs from your interaction with *Façade*, which are found at: C:\Façade\stageplays\ in text files.

Please open these text files and copy the contents into the space below. If there are multiple files, separate them with a horizontal line or dashes (-----). This information will be helpful for understanding how people play the game.

APPENDIX B: ONLINE SURVEY OF FAÇADE (RESULTS)

This appendix outlines the findings from the online survey of *Façade* provided in APPENDIX 1.

Demographics of Online Survey Respondents

As for demographics, 88 out of 105 (84%) were white, of which 81 were males (see Figure B.1). The overall gender breakdown was about 9 males to 1 female; although it is claimed that *Façade* appeals to females as much as males, so this is probably an indication that the online survey only reached a subset of the *Façade* fan population.

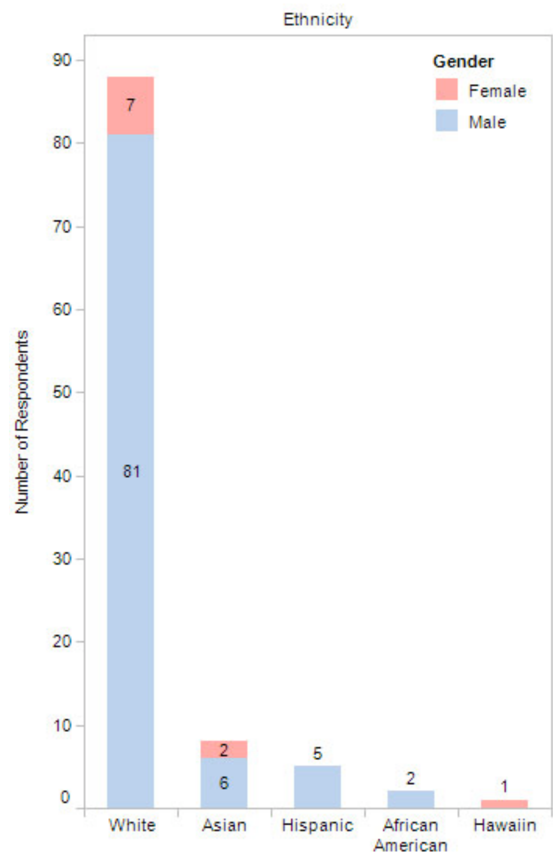


Figure B.1: Ethnicity and gender of online survey respondents

We had a range of occupations, although about 1 in 3 respondents reported they were students (34 total). Education level could be roughly broken down in thirds, with 28 completing a highschool diploma, 43 with an undergraduate (2-4 year programs), and 34 achieving a PhD, masters, or other terminal degrees.

We inquired about typing ability to get a baseline understanding of what effect the typing interface might have on their ability to communicate in the game. 89% claim to be at least an average typer—able to type 30 words per minute or better (see Figure B.2).

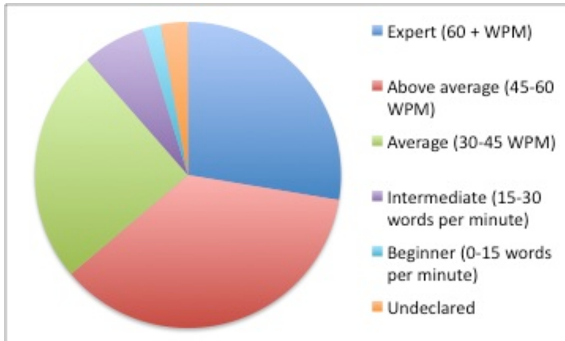
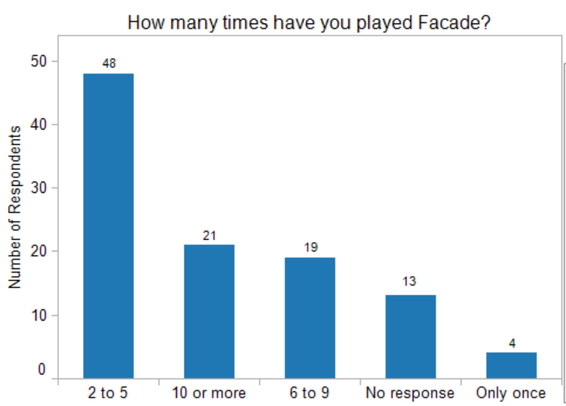


Figure B.2: Self-reported typing ability of online survey respondents

Our respondents play an average of 10.2 hours of video games per week (SE: 1.06, Max: 70) and 8.3 hours of TV/movies (SE: 0.96, Max: 84). Players' top two favorite types of games were Adventure games and Role-playing games, in that order. The top two preferred genre of TV/movies were Science Fiction and Comedy.

Responses to Game Play Questions

Looking at the *Façade* play habits, about half the players had tried *Façade* 2-5 times (48 of 105), but 84% of respondents claimed to have played *Façade* at least twice (see Figure B.3, left). Of the 77 people who answered the subsequent question whether they would play again, over two-thirds said they would play again (see Figure B.3, right). These findings point to *Façade*'s overall satisfaction and *replay-value*—players find it enjoyable to try different strategies or to try to achieve alternative endings.



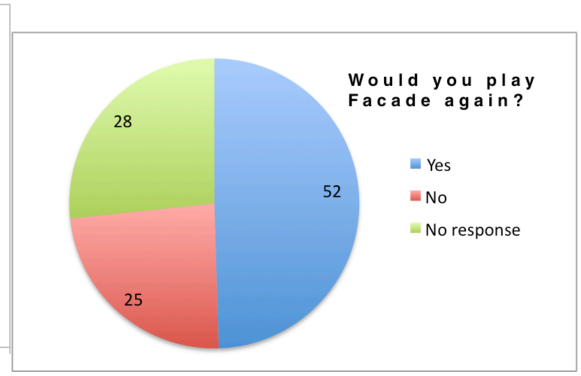


Figure B.3: Reports of playing quantity by online survey respondents (left) Number of times online survey respondents played *Façade*; (right) Percent of online survey respondents who would play *Façade* again.

One survey question asked players to select all the strategies they used while playing *Façade*. They could select or or multiple from the choices listed or list new one. The results were most interesting when comparing the genders (see Figure B.4). For female players, the most frequent strategy was to try to seduce Trip (7 instances), followed by trying to seduce Grace (6). For male players, the number one strategy was to seduce Grace (60), followed by trying to reconcile their differences (59). Males listed 6 other strategies before they would try to seduce with Trip (34 instances). Due to the relationship content in *Façade*, gender will be a recurrent theme, especially in our studies of *AR Façade* where we balanced the number in each gender and asked players to reflect on the relationship of Trip and Grace.

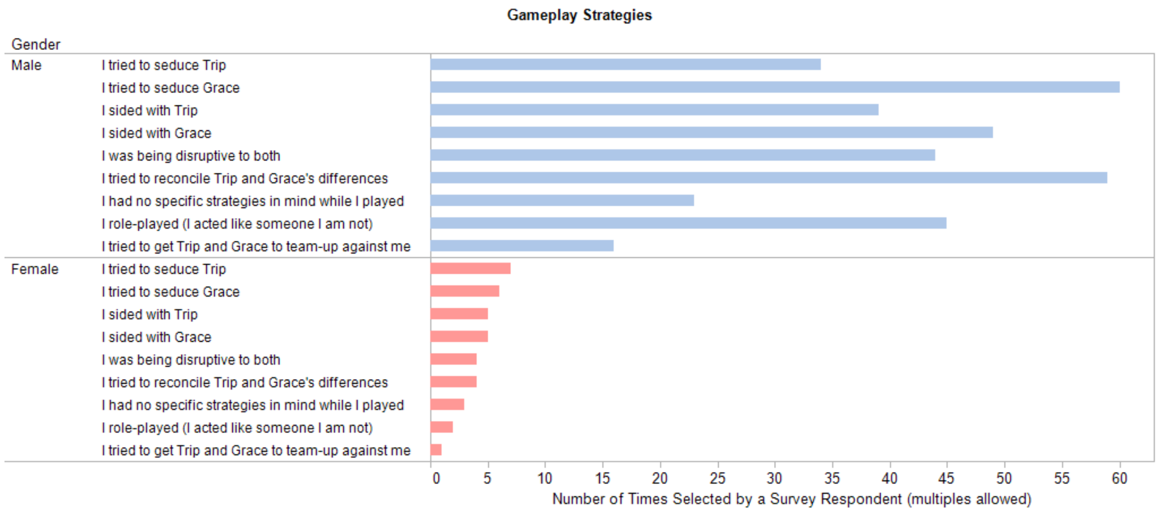


Figure B.4: Gameplay strategies used by online survey respondents (separated by gender)

We were curious how players felt about the conversational aspects of *Façade* and if the experience created and maintained a sense of player agency. Two survey probed how the players felt about the game's responsiveness over the course of the game episode (beginning, middle, and end): "When was it most difficult to hold a conversation with the characters?" and "When did you have the least influence on the story?" (see Figure B.5). The results would indicate that the ability to hold a conversation with Trip and Grace appears to fade over time with most respondents (44% of those that responded to this question) choosing the last third as the most challenging. Similarly, 60% believed they had the least influence over the story towards the end, followed by 24% who felt they had the least influence at the beginning of the game, and 16% who picked the middle. This data provides the first evidence—backed up by deeper episode analyses in Chapter 6—that *Façade* falters towards the end, failing to maintain conversation flow and overall sense of agency.

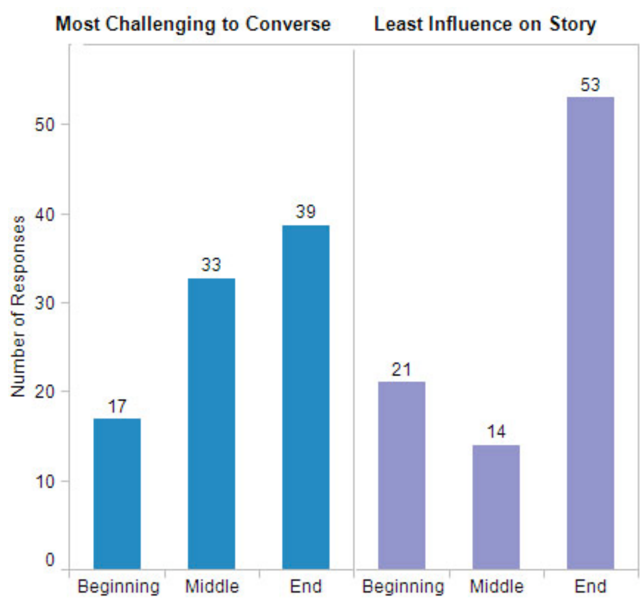


Figure B.5: Subjective ratings about conversation by online survey respondents

The survey also explored how players interpreted the increasing difficultly of communicating with Trip and Grace and their strategies for adapting. When asked to list the possible explanations for why communication breakdowns occurred (see Figure B.6), 68% had to do with socially-related rationale (e.g.

"Trip and Grace didn't understand me", "they didn't listen", "I didn't want to interrupt", etc.), while only 22% of the reasons dealt with interface issues (e.g. "I couldn't type fast enough", "I hit the buffer limit", "the computer has errors", etc.). Ten percent of folks listed "couldn't figure out what to say" as a reason for difficult communication with Trip and Grace.

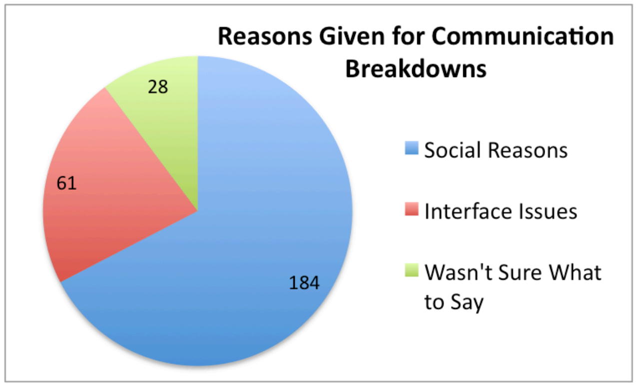


Figure B.6: Reasons given by online survey respondents for communication breakdowns

Survey respondents selected one or more communicative adaption strategies to deal with the breakdowns (see Figure B.7). The two most frequent responses—speaking in short sentences and limiting vocabulary usage— indicate that players' were not necessarily allowed to as expressive as they wished to be. The next two frequent adaption strategies speak to another recurrent theme: there are timing issues that require the player to adjust. Other strategies are revealed in the responses to the open-ended questions below.

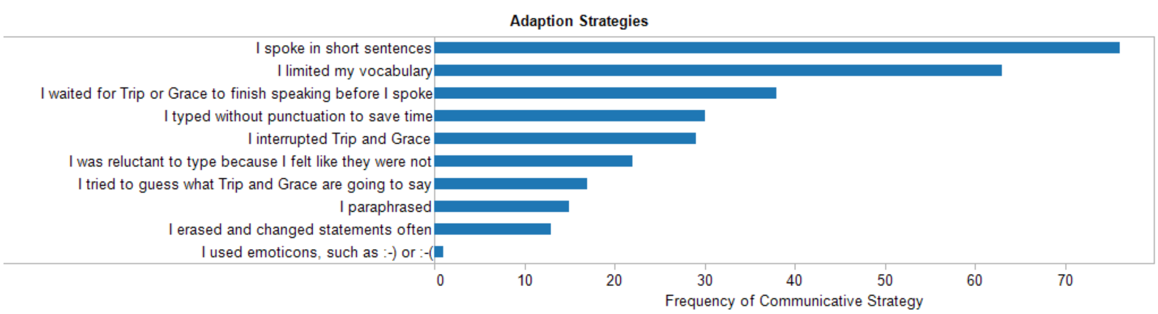


Figure B.7: Strategies used by online survey respondents to adapt to breakdowns

Online Survey Responses to Open-Ended Questions

The online survey left room for open-ended responses to two questions: "what other strategies did you deploy for communicating with Trip and Grace?" and "what are your overall impressions?" Some of the responses reveal an intimate knowledge about *Façade*'s inter-workings (referring to discourse acts for example), again calling into question the generalizability of these findings, but interesting nonetheless.

Some of the strategies were worth noting, and are reflective of the play types I introduce in Chapter 5:

> *"The first few times, I simply reacted as I would in a real situation- after a few different results, I tried to 'game' the system to achieve a specific result (which I achieved about once for every 5 or 6 tries)." … "I was very*

aware that there was a 'hotlist' of words that grace and trip would recognize" ... "I tried to figure out the proper noun verb combinations."

"I walked in the door and said 'WAZZUP BITCHES?!?!' and Trip basically kicked me right out. It would have been funnier if they had asked whether I was drunk. And man I have tried SO hard to seduce Grace it's not even funny." ... "I was abusive, aggressive, drunk, I stole their stuff and lied to them."

Overall impressions were both positive and negative:

"I found Façade an engaging experience to the extent that I felt like I somehow 'knew' Grace and Trip as people."... "It is still fascinating stuff." ... " 'relationship trouble' was a good non-violent way of creating conflict"... "I love how the characters reactions change."... "I've showed Façade to everyone I know, often resulting in tears of laughter."

"The experience is simply frustrating"... "I found the story and characters unlikeable and boring"..."I just see Façade as a failure. I won't play again unless all of it changes."... "the bickering couple did not maintain my curiosity. The outcomes were a little bit broad/cliched and predictable. Truth be told, simulating real life feels boring."

Some of the comments were constructive:

"I could not control whom I was talking to." ... "Is there any way you can supplement the interface to give the player clearer feedback on which discourse acts they're actually engaging?"...." I would have preferred to choose from a massively context-sensitive menu of options, each of which I knew it would understand."

APPENDIX C: INTERFACE COMPARISON STUDY (GAME INSTRUCTIONS)

These instructions will be read aloud to the participant after consent is obtained. We will make sure the participant understands these instructions before proceeding.
General instructions at the beginning:

Façade is an experiment in electronic interactive drama. In this experience you will interact with two fictitious characters, Trip and Grace, in their apartment. You can move around and touch any of the objects in the apartment, have a conversation with the characters, and even hug, kiss, and comfort Trip and Grace. We encourage you to explore the characters and the game space.

Keyboard-based Desktop *Façade* (instructions to be given before this variation):

In this variation of *Façade* you will sit at a traditional desktop machine with a monitor, keyboard and mouse. You will move around the apartment using the arrow keys (left and right to rotate and up/down to move forward and back). You can use the mouse to manipulate objects in the space, such as picking up a glass and drinking from it, and to physically interact with Trip and Grace (so you can mouse-over the characters to Kiss, Hug or Comfort them). You can type statements to the Trip and Grace using the keyboard. Each statement can be about 35 letters long (or what can fit in one line of text), and must be entered before the system can respond.

Speech-based Desktop *Façade* (instructions to be given before this variation):

In this variation of *Façade* you will sit at a traditional desktop machine with a monitor, keyboard and mouse. You will move around the apartment using the arrow keys (left and right to rotate and up/down to move forward and back). You can use the mouse to manipulate objects in the space, such as picking up a glass and drinking from it, and to physically interact with Trip and Grace (so you can mouse-over the characters to Kiss, Hug or Comfort them). You can communicate with the characters by speaking out loud directly into the microphone. You will experience a slight delay and then your words will appear at the bottom of the screen. And FYI, the system cannot understand long sentences.

Augmented Reality *Façade* (instructions to be given before this variation):

In this variation of *Façade* you will put on a head-mounted display and a backpack that weighs about 8-lbs (show the participant the backpack and display). You will move around a physical apartment wearing the equipment (show the room). You can touch and point to objects in the space to reference them. You can communicate with the characters by speaking out loud directly into the microphone. You will experience a slight delay and then your words will appear at the bottom of the screen. And FYI, the system cannot understand long sentences.
Finally, there are several gestures that you can use during the game:

- You can kiss Trip or Grace. (show kiss gesture)
- You can hug them. (show hug gesture)
- Or, you can comfort them. (show comfort)
- You can pick up a virtual drink or the magic 8-ball. (show picking)
- And, you can place them back on the bar. (show placing)
- Finally, if you have a virtual drink you can drink from it. (show drinking)

If you get exhausted and want to stop, just let us know and we will just stop right there. Or, if you simple want to sit down on the couch, feel free to do so. You will play for at least five minutes. If you get kicked out of the apartment before the 5 minutes is up, we will start again in this variation of the game play. Do you have any questions before we get started?

APPENDIX D: INTERFACE COMPARISON STUDY (QUESTIONNAIRE)

To be filled out by the participant just before the final interview.

1. Experience **using computers**:

Never use		Occasional Use		Everyday Use		Expert User
1	2	3	4	5	6	7

2. Estimated **hours using instant messaging** per week:

3. Estimated **hours playing video games** per week:

4. Favorite kind of **games** (check all that apply)

___ Action Adventure
___ Role-Playing
___ First Person Shooters
___ Strategy Games Including Real-Time Strategy
___ Adventure
___ Sports
___ Puzzle
___ Sim series games (e.g. Simcity, Sims roller-coaster)
___ Massively Multiplayer
___ Casual games (web based)
___ Other ________________________

5. Estimated **hours watching TV/movies** per week:

6. Favorite kind of **TV/movies** (check all that apply)

___ Action
___ Drama
___ Comedy
___ Mystery
___ Detective Stories
___ Documentaries
___ Love Stories
___ Science Fiction
___ Thriller
___ Art films
___ Westerns
___ Animated
___ Family
___ Musical
___ Reality TV
___ Soap-Opera
___ Other ________________________

7. What was your **overall rating** of Augmented Reality *Façade*?

Poor Average Excellent

1 2 3 4 5 6 7

8. What was your **overall rating** of the *Façade* **story**?

Poor Average Excellent
1 2 3 4 5 6 7

9. How **believable were the characters**, Trip and Grace?

Not believable Neutral Very Believable
1 2 3 4 5 6 7

10. Did you feel as if you were **there**, inside Trip and Grace's apartment?

Did not feel there Average Felt as if I was inside their apartment
1 2 3 4 5 6 7

11. How **connected** did you feel to the characters?

Did not feel connected Average Felt as if I was there with Trip and Grace
1 2 3 4 5 6 7

12. How much did you feel **part of the drama**?

Was not part of drama Average Felt as if I was part of the drama
1 2 3 4 5 6 7

13. How much did your interaction **influence the story**?

No effect Neutral Big influence
1 2 3 4 5 6 7

14. How inclined are you to **play again**?

No thanks Neutral Yes please
1 2 3 4 5 6 7

15. In which variation of the game did you feel the **most engaged** in the experience?

Keyboard-based Desktop Speech-based Desktop Augmented Reality

16. Which variation seemed most **like a real world** social setting?

Keyboard-based Desktop Speech-based Desktop Augmented Reality

17. Which interface variation was the **most challenging** to learn and use?

Keyboard-based Desktop Speech-based Desktop Augmented Reality

18. Which interface variation was the **easiest or most natural** to use?

Keyboard-based Desktop Speech-based Desktop Augmented Reality

APPENDIX E: INTERFACE COMPARISON STUDY (INTERVIEW GUIDES)

We will be asking open-ended questions at three different parts of the lab study. At each point, the questions listed here are to serve as an interview guide or a rough outline of questions, but we can diverge from these depending on how the participant answers the questions. We will record each interview so copious notes do not have to be take during the interview.

Post variation: After each variation of the game play
Retrospective: For each marked breakdown by the player or researcher
Overall: At the end of all three variations

Post-variation questions:

Tell me about your experience:
  So how did that feel?
Tell me about Trip and Grace:
  How believable were the characters to you?
  Which character do you prefer and why?
Throughout the experience how did you decide what to do?
Do you think you had a strategy?
Tell me about your strategy?
How much influence did you have over the story?
How did the characters react to your actions and statements?
What are your thoughts on the game interface [KB, SB, AR]?
Did you find the interface difficult or easy to use? How so?

Retrospective/breakdown questions:

Here in the video, I noticed you __(e.g.; kept looking around…)__:
Tell me what was happening at that point:
  Tell me what you were thinking here:
How did your strategy (from above) impact this moment?
  What were you trying to do?
What did (moment under discussion in player's own words) mean to you?

Overall/end of study questions:
These questions will come after we have reviewed all their answers to the questionnaire and the earlier discussions.

Now that you've played *Façade* three different times, what are your impressions about the game in general?
What did you think about the content of
  the game?
  the story?
  the characters?
You said that you felt like you [did not] have influence over the ending of the story.
  Tell me more about that.
  How did you feel about how it ended?
  Did it matter to you how it ended?
On the survey, you said that ___you thought Trip and Grace were [not] believable__.
  Tell me more about that.
On the survey, you marked that the social setting [did not feel or felt] real.
  Tell me more about that.
On the survey, you said that you most engaged during the [AR, KD, SD] variation.
  (Pause, wait)
Compare the three different ways that you played the game.

Compare how “real” the three experiences felt.
We are interested in the social aspects of *Façade*.
    Can you speak to that?
On the survey, you said that [AR, KD, SD] had a more challenging interface:
    Why?
On the survey, you said that [AR, KD, SD] has a more ‘natural’ interface.
    What does that mean to you?
On the survey, you said you would rather [speak or type] than [type or speak].
    Why?
    Compare these versions.
On the survey, you said you would rather [walk or sit] than [sit or walk].
    Why?
    Compare these versions.
You said on the survey you would [not] like to play [AR, KD, SD] again.
    Why or why not?
Towards future development:
    What would make the game more fun for you?
    What would make the game more engaging to you?
    What would make the content better to you?
    What could have made the augmented reality interface better?
Do you have any questions for me?

Reaffirming consent of your participation, is it still ok that we use your data?

APPENDIX F: INTERFACE COMPARISON STUDY (RESULTS)

This appendix includes raw data—questionnaire results (Table F.1) and episode statistics (Table F.2)—from interface comparison study conducted at Georgia Tech in the summer of 2006. This data is also summarized in various graphs throughout the dissertation. I also ran paired samples T-tests for key variables from the the episodes (Table F.3), which is part of a discussion in Section 6.4.1. Key demographics for these players are provided in Chapter 5 (see Figure 5.1).

Table F.1: Key survey questions from the interface comparison study (N=12)

Player	Most Realistic	Most Challenging	Easiest to Learn	Preferred Interaction
1	AR	KB	AR	AR
2	AR	AR	KB	KB
3	AR	AR	KB	KB
4	SB	AR	KB	KB
5	SB	AR	KB	SB
6	AR	KB	AR	AR
7	SB	AR	KB	AR
8	AR	SB	KB	AR
9	AR	AR	SB	AR
10	SB	KB	SB	KB
11	AR	AR	KB	KB
12	AR	AR	KB	AR
Totals	8 = AR 4 = SB	8=AR 1=SB 3=KB	2=AR 2=SB 8=KB	6=AR 1=SB 5=KB

Table F.2: Key episode statistics from the interface comparison study (N=12)

Player	Order	Lines of dialog per min			Episode time (minutes)			IPD with Trip			IPD with Grace		
		AR	SB	KB	AR	SB	KB	AR	SB	KB	AR	SB	KB
1	AR,SB,KB	7.10	7.18	4.46	16.4	18.7	20.8	1.37	2.26	2.61	1.79	2.67	2.21
2	SB,KB,AR	2.20	2.36	2.83	16.8	21.6	21.2	2.36	2.86	2.59	1.89	2.73	3.34
3	KB,AR,SB	2.39	2.73	1.34	19.2	27.8	25.3	1.75	2.29	2.13	1.89	1.82	2.39

Player	Order	Lines of dialog per min			Episode time (minutes)			IPD with Trip			IPD with Grace		
4	AR,KB,SB	5.27	N/A	2.66	20.8	13.9	20.7	2.30	N/A	3.01	2.28	N/A	3.41
5	SB,AR,KB	2.19	3.00	1.94	20.2	23.3	22.2	2.33	2.92	2.57	2.44	3.16	2.71
6	KB,SB,AR	4.80	2.72	3.51	16.7	22.8	24.5	1.63	1.79	1.97	1.44	2.14	2.48
7	AR,SB,KB	6.56	4.12	2.73	16.6	11.2	13.2	1.72	2.30	2.21	2.31	2.65	2.80
8	SB,KB,AR	5.57	3.69	3.69	7.00	11.7	13.0	2.61	2.42	2.39	2.41	2.60	2.78
9	KB,AR,SB	1.65	1.41	1.35	20.6	17.0	23.0	1.68	2.46	3.40	2.12	2.24	2.10
10	AR,KB,SB	1.89	2.82	2.68	20.7	22.0	25.0	2.19	2.34	2.93	2.60	2.18	2.69
11	SB,AR,KB	1.98	2.96	1.47	14.7	11.2	17.7	2.55	2.48	3.18	1.98	2.75	3.47
12	KB,SB,AR	3.52	2.41	2.47	19.6	18.7	20.7	2.17	2.98	2.83	2.46	3.20	2.17
Totals	**balanced**	**3.51**	**2.94**	**2.55**	**17.5**	**18.3**	**20.6**	**2.06**	**2.46**	**2.65**	**2.13**	**2.56**	**2.71**

Table F.3: Paired-samples T-test statistics for the interface comparison study (N=12) There were significant differences between pair 2, 8, 9, 10, 14, and 15— AR was greater than KB for text input (T=2.328, p=0.04), AR was less than KB for text erased (T=4.075, p=0.002), SB was less than KB for text erased (T=4.511, p=0.001), AR was less than SB for gestures used (T=2.216, p=0.051), and KB episodes were longer than both the AR and SB episodes (T=3.56, p=0.004 and T=2.681, p=0.21).

		Paired Differences					t	df	Sig. (2-tailed)
		Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference: Lower	Upper			
Pair 1	Chars_AR - Chars_SB	2.2108915	21.6677089	6.5330601	-12.3457	16.76746	.338	10	.742
Pair 2	Chars_AR - Chars_KB	14.60547	21.7375373	6.2750865	.7940977	28.41684	2.328	11	.040
Pair 3	Chars_SB - Chars_KB	9.0829954	18.3707623	5.5389932	-3.25865	21.42464	1.640	10	.132
Pair 4	STMT_AR - STMT_SB	-.5411001	3.0438559	.9177571	-2.58599	1.5037901	-.590	10	.569
Pair 5	STMT_AR - STMT_KB	-.9690267	3.4820307	1.0051757	-3.18140	1.2433500	-.964	11	.356
Pair 6	STMT_SB - STMT_KB	-.6325489	3.2663042	.9848278	-2.82688	1.5617841	-.642	10	.535
Pair 7	CharsErased_AR - CharsErased_SB	.1837074	.2973441	.0896526	-.0160511	.3834659	2.049	10	.068
Pair 8	CharsErased_AR - CharsErased_KB	-4.40002	3.7406275	1.0798261	-6.77670	-2.02334	-4.075	11	.002
Pair 9	CharsErased_SB - CharsErased_KB	-4.94316	3.6342978	1.0957820	-7.38472	-2.50161	-4.511	10	.001
Pair 10	Gestures_AR - Gestures_SB	-.1786032	.2672875	.0805902	-.3581694	.0009630	-2.216	10	.051
Pair 11	Gestures_AR - Grestures_KB	-.0736100	.2439080	.0704102	-.2285817	.0813617	-1.045	11	.318
Pair 12	Gestures_SB - Grestures_KB	.1246904	.4276024	.1289270	-.1625768	.4119576	.967	10	.356
Pair 13	Length_AR - Length_SB	-.8733333	4.9406751	1.4262501	-4.01249	2.2658219	-.612	11	.553
Pair 14	Length_AR - Length_KB	-3.15944	3.0741394	.8874276	-5.11266	-1.20623	-3.560	11	.004
Pair 15	Length_SB - Length_KB	-2.28611	2.9542798	.8528271	-4.16317	-.4090513	-2.681	11	.021

APPENDIX G: ELEVEN-WEEK GALLERY DEPLOYMENT

(OPTIONAL PAPER-BASED SURVEY)

Tell us about your experience in *AR Façade*!

We want to know your impressions of this research project, and your insights will help us make improvements for future interactive dramas. Please fill out this short 5-minute questionnaire. Your identity will not be collected on the survey, so feel free to be candid.

If you have any questions about the study, you may contact Steven Dow, the Co-Principle Investigator, by email at stevendow@gmail.com or if you have any questions about your rights as a research participant, you may contact Ms. Melanie Clark, Georgia Tech at (404) 894-6942.

Age: _______
Gender: ________
Occupation: __________

Experience using speech-based computer interfaces:
Never use Occasional Use Everyday Use Expert User
1 2 3 4 5 6 7

Favorite genre (circle one):
Drama Comedy Mystery Action Sci-fi Documentary
Other: ______

How would you best describe your personality?

(select one ->) Reserved Expressive
(select one ->) Pragmatic Instinctual
(select one ->) Follows logic Follows emotions
(select one ->) Structured Spontaneous

For you, which medium most closely resembles the *AR Façade* experience? And Why?
Television Video games Movies Theatre Improv
Other: ___________
And Why? ___________________________________

Is there a particular moment from the experience that really sticks out? Which one and why?

What is your rating of the interaction with Trip and Grace *AR Façade*?
Poor Average
Excellent
1 2 3 4 5 6 7

What is your overall rating of *AR Façade*?
Poor Average
Excellent
1 2 3 4 5 6 7

Please tell us anything else about your experience in your words:

APPENDIX H: ELEVEN-WEEK GALLERY DEPLOYMENT

(SURVEY RESULTS AND EPISODE DATA)

In total, we had 40 responses to a paper-based survey that players had the option to fill out after trying *AR Façade*. Considering we recorded 106 episodes of *AR Façade* throughout the course of the installation, that means about 38% of the *AR Façade* participants filled out the survey, although we cannot match surveys to episodes. Table H.1 provides the demographics of the 40 respondents. Most of the survey questions turned out to be uninteresting, but Figure H.1 shows the breakdown how players relate *AR Façade* to other media experiences, backing up what was found in the in-depth studies of *AR Façade*.

Table H.1: General demographics of 40 respondents to the optional paper survey

Average Age	Genders	Occupations	Interaction Rating	Overall Rating
23.6	23 females and 17 males	31 students	4.4 / 7.0	5.4 / 7.0

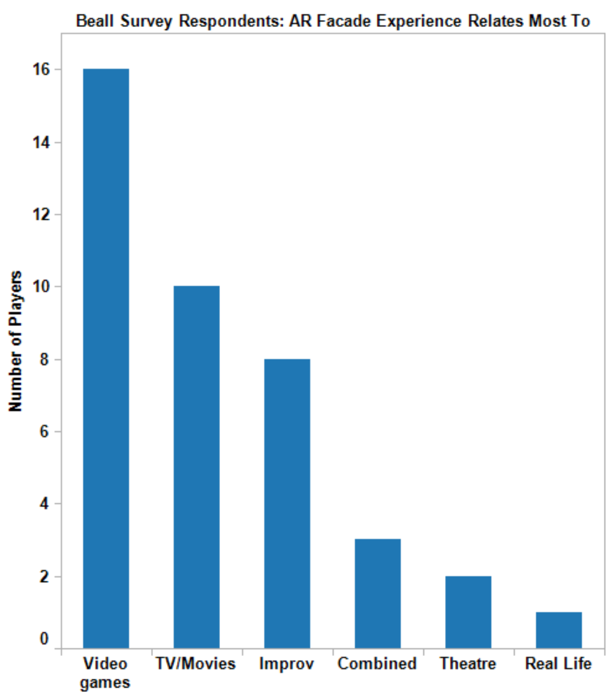


Figure H.1: Players relating *AR Façade* to other media experiences (opt. survey)
(40 respondents to optional paper survey)

As for episode data, during 11-weeks at the Beall Center we collected 126 data logs for the classic desktop interaction version of *Façade* and 106 logs for *AR Façade*. Although we did not capture any demographics for these players, there was one interesting point to make about the AR players: only 45 out of 106 (42%) episodes recorded female-gendered names, while the survey showed the percent of female players to be around 58% (see Table H.1).

The main point of collecting this data was to look at the quantitative differences between AR and KB from a long-term deployment. With such a large same size, I did discover some some significant differences the two versions of the game. The amount of dialog used per minute during the experience is greater in the AR version than it is in the KB; this finding is commensurate with numbers from the comparative study and with my analysis that players were less constrained in the AR version.

Table H.2: T-test comparing KB and AR for DialogPerMin at Beall Center (N=232) Dialog Per Minute is significantly greater in AR (t=3.753; p=.000)

Group Statistics

	InterfaceType	N	Mean	Std. Deviation	Std. Error Mean
DialogPerMin	KB	126	2.414748	1.3149019	.1171408
	AR	106	3.193620	1.8365600	.1783825

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means						
									95% Confidence Interval of the Difference	
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper
DialogPerMin	Equal variances assumed	13.139	.000	-3.753	230	.000	-.7788719	.2075326	-1.18778	-.3699639
	Equal variances not assumed			-3.650	186.026	.000	-.7788719	.2134063	-1.19988	-.3578643

The dialog per minute is calculated by summing the number of independently-entered lines of utterances (by the player in KB and by the wizard in AR) and then dividing by the number of minutes in that episode (see Table H.2). The average statement length of each independently-entered utterance was actually the same across both versions (12.5 average for KB; 13.0 for AR) (see Table H.3). This makes sense because the wizards also had to abide by the 35 character buffer limit—they just entered more dialogs to keep up with the player speech.

Table H.3: T-test comparing KB and AR for StmtLength at Beall Center (N=232)
Statement Length is not significantly different between AR and KB (t=0.849; p=.397)

Group Statistics

	InterfaceType	N	Mean	Std. Deviation	Std. Error Mean
StmtLength	KB	124	12.5311	4.14161	.37193
	AR	90	12.9948	3.65719	.38550

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means						
									95% Confidence Interval of the Difference	
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper
StmtLength	Equal variances assumed	.625	.430	-.849	212	.397	-.46366	.54636	-1.54065	.61333
	Equal variances not assumed			-.866	203.941	.388	-.46366	.53567	-1.51982	.59250

Another significant finding was that players used more gestures per minute in desktop *Façade* than in the AR version (see Table H.4). This is also aligns with the earlier finding that less gestures were registered in the AR version. My analysis is that fewer gestures occurred in AR because the wizard was not able to visually see all of them happening, and because the KB interface had much clearer affordances for touch things in the space.

Finally, there was no significant difference between AR and KB in terms of the length of episode, although the average time of episode for KB (9.87) was about a minute longer than AR (8.83) (see Table H.5).

Table H.4: T-test comparing KB and AR for GesturePerMin at Beall Center (N=232)
Number of Gestures Used Per Minute is significantly smaller in AR than in KB (t=4.172; p=.000)

Group Statistics

	InterfaceType	N	Mean	Std. Deviation	Std. Error Mean
GesturePerMin	KB	126	.855219	1.8737623	.1669280
	AR	106	.093230	.1614108	.0156776

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means						
									95% Confidence Interval of the Difference	
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper
GesturePerMin	Equal variances assumed	25.801	.000	4.172	230	.000	.7619889	.1826252	.4021566	1.1218211
	Equal variances not assumed			4.545	127.203	.000	.7619889	.1676626	.4302199	1.0937578

Table H.5: T-test comparing KB and AR for TimeOfPlay at Beall Center (N=232)
Time of Play is not significantly different between AR and KB (t=1.472; p=.142)

Group Statistics

	InterfaceType	N	Mean	Std. Deviation	Std. Error Mean
TimeOfPlay	KB	126	9.8652	5.89730	.52537
	AR	106	8.8319	4.54957	.44189

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means						
									95% Confidence Interval of the Difference	
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper
TimeOfPlay	Equal variances assumed	17.073	.000	1.472	230	.142	1.03327	.70176	-.34942	2.41596
	Equal variances not assumed			1.505	228.362	.134	1.03327	.68650	-.31942	2.38596

APPENDIX I: TWO-WEEK PLAYER INVESTIGATION

(INTERVIEW GUIDE)

Pre-questionnaire (brief):

First question: what is your best estimate of the amount of minutes you were playing *AR Façade*?
Tell me about your experience:
Is there a particular moment from the experience that really sticks out? Which one and why?
What are your impressions of Trip and Grace:
Throughout the experience how did you decide what to do? Did you feel as if your actions impacted the situation?
Did you follow any particular strategies? Why?
Tell me what was happening at [such and such] point:
    Tell me what you were thinking here:
        What were you trying to do?
    What did (moment under discussion in player's own words) mean to you?
Did you use the "pause" feature? What did you think of it?
What did you think about the genre of content (game, story, characters)?

Post-questionnaire interview (to clarify answers):

Read over the questionnaire and have the player tell us why they answered the way they did. For example… "Your overall rating of *AR Façade* was ___[long pause] ….Why?"
Wrap up: Now how long did that interview last (in minutes)?
Do you have any questions for me?

Reaffirming consent of your participation, is it still ok that we use your data?

APPENDIX J: TWO-WEEK PLAYER INVESTIGATION

(DEMOGRAPHIC QUESTIONS)

Have you played the interactive computer drama "*Façade*" before? (yes/no)
How many times? _________

Occupation / age / gender: ________________

Experience using computers:

Never use Occasional Use Everyday Use Expert User
1 2 3 4 5
6 7

Estimated hours playing video games per week:

Experience using Speech based Interfaces:

Never use Occasional Use Everyday Use Expert User
1 2 3 4 5
6 7

Rank your top three types of games (put 1, 2, 3 next to your top three):

___ Action Adventure
___ Role-Playing
___ First Person Shooters
___ Strategy Games Including Real-Time Strategy
___ Adventure
___ Sports
___ Puzzle
___ Sim series games (e.g. Simcity, Sims roller-coaster)
___ Massively Multiplayer
___ Casual games (web based)
___ Other _________________________

Estimated hours watching TV/movies per week:

Rank your top three genre of movies or TV (put 1, 2, 3 next to your top three):

___ Action
___ Drama
___ Comedy
___ Mystery
___ Detective Stories
___ Documentaries
___ Love Stories
___ Science Fiction
___ Thriller
___ Art films
___ Westerns
___ Animated
___ Family
___ Musical
___ Reality TV

___ Soap-Opera
___ Other ___________________________

Do you have prior experience with acting? Please describe:

Please select the set of statements that best describes your personality (if even just a little better). Neither one is better or worse than the other. Choose an entire column based on whom you really are, not how you wish you were, or have to be at work.

a: (choose either the left group or right group based on how well it matches your personality)

Have high energy Talk more than listen Think out loud Act, then think Like to be around people a lot Prefer a public role Can sometimes be easily distracted Prefer to do lots of things at once Are outgoing & enthusiastic	Have quiet energy Listen more than talk Think quietly inside my head Think, then act Feel comfortable being alone Prefer to work "behind-the-scenes" Have good powers of concentration Prefer to focus on one thing at a time Are self-contained and reserved

b: (choose one column below)

Focus on details & specifics Admire practical solutions Notice details & remember facts Are pragmatic - see what is Live in the here-and-now Trust actual experience Like to use established skills Like step-by-step instructions Work at a steady pace	Focus on the big picture & possibilities Admire creative ideas Notice anything new or different Are inventive - see what could be Think about future implications Trust their gut instincts Prefer to learn new skills Like to figure things out for themselves Work in bursts of energy

c: (choose one column below)

Make decisions objectively Appear cool and reserved Are most convinced by rational arguments Are honest and direct Value honesty and fairness Take few things personally Tend to see flaws Are motivated by achievement Argue or debate issues for fun	Decide based on their values & feelings Appear warm and friendly Are most convinced by how they feel Are diplomatic and tactful Value harmony and compassion Take many things personally Are quick to compliment others Are motivated by appreciation Avoid arguments and conflicts

d: (choose one column below)

Make most decisions pretty easily Are serious & conventional Pay attention to time & are prompt Prefer to finish projects Work first, play later Want things decided See the need for most rules Like to make & stick with plans Find comfort in schedules	May have difficulty making decisions Are playful & unconventional Are less aware of time & run late Prefer to start projects Play first, work later Want to keep their options open Question the need for many rules Like to keep plans flexible Want the freedom to be spontaneous

APPENDIX K: TWO-WEEK PLAYER INVESTIGATION (QUESTIONNAIRE)

AR Façade Quiz!

1) You introduced Grace and Trip during their:
   a. senior year of college
   b. first year of college
   c. senior year in highschool
   d. (I never learned about this)

2) How long have Trip and Grace been married?
   a. 10 years
   b. 5 years
   c. 15 years
   d. (I never learned about this)

3) Trip and Grace recently took a holiday to:
   a. Spain
   b. Italy
   c. France
   d. (I never learned about this)

4) Grace and Trip both work in:
   a. Sales
   b. Advertising
   c. Fashion
   d. (I never learned about this)

5) In college, Trip's part-time job was:
   a. Waiter
   b. House painter
   c. Bartender
   d. (I never learned about this)

For you, which medium most closely resembles the *AR Façade* experience? And Why?
Television Video games Movies Theatre Improv
Other: ___________

What is your overall feeling about the experience?
Hated it Neutral
Loved it
1 2 3 4 5
6 7

Which character did you prefer?
Trip Grace Hate them both equally Like them both equally

How interested were you in the outcome?
Didn't care Neutral Very
Curious
1 2 3 4 5
6 7

What was your overall rating of the *Façade* characters, Trip and Grace?
Poor Average Excellent
1 2 3 4 5 6 7

How engaged were you overall in the *Façade* experience?
Very bored Average Very engaged
1 2 3 4 5 6 7

How much did you feel part of the story?
Not at all Neutral Felt as if I was there
1 2 3 4 5 6 7

How much did your interaction influence the story?
No effect Neutral Big influence
1 2 3 4 5 6 7

How much did you feel like you were one of the characters?
Not at all Neutral Felt like I was one of them
1 2 3 4 5 6 7

To what extent did the characters appear to be there?
Not at all Neutral Felt like they were there
1 2 3 4 5 6 7

How much did your experiences seem consistent with other real-world experiences?
Not at all Neutral Felt like real life
1 2 3 4 5 6 7

How much did the technology impact your experience?
Not at all Neutral It made a big impact
1 2 3 4 5 6 7

APPENDIX L: EPISODE VIDEO CODING ANALYSIS (REFERENCE SHEET)

Category	**Possible Values/Codes**
Speech utterance: a sound that comes out of the player's mouth (phrases, words, sounds)	0 - No speech 1 - Speech utterance is appropriate for social situation, follows along with character dialogue, does not try to be silly (e.g. showing concern with characters, going through traditional greeting dynamics, responding to questions asked by the characters.) 2 - Speech utterance is overly-dramatic, goofy, playful, disruptive to the social scenario or trying to provoke the characters into acting silly. (e.g., yelling obscenities at the characters, asking about completely unrelated topics, telling Grace that you love her, talking about taboo topics, telling Trip to shut-up) 3 - Speech utterance is a tactic to experiment with the interface and interaction to understand its edges. (e.g. repeating statements slowly to make it "work" with the speech recognition, asking the researchers how to play, using the pause feature, speaking very slowly and deliberately "robot-voice")
Physical gesture: a bodily motion enacted by the player (could be conscious or subconscious)	0 - No gesture 1 - Physical gesture is appropriate for social situation, follows along with social dynamic, does not try to be silly (e.g. kissing or hugging to greet the character, comforting the characters during the fight, talking with hands but not exaggerated, walking to parts of the room when directed) 2 - Physical gesture is overly-dramatic, goofy, playful, disruptive to the social scenario or trying to provoke the characters into acting silly. (e.g., kissing and hugging after the initial greeting, making wild gestures towards the characters, trying to push Trip or Grace, pointing inappropriately) 3 - Physical gesture is a tactic to experiment with the interface and interaction to understand its edges. (e.g. playing with graphics, looking down on characters to see their "flatness", trying to feel the characters as a way to feel out the edges of the experience, exploring the objects in the room)
Characters do not respond to player speech	1 - Provide this code if the player says something that warrants some response from the characters, but they do NOT respond. This code is used in conjuncture with any type of speech utterance (e.g. player asks a question, player makes an observation that deserves recognition, any obvious verbal communication breakdowns). It should not be used if the player may not expect an explicit response (e.g. Trip says "come on in", player says "OK"—here the player spoke last, but that's not a breakdown.)
Characters do not respond to player gesture	1 - Provide this code if the player does a physical action that is explicit enough to warrant some reaction, but the characters do NOT respond. This code is used in conjuncture with any type of physical gesture, but it must be explicit enough to expect some reaction. ('Talking with hands' would not need this code, but if the player tries to touch the characters and they do not react, that would be considered a gestural breakdown)
Technical disturbance	1 - Anything that takes away from the illusion of a seamless experience. So-called "breaks in presence", tracking errors, misalignment of graphics, disappearance of characters, audio failures (on player, not audience), AI crashes, etc.

APPENDIX M: EPISODE VIDEO CODING ANALYSIS (KAPPA STATISTIC SUMMARY FOR INTER-CODER RELIABILITY)

A collaborator and I conducted the video coding method and I calculated the inter-coder reliability Kappa statistics for 5% of the overall video data. As I outline in Appendix LAPPENDIX, there were five separate categories of codes: type of speech, type of gesture, whether a response was provided to a speech utterance, whether a response was provided to a gesture, and whether a technical error occurred. The first two categories had 4 possible codes, where as the others were Boolean on/off codes. I manually calculated the Kappa statistic from 140 samples of inter-coded data (140 unique 15-second intervals of the video data). All Kappa statistic calculations are based on examples provided here: http://www.dmi.columbia.edu/homepages/chuangj/kappa/ (Access 9/22/08)

Category 1: Type of Speech

Rater 1 Rater 2	0	1	2	3	Total
0	23	3	0	0	0.186
1	6	73	2	1	0.586
2	0	2	23	0	0.179
3	0	0	1	6	0.050
Total	0.207	0.564	0.179	0.050	

Observed = 0.8929
Chance = 0.4034
Kappa = 0.8204

Category 2: Type of Gesture

Rater 1 Rater 2	0	1	2	3	Total
0	99	2	2	0	0.736
1	2	11	0	0	0.093
2	4	0	19	0	0.164
3	3	1	0	0	0.007
Total	0.757	0.093	0.150	0.000	

Observed = 0.9214
Chance = 0.5903
Kappa = 0.8082

Category 3: No Speech Response from Characters

Rater 1 / Rater 2	0	1	Total
0	70	1	0.507
1	2	67	0.493
Total	0.514	0.496	

Observed = 0.9786
Chance = 0.5002
Kappa = 0.9571

Category 4: No Gesture Response from Characters

Rater 1 / Rater 2	0	1	Total
0	109	1	0.786
1	6	24	0.214
Total	0.821	0.179	

Observed = 0.9500
Chance = 0.6837
Kappa = 0.8419

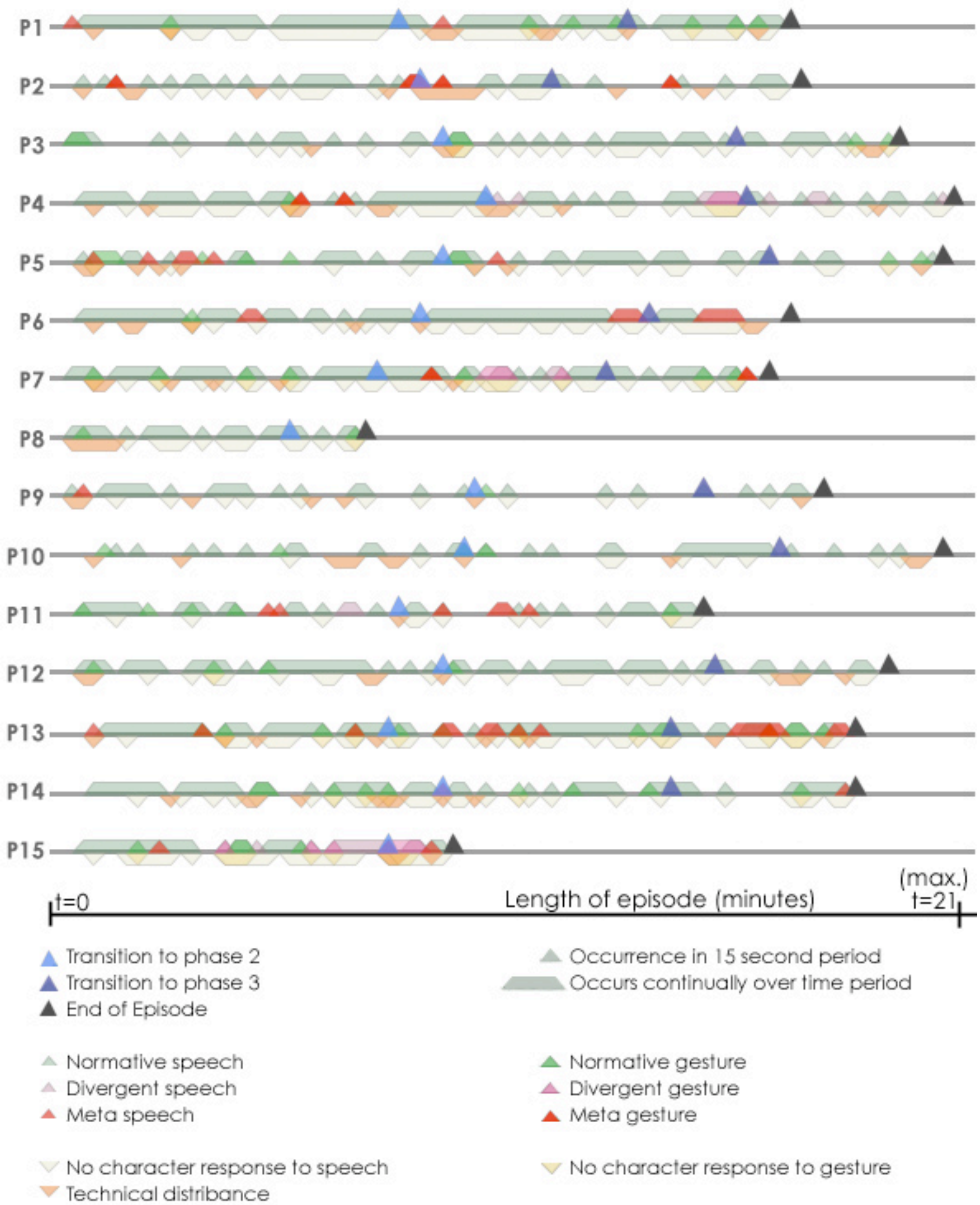
Category 5: Technical Error Occurred

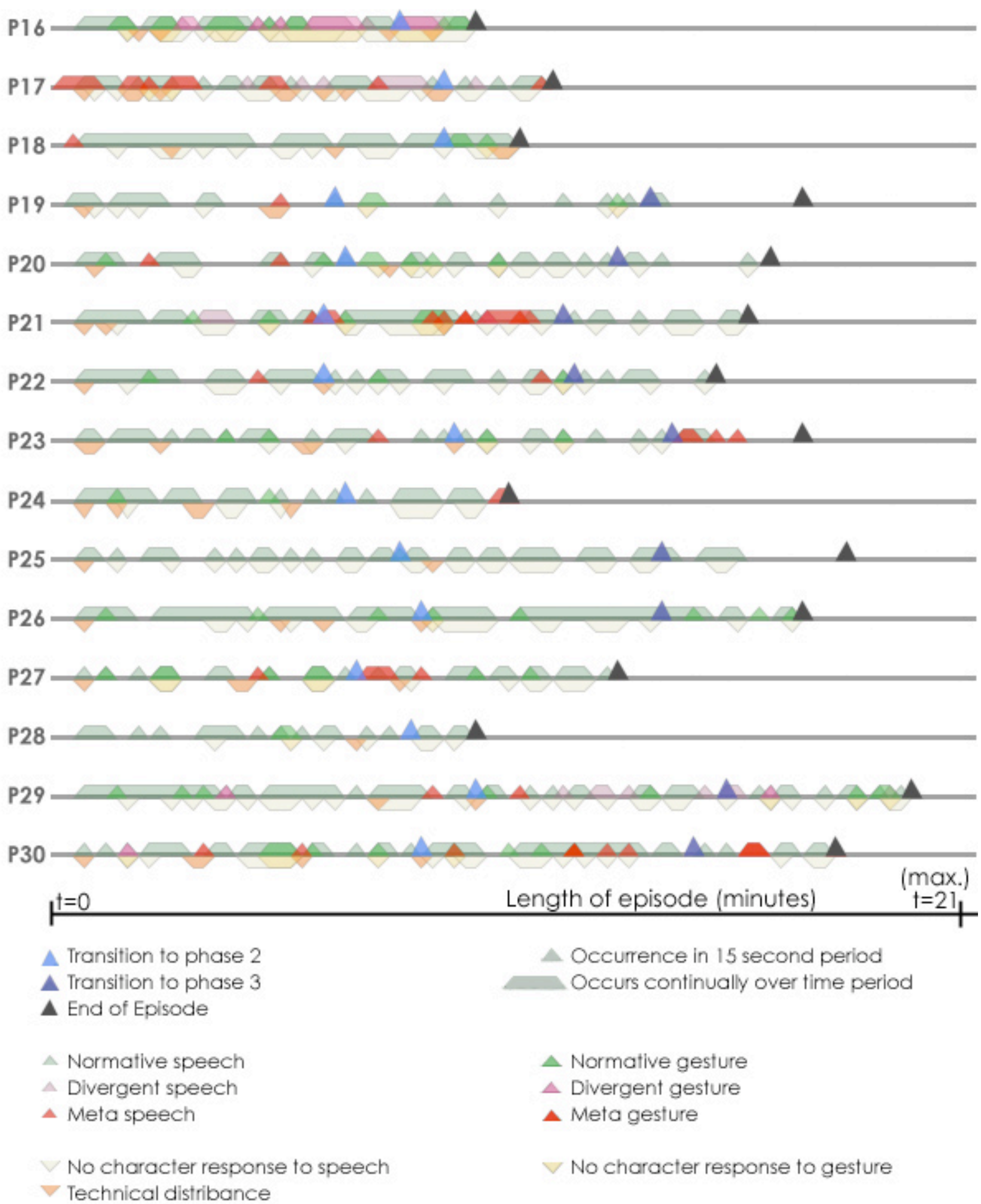
Rater 1 / Rater 2	0	1	Total
0	126	0	0.900
1	3	11	0.100
Total	0.921	0.079	

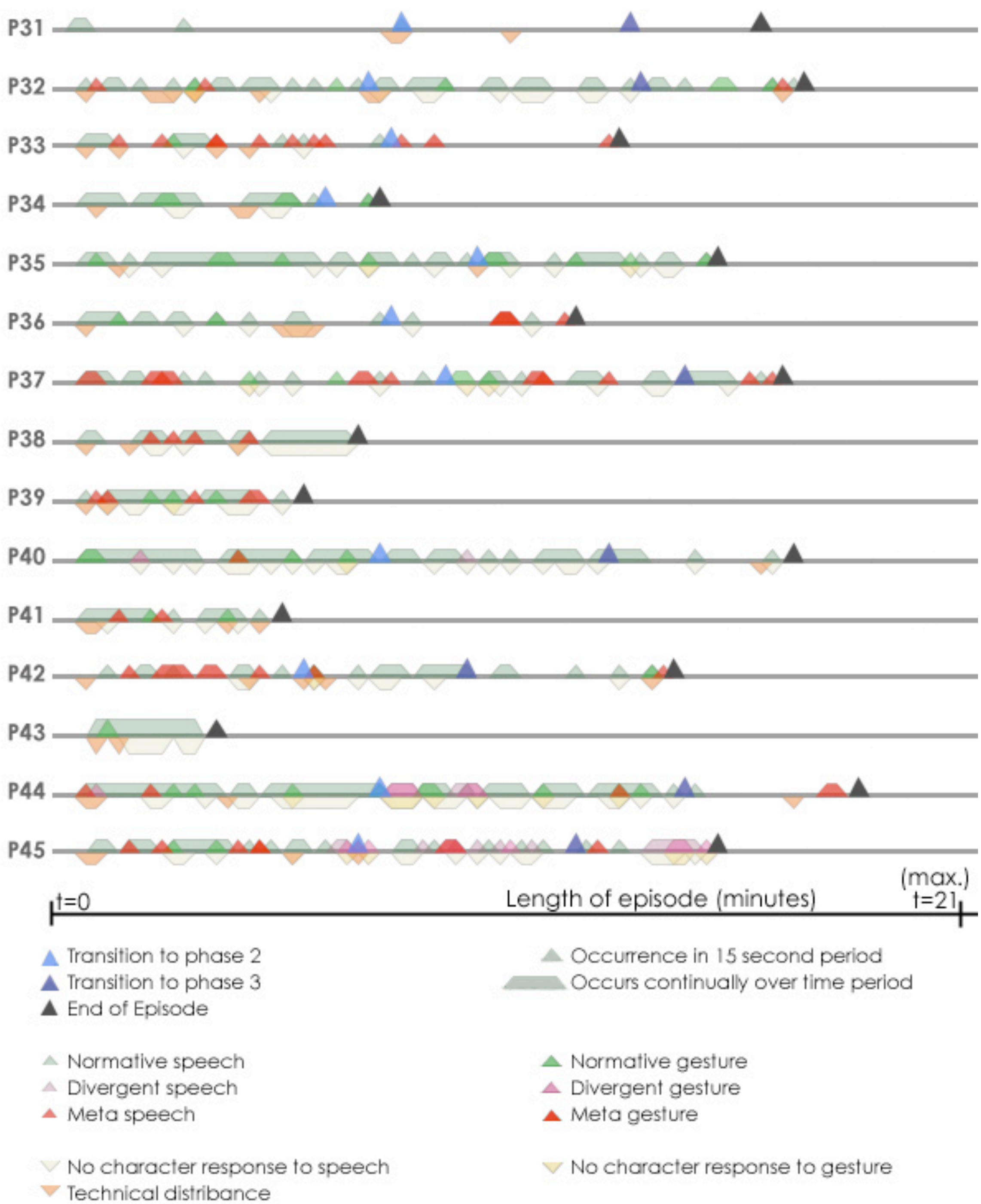
Observed = 0.9786
Chance = 0.8371
Kappa = 0.8684

APPENDIX N: EPISODE VIDEO CODING ANALYSIS (PLAYER VISUALIZATIONS)

Episode visualizations for Players 1-45 spread across three pages:




P16
P17
P18
P19
P20
P21
P22
P23
P24
P25
P26
P27
P28
P29
P30
t=0
Length of episode (minutes)
(max.) t=21
Transition to phase 2
Transition to phase 3
End of Episode
Occurrence in 15 second period
Occurs continually over time period
Normative speech
Divergent speech
Meta speech
Normative gesture
Divergent gesture
Meta gesture
No character response to speech
Technical distribance
No character response to gesture


P31
P32
P33
P34
P35
P36
P37
P38
P39
P40
P41
P42
P43
P44
P45
t=0
Length of episode (minutes)
(max.)
t=21
Transition to phase 2
Transition to phase 3
End of Episode
Occurrence in 15 second period
Occurs continually over time period
Normative speech
Divergent speech
Meta speech
Normative gesture
Divergent gesture
Meta gesture
No character response to speech
Technical distribance
No character response to gesture

APPENDIX O: POSSIBLE CORRELATIONS ACROSS BOTH INSTALLATIONS

In this appendix, I explore some of the possible correlations between game statistics, player demographics, and the subjective ratings from the players. I ran a Pearson Correlation analysis among the forty-five players (N=45) from the Atlanta (just AR data) and Beall Center (see Table O.1).

Table O.1: Pearson Correlation analysis for all 45 players: Includes length of play, number of gestures, IPD between the player and Trip, IPD between the player and Grace, player age, player's experience with computers (subj rating), player's number of hours of video games per week, player's number of hours of TV per week, and player's overall rating. Variables with significant correlations are marked with a * (.05 level) or ** (0.01).

		LengthOfPlay	Gesture	IPDTrip	IPDGrace	Age	ComputerExp	VideoExp	TVHours	SuperRating
LengthOfPlay	Pearson Correlation	1	.322*	-.145	.190	-.315*	.033	.184	.190	.171
	Sig. (2-tailed)		.031	.342	.210	.035	.835	.243	.227	.280
	N	45	45	45	45	45	42	42	42	42
Gesture	Pearson Correlation	.322*	1	-.304*	-.056	-.314*	-.073	-.200	.003	.286
	Sig. (2-tailed)	.031		.042	.716	.036	.646	.205	.985	.067
	N	45	45	45	45	45	42	42	42	42
IPDTrip	Pearson Correlation	-.145	-.304*	1	.207	-.040	-.077	-.206	-.101	-.240
	Sig. (2-tailed)	.342	.042		.173	.794	.627	.190	.527	.126
	N	45	45	45	45	45	42	42	42	42
IPDGrace	Pearson Correlation	.190	-.056	.207	1	-.198	.011	-.067	.108	.081
	Sig. (2-tailed)	.210	.716	.173		.193	.943	.674	.497	.609
	N	45	45	45	45	45	42	42	42	42
Age	Pearson Correlation	-.315*	-.314*	-.040	-.198	1	.180	.140	-.068	-.435**
	Sig. (2-tailed)	.035	.036	.794	.193		.255	.375	.667	.004
	N	45	45	45	45	45	42	42	42	42
ComputerExp	Pearson Correlation	.033	-.073	-.077	.011	.180	1	.343*	.048	.184
	Sig. (2-tailed)	.835	.646	.627	.943	.255		.026	.763	.243
	N	42	42	42	42	42	42	42	42	42
VideoExp	Pearson Correlation	.184	-.200	-.206	-.067	.140	.343*	1	.422**	.229
	Sig. (2-tailed)	.243	.205	.190	.674	.375	.026		.005	.144
	N	42	42	42	42	42	42	42	42	42
TVHours	Pearson Correlation	.190	.003	-.101	.108	-.068	.048	.422**	1	.397**
	Sig. (2-tailed)	.227	.985	.527	.497	.667	.763	.005		.009
	N	42	42	42	42	42	42	42	42	42
SuperRating	Pearson Correlation	.171	.286	-.240	.081	-.435**	.184	.229	.397**	1
	Sig. (2-tailed)	.280	.067	.126	.609	.004	.243	.144	.009	
	N	42	42	42	42	42	42	42	42	42

*. Correlation is significant at the 0.05 level (2-tailed).

**. Correlation is significant at the 0.01 level (2-tailed).

While a number of the variables are statistically correlated, they are not necessarily meaningful. I dug deeper with a couple of these potentially correlated values to see how they look visually. For example players' overall rating seems to be positively correlated to the amount of hours spent watching Movies/TV per week, but the visual plot of this data was not compelling. Players' overall rating appears to be inversely correlated with the players age (young players rated *AR Façade* more favorably) (see Figure O.1). The trend line indicates a relationship—the highest ratings come from younger players and the poorest ratings come from older players—although looking at the 20-30 year old range, it is not a statistically tight relationship. The R-squared value for the trend line is only 18.9%, so more data would have to be collected to make strong claims about the relationship between age and overall rating.

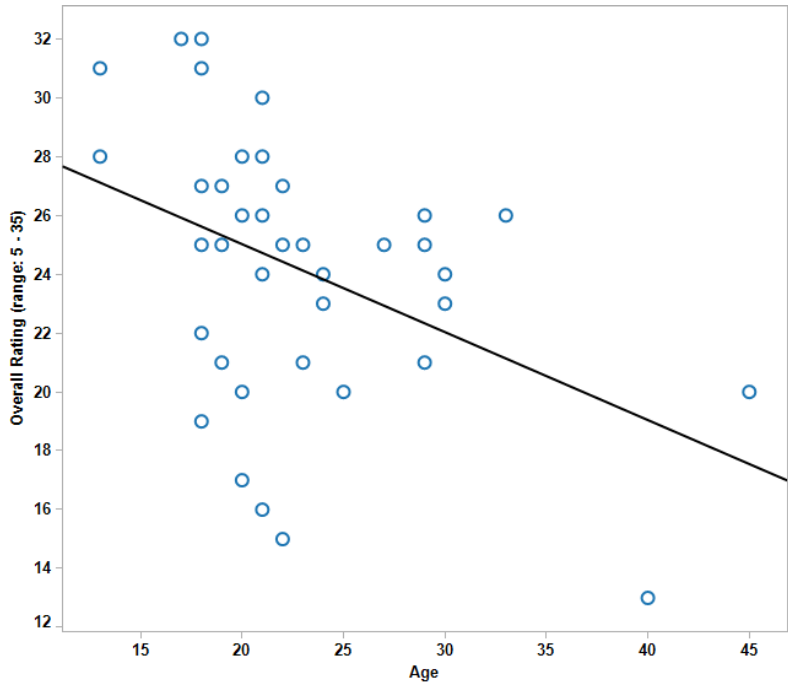


Figure O.1: Players' overall rating vs. players' age (N=45) (with inverse trend line, R^2=18.9%)

Finally, not surprisingly if players quit the experience early, they gave the experience the least favorable overall rating (see Figure O.2).

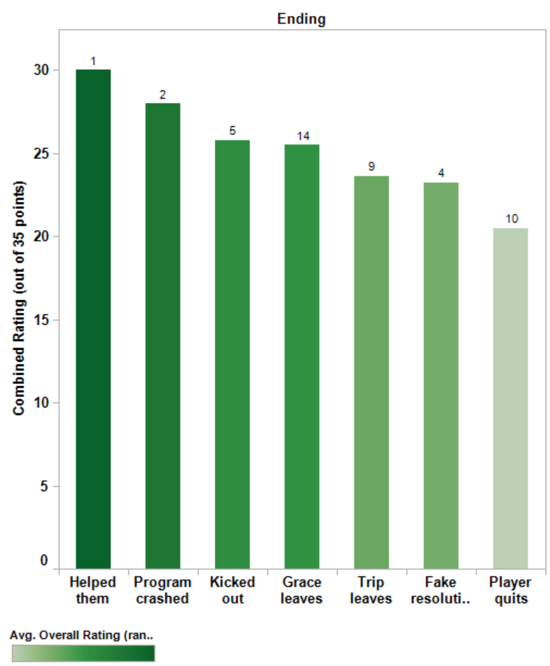


Figure O.2: Players' overall rating vs. players' episode ending (N=45): Least favorable ratings (20.5) came from the ten players who quit the experience early.

APPENDIX P: MYERS-BRIGGS PERSONALITY TESTS

With only 30 participants for this portion of the study, it is difficult to make strong conclusions about the influence of personality on the player experience. Some authors have been able to make links between personality and style of play, but it required long-term survey and an analysis of a large number of players (Bateman and Boon, 2005). The results of my data collection is presented in Figure P.1 (left), showing the number of participants in each of the 16 personality groups (six personalities had no representatives). Players' personalities types were not evenly distributed. Of the 30 participants who did the Myers-Briggs Personality test, 9 of them are ENFP; 6 of them are INTJ; 4 are ENTP; 2 of ESTP; 2 of ISTJ, 2 of ISFJ; 2 of INFP; 1 of ENTJ; 1 of ENFJ; 1 of ISTP.

Looking at specific dichotomies, there were no significant differences along any of the dimensions. One theory was that extroverts would ranks the experience higher than introverts, but this was not the case in our small sampling. There appeared to be a slight favoring by "feelers" over "thinkers", but not big enough to claim statistical significance (see Figure P.1, right).

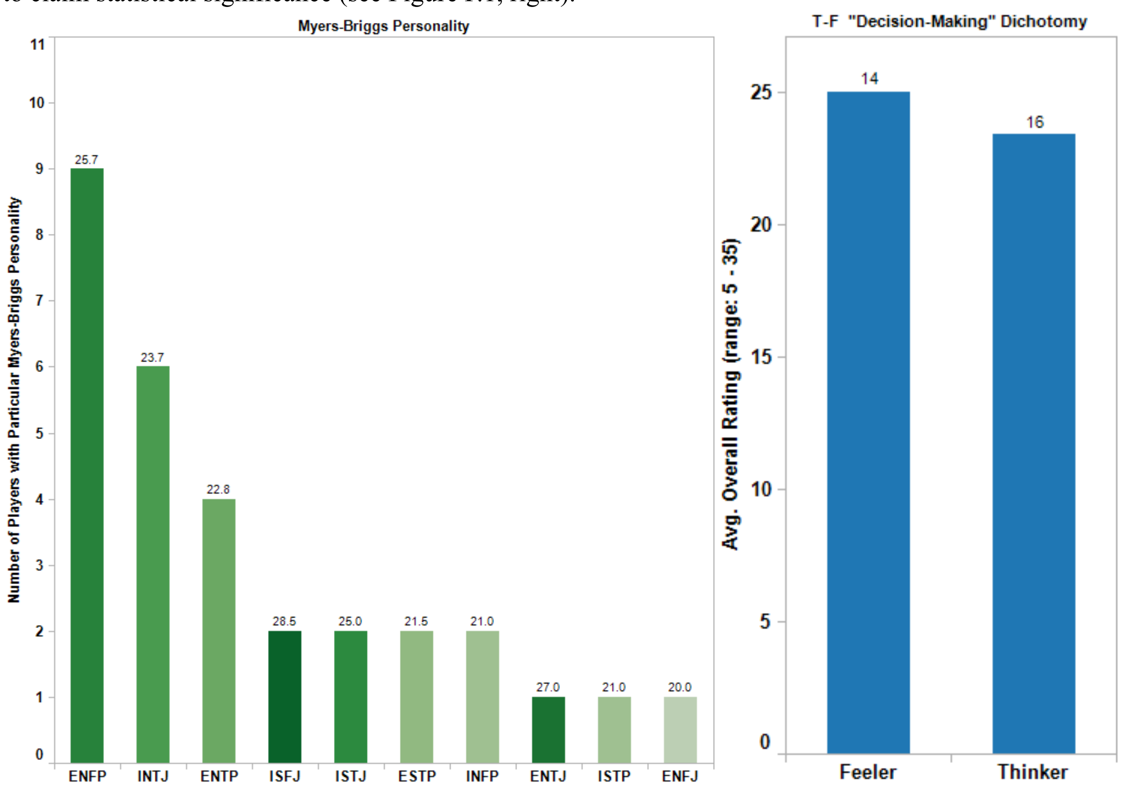


Figure P.1: Personality test results from Beall study (N=33) (left) Myers Brigs Personality Types for thirty players (N=30) with overall rating marked above the bar and with color contrast. (right) The overall rating comparison between "feelers" and "thinkers", one of four dichotomies in the Myers Briggs personality test.

APPENDIX Q: INTERPERSONAL DISTANCE

In this appendix I provide the statistical analysis for the interpersonal distance (IPD) analyses referenced in Chapter 7. First, I outline the statistical analyses of the interface IPD differences from the Atlanta interface comparison study (paired-samples T-tests) and the Beall Center (independent-samples T-test). Then, I include data regarding the difference in IPD between Atlanta (narrow) and the Beall Center (wide) for the desktop-based interaction. Finally, I provide data for the gender differences in IPD.

IPD differences among three versions of *Façade* in the interface comparison study (Atlanta):
Participants: 12 participants in Atlanta
Result: there is a significant difference in IPD (for both Trip and Grace) between AR and SB (For IPD-Grace, t=-3.348, sig=.007; For IPD-Trip, t=-3.951, sig=.003), and between AR and KB (For IPD-Grace, t=-3.490, sig=.005; For IPD-Trip, t=-4.131, sig=.002). There is no significant difference of IPD between SB and KB.
(Note: t is negative because people in the AR version are closer...)

Table Q.1: Paired-samples T-test for IPD in three versions of *Façade* (N=12)

		Paired Differences							
					95% Confidence Interval of the Difference				
		Mean	Std. Deviation	Std. Error Mean	Lower	Upper	t	df	Sig. (2-tailed)
Pair 1	AR_IPD_G - SB_IPD_G	-.43641	.43235	.13036	-.72687	-.14595	-3.348	10	.007
Pair 2	SB_IPD_G - KB_IPD_G	-.09199	.54928	.16561	-.46100	.27702	-.555	10	.591
Pair 3	AR_IPD_G - KB_IPD_G	-.57873	.57437	.16581	-.94367	-.21380	-3.490	11	.005
Pair 4	AR_IPD_T - KB_IPD_T	-.59795	.50143	.14475	-.91655	-.27936	-4.131	11	.002
Pair 5	AR_IPD_T - SB_IPD_T	-.43165	.36237	.10926	-.67509	-.18821	-3.951	10	.003
Pair 6	SB_IPD_T - KB_IPD_T	-.15629	.43172	.13017	-.44632	.13374	-1.201	10	.258

IPD differences among two versions of *Façade* during the Beall Center installation:
Participants: 126 participants in KB and 106 participants in AR at the Beall
Result: there is a significant difference in IPD (for both Trip and Grace) between AR and KB (For IPD-Grace, t=5.563, p=0.000; For IPD-Trip, t=2.854, p=0.005).

Table Q.2: Independent-samples T-test for IPD at the Beall, comparing AR (N=106) and KB (N=126) *Façade*

		Levene's Test for Equality of Variances		t-test for Equality of Means						
									95% Confidence Interval of the Difference	
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper
IPDTrip	Equal variances assumed	6.091	.014	2.854	230	.005	.19983	.07001	.06190	.33777
	Equal variances not assumed			2.929	226.192	.004	.19983	.06823	.06538	.33428
IPDGrace	Equal variances assumed	33.346	.000	5.563	230	.000	.44323	.07968	.28624	.60023
	Equal variances not assumed			5.843	195.540	.000	.44323	.07586	.29362	.59284

IPD differences between KB version of *Façade* in Atlanta and KB version at the Beall Center (FOV test)
Participants: 12 participants in KB in Atlanta and 126 participants in KB at the Beall
Result: there is no significant difference in IPD (for both Trip and Grace) between KB Atlanta and KB Beall, despite the change in field of view. (Note: t is negative because people in the AR version are closer...)

Table Q.3: Independent-samples T-test for IPD (FOV change), comparing KB *Façade* in Atlanta (N=12) and KB *Façade* at the Beall (N=126)

		Levene's Test for Equality of Variances		t-test for Equality of Means						
									95% Confidence Interval of the Difference	
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper
IPDTrip	Equal variances assumed	.667	.416	-1.327	136	.187	-.23478	.17699	-.58478	.11523
	Equal variances not assumed			-1.719	15.243	.106	-.23478	.13661	-.52554	.05599
IPDGrace	Equal variances assumed	2.040	.155	-.689	136	.492	-.15014	.21792	-.58108	.28080
	Equal variances not assumed			-.977	16.420	.343	-.15014	.15370	-.47529	.17502

IPD differences among gender across both installations of *AR Façade*:
Participants: 45 total participants (12 AR players in Atlanta, 33 AR players at the Beall); 22 males, 23 females.

Result: For IPD Grace, there is a significant different between genders (t=-2.133, p=0.039). For IPD Trip, there is no significant different between genders (t=-1.661, p=0.104).
(Note: t is negative because people in the AR version are closer...)

Table Q.4: Independent-samples T-test for IPD across genders, from both installations of *AR Façade* (N=45)

		Levene's Test for Equality of Variances		t-test for Equality of Means						
									95% Confidence Interval of the Difference	
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper
IPDTrip	Equal variances assumed	.004	.952	-1.661	43	.104	-.17561	.10571	-.38879	.03757
	Equal variances not assumed			-1.662	42.943	.104	-.17561	.10568	-.38875	.03753
IPDGrace	Equal variances assumed	.844	.363	-2.133	43	.039	-.26530	.12441	-.51619	-.01441
	Equal variances not assumed			-2.149	39.753	.038	-.26530	.12348	-.51491	-.01569

APPENDIX R: WIZARD DOCENT INVESTIGATIONS (INTERVIEW GUIDES)

Interview Guide (Before 11-week installation)

Have you ever performed as a wizard like this before? If so, when?

Do you understand the tasks you are asked to perform?

Do you have any problems with either of the specific interfaces ("speech recognition" or "discourse")?

Which interface ("speech recognition" or "discourse") do you think will be easier to use?

Which one do you think you will use more often?

Describe your understanding of speech recognition software:

Describe your understanding of vision-based gesture recognition:

Interview Guide (At the end of the 11-week installation)

Tell me about your experience as a wizard:

How would you compare the speech recognition and discourse act version of the wizard interface? Which did you prefer (and why?)?

How did you perform your task? What was your strategy? How did this change over time?

Was the wizard performance difficult or easy? Explain any issues that occurred during the course of the experiment:

How (if at all) did your actions as a wizard performer impact the player's experience in *AR Façade*? How do you feel you performed on the wizard task?

Describe your understanding of speech recognition software:

Describe your understanding of vision-based gesture recognition:

Do you have any questions for me?

Reaffirming consent of your participation, is it still ok that we use your data?

REFERENCES

Aarseth, E. (2004). Genre Trouble: Narrativism and the Art of Simulation. *First Person: New Media as Story, Performance, and Game.* MIT Press.

Aarseth, E. J. (1997). Cybertext: perspectives on ergodic literature. Johns Hopkins University Press. Baltimore, MD, USA.

Adams, E. (1999). Three Problems for Interactive Storytellers. *Gamasutra.* December.

Adams, E. (2006). A New Vision for Interactive Stories (Lecture). *Game Developers' Conference.*

Agre, P. (1997). Computation and Human Experience. Oxford University Press.

Aoki, P. M., Grinter, R. E., Hurst, A., Szymanski, M. H., Thornton, J. D., and Woodruff, A. (2002). Sotto voce: exploring the interplay of conversation and mobile audio spaces. *CHI '02: Proceedings of the SIGCHI conference on Human factors in computing systems.* ACM Press. New York, NY, USA. 431–438.

Argyle, M. (1988). Bodily Communication. New York: Methuen & Co.

Aristotle (330 BC). Poetics. Penguin Books, London, England, 1996.

Auslander, P. (1999). Liveness: Performance in a Mediatized Culture. Routledge.

Azuma, R. (1997). Making Direct Manipulation Work in Virtual Reality (Course). *ACM SIGGRAPH.*

Bailenson, J. N., Blascovich, J., Beall, A. C., and Loomis, J. M. (2003). Interpersonal distance in immersive virtual environments. *Personality and Social Psychology Bulletin.* 29 1-15.

Bailenson, J. N., Blascovich, J., Beall, A. C., and Loomis, J. M. (2001). Equilibrium revisited: Mutual gaze and personal space in virtual environments. *Presence: Teleoperators and Virtual Environments.* 10 583-598.

Baker, P. M., Stein, R. J., Khakshour, A., and Sherman, W. B. (1998). Virtual Environments For Army Tasks. *ARL Federated Lab Report on Advanced and Interactive Display.*

Barfield, W., Zeltzer, D., Sheridan, T., and Slater, M. (1995). Presence and performance within virtual environments. *Virtual environments and advanced interface design.* Oxford University Press, Inc.. New York, NY, USA. 473–513.

Bartle, R. (1996). Hearts, Clubs, Diamonds, Spades: Players Who Suit MUDs. http://www.brandeis.edu/pubs/jove/HTML/v1/bartle.html (accessed on 9/22/08)

Bateman, C. and Boon, R. (2005). 21st Century Game Design. Charles River Media.

Bates, J. (1992). The Nature of Character in Interactive Worlds and The Oz Project. *School of Computer Science, Carnegie Mellon University.*

Bateson, G. (1972). Steps to an Ecology of Mind. Univ. of Chicago Press.

Baudrillard, J. (1981). Simulacra and Simulation. University of Michigan Press.

Baum, F. (1900). The Wonderful Wizard of Oz. George Hill, Chicago.

Baxter, J. C. (1970). Interpersonal spacing in natural settings. *Sociometry.* 33 (4), 444-456.

Bederson, B. (2004). Interfaces for Staying in the Flow. *Ubiquity.* 5 (27).

Benford, S., Crabtree, A., Flintham, M., Drozd, A., Anastasi, R., Paxton, M., Tandavanitj, N., Adams, M., and Row-Farr, J. (2006). Can you see me now?. *ACM Trans. Comput.-Hum. Interact.*. ACM Press. New York, NY, USA. 13 (1), 100–133.

Benjamin, W. (1936). The Work of Art in the Age of Mechanical Reproduction. *Film Theory and Criticism: Introductory Readings.* New York: Oxford University Press. 731-751.

Billinghurst, M., Kato, H., and Poupyrev, I. (2001). The MagicBook: A Transitional AR Interface. *Computers and Graphics.* November 745-753.

Biocca, F. and Delaney, B. (1995). Immersive virtual reality technology. *Communication in the age of virtual reality.* Lawrance Erlbaum Associates. 57-124.

Biocca, F. (1997). The cyborg's dilemma: Progressive embodiment in virtual environments. *Journal of Computer-Mediated Communication.* 3 (2).

Boal, A. (1979). Theatre of the Oppressed. Theatre Communications Group.

Bobick, A. F., Intille, S. S., Davis, J. W., Baird, F., Pinhanez, C. S., Campbell, L. W., Ivanov, Y. A., Schũtte, A., and Wilson, A. (2000). Perceptual user interfaces: the KidsRoom. *Commun. ACM.* ACM Press. New York, NY, USA. 43 (3), 60–61.

Bolter, J. D. and Grusin, R. (1999). Remediation: Understanding New Media. Cambridge, MA: MIT Press.

Bolter, J. D., MacIntyre, B., Gandy, M., and Schweitzer, P. (2006). New Media and the Permanent Crisis of Aura. *Convergence: The International Journal of Research into New Media Technologies.* 12 (1), 21-39.

Bolter, J. D. and Gromala, D. (2003). Windows and Mirrors. Interaction Design, Digital Art, and the Myth of Transparency. MIT Press.

Bolter, J. D. (1992). Literature in Electronic Writing Space. *Literacy Online: The Promise (and Peril) of Reading (and Writing) with Computers.* U of Pittsburgh P. 19-42.

Brannon, R. (1976). The male sex role: Our culture's blue-print for manhood, what it's done for us lately. *The forty-nine percent majority: The male sex role.* Reading, MA: Addison-Wesley Publishing Company.

Brecht, B. and Willett, J. (1964). Brecht on Theatre: The Development of an Aesthetic. Hill and Wang.

Brennan, S. (1990). Conversation as direct manipulation: An iconoclastic view. *The Art of Human-Computer Interface Design.* Addison-Wesley Publishing Co., Inc.

Cassell, J., Bickmore, T., Billinghurst, M., Campbell, L., Chang, K., Vilhjlmsson, H., and Yan, H. (1999). Embodiment in conversational interfaces: Rea. *CHI '99: Proceedings of the SIGCHI conference on Human factors in computing systems.* ACM Press. New York, NY, USA. 520–527.

Cassell, J., Bickmore, T. W., Vilhjalmsson, H. H., and Yan, H. (2000). More than just a pretty face: affordances of embodiment. *Intelligent User Interfaces.* 52-59.

de Certeau, M. (2002). Practice of Everyday Life. University of California Press.

Chastine, J. W., Nagel, K., Zhu, Y., and Yearsovich, L. (2007). Understanding the Design Space of Referencing in Collaborative Augmented Reality Environments. *Graphics Interface.*

Chen, J. (2007). Flow in games (and everything else). *Communications of the ACM.* ACM. New York, NY, USA. 50 (4), 31–34.

Cheok, A. D., Goh, K. H., Liu, W., Farbiz, F., Fong, S. W., Teo, S. L., Li, Y., and Yang, X. (2004). Human Pacman: a mobile, wide-area entertainment system based on physical, social, and ubiquitous computing. *Personal Ubiquitous Comput..* Springer-Verlag. London, UK. 8 (2), 71–81.

Cheverst, K., Davies, N., Mitchell, K., Friday, A., and Efstratiou, C. (2000). Developing a context-aware electronic tourist guide: some issues and experiences. *CHI '00: Proceedings of the SIGCHI conference on Human factors in computing systems.* ACM Press. New York, NY, USA. 17-24.

Cohen, J. (1960). A coefficient of agreement for nominal scale. *Journal of Educational Psychology Measure.* 20, 37-46.

Coleridge, S. T. (1817). Biographia Literaria.

Corradini, A., Mehta, M., Bernsen, N., and Charfuelan, M. (2005). Animating an interactive conversational character for an educational game system. *IUI '05: Proceedings of the 10th international conference on Intelligent user interfaces.* ACM Press. New York, NY, USA. 183–190.

Costikyan, G. (2000). Where Stories End and Games Begin. *Game Developer.* 7 (9), 44-53.

Cox, C. and Fu, D. (2005). AI for automated combatants in a training application. *IE2005: Proceedings of the second Australasian conference on Interactive entertainment.* Creativity & Cognition Studios Press. Sydney, Australia, Australia. 57-64.

Crowston, K. and Malone, T. W. (1988). Intelligent software agents. *Byte.* December, 267-271.

Csikszentmihalyi, M. (1990). Flow: The Psychology of Optimal Experience. New York: Perennial.

Csikszentmihalyi, M. (1975). Beyond Boredom and Anxiety. Jossey-Bass: San Francisco, CA.

Czerwinski, M., Horvitz, E., and Cutrell, E. (2001). Subjective Duration Assessment: An Implicit Probe for Software Usability. *Human-Computer Interaction (IHM-HCI).* 167-170.

Dahlback, N., Jonsson, A., and Ahrenberg, L. (1993). Wizard of Oz studies: why and how. *IUI '93: Proceedings of the 1st international conference on Intelligent user interfaces.* ACM Press. New York, NY, USA. 193–200.

Dennett, D. (1978). Where Am I?. *Brainstorms: Philosophical Essays on Mind and Psychology.* Bradford Books.

Dourish, P. (2001). Where the action is. MIT Press. Cambridge.

Dow, S., MacIntyre, B., Lee, J., Oezbek, C., Bolter, J. D., and Gandy, M. (2005). Wizard of Oz Support throughout an Iterative Design Process. *IEEE Pervasive Computing (Special Issue on Rapid Prototyping).* November.

Dow, S., Lee, J., Oezbek, C., MacIntyre, B., Bolter, J. D., and Gandy, M. (2005). Exploring Spatial Narratives and Mixed Reality Experiences in Oakland Cemetery. *ACM SIGCHI Conference on Advances in Computer Entertainment (ACE'05).*

Dow, S., Mehta, M., Harmon, E., MacIntyre, B., and Mateas, M. (2007). Presence and Engagement in an Interactive Drama. *In Conference on Human Factors in Computing Systems (CHI'07), San Jose, CA, USA.*

Dow, S., Mehta, M., Lausier, A., MacIntyre, B., and Mateas, M. (2006). Initial lessons from AR Façade, an interactive augmented reality drama. *ACE '06: Proceedings of the 2006 ACM SIGCHI international conference on Advances in computer entertainment technology.* ACM Press. New York, NY, USA.

El-Nasr, M. S. (2007). Interaction, Narrative, and Drama: Creating an Adaptive Interactive Narrative using Performance Arts Theories. *Interaction Studies.* 8 (2).

Ellis, M. (1973). Why People Play. Prentice Hall.

Ellison, B. (2008). Defining Dialogue Systems. *Gamasutra.*

Feiner, S., MacIntyre, B., Höllerer, T., and Webster, A. (1997). A touring machine: Prototyping 3D mobile augmented reality systems for exploring the urban environment. *International Symposium on Wearable Computers.*

Feiner, S., MacIntyre, B., and Seligmann, D. (1993). Knowledge-based augmented reality. *Communications of the ACM.* 36 (7), 52-62.

Fischer, J. and Bartz, D. (2005). Stylized Augmented Reality for Improved Immersion. *VR '05: Proceedings of the 2005 IEEE Conference 2005 on Virtual Reality.* IEEE Computer Society. Washington, DC, USA. 195–202.

Freytag, G. (1863). Die Technik des Dramas.

G. Thomas, L. Johnson, S. D. and Stanford, C. (2001). The Design and Testing of a Force Feedback Dental Simulator. *Journal of Computer Methods and Programs in Biomedicine.*

Garfinkel, H. (1967). Studies in Ethnomethodology. *Studies in Ethnomethodology.* Prentice Hall.

Gerrig, R. J. (1993). Experiencing Narrative Worlds: On the Psychological Activities of Reading. New Haven: Yale UP.

Goffman, E. (1956). Presentation of Self in Everyday Life. Doubleday. New York.

Grasset, R., Dünser, A., Seichter, H., and Billinghurst, M. (2007). The Mixed Reality Book: A New Multimedia Reading Experience. *CHI'2007 Interactivity.*

Green, M. C. and Brock, T. C. (2000). The role of transportation in the persuasiveness of public narratives. *Journal of Personality and Social Psychology.* 79, 701-721.

Green, M. C. (2004). Transportation Into Narrative Worlds: The Role of Prior Knowledge and Perceived Realism. *Discourse Processes.* 38 (2), 247-266.

Hall, E. T. (1966). The Hidden Dimension. Doubleday.

Haller, M., Landerl, F., and Billinghurst, M. (2005). A loose and sketchy approach in a mediated reality environment. *GRAPHITE '05: Proceedings of the 3rd international conference on Computer graphics and interactive techniques in Australasia and South East Asia.* ACM. New York, NY, USA. 371–379.

Hammon, S., P. H. and Smith, T. J. (2007). Player Agency in Interactive Narrative: Audience, Actor & Author. *Artificial and Ambient Intelligence.*

Hayes, G. R., Gardere, L. M., Abowd, G. D., and Truong, K. N. (2008). CareLog: a selective archiving tool for behavior management in schools. *CHI '08: Proceeding of the twenty-sixth annual SIGCHI conference on Human factors in computing systems.* ACM. New York, NY, USA. 685–694.

Heeter, C. (1992). Being there: the subjective experience of presence. *Presence: Teleoper. Virtual Environ.*. MIT Press. Cambridge, MA, USA. 1 (2), 262–271.

Hodges, L. F., Kooper, R., Meyer, T. C., Rothbaum, B. O., Opdyke, D., de Graaff, J. J., Williford, J. S., and North, M. M. (1995). Virtual Environments for Treating the Fear of Heights. *Computer.* IEEE Computer Society Press. Los Alamitos, CA, USA. 28 (7)27–34.

Hodges, L. F., Watson, B., Kessler, G. D., Opdyke, D., and Rothbaum, B. O. (1996). A Virtual Airplane for Fear of Flying Therapy. *VRAIS '96: Proceedings of the 1996 Virtual Reality Annual International Symposium (VRAIS 96).* IEEE Computer Society. Washington, DC, USA.

Hoffman, D. L., Novak, T. P., and Duhachek, A. (2002). The Influence of Goal-Directed and Experiential Activities on Online Flow Experiences. *Journal of Consumer Psychology,* 13.

Holman, D. and Vertegaal, R. (2008). Organic user interfaces: designing computers in any way, shape, or form. *Commun. ACM.* ACM. New York, NY, USA. 51 (6), 48–55.

Hughes, C. E., Stapleton, C. B., Hughes, D. E., and Smith, E. M. (2005). Mixed Reality in Education, Entertainment, and Training. *IEEE Computer Graphics and Applications.* IEEE Computer Society. Los Alamitos, CA, USA. 25 (6), 24-30.

Huizinga, J. (1971). Homo Ludens: A Study of the Play-Element in Culture. Beacon Press.

Hutchins, E., Hollan, J., and Norman, D. A. (1986). Direct manipulation interfaces. *User centered system design: New perspectives on Human-Computer Interaction.* Lawrance Erlbaum. 87–124.

Isbell, C., Omojokun, O., and Pierce, J. (2004). {From Devices to Tasks: Automatic Task Prediction for Personalized Appliance Control}. *Personal and Ubiquitous Computing.* 8 (3), 146–153.

Isbister, K., Hook, K., Sharp, M., and Laaksolahti, J. (2006). The sensual evaluation instrument: developing an affective evaluation tool. *CHI '06: Proceedings of the SIGCHI conference on Human Factors in computing systems.* ACM Press. New York, NY, USA. 1163-1172.

Ishii, H. and Ullmer, B. (1997). Tangible bits: towards seamless interfaces between people, bits and atoms. *CHI '97: Proceedings of the SIGCHI conference on Human factors in computing systems.* ACM Press. New York, NY, USA. 234–241.

Jacob, R. J. K., Girouard, A., Hirshfield, L. M., Horn, M. S., Shaer, O., Solovey, E. T., and Zigelbaum, J. (2007). Reality-based interaction: unifying the new generation of interaction styles. *CHI '07: CHI '07 extended abstracts on Human factors in computing systems.* ACM. New York, NY, USA., 2465–2470.

Johnson, D. and Wiles, J. (2003). Effective Affective User Interface Design in Games. *Ergonomics.* 46 (13-14), 1332-1345.

Johnson, W. L. and Rickel, J. W. (2000). Animated Pedagogical Agents: Face-to-Face Interaction in Interactive Learning Environments. *International Journal of Artificial Intelligence in Education.* 11, 47-78.

Juul, J. (March, 2008). Hardcore players of casual games: Locating the ``casual'' in casual games. Lecture at Georgia Tech.

Katz, S. (1991). Shot by Shot Visualizing from Concept to Screen. Michael Wiese Productions.

Kelso, M. T., Weyhrauch, P., and Bates, J. (1992). Dramatic Presence. *PRESENCE: The Journal of Teleoperators and Virtual Environments, Vol 2, No. 1, MIT Press. Also Technical Report CMU-CS-92-195, School of Computer Science, Carnegie Mellon University, Pittsburgh, PA, December 1992.*

Klemmer, S. R., Sinha, A. K., Chen, J., Landay, J. A., Aboobaker, N., and Wang, A. (2000). Suede: a Wizard of Oz prototyping tool for speech user interfaces. *UIST '00: Proceedings of the 13th annual ACM symposium on User interface software and technology.* ACM Press. New York, NY, USA. 1–10.

Knickmeyer, R. L. and Mateas, M. (2005). Preliminary evaluation of the interactive drama facade. *CHI '05: CHI '05 extended abstracts on Human factors in computing systems.* ACM Press. New York, NY, USA. 1549–1552.

Koleva, B., Taylor, I., Benford, S., Fraser, M., Greenhalgh, C., Schnadelbach, H., vom Lehn, D., Heath, C., Row-Farr, J., and Adams, M. (2001). Orchestrating a mixed reality performance. *CHI '01: Proceedings of the SIGCHI conference on Human factors in computing systems.* ACM Press. New York, NY, USA. 38–45.

Kolodner, J. L. (1993). Case-Based Reasoning. Morgan Kaufmann Publishers Inc.. San Francisco, CA, USA.

Kramer, G. (1995). Sound and Communication in Virtual Reality. *Communication in the age of virtual reality.* Lawrance Erlbaum Associates. 57-124.

Krueger, M. (1990). Artificial Reality II. Addison-Wesley Publishing Co., Inc..

Lang, A., Dillon, K., and Dong, Q. (1995). The Effects of Emotional Arousal and Valence on Television Viewers' Cognitive Capacity and Memory. *Journal of Broadcasting & Electronic Media.*

Lang, T., MacIntyre, B., Zugaza, I.J. (2008). Massively Multiplayer Online Worlds as a Platform for Augmented Reality Experiences. In Ext. Abstracts of Conference of Virtual Reality, 67-70.

Lanier, J. and Biocca, F. (1992). An Insider's View of the Future of Virtual Reality. *Journal of Communication.* 42 (4), 150-172.

Laurel, B. (1993). Computers as Theatre. Reading, MA: Addison-Wesley Publishing Company.

Laurel, B., Strickland, R., and Tow, R. (1994). Placeholder: landscape and narrative in virtual environments. *SIGGRAPH Comput. Graph.*. ACM Press. New York, NY, USA. 28 (2), 118–126.

Laurel, B. (2004). Response to "A Preliminary Poetics for Interactive Drama and Games". *First Person: New Media as Story, Performance, and Game.* MIT Press.

Lave, J. and Wenger, E. (1991). Situated Learning: Legitimate Peripheral Participation. Cambridge University Press.

Laws, R. (2001). Robin's Laws of Good Game Mastering. Steve Jackson Games. http://www.sjgames.com/robinslaws/. (accessed on 9/22/08)

Lazzaro, N. (2004). Why We Play Games: Four Keys to More Emotion Without Story.

Lee, K. M. and Nass, C. (2003). Designing social presence of social actors in human computer interaction. *CHI '03: Proceedings of the SIGCHI conference on Human factors in computing systems.* ACM Press. New York, NY, USA. 289–296.

Levant, R. F. (1992). Toward the reconstruction of masculinity. *Journal of Family Psychology.* 5, 379-402.

Liebler, C. A. and Sandefur, G. D. (2002). Gender differences in the exchange of social support with friends, neighbors, and co-workers at midlife. *Social Science Research.* 31(3), 364-391.

Loizos, C. (1969). Play Behavior in Higher Primates. *Primate Ethology.* 176-218.

Lombard, M. and Ditton, T. B. (1997). At the heart of it all: The concept of presence. *Journal of Computer-Mediated-Communication.* 3 (2).

MacIntyre, B. and Bolter, J. D. (2003). Single-Narrative, Multiple Point-of-View Dramatic Experiences in Augmented Reality. *Journal of Virtual Reality.*

MacIntyre, B., Gandy, M., Dow, S., and Bolter, J. D. (2004). DART: a toolkit for rapid design exploration of augmented reality experiences. *UIST '04: Proceedings of the 17th annual ACM symposium on User interface software and technology.* ACM Press. New York, NY, USA. 197–206.

MacIntyre, B., Bolter. J., Moreno. E. and Hannigan, B. (2001). Augmented Reality as a New Media Experience. *In International Symposium on Augmented Reality (ISAR 2001).*

Magerko, B. S. (2006). Player modeling in the interactive drama architecture. University of Michigan. Ann Arbor, MI, USA.

Mandryk, R. L., Atkins, M. S., and Inkpen, K. M. (2006). A continuous and objective evaluation of emotional experience with interactive play environments. *CHI '06: Proceedings of the SIGCHI conference on Human Factors in computing systems.* ACM Press. New York, NY, USA. 1027–1036.

Marsh, T., Wright, P., and Smith, S. (2001). Evaluation for the design of experience in virtual environments: Modeling breakdown of interaction and illusion. *Journal of CyberPsychology and Behavior, Special Issue on Presence.* 4 (2) 225-238.

Marsh, T. (2003). Presence as experience: film informing ways of staying there. *Presence: Teleoper. Virtual Environ.*. MIT Press. Cambridge, MA, USA. 12 (5) 538–549.

Marsh, T., Yang, K., and Shahabi, C. (2006). Game development for experience through staying there. *sandbox '06: Proceedings of the 2006 ACM SIGGRAPH symposium on Videogames.* ACM Press. New York, NY, USA. 83–89.

Marsh, T., Yang, K., Shahabi, C., Wong, W. L., Nocera, L., Carriazo, E., Varma, A., Yoon, H., and Kyriakakis, C. (2005). Automating the detection of breaks in continuous user experience with computer games. *CHI '05: CHI '05 extended abstracts on Human factors in computing systems.* ACM Press. New York, NY, USA. 1629–1632.

Marsh, T. (2003). Staying there: An activity-based approach to narrative design and evaluation as an antidote to virtual corpsing. *Being there: Concepts, effects and measurements of user presence in synthetic environments.* Amsterdam, The Netherlands: IOS Press.

Mateas, M. and Stern, A. (2000). Towards Integrating Plot and Character for Interactive drama.

Mateas, M. and Stern, A. (2004). A Behavior Language: Joint Action and Behavioral Idioms. *Life-like Characters. Tools, Affective Functions and Applications.* Springer.

Mateas, M. and Stern, A. (2003). Façade: An Experiment in Building a Fully-Realized Interactive Drama. *Game Developer's Conference: Game Design Track.*

Mateas, M. and Stern, A. (2004). Natural language understanding in Facade: Surface-text processing. *Technologies for Interactive Digital Storytelling and Entertainment (TIDSE).*

Mateas, M. (2001). A Preliminary Poetics for Interactive Drama and Games. *Digital Creativity.* 12 (3) 140-152.

Mateas, M. (2002). Interactive Drama, Art, and Artificial Intelligence. PhD Thesis. Carnegie Mellon University.

Mateas, M. (2005). Procedural Literacy: Educating the New Media Practitioner. *On The Horizon. Special Issue. Future of Games, Simulations and Interactive Media in Learning Contexts.* 13 (1).

Maulsby, D., Greenberg, S., and Mander, R. (1993). Prototyping an intelligent agent through Wizard of Oz. *CHI '93: Proceedings of the SIGCHI conference on Human factors in computing systems.* ACM Press. New York, NY, USA. 277–284.

Mazalek, A., Wood, A., and Ishii, H. (2001). genieBottles: An Interactive Narrative in Bottles.

Mazalek, A., Davenport, G., and Ishii, H. (2002). Tangible Viewpoints: A Physical Interface for Exploring CharacterDriven Narratives.

McMahon, A. (2003). Immersion, Engagement, and Presence. *The Video Game Theory Reader.* Routledge.

Meehan, J. R. (1976). The metanovel: writing stories by computer. Yale University. New Haven, CT, USA.

Meehan, M., Insko, B., Whitton, M., and Frederick P. Brooks, J. (2002). Physiological measures of presence in stressful virtual environments. *SIGGRAPH '02: Proceedings of the 29th annual conference on Computer graphics and interactive techniques.* ACM Press. New York, NY, USA. 645–652.

Meehan, M. J. (2001). Physiological reaction as an objective measure of presence in virtual environments.

Mehta, M., Dow, S., MacIntyre, B., and Mateas, M. (2007). Evaluating a Conversation-centered Interactive Drama. *Conference on Autonomous Agents and Multiagent Systems.*

Merleau-Ponty, M. (2005). Phenomenology of Perception. Routledge.

Milgram, P. and Kishino, F. (1994). A Taxonomy of Mixed Reality Visual Displays. *IEICE Transactions on Information Systems.* E77-D (12)

Moreno, E., MacIntyre, B., and Bolter, J. D. (2001). Alice's Adventures in New Media - An exploration of interactive narratives in augmented reality. *CAST.*

Mori, M. (2005). On The Uncanny Valley. *Humanoids-2005 Workshop.*

Murray, J. (1997). Hamlet on the Holodeck. Cambridge, MA: MIT Press.

Murray, J. (2004). From Game-Story to Cyberdrama. *First Person: New Media as Story, Performance, and Game.* MIT Press .

Nell, V. (1988). Lost in a Book: The Psychology of Reading for Pleasure. New Haven: Yale UP.

Nielsen, J. (1994). Heuristic Evaluation. *Usability Inspection Methods.* John Wiley & Sons.

Norman, D. A. (1998). The Invisible Computer. *The Invisible Computer.* MIT Press.

O'Neil, J. M. (1981). Patterns of gender role conflict and strain: Sexism and fear of femininity in men's lives. *The Personnel and Guidance Journal.* 60 203-210.

Oviatt, S. (1996). User-Centered Modeling for Spoken Language and Multimodal Interfaces. *IEEE MultiMedia.* 3 (4) 26–35.

Oviatt, S. L., Cohen, P. R., and Wang, M. (1994). Toward interface design for human language technology: modality and structure as determinants of linguistic complexity. *ISSD-93: Selected papers presented at the international symposium on Spoken dialogue.* Elsevier North-Holland, Inc.. New York, NY, USA. 283–300.

Pace, S. (2004). A grounded theory of the flow experiences of web users. *Int. J. Human-Computer Studies.* Academic Press, Inc.. Duluth, MN, USA. 60 (3) 327–363.

Paiva, A., Dias, J., Sobral, D., Aylett, R., Sobreperez, P., Woods, S., Zoll, C., and Hall, L. (2004). Caring for Agents and Agents that Care: Building Empathic Relations with Synthetic Agents. *AAMAS '04: Proceedings of the Third International Joint Conference on Autonomous Agents and Multiagent Systems.* IEEE Computer Society. Washington, DC, USA. 194–201.

Parsons, H. M. (1974). What happened at Hawthorne? *Science.* 183 (93).

Parsons, T. D. and Rizzo, A. A. (2008). Affective Outcomes of Virtual Reality Exposure Therapy for Anxiety and Specific Phobias. *Journal of Behavior Therapy and Experimental Psychiatry.* 39, 250-261.

Patten, J. and Ishii, H. (2007). Mechanical constraints as computational constraints in tabletop tangible interfaces. *CHI '07: Proceedings of the SIGCHI conference on Human factors in computing systems.* ACM. New York, NY, USA. 809–818.

Paulos, E. (2007). HCI cannot be used to evaluate art (Workshop on Evaluation in Media Experiences). *CHI '07 extended abstracts on Human factors in computing systems.*

Pausch, R., Snoddy, J., Taylor, R., Watson, S., and Haseltine, E. (1996). Disney's Aladdin: first steps toward storytelling in virtual reality. *SIGGRAPH '96: Proceedings of the 23rd annual conference on Computer graphics and interactive techniques.* ACM Press. New York, NY, USA. 193–203.

Pearce, C. (2004). Towards a Game Theory of Game. *First Person: New Media as Story, Performance, and Game.* MIT Press.

Penny, S. (2004). Representation, Enaction, and the Ethics of Simulation. *First Person: New Media as Story, Performance, and Game.* MIT Press.

Perlin, K. (2004). Can there be a form between a game and a story?. *First Person: New Media as Story, Performance, and Game.* MIT Press.

Pimentel, K. and Teixeira, K. (1993). Virtual Reality: Through the New Looking Glass. Windcrest McCraw Hill.

Pinelle, D., Wong, N., and Stach, T. (2008). Heuristic evaluation for games: usability principles for video game design. *CHI '08: Proceeding of the twenty-sixth annual SIGCHI conference on Human factors in computing systems.* ACM. New York, NY, USA. 1453–1462.

Pinhanez, C. S., Davis, J. W., Intille, S., Johnson, M. P., Wilson, A. D., Bobick, A. F., and Blumberg, B. (2000). Physically interactive story environments. *IBM Syst. J.*. IBM Corp.. Riverton, NJ, USA. 39 (3-4)438–455.

Prendinger, H., Mori, J., and Ishizuka, M. (2005). Recognizing, modeling, and responding to users affective states. *In User Modeling'05.*

Putnam, H. (1982). Reason, Truth and History. Cambridge: Cambridge University Press.

Reeves, B. and Nass, C. (1996). The media equation: how people treat computers, television, and new media like real people and places. Cambridge University Press. New York, NY, USA.

Rekimoto, J. (2002). SmartSkin: an infrastructure for freehand manipulation on interactive surfaces. *CHI '02: Proceedings of the SIGCHI conference on Human factors in computing systems.* ACM. New York, NY, USA. 113–120.

Rekimoto, J. (2008). Organic interaction technologies: from stone to skin. *Commun. ACM.* ACM. New York, NY, USA. 51 (6) 38–44.

Rheingold, H. (1991). Virtual Reality. Simon and Schuster.

Rice, R. E. (1992). Task analyzability, use of new medium and effectiveness: A multi-site exploration of media richness. *Organization Science.* 3 (4) 475-500.

Riedl, M., Saretto, C. J., and Young, R. M. (2003). Managing interaction between users and agents in a multi-agent storytelling environment. *AAMAS '03: Proceedings of the second international joint conference on Autonomous agents and multiagent systems.* ACM Press. New York, NY, USA. 741–748.

Riedl, M. O. and Young, R. M. (2004). An Intent-Driven Planner for Multi-Agent Story Generation. *AAMAS '04: Proceedings of the Third International Joint Conference on Autonomous Agents and Multiagent Systems.* IEEE Computer Society. Washington, DC, USA. 186–193.

Robinett, W. (1994). Interactivity and Individual Viewpoint in Shared Virtual Worlds: The Big Screen vs. Networked Personal Displays. *Computer Graphics.* 28 (2) 127.

Romero, M. and Mateas, M. (2005). A Preliminary Investigation of Alien Presence. *Proceedings of Human-Computer Interaction International (HCII 2005).*

Roush, W. (2007). Second Earth. *MIT Technology Review.* July/Aug.

Roussou, M. (2004). Learning by doing and learning through play: an exploration of interactivity in virtual environments for children. *Comput. Entertain.*. ACM Press. New York, NY, USA. 2 (1) 1-10.

Russell, J. A., Weiss, A., and Mendelsohn, G. A. (1989). Affect grid: A single-item scale of pleasure and arousal. *Journal of Personality and Social Psychology.* 57 (3) 493-502.

Ryan, M. (2001). Narrative as Virtual Reality. Baltimore: The Johns Hopkins University Press.

Ryan, M. (1997). Interactive Drama: Narrativity in a Highly Interactive Environment. *Modern Fiction Studies.* 43 (3) 677-707.

Ryan, M. (2001). Beyond Myth and Metaphor: The Case of Narrative in Digital Media. *Gamestudies.org.* 1 (1).

Ryokai, K., Marti, S., and Ishii, H. (2004). I/O brush: drawing with everyday objects as ink. *CHI '04: Proceedings of the SIGCHI conference on Human factors in computing systems.* ACM. New York, NY, USA. 303–310.

Salen, K. and Zimmerman, E. (2003). Rules of Play. MIT Press.

Schechner, R. (2002). Performance Studies: An Introduction. Routledge.

Schegloff, E. (1991). Converation Analysis and Socially Shared Cognition. *Perspectives on Socially Shared Cognition.* American Psychological Association.

Schell, J. and Shochet, J. (2001). Designing Interactive Theme Park Rides: Lessons from Disney's Battle for Buccaneer Gold. *Gamasutra.* July

Schubert, T. and Crusius, J. (2002). Five theses on the book problem: Presence in books, film and VR. *Proceedings of the Fifth Annual International Workshop Presence.*

Schutz, A. (1932). Der sinnhafte Aufbau der sozialen Welt: eine Einleitung in die verstehende Soziologie. Wein: Springer.

Sha, X. W., Iachello, G., Dow, S., Serita, Y., St. Julien, T., and Fistre, J. (2003). Continuous sensing of gesture for control of audio-visual media. *In Extended Abstracts of the International Symposium on Wearable Computing (ISWC'03).*

Shank, R., K. A. and Riesbeck, C. (1994). Inside Case-Based Reasoning. Lawrance Erlbaum Associates. New Jersey.

Sharma, M., Mehta, M., Ontañón, S., and Ram, A. (2007). Player Modeling Evaluation for Interactive Fiction. *In Artificial Intelligence and Interactive Digital Entertainment: Workshop on Optimizing Player Satisfaction.*

Sheridan, T. B. (1992). Musings on telepresence and virtual presence. *Presence: Teleoper. Virtual Environ..* MIT Press. Cambridge, MA, USA. 1 (1) 120–126.

Shin, N. (2006). Online learner's 'flow' experience: an empirical study. *British Journal of Educational Technology.* 37 (5) 705–720.

Shneiderman, B. (1983). Direct manipulation: A step beyond programming languages. *Computer.* 16 (8) 57-69.

Shneiderman, B. (1988). A nonanthropomorphic style guide: overcoming the humpty-dumpty syndrome. *The Computing Teacher.* October 9-10.

Shneiderman, B. and Maes, P. (1997). Direct manipulation vs. interface agents. *Interactions.* ACM. New York, NY, USA. 4 (6) 42–61.

Short J., W. E. and B., C. (1976). The social psychology of telecommunications. Wiley. Toronto, ON.

Si, M., Marsella, S. C., and Pynadath, D. V. (2005). Thespian: using multi-agent fitting to craft interactive drama. *AAMAS '05: Proceedings of the fourth international joint conference on Autonomous agents and multiagent systems.* ACM. New York, NY, USA. 21–28.

Slater, M., Howell, J., Steed, A., Pertaub, D., and Garau, M. (2000). Acting in virtual reality. *CVE '00: Proceedings of the third international conference on Collaborative virtual environments.* ACM Press. New York, NY, USA. 103–110.

Slater, M., Brogni, A., and Steed, A. (2003). Physiological Responses to Breaks in Presence: A Pilot Study. *6th Annual International Workshop on Presence.*

Slater, M. (2006). How Colorful was your day? Why questionnaires cannot assess presence in virtual environments. *Presence.* 13 (4) 484-493.

Slater, M., Steed, A., McCarthy, J., and Maringelli, F. (1998). The influence of body movement on presence in virtual environments. *Human Factors.* 40, 469-477.

Smith, D. W. (2006). Husserl. Routledge.

Stanislavski, C. (1936). An Actor Prepares. Theatre Arts Books.

Stapleton, C. and Hughes, C. (2003). Interactive Imagination: Tapping the Emotions through Interactive Story for Compelling Simulations. *IEEE Comput. Graph. Appl..* IEEE Computer Society Press. Los Alamitos, CA, USA. 23 (5) 11–15.

State, A., Livingston, M. A., Garrett, W. F., Hirota, G., Whitton, M. C., Pisano, E. D., and Fuchs, H. (1996). Technologies for augmented reality systems: realizing ultrasound-guided needle biopsies. *SIGGRAPH '96: Proceedings of the 23rd annual conference on Computer graphics and interactive techniques.* ACM Press. New York, NY, USA. 439–446.

Stern, A. and Mateas, M. (2005). Build It to Understand It: Ludology Meets Narratology in Game Design Space. *the International DiGRA Conference, June 16th - 20th, 2005, Vancouver, British Columbia, Canada.*

Steuer, J. (1995). Defining virtual reality: Dimensions determining telepresence. *Communication in the age of virtual reality.* Lawrance Erlbaum Associates. 57-124.

Stokes, J. (1980). Gender Differences in Self-Disclosure to Various Target Persons. *Journal of Counseling Psychology.* 27 (2) 192-198.

Strasberg, L. (1988). A Dream of Passion: The Development of the Method . Plume.

Strauss, A. and Corbin, J. (1990). Basics of Qualitative Research: Grounded Theory Procedures and Techniques. Sage.

Suchman, L. A. (1987). Plans and situated actions: the problem of human-machine communication. Cambridge University Press. New York, NY, USA.

Sutherland, I. (1968). A head-mounted three-dimensional display. *Fall Joint Computer Conference.* 757-764.

Sutton-Smith, B. (1997). The Ambiguity of Play. Harvard University Press.

Swartout, W., Hill, R., Gratch, J., Johnson, W. L., Kyriakakis, C., LaBore, C., Lindheim, R., Marsella, S., Miraglia, D., Moore, B., Morie, J., Rickel, J., Thibaux, M., Tuch, L., Whitney, R., and Douglas, J. (2001). Toward the holodeck: integrating graphics, sound, character and story. *AGENTS '01: Proceedings of the fifth international conference on Autonomous agents.* ACM Press. New York, NY, USA. 409–416.

Sweetser, P., Johnson, D., Sweetser, J., and Wiles, J. (2003). Creating engaging artificial characters for games. *ICEC '03: Proceedings of the second international conference on Entertainment computing.* Carnegie Mellon University. Pittsburgh, PA, USA. 1–8.

Sweetser, P. and Wyeth, P. (2005). GameFlow: a model for evaluating player enjoyment in games. *Computers and Entertainment.* ACM. New York, NY, USA. 3 (3).

Sweetser, P. and Johnson, D. (2004). Player-Centered Game Environments: Assessing Player Opinions, Experiences and Issues. *LECTURE NOTES IN COMPUTER SCIENCE.* 3166 321-332.

Tan, E. S. (1996). Emotion and the structure of narrative film. Film as an emotion machine. Mahwah: Lawrence Erlbaum.

Tang, A., Biocca, F., and Lim, L. (2004). Comparing Differences in Presence during Social Interaction in Augmented Reality versus Virtual Reality Environments: An Exploratory Study. *International Workshop on Presence.*

Thomas, B., Close, B., Donoghue, J., Squires, J., de Bondi, P., Morris, M., and Piekarski, W. (2000). ARQuake: An Outdoor/Indoor Augmented Reality First Person Application. *ISWC '00: Proceedings of the 4th IEEE International Symposium on Wearable Computers.* IEEE Computer Society. Washington, DC, USA. 139.

Thomas, B., Close, B., Donoghue, J., Squires, J., Bondi, P. D., and Piekarski, W. (2002). First Person Indoor/Outdoor Augmented Reality Application: ARQuake. *Personal Ubiquitous Comput..* Springer-Verlag. London, UK. 6 (1) 75–86.

Thomaz, A. L., H. G. and Breazeal, C. (2006). Experiments in Socially Guided Machine Learning: Understanding How Humans Teach. *In Proceedings of the 1st Annual conference on Human-Robot Interaction (HRI).*

Thompson, C. (2007). Halo 3: How Microsoft Labs Invented a New Science of Play. *Wired Magazine.* August.

Thue, D., Bulitko, V., and Spetch, M. (2008). PaSSAGE: A Demonstration of Player Modeling in Interactive Storytelling. *In Artificial Intelligence and Interactive Digital Entertainment (AIIDE'08).*

Traum, D., Roque, A., Leuski, A., Georgiou, P., Gerten, J., Martinovski, B., Narayanan, S., Robinson, S., and Vaswani, A. (2007). Hassan: A Virtual Human for Tactical Questioning. *Proceedings of the 8th SIGdial Workshop on Discourse and Dialoque.*

Turkle, S. (1995). Life on the Screen: Identity in the Age of the Internet. Simon & Schuster Trade.

Ullmer, B. and Ishii, H. (1997). The metaDESK: models and prototypes for tangible user interfaces. *UIST '97: Proceedings of the 10th annual ACM symposium on User interface software and technology.* ACM. New York, NY, USA. 223–232.

Utterback, C. (2004). Unusual Positions -- Embodied Interaction with Symbolic Spaces. *First Person: New Media as Story, Performance, and Game.* MIT Press.

Vlahakis, V., Karigiannis, J., Tsotros, M., Gounaris, M., Almeida, L., Stricker, D., Gleue, T., Christou, I. T., Carlucci, R., and Ioannidis, N. (2001). Archeoguide: first results of an augmented reality, mobile computing system in cultural heritage sites. *VAST '01: Proceedings of the 2001 conference on Virtual reality, archeology, and cultural heritage.* ACM Press. New York, NY, USA. 131–140.

Voida, A. (2008). Exploring a technological hermeneutic: Understanding the interpretation of computer-mediated messaging systems. PhD Thesis. Georgia Institute of Technology.

Vygotsky, L. S. (1978). Mind and society: The development of higher mental processes. Cambridge, MA: Harvard UP.

Wagner, D., Pintaric, T., Ledermann, F., and Schmalstieg, D. (2005). Towards Massively Multi-user Augmented Reality on Handheld Devices. *Pervasive Computing.* 208–219.

Walker, M. A., Fromer, J., Fabbrizio, G. D., Mestel, C., and Hindle, D. (1998). What can I say?: evaluating a spoken language interface to Email. *CHI '98: Proceedings of the SIGCHI conference on Human factors in computing systems.* ACM Press/Addison-Wesley Publishing Co.. New York, NY, USA. 582–589.

Walton, K. (1990). Mimesis as Make-Believe: On the Foundations of the Representational Arts. Cambridge, MA: Harvard UP.

Webster, J. and Ho, H. (1997). Audience engagement in multimedia presentations. *SIGMIS Database.* ACM Press. New York, NY, USA. 28 (2) 63–77.

Weiser, M. (1991). The Computer for the 21st Century. *Scientific American.* 94-104.

Weiser, M. (1994). Creating the Invisible Interface (Invited talk).

Weizenbaum, J. (1966). ELIZA\---a computer program for the study of natural language communication between man and machine. *Commun. ACM.* ACM Press. New York, NY, USA. 9 (1) 36–45.

Weybrew, B. B. (1984). The Zeigarnik Phenomenon Revisited: Implications for Enhancement of Morale. *Perceptual and Motor Skills.* 58, 223-226.

Witmer, B. G. and Singer, M. J. (1998). Measuring presence in virtual environments: A presence questionnaire. *Presence: Teleoperators and Virtual Environments.* 7 (3) 225-240.

Xiao, J. (2001). Understanding the use and utility of anthropomorphic interface agents. *CHI '01: CHI '01 extended abstracts on Human factors in computing systems.* ACM. New York, NY, USA. 409–410.

Yee, N., Bailenson, J. N., Urbanek, M., Chang, F., and Merget, D. (2007). The Unbearable Likeness of Being Digital: The Persistence of Nonverbal Social Norms in Online Virtual Environments. *CyberPsychology & Behavior.* 10 (1), 115–121.

Zimmerman, E. (2004). Narrative, Interactivity, Play, and Games. *First Person: New Media as Story, Performance, and Game.* MIT Press.