

NONVERBAL COMMUNICATION



IN VIRTUAL WORLDS

Understanding and Designing Expressive Characters

EDITED BY: THERESA JEAN TANENBAUM MAGY SEIF EL-NASR MICHAEL NIXON

PRAISE FOR NONVERBAL COMMUNICATION IN VIRTUAL WORLDS

“Tanenbaum, Serif El-Nasr, and Nixon have done an outstanding job of assembling scholars from a wide range of disciplines to explore an essential yet under examined topic in digital media studies. The result is a book in which the authors draw upon a powerfully diverse array of theoretical frameworks from the arts, sciences, and humanities to help us begin to understand how we communicate in these technologically mediated spaces. While not every chapter in this book will prove essential for every digital media scholar, there is clearly something here for everyone with a scholarly interest in the array of related fields concerned with virtual worlds, virtual reality, video games, and other spaces for online interaction. For scholars concerned with video games and play, where so many of our digital landscapes are occupied by individuals using humanoid avatars to interact with both non-human and human actors, this book is a definite must read. *Nonverbal Communication in Virtual Worlds* certainly isn’t the definitive tome on the topic, for as the authors themselves note, the area of study is itself too new and the technology that shapes these spaces continues to advance. Rather, this book serves as an excellent point of departure for continuing work that can help us better understand how human interaction is re-mediated through digital worlds, and how those mechanisms of communication will continue to evolve with the ongoing development of technology.”

– MOSES WOLFENSTEIN, PHD

Associate Director of Research at the Academic Advanced Distributed Learning Co-Laboratory

“Virtual worlds, for both play or work, have come to be an important part of online experience and this collection offers fantastic insight into the role of nonverbal communication in these environments. From issues of avatar appearance and animation, all the way to considerations of empathy and gaze, this is a must read book for anyone wanting to better understand the rich potential of digital spaces.”

– T.L.TAYLOR, PHD

*Associate Professor of Comparative Media Studies at MIT and Co-Author of *Ethnography in Virtual Worlds: A Handbook of Method**

“The world, as reflected by popular media and various technologists, seems stuck in time, continually asking whether we’re losing our humanity whilst we engage in new forms of media. This has gone on since at least print media, but the latest fear and lament is on how we get so lost in our digital screens that interpersonal communications and relationships suffer. *Nonverbal Communication in Virtual Worlds* does a good job of dispelling that myth; in fact, it’s so far ahead of our stuck moment that it’s moved beyond our fears and into the realm of possibilities and practical applications. The chapter authors do a fantastic job of actually exploring what new media--in this case, virtual worlds and 3D online games--afford us, even going so far as to help designers plan for meaningful interaction. Amazingly, there’s so much richness to what’s possible that the book editors, Theresa Tanenbaum, Magy Seif El-Nasr, and Michael Nixon, limited it to just non-verbal communication. Indeed, the various authors look at how we communicate in these spaces through gestures, gaze, appearance, and performance--all things we feared were being lost to us by engaging in screen life--all things that define us as humans living in cultural contexts. I cannot recommend a better book that so nonchalantly dismisses our technophobic inclinations, giving us instead frameworks and heuristics for realizing new media’s potential.”

– MARK CHEN, PHD

Spare-time Game Designer, Part-time Professor at Pepperdine University, UW Bothell, and UOIT, and Author of Leet Noobs: The Life and Death of an Expert Player Group in World of Warcraft

Comprehensive and practical! This book provides an all-inclusive review of the theories, history and current practice of creating non-verbal behaviors for digital avatars. A great guide to anyone who wants to study in this young and fast-growing interdisciplinary field!

– MEI SI, PHD

Assistant Professor of Cognitive Science, Rensselaer Polytechnic Institute

NONVERBAL COMMUNICATION IN VIRTUAL WORLDS:

UNDERSTANDING AND DESIGNING EXPRESSIVE CHARACTERS

Edited by

Theresa Jean Tanenbaum, Magy Seif El-Nasr, & Michael Nixon

**Nonverbal Communication in Virtual Worlds:
Understanding and Designing Expressive Characters**

Copyright © by
Theresa Jean Tanenbaum, Magy Seif El-Nasr, Michael Nixon
and ETC Press 2014
<http://press.etc.cmu.edu/>

Design Direction by Shirley Yee

ISBN: 978-1-304-81204-9
Library of Congress Control Number: 2014931252

TEXT: The text of this work is licensed under a Creative Commons
Attribution-NonCommerical-NonDerivative 2.5 License
(<http://creativecommons.org/licenses/by-nc-nd/2.5/>)



IMAGES: All images appearing in this work are property
of the respective copyright owners,
and are not released into the Creative Commons.
The respective owners reserve all rights.

ACKNOWLEDGEMENTS

The editors would like to acknowledge IARPA for initial support of the social signaling in virtual worlds project, which was the beginning of a bigger investigation on nonverbal behavior and puppeteering in the virtual environment, now manifested in this book.

Theresa would like to thank Jim and Justine Bizzocchi, and Alissa N. Antle for their patience with him while he distracted himself from his dissertation with this project. He also would like to thank his brilliant wife, Karen, who kept him sane during an insane undertaking, and his parents – Annette and Ronald – who have supported him unconditionally.

Michael would like to thank his supervisors Steve DiPaola and Magy Seif El-Nasr. He would also like to thank his wife Patricia for her ongoing support of his graduate career.

Magy would like to thank Theresa and Michael for all their help putting this together. She would also like to thank Suzanne deCastle and Jennifer Jenson for all their encouragement and constructive feedback on the social signaling project which started this piece of research and resulted in the development of the idea behind this book. She would also like to thank Katherine Isbister, Jeffery Ventrella, and Bardia Aghabeigi who worked with her on the non-verbal and puppeteering project, their thoughts and ideas were also what sparked the idea for this book. Lastly, she would like to thank her family, parents Wafika and Safwat and brother Hossam for their tremendous support for every idea that comes up in her mind, even the very crazy ones.

We would also like to acknowledge PLAIT (Playable Innovative Technologies) lab and the Game Design Program, the College of Computer and Information Sciences and the College of Arts, Media and Design at Northeastern University for helping support the work needed to finish this book. Also, a big thanks to Sean Elks, an undergraduate in the Digital Arts and Game Design program at Northeastern University, for stepping up to help put together the video figure resources needed for some chapters in the book.

We also would like to acknowledge the incredible patience and support of Drew Davidson and his team at ETC-Press for sticking with this book even when it didn't look like it was going to come together.

We would also like to thank the reviewers for their help shaping this book; as well as the authors who spent considerable amounts of time developing their content and revising it.

TABLE OF CONTENTS

Acknowledgements	v
Table of Contents	vii
1. Introduction to this Collection	1
By Theresa Jean Tanenbaum.....	1
1.1. What You Will Find in this Book.....	1
2. Author and Editor Bios	5
2.1. Author Bios.....	5
2.2. Editor Bios.....	12
Section I – Introduction to the History and Theory of NVC for VWs	15
3. Basics of Nonverbal Communication in the Physical World	17
By Theresa Jean Tanenbaum, Michael Nixon, & Magy Seif El-Nasr.....	17
3.1. The Foundations of Nonverbal Communication Research.....	17
3.2. Contemporary Systems of Nonverbal Communication.....	20
3.3. Movement Analysis Systems from Theater and Dance.....	23
3.4. Chapter Review.....	30
3.5. Bibliography.....	30
4. Basics of Nonverbal Communication in Virtual Worlds	33
By Theresa Jean Tanenbaum, Michael Nixon, & Magy Seif El-Nasr.....	33
4.1. Studies of NVC in Digital Environments and Virtual Reality Studies.....	33
4.2. of NVC in Virtual Worlds.....	37
4.3. Chapter Review.....	41
4.4. Bibliography.....	41
Section II – Identity and Communication in Virtual Worlds	45
5. Our Empathic Experience of Believable Characters	47
By Leslie Bishko.....	47
5.1. Introduction.....	47
5.2. Empathy, Authenticity, and Believability.....	50
5.3. Empathy, Interaction and Metaphor.....	54
5.4. Summary.....	55
5.5. Bibliography.....	56
5.6. Appendix: Laban Movement Analysis Resources.....	56
6. Virtual Gaze: The Communicative Energy Between Avatar Faces	61
By Jeffrey Ventrella.....	61
6.1. Introduction.....	61
6.2. Concepts.....	63
6.3. Techniques.....	66

6.4.	The Social Coordinate System.....	69
6.5.	Closing Thoughts.....	73
6.6.	Bibliography.....	75
7.	Avatar Appearance as Prima Facie Non-Verbal Communication.....	77
	By Jacquelyn Ford Morie.....	77
7.1.	Introduction.....	77
7.2.	Why Appearance is so Important.....	82
7.3.	On Looking at One’s Self.....	86
7.4.	Surveys and Hints of Meaning.....	89
7.5.	Creating an Avatar.....	91
7.6.	Discussion.....	97
7.7.	Conclusion.....	99
7.8.	Bibliography.....	100
8.	TimeTraveller™: First Nations Nonverbal Communication in <i>Second Life</i>.....	105
	By Elizabeth LaPensée and Jason Edward Lewis.....	105
8.1.	Nonverbal Communication.....	106
8.2.	TimeTraveller™ Case Study.....	108
8.3.	Nonverbal Communication In TimeTraveller™.....	110
8.4.	Future Work.....	116
8.5.	Conclusions.....	116
8.6.	Bibliography.....	117
	Section III – Virtual Performance and Theater.....	121
9.	Lessons from the Arts: What the Performing Arts Literature Can Teach Us About Creating Expressive Character Movement.....	123
	By Michael Neff.....	123
9.1.	Sidebar: A Brief Introduction to Computer Animation.....	125
9.2.	Key Sources.....	125
9.2.2.	Transitions.....	133
9.3.	Sidebar: Simple Techniques for Improving Character Motion.....	142
9.4.	Lost in Translation: Challenges in Applying Arts Material Computationally.....	143
9.5.	Steps Forward.....	145
9.6.	Bibliography.....	146
10.	Theater as Virtual Reality.....	151
	By Jim R. Parker.....	151
10.1.	Virtual Theatre.....	152
10.2.	Mounting a Virtual Performance.....	167
10.3.	Other Virtual Spaces For Theatre.....	170
10.4.	The Potential of Virtual Theatre.....	172
10.5.	Bibliography.....	173
11.	Animation Principles and Laban Movement Analysis: Movement Frameworks for Creating Empathic Character Performances.....	177
	By Leslie Bishko.....	177
11.1.	The Principles of Animation.....	177

11.2.	Laban Movement Analysis.....	178
11.3.	Applying LMA.....	197
11.4.	Bibliography.....	203
12.	Loss of Agency as Expression in Avatar Performance.....	205
	By Ben Unterman and Jeremy Owen Turner.....	205
12.1.	Introduction.....	205
12.2.	The Role of Non-Mimetic Gesture and Agency in RL Performance.....	206
12.3.	A Short History of Textual and 2D Avatar Performance.....	208
12.4.	Case Study 1: Lines.....	210
12.5.	Avatar Performance in 3D Virtual Environments.....	212
12.6.	Case Study 2: Spawn of the Surreal.....	214
12.7.	Loss of Agency as Non-Verbal Communication.....	216
12.8.	Bibliography.....	217
	Section IV – Animating and Puppeteering.....	221
13.	Empathy in virtual worlds: Making characters believable with Laban Movement Analysis.....	223
	By Leslie Bishko.....	223
13.1.	Character Movement.....	223
13.2.	An Empathic Toolset.....	230
13.3.	Conclusion.....	233
13.4.	Bibliography.....	234
14.	Avatar Puppeteering: Direct Manipulation of Avatar Joints for Spontaneous Body Language.....	237
	By Jeffrey Ventrella.....	237
14.1.	Introduction.....	237
14.2.	Case Study: Puppeteering the <i>Second Life</i> Avatar.....	241
14.3.	Applications.....	246
14.4.	Challenges.....	247
14.5.	A Thought About Future Puppeteering Interfaces.....	248
14.6.	Conclusion.....	248
14.7.	Bibliography.....	249
15.	Automation of Avatar Behavior.....	251
	By Hannes Högni Vilhjálmsson.....	251
15.1.	Chat vs. Avatars.....	251
15.2.	Idle Animation.....	252
15.3.	Locomotion Animation.....	253
15.4.	Engagement without Micromanagement.....	254
15.5.	Social Animation with BodyChat.....	254
15.6.	More Control with Automation?.....	256
15.7.	Face-to-Face Communication.....	257
15.8.	Are Communicative Functions Supported in Mediated Environments?.....	258
15.9.	Automating Communicative Behavior.....	260
15.10.	The Spark System.....	261
15.11.	Impact on Communication.....	211

15.12.	Balance of Control.....	263
15.13.	The Future of Automated Avatars.....	264
15.14.	Bibliography.....	266
16.	Synthesizing Virtual Character Behaviors from Interactive Digital Puppetry.....	269
	By Elena Erbiceanu, Daniel Mapes, and Charles E. Hughes.....	269
16.1.	Introduction.....	269
16.2.	Virtual Character Behaviors.....	271
16.3.	Transferring NVC from a Human to a Virtual Character: Digital Puppetry for Interactive Performance.....	272
16.4.	The Geppetto Digital Puppetry System.....	273
16.5.	Can We Automate Behaviors for Virtual Characters?.....	280
16.6.	Conclusions.....	285
16.7.	Bibliography.....	286
	Section V – Studying Nonverbal Communication in Virtual Worlds.....	289
17.	A Few Choice Animations: Nonverbal Communication Through Production and Consumption in <i>Second Life</i>.....	291
	By Jennifer Martin.....	291
17.1.	Nonverbal Communication in <i>Second Life</i>	292
17.2.	Theoretical and Methodological Considerations.....	293
17.3.	Creating and Using Communication.....	295
17.4.	Production, Consumption, and Communication.....	296
17.5.	Communication and Community.....	299
17.6.	Additional Considerations.....	300
17.7.	Conclusions.....	301
17.8.	Bibliography.....	302
18.	A Microsociological Perspective on Non-Verbal Communicative Strategies in MMORPGs.....	307
	By David Kirschner and J. Patrick Williams.....	307
18.1.	Raiding in the World of Warcraft.....	309
18.2.	The User Interface.....	311
18.3.	From User Interface to Coordinated Action.....	315
18.4.	Computer-mediated Communication, Virtual Worlds, and Coordinated Action.....	317
18.5.	Conclusion.....	319
18.6.	Bibliography.....	320
19.	The Uncanny Valley and Nonverbal Communication in Virtual Characters.....	325
	By Angela Tinwell, Mark Grimshaw, and Debbie Abdel-Nabi.....	325
19.1.	Introduction.....	325
19.2.	The Uncanny Valley.....	327
19.3.	Facial Expression of Emotion and the Uncanny.....	330
19.4.	Nonverbal Communication (NVC).....	332
19.5.	A Lack of NVC in Virtual Worlds.....	334
19.6.	Limitations and Future Work.....	336

19.7.	Conclusion.....	337
19.8.	Bibliography.....	338

Section VI – New Directions for NVC in VWs..... 343

20.	The Future of Avatar Expression: Body Language Evolves on the Internet.....	345
	By Jeffrey Ventrella.....	345
20.1.	Introduction.....	345
20.2.	Trans-World Body Language.....	347
20.3.	The Problem of Input.....	348
20.4.	From Gesture to Speech, and then Back to Gesture.....	349
20.5.	Conclusions.....	350
20.6.	References.....	351
21.	Challenges and Opportunities for the Ongoing Study of Nonverbal Communication in Virtual Worlds.....	355
	By Theresa Jean Tanenbaum, Magy Seif El-Nasr, & Michael Nixon.....	355
21.1.	Factors in the Evolution of NVC for VWs.....	355
21.2.	Future Directions for Research.....	358
21.3.	Final Thoughts.....	360
21.4.	Bibliography.....	361

1

INTRODUCTION TO THIS COLLECTION

By Theresa Jean Tanenbaum

Over the last 20 years we have seen an expansion of network mediated social activities. Where once socialization online was limited to typing in various terminal windows, the evolution of the web and of shared virtual environments has opened up new possibilities for human communication at a distance. Of particular interest is the rise of what have come to be known as “virtual worlds”: persistent graphical environments populated (and often partially authored) by large communities of individual users. Virtual worlds have their technical roots in multi-user domains and their variants (MUDs, MUCKs, MOOs, MUSHs etc.): textually mediated environments in which written language was the primary means of navigation, exploration, expression, and communication¹. Their spiritual roots can be found in the science fictional imaginings of the cyperpunk authors of the 1980s, perhaps most notably William Gibson, whose *Neuromancer* envisioned a digital landscape of metaphorically embodied computer code into which hackers immersed themselves. In *Snow Crash*, Neal Stephenson described a virtual environment called the *Metaverse*: a fictional virtual world which continues to inform our desires and imaginations for what virtual worlds might be. In both of these examples, virtual environments are rendered in a sensorially immersive fashion, often using visual metaphors to represent abstract computational structures and functions. Interactors in these worlds are *embodied* as *avatars*: digital puppets or representations through which the user exerts his or her will on the environment. It is this virtual embodiment that makes today’s virtual worlds so interesting. Virtual worlds such as *Second Life*, the now defunct *There.com*, *Active Worlds*, *Traveler*, and *Habbo Hotel* provide users with customizable avatars in graphical environments with a range of communicative affordances including text and voice chat. With virtual embodiment comes a host of new and important communicative possibilities, and an assortment of new challenges and literacies including a wide range of nonverbal communication behaviors and non-linguistic social signaling options. In this book, we begin the work of articulating the challenges and possibilities for non-verbal communication in virtual worlds.

1. WHAT YOU WILL FIND IN THIS BOOK

This short introductory chapter can be used as a guide to the book as a whole. Chapter 2 introduces the authors and editors: a multi-disciplinary collection of experts who are pushing the boundaries of theory and design in the area of nonverbal communication and virtual worlds.

¹Legendary game designer Raph Koster has compiled stories from many of the central figures in the rise of online worlds and MUDs into a timeline of the medium that explores this history in much greater detail: <http://www.raphkoster.com/gaming/mudtimeline.shtml>

Section I – Introduction to the History and Theory of NVC for VWs

These chapters provide a broad survey of the history of nonverbal communications research.

- Chapter 3: A crash course introduction to the history of nonverbal communication research in the social sciences, performing arts, and psychology.
- Chapter 4: We follow this with a whirlwind review of the history of nonverbal communication research in virtual worlds.

These are by no means comprehensive: their purpose is simply to ground the reader in some of the most significant literature and research undertaken in this field.

Section II – Identity and Communication in Virtual Worlds

This second section of the book deals with complex issues of identity, meaning, and culture that must be grappled with when communicating in virtual worlds.

- Chapter 5: In the first of three installments, Leslie Bishko discusses the nature of empathy and believability in virtual characters.
- Chapter 6: Designer and co-founder of *There.com*, Jeffrey Ventrella argues for the importance of gaze in nonverbal communication and describes his own experiences designing systems for mutual gaze.
- Chapter 7: Jacquelyn Morie argues that avatar appearance is the most important channel for the creation of virtual identities and nonverbal communication.
- Chapter 8: Elisabeth LaPensée and Jason Edward Lewis present their work developing a Machinima and Alternate Reality Game project in *Second Life* that grapples with issues of identity and representation within First Nations and Indigenous communities.

Section III – Virtual Performance and Theater

Theater, Dance, and Performance art have long traditions of exploring the communicative power of the human body. It should come as no surprise that these traditions have led to some of the most important explorations of bodily communication in virtual worlds.

- Chapter 9: Michael Neff provides a deep look at how techniques from live theatrical performance can be used to support nonverbal communication in virtual worlds.
- Chapter 10: Jim Parker relates his experiences designing and mounting virtual theatrical performances.
- Chapter 11: Leslie Bishko provides an in depth guide to Laban's Movement Analysis system for animators and designers.
- Chapter 12: Jeremy Owen Turner and Ben Unterman look at how modulation of an audience's experience of agency and embodiment in virtual performances can function as a form of nonverbal communication, which they illustrate in through the lens of two different virtual performances.

Section IV – Animating and Puppeteering

The chapters in this section deal with the design of interfaces and characters in virtual worlds. As designers, how do we create systems to support nuanced NVC, and as users how do we express ourselves through those systems?

- Chapter 13: Leslie Bishko takes the lessons from her first two chapters and synthesizes them into a holistic approach to designing empathic characters for virtual worlds grounded in Laban Movement Analysis.
- Chapter 14: Jeffrey Ventrella discusses the relevant issues and technologies needed to understand how to design systems for avatar puppeteering.
- Chapter 15: Hannes Högni Vilhjálmsson considers the importance of automated character animations.
- Chapter 16: Elena Erbiceanu and her colleagues describe a different approach to virtual puppeteering rooted in studying the actions of trained professional puppeteers.

Section V – Studying Nonverbal Communication in Virtual Worlds

This section includes several examples of recent research that is taking place within virtual worlds as field sites, which illustrates the range of methodologies and approaches which can be applied to them.

- Chapter 17: Jennifer Martin looks at practices of production and consumption of custom nonverbal communication assets within *Second Life*.
- Chapter 18: David Kirschner and J. Patrick Williams use a “microsociological” analysis of custom user interfaces within World of Warcraft to explore how user communities adapt interfaces to extremely complicated communicative tasks.
- Chapter 19: Angela Tinwell and her colleagues describe an empirical study of facial expressions and the perception of the “uncanny valley” through a controlled experiment involving variably animated character faces.

Section VI – New Directions for NVC in VWs

This final section deals with broader speculations about the future of virtual worlds, and about the future of research into nonverbal communications within them.

- Chapter 20: Jeffrey Ventrella draws on his years of experience as a designer of virtual worlds to speculate on how expressive avatars will evolve.
- Chapter 21: We close the collection with a blanket analysis of where we see scholarly research progressing in this field, and where we believe so unexplored opportunities remain.

Taken as a whole, we hope this book reflects the interdisciplinary breadth of this still young field, while providing insight into the challenges and opportunities faced by designers and researchers of virtual worlds.

2

AUTHOR AND EDITOR BIOS

1. AUTHORS BIOS

LESLIE BISHKO

Emily Carr University of Art + Design
<http://www.ecuad.ca/people/profile/14460>
<http://labanforanimators.wordpress.com>
lbishko@ecuad.ca

Leslie Bishko's work is centered on the dynamics of expression through movement in the medium of animation, influenced by abstract experimental animation and contemporary modern dance. Exploration of these themes led her to integrate 3D computer animation with Laban Movement Analysis (LMA), which are the focus of her research and teaching. Recent work includes 3D computer animation tutorials for the Autodesk Industry Careers Framework, and Guest Faculty for the Laban/Bartenieff and Somatic Studies Canada professional certification program in LMA. Her award winning 3D computer animation film *Gasping for Air* (1993) has screened at numerous international film festivals.

Leslie would like to acknowledge the following people who supported her work for this project: Diane Allen, Andrew Bailey, Heather Conn, Leslie Fedorchuk, Heather Fitzgerald, Travaldo Farrington, Michael Neff, Jay White

ELENA ERBICEANU

University of Central Florida
ee@knights.ucf.edu

Elena Erbiceanu is a PhD student in Computer Science at the University of Central Florida. She earned her Masters degree in Computer Science at the same university in 2008, and her BSc in Computer Science from "Al. I. Cuza" University in Romania. She is currently a member of the Synthetic Reality Laboratory (SREAL) and her research interests include the study of automating behaviors for virtual characters in interactive experiences, the use of machine learning and AI techniques to enhance virtual reality experiences, evolutionary computation, and human-computer interaction. She was awarded a four year Presidential Fellowship award from UCF and became one of two recipients for the 2009-2010 Link Foundation Fellowship Award. She was a software engineering intern at Google New York/Mountain View in the summers of 2008-2010, and accepted a full-time position at Google in 2012.

MARK GRIMSHAW

Obel Professor of Music at Aalborg University, Denmark
<http://personprofil.aau.dk/profil/126217>
grimshaw@hum.aau.dk

Mark Grimshaw is the Obel Professor of Music at Aalborg University, Denmark where he is co-chair of the research centre RELATE and chair of the Music & Sound Knowledge Group. He has a BMus (Hons) at the University of Natal, South Africa, an MSc in Music Technology from the University of York, UK, and a PhD on sound in the First-Person Shooter from the University of Waikato, New Zealand. He writes extensively on sound in computer games with a particular interest in emotioneering and the use of biofeedback for the real-time synthesis of game sound. He also writes free, open source software for virtual research environments (WIKINDX) and is investigating the uses of sonification to facilitate creativity in the context of such knowledge tools. His last book was an anthology on computer game audio published in 2011 and he is currently editing the Oxford Handbook of Virtuality for OUP.

CHARLES E. HUGHES

Synthetic Reality Laboratory, University of Central Florida
<http://www.cs.ucf.edu/~ceh>
ceh@cs.ucf.edu

Dr. Charles Hughes is a Pegasus Professor in the Department of Electrical Engineering and Computer Science, Computer Science Division, at the University of Central Florida. He also holds appointments in the School of Visual Arts and Design and the Institute for Simulation and Training (IST), is a Fellow of the UCF Academy for Teaching, Learning and Leadership, and holds an IPA appointment with the US Department of Veterans Affairs. He has a B.A. in Mathematics from Northeastern University, and a M.S. and Ph.D. in Computer Science from Pennsylvania State University. He is director of the UCF Synthetic Reality Laboratory (<http://sreal.ucf.edu>). His research is in augmented reality environments with a specialization on networked digital puppetry (the remote control by humans of virtual and physical-virtual avatars). He conducts research on the use of digital puppetry-based experiences in cross cultural and situational awareness training, teacher and trainer education, social and interpersonal skills development, and physical and cognitive assessment and rehabilitation. Dr. Hughes is author or co-author of over 170 refereed publications. He is an Associate Editor of Entertainment Computing and the Journal of Cybertherapy and Rehabilitation, and a member of the Program Committee and co-chair of Research Exhibits for IEEE VR 2013. He has active funding to support his research from the National Endowments for the Humanities, the National Institutes of Health, the National Science Foundation, the Office of Naval Research, US Department of Veterans Affairs and the Bill & Melinda Gates Foundation.

DAVID KIRSCHNER

PhD Candidate, Nanyang Technological University in Singapore, School of Sociology
davi0017@e.ntu.edu.sg

David Kirschner is a Ph.D candidate in Sociology at Nanyang Technological University in Singapore. He holds a B.A. in Sociology and an M.Ed. in Social Science Education from the University of Georgia. He is currently writing his dissertation on how players make meaning in video games and experience becoming socialized into gameplay

ELISABETH LAPENSÉE

PhD Candidate, Simon Fraser University, School of Interactive Arts & Technology
<http://www.bethaileen.com>
beth@bethaileen.com

Elisabeth LaPensée, Ph.D. (ABD) in Interactive Arts and Technology at Simon Fraser University in British Columbia is an Irish, Anishinaabe, and Métis border-crossing writer whose work addresses Indigenous determination in media such as games, animation, and web comics. Most recently, she is focusing on her dissertation, which looks at the impact of the health and wellness game for change *Survivance* (2011) on the Portland urban Native American community. Her work is supported by the Aboriginal Territories in Cyberspace research network.

JASON EDWARD LEWIS

Aboriginal Territories in Cyberspace/Concordia University
www.abtec.org
Jason.lewis@concordia.ca

Jason Edward Lewis is a digital media poet, artist, and software designer. He founded Obx Laboratory for Experimental Media, where he directs research/creation projects devising new means of creating and reading digital texts, developing systems for creative use of mobile technology and using virtual environments to assist Aboriginal communities in preserving, interpreting and communicating cultural histories. Along with Skawennati, he co-directs Aboriginal Territories in Cyberspace and the Skins Workshops on Aboriginal Storytelling and Video Game Design. His other interests include computation as a creative material, emergent media theory and history, and methodologies for conducting art-led technology research. Lewis' creative work has been featured at Ars Electronica, Mobilefest, Urban Screens, ISEA, SIGGRAPH, and FILE, among other venues, and has been recognized by a Prix Ars Electronic Honorable Mention, several imagineNATIVE Best New Media awards and four solo exhibitions. He's the author or co-author of five chapters in collected editions covering mobile media, video game design, machinima and experimental pedagogy with Native American communities, as well as numerous journal articles and conference papers. He has worked in a range of industrial research settings, including Interval Research, US West's Advanced Technology Group, and the Institute for Research on Learning, and, at the turn of the century, he founded and ran a research studio for the venture capital firm Arts Alliance. Born and raised in California, he is currently Associate Professor and Program Director of Computation Arts at Concordia University, Montreal.

DANIEL P. MAPES

Daniel P. Mapes's advanced degrees include a B.S. in Physics from Carnegie Mellon University and an M.S. in Computer Science from the University of Central Florida. Between degrees, he gained formative entrepreneurial experience as Vice President of FGR Inc., a small publishing startup. Concluding his Masters thesis in 1994, he was awarded the student research award for his PolyShop testbed at IST/UCF in Orlando. In a move to commercialize PolyShop he managed the software engineering team that developed Multigen's breakthrough SmartScene product. In pursuit of deeper industry adoption of SmartScene he accepted a research position within LEGO's advance toy research division in Denmark. He then established RoninWorks BV in the Netherlands to pursue commercialization of consumer solutions in professional animation and game applications. After a short sabbatical teaching at the School of Film and Digital Media at UCF in Orlando and a sub-contracting he joined the Synthetic Reality Laboratory (SREAL) at IST/UCF as research faculty and continues making toys into tools and tools into toys.

JENNIFER MARTIN

University of Western Ontario, Faculty of Information and Media Studies
www.jenmartin.com
jmart9@uwo.ca

Jennifer Martin is an instructor in the Faculty of Information and Media Studies at the University of Western Ontario where she teaches about digital communication, video games, and virtual worlds. Her research interests include digital cultures, identity exploration and play, social interaction in online environments, and virtual consumption and economics.

JACQUELYN FORD MORIE

University of Southern California, Institute for Creative Technologies
www.ict.usc.edu/~morie
morie@ict.usc.edu

Jacquelyn Ford Morie is an expert in immersive worlds and avatar technologies at the USC Institute for Creative Technologies. Her recent "Coming Home" research developed a healing center for veterans in the online world *Second Life* and led to the adoption of virtual worlds as an advanced form of telehealth care at the National Intrepid Center for Excellence in Psychological Health and Traumatic Brain Injury (NICoE) in Bethesda, MD. This continuing effort will be the first to bring the power of the virtual world to military patients as a part of their healing.

Before ICT, Morie spent six years in the animation and effects industry at Disney Feature Animation, Blue Sky|VIFX and the award winning Rhythm & Hues, creating specialized training programs for both traditional and digital aspects of production. From 1990 to 1994, Morie worked as an artist, researcher and educator at the University of Central Florida in Orlando. Her Virtual Reality work at UCF's Institute for Simulation and Training involved creating scenarios and environments designed to evoke emotional responses from their participants.

Morie has been an active member of the games and serious games communities, both as a theorist and an active developer. She is a member of ACM, ACM SIGGRAPH, and currently serves as Director at Large on the Executive Council for SIGGRAPH. Morie was recently invited (2012) to DARPA's Information and Science Technology (ISAT) Board. She is an active speaker on the potential of VWs for real world good, and has authored over 50 relevant papers and book chapters. Her PhD from the University of East London defined ways to create meaningful immersive virtual environments.

DR DEBORAH ABDEL NABI

Lecturer in Psychology at Bolton University, UK.
<http://www.bolton.ac.uk/Subjects/Psychology/Staff/DrDebbieAbdelNabi.aspx>
A.Nabi@bolton.ac.uk

Deborah Abdel Nabi is a Senior Lecturer in Psychology at Bolton University, UK. With a PhD in human brain electrophysiology /pre-attentive visual processing from The University of Manchester, she later evolved an interest in the topic of cyberpsychology. Her work with the University of Bolton Computer and Cyberpsychology Research Unit includes projects on the psychological factors underlying perception of the uncanny valley in virtual characters, impression formation and identity management in cyberspace; and e-learning, including innovative uses of emergent information technology to facilitate conversational models of teaching and learning. Her current research centres around the BARTA (Bolton Affect Recognition Tri-stimulus Approach), a new database of synthetic and human images to improve emotion recognition training in children with Autistic Spectrum Disorders; a paper on the work has been submitted to The Journal of Nonverbal Communication. She is also currently assessing the EEG correlates of uncanny characters expressing the six basic emotions.

MICHAEL NEFF

Department of Computer Science and Program for Cinema and Technocultural Studies, University of California – Davis

<http://www.cs.ucdavis.edu/~neff/>

neff@cs.ucdavis.edu

Michael Neff is an associate professor in Computer Science and Cinema & Technocultural Studies at the University of California, Davis where he directs the Motion Lab, an interdisciplinary research effort in character animation and embodied input. His Ph.D. is from the University of Toronto and he is also a Certified Laban Movement Analyst. His interests include character animation tools, especially modeling expressive movement, physics-based animation, gesture and applying performing arts knowledge to animation. At Davis, he is working to bridge the art and technology communities on campus, collaborating with computer scientists, dancers, choreographers and geologists. He received an NSF CAREER Award (2009-14), the Alain Fournier Award for his dissertation (2005), a best paper award from Intelligent Virtual Agents (2007) and the Isadora Duncan Award for Visual Design (2009).

JIM (J.R.) PARKER

Digital Media Laboratory, Faculty of Arts, University of Calgary

www.ucalgary.ca/~jparker

jparker@ucalgary.ca

With a Ph.D. in Informatics from the State University of Gent, Belgium ‘met grootste onderscheiding’ (with greatest distinction Jim nonetheless seems unable to hold a steady job. He has been a full professor of computer science, a Drama professor, and is now working as a professor in the Art department at University of Calgary.

Formerly known for work in machine vision, his research interests now include: multimedia, computer games (especially serious games), discrete event simulation, and theatre in non-traditional spaces. Major research projects include virtual theatre projects, the design and construction of kinetic games (including human activity feedback), and the portrayal of Canadian (military) history and achievements in digital and online form, both as games and animations.

Dr. Parker has written over 150 technical articles, developed video games including *The Booze Cruise* and *OceanQuest*, has directed theatrical productions in *Second Life*, and has five books to his credit including ‘The Guide to Simulations and Games’ (Wiley, 2011).

ANGELA TINWELL

Programme Leader for the BA (Hons) Animation and Illustration Course at Bolton University, UK

<http://www.bolton.ac.uk/Subjects/Games-Computing-IT/Staff/Angela-Tinwell.aspx>

A.Tinwell@bolton.ac.uk

Angela Tinwell is Programme Leader for the BA (Hons) Animation and Illustration course and a Senior Lecturer in Games and Creative Technologies at Bolton University, UK. With research interests in video games art and design, animation, illustration, character design, human computer interaction, cyberpsychology and research methods and practice, much of my research to date has centred on the Uncanny Valley phenomenon in realistic, human-like characters featured in animation and video games. Angela achieved her PhD on the subject of facial expression and speech and viewer perception of the Uncanny Valley in virtual characters at the University of Bolton that was passed with no amendments. She is currently authoring the book *The Uncanny Valley in Games and Animation* for A K Peters/CRC Press that is due to be published in April, 2014.

JEREMY OWEN TURNER

PhD Candidate, Simon Fraser University, School of Interactive Arts and Technology
<http://www.siat.sfu.ca/graduate-students/profile/jeremy-turner>
ideatron@gmail.com

Jeremy Owen Turner is currently a PhD Candidate at SIAT. Since 1996, Turner has developed an international portfolio as a performance artist, media-arts historian, music composer and art-critic within virtual worlds and video games. Turner's academic history includes an MA about avatar-design in *Second Life* and an interdisciplinary BA that focused on both Art-History and Music Composition. Turner's current PhD research explores developing audio-visual and behavioral design heuristics for virtual agents (automated characters/NPCs) in virtual worlds and video games. These heuristics are derived from some key scholars in the field of Aesthetic Philosophy, including Immanuel Kant.

BENJAMIN UNTERMAN

Simon Fraser University, School of Interactive Arts + Technology
bau@sfu.ca
www.sfu.ca/~bau/

Benjamin Unterman is a PhD candidate and lecturer in the School of Interactive Art and Technology at Simon Fraser University. He researches audience copresence and interaction in avatar performance, focusing on the conventions and social cues which create and encourage a feeling of connection between the audience and performers. Artistically, he creates live interactive stories using a variety of communication platforms including chat-based interfaces, virtual worlds and multi-player game environments. Benjamin is interested more broadly in the recent explosion of live mediated storytelling, which pushes the boundaries of technical expression and channel the theatrical impulse into yet another medium.

JEFFREY VENTRELLA

Independent Artist/Researcher
<http://www.virtualbodylanguage.com/>
jeffrey@ventrella.com

Jeffrey Ventrella is an artist, computer programmer, and expert on avatar expression. He was the primary designer and engineer of Avatar-Centric Communication for There.com - a virtual world recognized for its socially-focused avatars. Jeffrey was also employed at Linden Lab, makers of *Second Life*, where he invented the celebrated "flexiprims" and developed avatar-related designs, including Avatar Puppeteering. Jeffrey is the author of the book *Virtual Body Language* (ETC Press, 2011), which addresses questions of how we project (and invent) body language in new media, with implications for the future of internet life. Jeffrey has a master's degree for the MIT Media Lab, and has lectured in Europe and North America on artificial life, digital art, and virtual worlds. He is the creator of Gene Pool, an artificial life game, and PeckPeck, a virtual pet - both of which are available as iPad apps.

HANNES HÖGNI VILHJÁLMSOHN

Associate Professor of Computer Science at Reykjavik University
<http://www.ru.is/~hannes>
hannes@ru.is

Hannes is an Associate Professor of Computer Science at Reykjavik University where he leads the Socially Expressive Computing group at the Center for Analysis and Design of Intelligent Agents (CADIA). He has been doing research on the automatic generation of social and linguistic nonverbal behavior in online avatars and autonomous agents since the mid-90s. His focus has been on making embodied communication and

socialization in virtual environments both effective and intuitive based on a deep understanding of both the sociological and technical aspects. Currently he is working on a social engine that could bring socially aware nonverbal behavior to any avatar-based 3D virtual world. Prior to joining Reykjavik University, Hannes was the technical director on the Tactical Language and Culture Training project at the Center for Advanced Research into Technology for Education at University of Southern California. The resulting interactive social simulation technology is still in wide use by the US Government. To continue that work, Hannes co-founded Los Angeles based Alelo Inc, that builds serious games for language and culture learning. Hannes is also a co-founder of Reykjavik based MindGames ehf., the first company to release mind training games for the iPhone that incorporate consumer-level brainwave headsets. Hannes is a founding member of the Icelandic Cognitive Studies Society and of LornaLAB, a creative technology forum and media lab in Reykjavik. He received his Ph.D. in Media Arts and Sciences from the MIT Media Lab in 2003 and M.Sc. from the same place in 1997.

JAY WHITE

Emily Carr Univerist of Art + Design

<http://www.ecuad.ca/people/profile/14430>, <http://www.gogomax49.com>

jwhite@ecuad.ca, gogomax49@gmail.com

Jay White's practice explores a wide range of topics through short film animation, interdisciplinary collaboration, and process-based work.

His career runs the gamut from creating music videos, animated television shows and video games to multimedia exhibitions, and collaborations with theatre, dance companies and independent filmmakers.

His work as Art Director on the animated series *Dragon Booster* helped it garner a Gemini Award for Best Animated Series in 2005. His animated short *Boar Attack* premiered at the Toronto International Film Festival, won many Best Animated Short awards at festivals internationally, and was longlisted for Academy Award nomination in 2009.

J. PATRICK WILLIAMS

Division of Sociology, Nanyang Technological University

www.jpatrickwilliams.net

subcultures@gmail.com

Patrick Williams is Assistant Professor of Sociology at Nanyang Technological University, where he teaches courses on social psychology, culture, and media. His research interests revolve around young people, interaction, and culture with specific emphasis on the roles that new media technologies play in mediating personal and cultural experiences. He is trained in the symbolic interactionist tradition of sociology, a social-psychological perspective that foregrounds language and meaning as key dimensions of understanding the everyday life. Patrick has published studies on various forms of virtual gaming worlds, including table-top fantasy games and massively-multiplayer online games. His latest funded project (2009-2011) was a multi-method study on *World of Warcraft*, the emphasis of which was on players' motivational and immersive experiences. He has edited two books on games and gaming culture: *Gaming as Culture: Essays in Social Reality, Identity and Experience in Fantasy Games* (McFarland 2006) and *The Players' Realm: Studies on the Culture of Video Games and Gaming* (McFarland 2007).

2. EDITORS BIOS

THERESA JEAN TANENBAUM

Simon Fraser University, School of Interactive Arts + Technology
tess.tanen@gmail.com

Theresa Jean Tanenbaum is a PhD Candidate at Simon Fraser University in the School of Interactive Arts + Technology. Her research spans a broad cross section of digital games, new media, tangible and embodied interaction, and human computer interaction. She is primarily interested in the poetics of digital narratives: her dissertation work draws on theories and practice from the performing arts to investigate how digital narratives can support transformative experiences. Her other research includes writing on the intersection of Design Fiction, Maker culture, Steampunk, and Democracy, and research into the uses of serious games to engage broad publics in intergenerational dialogues about local sustainability issues. She comes to the study of nonverbal communications and virtual worlds from the perspective of embodied and gestural interfaces for games.

MAGY SEIF EL-NASR

Northeastern University, Colleges of Computer and Information Sciences, and Arts, Media and Design.
<http://nuweb.neu.edu/magy/home.html>
magy@northeastern.edu

Magy Seif El-Nasr, Ph.D. is an Associate Professor in the Colleges of Computer and Information Sciences and Arts, Media and Design at Northeastern University. She is also the Director of the Game Educational Programs and Research at Northeastern and the Director of Game Design Program in College of Arts, Media and Design. She also directs the Game User Experience and Design Research Lab. Dr. Seif El-Nasr earned her Ph.D. degree from Northwestern University in Computer Science. Magy's research focuses on interactive narrative, enhancing game designs by developing tools and methods for evaluating and adapting game experiences. Her work is internationally known and cited in several game industry books, including *Programming Believable Characters for Computer Games (Game Development Series)* and *Real-time Cinematography for Games*. In addition, she has received several best paper awards for her work. Magy worked collaboratively with Electronic Arts, Bardel Entertainment, and Pixel Ante.

MICHAEL NIXON

Simon Fraser University, School of Interactive Arts + Technology
<http://www.siat.sfu.ca/graduate-students/profile/michael-nixon>
Mna32@sfu.ca

Michael Nixon is a PhD Candidate at the School of Interactive Arts & Technology at Simon Fraser University. He researches how to make characters in digital environments more believable through the use of better cognitive models and non-verbal behavior within a social context. Michael's M.Sc. thesis describes an investigation of the suitability of Delsarte's system of movement as a framework for believable characters. His dissertation research focuses on the use of social signals as cues in the creation of unique identities.

SECTION I

*INTRODUCTION TO THE HISTORY
AND THEORY OF NVC FOR VWS*

3

BASICS OF NONVERBAL COMMUNICATION IN THE PHYSICAL WORLD

By Theresa Jean Tanenbaum, Michael Nixon, and Magy Seif El-Nasr

Research into nonverbal communication (NVC) has a broadly interdisciplinary history spanning social science research, performing arts practice, and even popular culture and self-help manuals. The formal study of NVC has its roots in the Victorian Era. Charles Darwin is perhaps the first scholar to systematically study how we use our bodies to communicate in *The Expression of the Emotions in Man and Animals* (Darwin, 1874). An interest in analyzing movement arose during the Industrial Revolution, and the mechanization of labor influenced a scientific, analytic approach to efficiency in the workplace based on the photographic study of movement (Moore, 2005). The camera allowed artists and scientists to capture and study deep nuances of motion, pose, and gait, as exemplified by the photographic sequences of Eadweard Muybridge of the late 19th century (Muybridge, 1979). The early 20th century saw a decline in scientific interest in bodily communication, but the performing arts picked up the torch with dance researchers like Rudolf Laban and developing sophisticated systems for the annotation and analysis of human movement (Maletic, 1987; Stebbins, 1977). In the social sciences a new wave of systematic research into nonverbal communication was kicked off by Ray Birdwhistell's work on *Kinesics* in the 1950s and Edward Hall's work on *Proxemics* in the 1960s, which in turn led to a surge of public interest in "body language" with sensationalist works like Julius Fast's *Body Language* promising to teach readers "how to penetrate the personal secrets of strangers, friends and lovers by interpreting their body movements, and how to make use of [these] powers (Fast, 1970). In spite of (or perhaps, because of) the rise of populist interest in body language, the last 40 years have seen the development of a number of rigorously grounded models of nonverbal communication. In this chapter we provide a brief survey of some of the foundational systems of NVC from the last century, spanning both the social science and dance research communities.

1. THE FOUNDATIONS OF NONVERBAL COMMUNICATION RESEARCH

Early work into Nonverbal Communication (NVC) can be broken into two broad categories: Proxemics and Kinesics¹. In this chapter we give a high-level overview of these categories: sufficient to provide some grounding in the terminology and concepts discussed throughout this book.

¹A third core topic in NVC is "Paralanguage", or the study of tone of voice, speech fluency, and non-language vocalizations and sounds such as laughter, and grunting, as described by (Duncan, 1969). In virtual worlds research, the term "paralanguage" has been used to describe textual expressions such as emoticons and acronyms (Joinson, 2003). In this book we will be focusing primarily on aspects of NVC directly related to representations of the body, and will leave paralinguistic aspects of NVC for future work.

For more extensive analysis of NVC we recommend (Duncan, 1969), which undertakes a survey of the field and an analysis of the methods and approaches in use during the rise of NVC research in the social sciences. For a more recent survey of the state of the field we suggest (Calero, 2005), which is written to be broadly accessible (although it suffers from much of the same sensationalism as the Body Language manuals of the 1970s).

PROXEMICS

Some of the earliest work on NVC in the social sciences was that of Edward T. Hall, who coined the term “proxemics” in his 1966 book *The Hidden Dimension* (Hall, 1966). Proxemics is the study of the relationships between human bodies in space; it has been used to describe how people position themselves in space relative to each other, and how different demographic factors like age and gender alter these spacing behaviors. Hall’s proxemics outlines the notion of personal space, describing several zones of intimacy around the body. He developed this into four “personal reaction bubbles” [Figure 3 1] including: Intimate, Personal, Social, and Public (Hall, 1968).

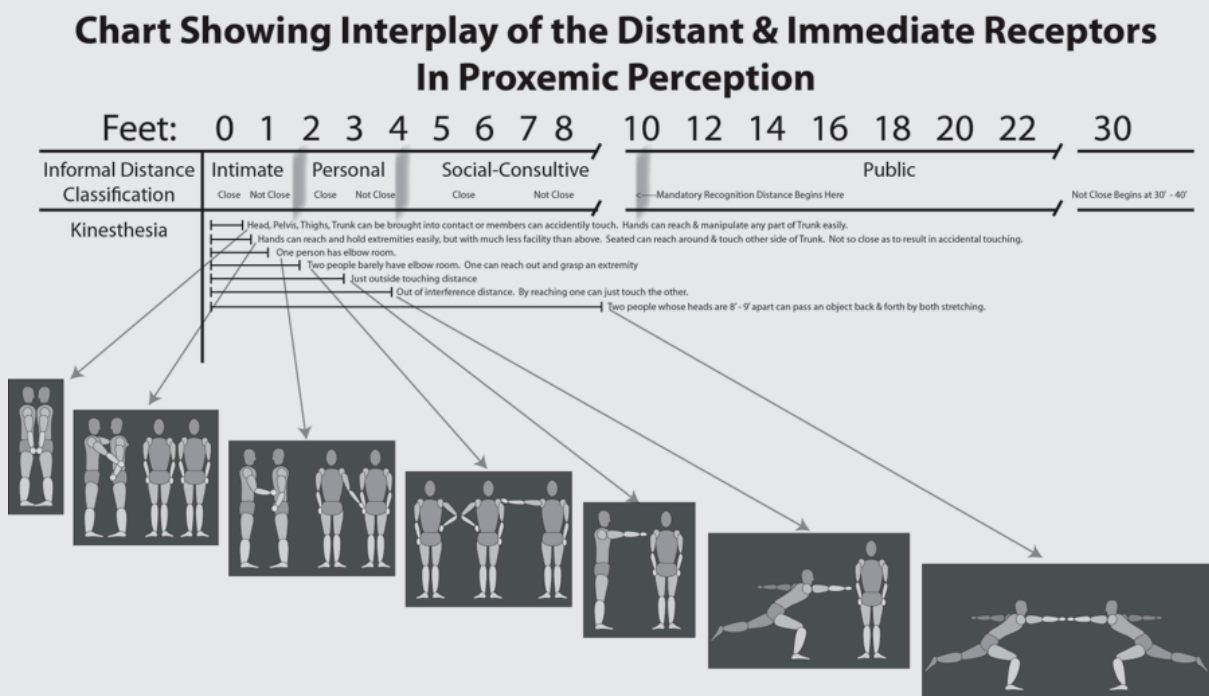


Figure 3-1: Personal space in Hall’s Proxemics. (adapted from Hall, 1968)

These zones of interpersonal space vary depending on a wide range of factors including the age, gender, level of intimacy and cultural background of the interactants (Burgess, 1983). Burgess tested some of Hall’s claims by observing groups of people in public spaces. He photographed crowds in a shopping mall, and measured the spacing between people in groups. He identified specific patterns of distance for groups of different ages, observing that “teenage children stayed further apart than grade school children” and “middle adults stayed further apart than young adults, but maintained closer distances to their companions than senior adults” (Burgess, 1983). It has been suggested that it is possible to build a predictive model of proxemic behavior, however there is currently no clear consensus on whether this is achievable (Eastman & Harper, 1971).

Proxemic behaviors are only “semi-conscious”, arising from social mediation, cultural influences, and the level of intimacy of the participants. Because most people are unaware that they are spacing themselves out according to these factors, they provide a valuable mechanism for analyzing groups of people in both physical and virtual worlds. Perhaps more importantly, proxemic behaviors do not require sophisticated puppeteering interfaces or complex technological affordances in order for them to be reproduced in a virtual world: any form of avatar embodiment – no matter how minimal – affords rudimentary proxemic spacing behaviors.

EQUILIBRIUM THEORY AND GAZE

In addition to the arrangement of bodies in space, proxemics is concerned with the orientation of those bodies to each other: particularly where gaze is concerned. One place where gaze and proximity converge is in “equilibrium theory”. First proposed and explored by Argyle and Dean in 1965, equilibrium theory deals with the relationship between mutual gaze and proxemic distance. In essence, equilibrium theory proposes that proxemic spacing and mutual gaze may both be used to indicate intimacy, and that people reach “equilibrium” comprised of a comfortable distance, and comfortable meeting of gaze that varies with their interpersonal comfort. If either of these factors is varied – if gaze is prolonged, or interpersonal distance is lessened – the interactants often vary the other factor in order to preserve this equilibrium (Argyle & Dean, 1965). For example, one of Argyle and Dean’s experiments determined that subjects would stand closer to another person if that person’s eyes were closed than they would if the person’s eyes were open. As with other proxemic spacing behaviors, equilibrium theory is subject to variances along demographic lines such as age, gender, cultural background, and familiarity (Argyle & Dean, 1965).

KINESICS

Kinesics is the aspect of NVC that deals with posture and gesture – those things that have been termed “body language”. Ray Birdwhistell coined the term kinesics in his book *Introduction to Kinesics* (Birdwhistell, 1952). Situating his work in the (then popular) study of Structural Linguistics, Birdwhistell tried to find the basic unit in movement comparable to the *morpheme* in linguistics. Unfortunately, Birdwhistell never published a definitive work on his system, and so much of his work remains inaccessible and riddled with problems. In spite of his position as progenitor of the field, Birdwhistell’s work on kinesics is of questionable value to present-day scholars. In his critique and analysis of Birdwhistell’s writings, Stephen Jolly provides a high-level evaluation of Birdwhistell’s work, ultimately concluding that:

“Birdwhistell’s theory of kinesics is not an adequate theory for the explanation of body motion as an interactional modality. Although his work marks an important beginning in the study of nonverbal phenomena and represents a first step toward a wider human communicative science, it suffers from a number of flaws which hamper its development and invalidate its results.” (Jolly, 2000).

Jolly identifies three core failings in Birdwhistell’s work. First, he identifies a “lack of systematic order” coupled with “inconsistent repetitiveness of views and their often unsubstantiated presentation” that makes the work difficult to assess (Jolly, 2000). He also evaluates kinesics alongside other systems of movement analysis and notation from dance, ultimately concluding that Birdwhistell’s system is neither efficient nor accurate. Finally, he criticizes it for its fundamental reliance on the principles of Structural Linguistics, which he claims is ultimately unfounded.

Kinesics as a system for notating and understanding nonverbal communication may lack descriptive and explanatory power, but it established the study of gesture and posture as legitimate areas of research, giving rise to more useful approaches to the domain. Following Birdwhistell’s coining of the term “kinesics”, a

number of researchers undertook to develop their own systems for evaluating nonverbal communication, including Paul Ekman and Wallace Friesen whose framework is the basis for most contemporary systems of NVC.

2. CONTEMPORARY SYSTEMS OF NONVERBAL COMMUNICATION

EKMAN AND FRIESEN'S FRAMEWORK

Ekman and Friesen developed one of the most comprehensive descriptive systems for categorizing and coding kinesic NVC. Their "Repertoire of Nonverbal Behavior" addresses three fundamental issues in the field: a behavior's usage, its *origin*, and its *coding* (Ekman & Friesen, 1981).

Ekman and Friesen use the term *usage* to refer to the "regular and consistent circumstances surrounding the occurrence of a nonverbal act" (Ekman & Friesen, 1981). This can be understood as the *context* in which a nonverbal behavior occurs, such as a gesture coinciding with a verbal behavior. Central to their understanding of *usage* is issues around *awareness* and *intentionality*: how conscious was the use of a particular nonverbal act and how deliberate was the nonverbal behavior? Usage is also concerned with the type of information that a nonverbal act may convey. Ekman and Friesen describe six categories of information that may be conveyed via nonverbal acts: Idiosyncratic Information, Shared Information, Encoded, Decoded, Informative, Communicative, and Interactive.

The second portion of Ekman and Friesen's framework, *origin*, deals with how people acquire nonverbal behaviors. They describe three ways in which people learn behaviors. First, they argue that there is an inherited neurological response to stimuli that is shared by all people; for example, the reflexive jump when startled by a sudden noise. A second source of nonverbal behaviors is the non-inherited (but shared) experience of embodiment in the world. They give the culturally independent example of humans using hands to convey food to the mouth (with or without implements). The final origin of nonverbal behavior is the specific experiences of individuals within the world, as dictated by many factors, including (but not limited to): culture, class, family, personality, and education (Ekman & Friesen, 1981).

Coding deals with the relationship between a nonverbal act and the *meaning* that it communicates. Ekman and Friesen describe three principles of coding: Arbitrary, Iconic, and Intrinsic. I find it helpful to think of these principles in terms of varying degrees of abstraction between the nonverbal act and the meaning signified. Arbitrarily coded acts bear no resemblance to the meaning that they signify, such as when someone opens and closes their hand to signify a greeting. Iconically coded acts resemble the meaning they signify in some way, such as when a person runs a finger under his throat to signify having one's throat cut. Intrinsically coded nonverbal acts have no separation between the nonverbal behavior and meaning signified such as when one person punches another to signify intent to do harm. Ekman and Friesen also identify five visual relationships between nonverbal acts and their meaning. These include:

1. Pictorial Relationships: These are iconically coded relationships where the nonverbal movement illustrates an event object or person, such as when someone uses two hands to show the size and shape of an object.
2. Spatial Relationships: These are also Iconic, but they use movement to indicate distance between people, objects or ideas such as when one places hands close together to indicate how close a car came to hitting a pedestrian.

3. Rhythmic Relationships: These are Iconic relationships that use movement to accent or describe the rate of some activity, such as tapping a foot or a finger to mark time in music.
4. Kinetic Relationships: These can be either iconic or intrinsically coded relationships in which a nonverbal behavior executes all or part of an action where the performance and the meaning overlap, such as when someone threatens another person with his fist (iconic) or actually strikes that person (intrinsic).
5. Pointing Relationships: These are intrinsically coded relationships where some part of the body (usually the fingers or hand) are used to point to something else.

In addition to their low-level descriptive schema for nonverbal behavior, Ekman and Friesen describe five different high-level categories of nonverbal behavior. These include Emblems, Illustrators, Regulators, Affect Displays, and Adaptors.

- a) *Emblems*, which have a direct dictionary definition;
- b) *Illustrators*, which help visualize a spoken concept;
- c) *Regulators*, which direct conversation;
- d) *Affect displays*, which show emotion; and
- e) *Adaptors*, which are fragmented portions of sequences for caretaking, grooming, and tool-use.

These categories, along with descriptions of a behavior's usage, origin, and coding, provide a thorough framework for cataloguing NVC.

MCNEILL'S GESTURAL CATEGORIES AND "THINKING-FOR-SPEAKING"

David McNeill provided nuance to this categorization by focusing solely on gesture as "movements of the arms and hands which are closely synchronized with the flow of speech" (McNeill, 1992). In McNeill's view different motions take on varying degrees of significance depending on whether you're focusing on personality and emotion or on speech and communication. McNeill and later Cassell (Cassell, 1998) explored the use of communicative gestures by observing and analyzing many cases of people talking about specific subjects, such as real estate, etc. They categorized gestures into the following five categories:

- a) *Iconic gestures*: these represent some features of the subject that a person is speaking about, such as space or shape;
- b) *Metaphoric gestures*: these represent an abstract feature of the subject that a person is speaking about, such as exchange or use;
- c) *Deictic gestures*: these indicate or refer to some point in space;
- d) *Beat gestures*: these are hand movements that occur to accent spoken words; and
- e) *Emblem gestures*: these are gestural patterns that have specific meanings within a given culture, such as "hello" or "ok".

McNeill proposes that gestures do not just reflect our thoughts but that they have a direct impact on the thinking process. He argues that language and gesture work together to *constitute thought*. McNeill writes:

"Such an argument helps to explain why gestures occur in the first place. Gestures occur, according to this way of thinking, because they are part of the speaker's ongoing thought process. Without them thought would be altered or incomplete. The argument also explains why gestures get more complex when the thematic discontinuity from the context is greater. Discontinuity implies the inauguration of a new dimension of thought; since each aspect of the gesture is a possible departure from the preceding context, the gesture can bring in new dimensions by adding complexity." (p. 245)

McNeill's view positions gesture and thought in a dialectic relationship. He connects this to Lakoff and Johnson's (1980)² work on image schema and embodied metaphor.

“Basically, metaphoric gestures permit thinking in terms of concrete objects and space when meaning is abstract. While the human mind may prefer imagistic content, it also has the crucial complementary metaphoric capacity to think of the abstract in terms of the concrete, and this capacity we observe at work in metaphoric gestures.” (p.263)

He also argues that gestures play a role in what Jerome Bruner (1986, 1990)³ has described as the narrative mode of thinking, which he regards as a distinct (and fundamental) mode of cognition that extends into our perception of reality. McNeill adds to this, writing:

“We can go a step further and conceive of gestures playing a role in thought in this narrative mode. Particularly inasmuch as the gesture more directly than the speech, can reflect the mind's own narrative structure, the breaking edge of our internal narrative may be the gesture itself; the words follow.” (p.266)

In a more recent article McNeill and Duncan assess gesture and speech together from a perspective of “thinking for speaking”: a theory related to the strong Whorfian hypothesis that language shapes habitual thought patterns, thus determining what an individual is capable of thinking. Thinking-for-speaking takes a less relativistic approach to language and thought, theorizing that “speakers organize their thinking to meet the demands of linguistic encoding on-line, during acts of speaking” (McNeill & Duncan, 2000). Using this milder form of linguistic relativity, the authors contend that language and gesture comprise a form of external and embodied cognition.

Drawing on Vygotsky and Heidegger they argue that gestures are *material carriers* of thinking: that gestures (and words) are forms of thinking and being in the world which provide windows into the inseparable cognitive and embodied processes they enact. They propose that gestures are used to materialize or concretize thought, and that the “greater the felt departure of the thought from the immediate context, the more likely its materialization in a gesture, because of this contribution to being” (McNeill & Duncan, 2000). This perspective is one of the most philosophically rich approaches to NVC, and its implications for embodied cognition are especially significant when transposed into a virtual worlds context.

VINCIARELLI'S SOCIAL SIGNALING FRAMEWORK

Recently, animators and computer scientists have grown interested in how to classify, categorize, and identify social signaling behaviors in digitally mediated contexts. Vinciarelli et al. discuss the field of “Social Signal Processing” (SSP) as a broad reaching project that aims to analyze social behavior in both human-human and human-computer interactions (Vinciarelli, Pantic, Bourlard, & Pentland, 2008). They describe social signals as: “...complex aggregates of behavioral cues accounting for our attitudes towards other human (and virtual) participants in the current social context. Social signals include phenomena such as attention, empathy, politeness, flirting, and (dis)agreement, and are conveyed through multiple behavioral cues including posture, facial expression, voice quality, gestures, etc.” (Vinciarelli et al., 2008)

Vinciarelli's system of SSP incorporates those elements of proxemics, kinesics, and paralinguistics that lend themselves to computational analysis. They identify five codes for behavioral cues, and discuss their communicative function as shown in Figure 3.2. The codes that they identify are those that lend themselves to automatic signal recognition using a computational system.

²Lakoff, G., and Johnson, M. 1980. *Metaphors we live by*. Chicago: University of Chicago Press.

³Bruner, J. 1986. *Actual Minds, possible worlds*. Cambridge, MA: Harvard University Press

Bruner, J. 1990. *Acts of meaning*. Cambridge, MA: Harvard University Press

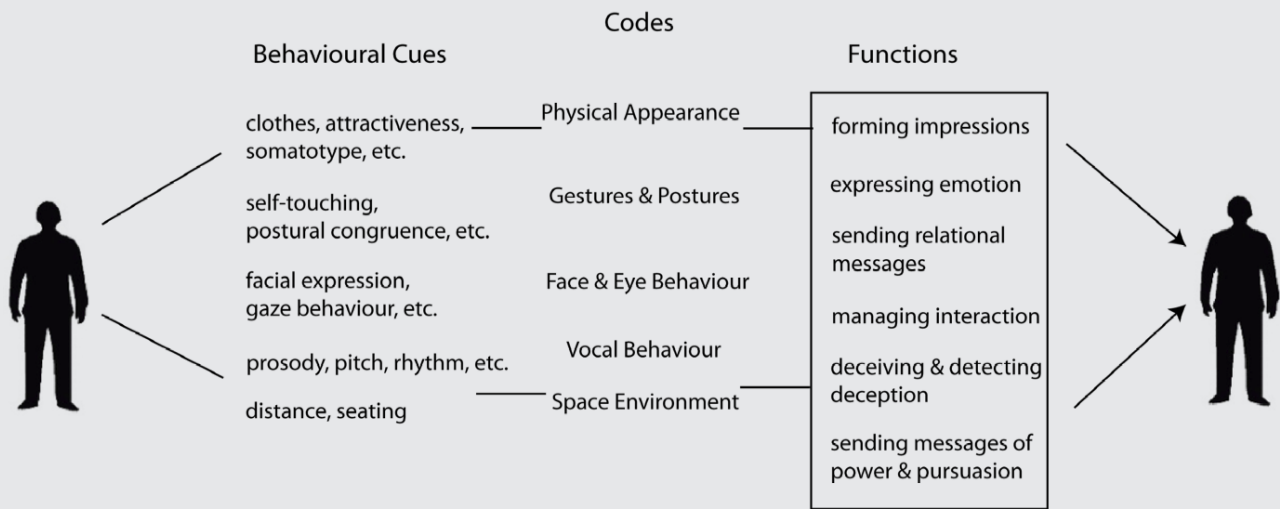


Figure 3-2: Codes, cues, and functions within Vinciarelli et al’s framework (2008)

These categories of behavioral cues can be used to communicate a number of common social signals including emotion, personality, status, dominance, persuasion, regulation, and rapport, as shown in [Table 3-1].

These contemporary systems of classifying and analyzing Nonverbal Communication and Social Signaling Processing provide some leverage on the challenges of communication in virtual environments and animation that the authors in this book grapple with.

3. MOVEMENT ANALYSIS SYSTEMS FROM THEATER AND DANCE

As we have suggested above, the other field that has rigorously explored bodily movements and expression is the Performing Arts. Some of these, such as the Delsarte’s system of movement and vocal expression and Laban’s dance notation, predate the pioneering work of Hall and Birdwhistell. Recent research into animation has given these systems new life and new relevance to the study of communication in virtual worlds.

Table 3-1: Social cues and signals from Vinciarelli et al. (2008)

SOCIAL CUES	EMOTION	PERSONALITY	STATUS	DOMINANCE	PERSUASION	REGULATION	RAPPORT
PHYSICAL APPEARANCE							
Height			✓	✓			
Attractiveness		✓	✓	✓	✓		✓
Body Shape		✓		✓			
GESTURE AND POSTURE							
Hand Gesture	✓				✓	✓	✓
Posture	✓	✓	✓	✓	✓	✓	✓
Walking		✓	✓	✓			
FACE AND EYES BEHAVIOR							
Facial Expressions	✓	✓	✓	✓	✓	✓	✓
Gaze Behavior	✓	✓	✓	✓	✓	✓	✓
Focus of Attention	✓	✓			✓	✓	✓
VOCAL BEHAVIOR							
Prosody	✓	✓			✓		
Turn Taking			✓	✓		✓	✓
Vocalizations	✓	✓		✓	✓	✓	✓
Silence							✓
SPACE AND ENVIRONMENT							
Distance		✓	✓		✓		✓
Seating Arrangement				✓			✓

LABAN MOVEMENT ANALYSIS

Laban Movement Analysis (LMA) is a framework for describing human movement and expression (Maletic, 1987). It is an expansion of Laban's original theories, which are spread across a number of books that he wrote between 1920 and 1984 – including some never translated from their original German. While Laban's practice and writings form the basis for this framework, it is a synthesis of his work and later adaptations of it, including the work of (Bartenieff, 1980), (Lamb & Watson, 1987), (Kestenberg, 1979) and Bainbridge-Cohen (Eddy, 2009). Besides his movement theories, Laban also developed a complex notation system similar to a musical score for recording movement. Laban's observations cover the following general principles (Moore & Yamamoto, 1988):

1. *Movement is a process of change*: it is a “fluid, dynamic transiency of simultaneous change in spatial positioning, body activation, and energy usage” (Moore & Yamamoto, 1988)
2. *The change is patterned and orderly*: due to the anatomical structure of our body, the sequences that movements follow are natural and logical.
3. *Human movement is intentional*: people move to satisfy a need, and therefore actions are guided and purposeful. As a result, intentions are made clear through our movement.
4. *The basic elements of human motion may be articulated and studied*: Laban states that there is a compositional alphabet of movement.
5. *Movement must be approached at multiple levels if it is to be properly understood*: in order to capture the dynamic processes of movement, observers must indicate the various components as well as how they are combined and sequenced.

These general principles are manifest in five categories of movement that comprise the full spectrum of LMA's movement parameters: Body, Effort, Shape, Space and Phrasing. The Body category is concerned with which body parts move and how movement starts and spreads throughout the body. Space describes the size occupied by gesture and the spatial pathways it follows. Shape describes changes in the body. Effort involves the qualities of movement and the energy used by it. Phrasing indicates the transitions that take place between movements.

The Effort category has become the most widely used due to its direct applicability and practice within theatre. Effort represents a broad parameter space that includes Flow, Weight, Space, and Time, and indicates how intent influences the quality of gestures (Dell, 1970). These four parameters form a continuum of opposing polarities, as seen in Table 3-2. Each motion either “indulges” in the quality or “fights” against it (Badler, Allbeck, Zhao, & Byun, 2002).

Table 3-2: Laban's Effort Parameters (Bishko, 1992)

PARAMETER	POLARITIES	MEANING
Flow	Free/Bound	Feeling, Progression, “How” Free is external and releases energy Bound is contained and inward.
Weight	Light/Strong	Sensing, Intention, “What” Moving with lightness is delicate, sensitive, and easy Moving with strength is bold, forceful, and determined.
Space	Flexible/Direct	Thinking, Attention, “Where” Flexibility means being open and broadly aware of spatial possibilities. Directness is focused and specific, paying attention to a singular spatial possibility.
Time	Sustained/Sudden	Intuition, Decision, “When” Sustained movement is continuous, linger, or indulgent. Sudden movement is unexpected, isolated, or surprising.

Space refers to the spatial directions movements follow, and Laban views it as the most significant element (Maletic, 1987). Movement can be reduced to basic directions relative to one's orientation in space. Laban visualizes these possibilities in terms of an icosahedron containing an octahedron that diagrammed the six dimensional directions and a cube diagramming the eight diagonal directions, as shown in Figure 3-3.

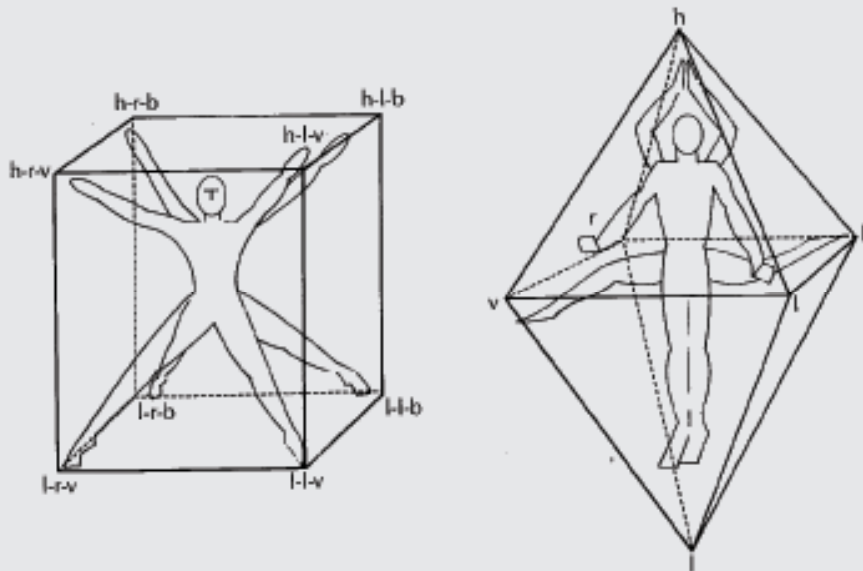


Figure 3-3: Left: The cube; Right: The octahedron (Maletic, 1987)

Laban terms this overall reach space the *kinesphere* (shown in Figure 3-4), and describes movement as geometrical shapes which connect points, including straight, curved, rounded, and twisted ones. He also describes the nature of the transitions that limbs take throughout this space. As the body moves, it constructs “trace-forms” within its kinesphere. These can be linear (straight), planar (curved), or even form volumes (three-dimensional twists or spirals).

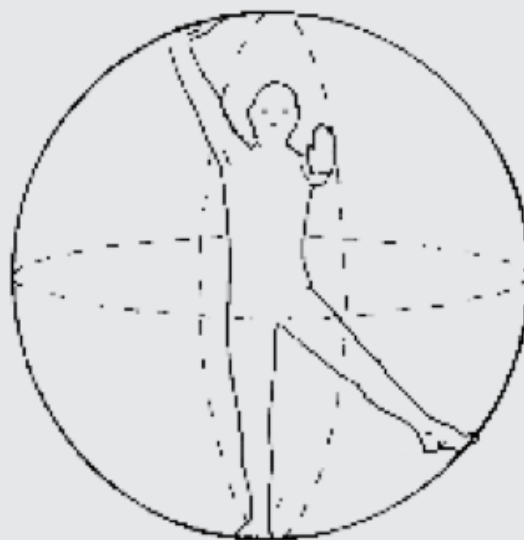


Figure 3-4: The kinesphere (Bartenieff, 1980)

Body describes which body parts are in motion. This can be divided into two broad categories: gestures and postures. Gestures are actions confined to a part of the body, although different parts of the body can make different gestures simultaneously. This ability greatly contributes to human expression and function. On the other hand, postures are positioned assumed by, or involving, the whole body. Description of these actions includes where in the body they start and how body part involvement is sequenced.

Shape is closely involved in the body's movement through space, and is often used as an integrating factor for combining the categories. This category was mostly developed after Laban's death by Warren Lamb. Essentially, movement can be described by the shape that the body takes on, as well as the way the body moves through space (Dell, 1970). This includes shape forms such "wall-like", "ball-like", and "pin-like." Shape qualities are used to describe the way the body is changing: either Opening (growing larger, extending) or Closing (growing smaller). When spatial orientation is introduced, more specific terms are used – rising, sinking, spreading, enclosing, advancing, and retreating. There are also modes of Shape change: Shape Flow – where the form results from changes within body parts, Directional – where the form results from a clear path through the environment, and Shaping/Carving – where the form results from the body actively and three-dimensionally interacting with the environment.

Phrasing describes the qualitative rhythm of movement and how we sequence and layer movements. LMA considers a movement phrase to be a complete idea or theme, represented by movement. A phrase is first prepared, then acted out, then recovered from. People's unique characteristics are expressed through how they construct rhythmic patterns of movement phrases and preferences towards the other four main categories described by LMA. Phrasing is especially concerned with the transitions of Effort qualities during movement, and can be classified in the following ways (Bishko, 1992):

- Even: continuous, unchanging
- Increasing: from lesser to greater intensity
- Decreasing: from greater to lesser intensity
- Accented: series of accents
- Vibratory: series of sudden, repetitive movements
- Resilient: series of rebounding, resilient movements

Laban's principles have been turned into a broadly useful framework for describing movement, especially in the field of dance. However, they have also been used within a variety of observational frameworks for nonverbal behaviour, including work efficiency, theatre, and physical therapy. The LMA system provides a thorough vocabulary for talking about the movements people make and framing one's conclusions. Before use, one should be careful to constrain what elements of LMA are being used, however. Since Laban's works spanned several decades and some weren't translated from German, some on what principles are most useful. As a result, his successors developed his principles in different directions and emphasized certain aspects depending on their interests. Nonetheless, by being selective and precise, one obtains a powerful tool in LMA.

DELSARTE

François Delsarte was a singer and actor that lived from 1811 to 1871, mainly in Paris, France. After having his voice ruined by bad singing coaches (Shawn, 1954), he turned to teaching. In order to train actors well, Delsarte performed systematic observations of human action and interaction. Based on this, he created a system of expression: an artistic aesthetic system that describes the link between meaning and motion.

Delsarte's system inspired dancers who changed the state of their art around the turn of the 20th century, and introduced a completely new technique and vocabulary that ushered in a new era of modern dance (Shawn, 1954). Though "couched in a language and terminology from the 1800s that strikes a 21st

century reader as perhaps quaint and metaphysical”, this technique structurally describes how attitude and personality are conveyed by body postures and gestures (Marsella, Carnicke, Gratch, Okhmatovskaia, & Rizzo, 2006). Since Delsarte’s work has inspired generations of dancers and has influenced other artistic systems, it provides a useful starting point in the study of meaningful movement.

Delsarte grounded his work in systematic observations and found that there are three forms of expression for gesture (Stebbins, 1977):

1. The habitual bearing of the agent
2. The emotional attitudes of the agent
3. The passing inflections of the agent

These forms roughly correspond to personality traits, emotions, and conversational gesticulations. Delsarte ties meaning in his system back to some combination of these forms. He also based several laws on them: the doctrine of special organs, the three great orders of movement, and the nine laws of motion. Delsarte’s system divides the body into zones, which he further subdivided into three parts, the mental, moral, and vital subsections (Stebbins, 1977). His doctrine of special organs indicates that these zones become significant points of arrival or departure for a meaning. For example, the head zone is associated with mental or intellectual meaning, and modifies gestures that start or end near the head accordingly.

Delsarte’s primary law is that of correspondence, because he believed that each gesture is expressive of something. This forms the foundation for connecting emotions and traits to motion within his system. Delsarte also provides several principles of motion, including the meaning of certain zones of the body, and the meaning of different directions of movement. Each part of the body corresponds to a different meaning, and actions are coloured by the zones in which they start and finish.

Motions away from the centre (e.g. the body) are termed “excentric” and have relation to the exterior world. Motions towards the centre are termed “concentric” and have relation to the interior. Balanced motion is “normal” and moderates between the two. The system also includes a variety of laws indicating the meaning of parallel motion and successions of movements along the body. Finally, it also describes nine possible poses per body zone (combinations of excentric, normal, and concentric) and their meaning.

Figure 3-5 depicts each of the nine pose combinations. In it, the excentric grand division is shown in the left column, demonstrating the head turned away from the object of its emotion. The right column shows the concentric grand division, with the head turned towards the object. The top row shows the excentric sub-division, with the head raised. The bottom row shows the concentric sub-division, with the head lowered. The middle column and row show the normal grand and sub-division, respectively.

The three orders of movement identified by Delsarte are “Oppositions”, “Parallelisms”, and “Successions”. Oppositions occur when two body parts move in opposite directions simultaneously, indicating strength. Parallelisms occur when two body parts move in the same direction simultaneously and indicate weakness, or possibly stylized movement such as dance. Successions occur when movement passes through the body from an origin through each connecting part in turn. True successions begin at the centre, work outwards, and indicate good and true motivations. Reverse successions begin at the extremity, work inwards, and indicate evil and false motivations (Shawn, 1954).

The nine laws of motion are attitude, force, motion (expansion and contraction), sequence, direction, form, velocity, reaction, and extension. These laws further modify what each movement means.

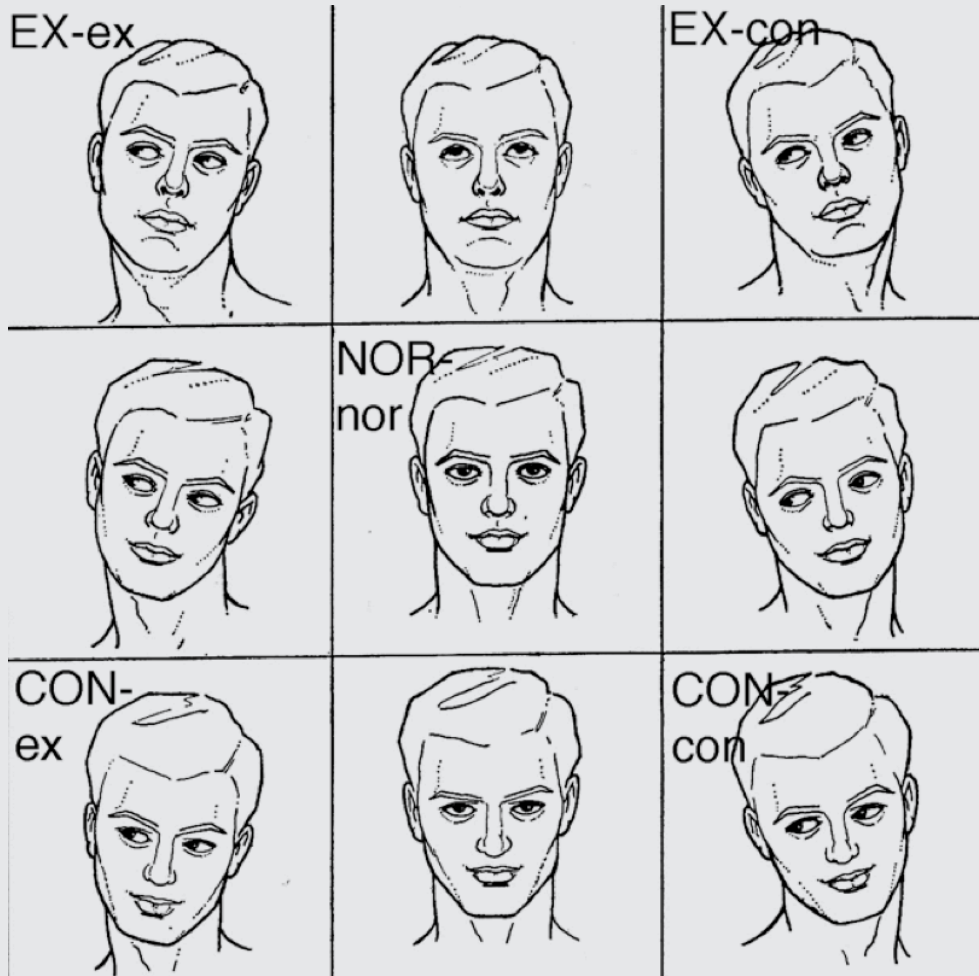


Figure 3-5: Delsarte’s nine-fold poses for the head (Shawn, 1954)

LAW	MEANING
Attitude	The vertical direction of a motion affects the meaning. Positive assertion rises, negative falls. Upwards, or forward motion is constructive or positive. Actions that move down or backwards are destructive or negative.
Force	Conscious strength assumes weak attitudes. Conscious weakness assumes strong attitudes.
Motion (expansion & contraction)	Excitement expands, thought contracts, love and affection moderate
Sequence	Thought leads to facial expression leads to the attitude of the body leads to gesture leads to speech. This is the proper sequence of events and true successions demonstrate them.
Direction	Movement in heights and depths (up and down) are intellectual. Movement in lengths (front and back) are passionate (emotional). Movement in breadths (side to side) are volitional (demonstrate the will). Diagonals are conflicted.
Form	Straight movement is vital. Circular movement is mental. Spiral movement is moral.
Velocity	The rhythm and tempo of the movement is proportionate to the mass (including emotion) to be moved.
Reaction	Everything surprising makes the body recoil. Degree should be proportionate to the degree of emotion.
Extension	The extension of the gesture is in proportion to the surrender of the will in emotion. Extension beyond the body achieved by holding the body at its culmination, with held breath.

Table 3-3 : Delsarte’s nine laws (Shawn, 1954)

Delsarte’s system has been the basis and inspiration for artists, but has not been rigorously and thoroughly validated across the entire body. His work on the attitudes of the hand has been studied (Marsella et al., 2006) through the use of participant observation. Results showed that people interpret the animation of an un-textured hand according to Delsarte’s mapping “remarkably consistently.” The impact of Delsarte’s system on the artistic expression of animators (Nixon, 2009) has also been examined and indicates that animators working under the direction of his system produce characters that are somewhat more likely to express similar emotions than when relying on their personal style. Furthermore, these expressions are somewhat more likely to convey the intended emotion to viewers. These explorations of Delarte’s work show that it has some potential for guiding the development of nonverbal animation in virtual environments; however, there is still work to be done formalizing and validating this approach if it is to be successful, particularly in regards to how it links psychological states and movement.

BOGART AND LANDAU’S VIEWPOINTS

“Viewpoints” is a framework of motion and performance that was developed in dance by choreographer Mary Overlie and later transposed into theatrical practice. Although originally conceived of as a six-part framework (Comprised of Space, Shape, Time, Emotion, Movement, and Story) various practitioners and collaborators have expanded it into a series of viewpoints that govern the composition of both motion and vocalization in performance. *The Viewpoints Book*, by Anne Bogart and Tina Landau is perhaps the most coherent guide to Viewpoints in theatre practice today (Bogart & Landau, 2005).

Bogart and Landau propose nine physical viewpoints divided into the categories of Time and Space. They provide several definitions of Viewpoints, writing:

- “Viewpoints is a philosophy translated into a technique for (1) training performers; (2) building ensemble; and (3) creating movement for the stage.
- Viewpoints is a set of names given to certain principles of movement through time and space; these names constitute a language for talking about what happens onstage.
- Viewpoints is points of awareness that a performer or creator makes use of while working.” (Bogart & Landau, 2005)

Viewpoints thus can be used as tool for creating theatrical “compositions” and also as a lens for evaluating motion in any performative context. The nine physical viewpoints described by Bogart and Landau include:

Table 3-3 : Delsarte’s nine laws (Shawn, 1954)

VIEWPOINTS OF TIME	
TEMPO	“The rate of speed at which a movement occurs; how fast or slow something happens onstage.”
DURATION	“How long a movement or sequence of movements continues. Duration, in terms of Viewpoints work, specifically relates to how long a group of people working together stay inside a certain section of movement before it changes.”
KINESTHETIC RESPONSE	“A spontaneous reaction to motion which occurs outside you; the timing in which you respond to the external events of movement or sound; the impulsive movement that occurs from a stimulation of the senses. An example: someone claps in front of your eyes and you blink in response; or someone slams a door and you impulsively stand up from your chair.”
REPETITION	“The repeating of something onstage. Repetition includes (1) <i>Internal Repetition</i> (repeating a movement within your own body); (2) <i>External Repetition</i> (repeating the shape, tempo, gesture, etc. of something outside your own body).”

(con’t. on next page)

VIEWPOINTS OF SPACE
SHAPE
“The contour or outline the body (or bodies) makes in space. All Shape can be broken down into either (1) <i>lines</i> ; (2) <i>curves</i> ; (3) a <i>combination</i> of lines and curves... In addition, Shape can either be (1) stationary; (2) <i>moving</i> through space. Lastly Shape can be made in one of three forms: (1) the body in space; (2) the body in relationship to architecture making a shape; (3) the body in relationship to other bodies making a shape.”
GESTURE
“A movement involving a part or parts of the body; Gesture is Shape with a beginning, middle and end...Gesture is broken down into: 1. BEHAVIORAL GESTURE. Belongs to the concrete, physical world of human behavior as we observe it...scratching, pointing, waving, sniffing, bowing, saluting. A Behavioral Gesture can give information about character, time period, physical health, circumstance, weather, clothes, etc...[it] can be further broken down...in terms of <i>Public Gesture</i> and <i>Private Gesture</i> . 2. EXPRESSIVE GESTURE. Expresses an inner state, an emotion, a desire, an idea, or a value. It is abstract and symbolic rather than representational.”
ARCHITECTURE
“The physical environment in which you are working and how awareness of it affects movement...Architecture is broken down into: 1. SOLID MASS. Walls, floors, ceilings, furniture, windows, doors, etc. 2. TEXTURE. Whether the solid mass is wood, or metal, or fabric. [etc.] 3. LIGHT. The sources of light in the room, the shadows we make in relationship to these sources, etc. 4. COLOR. Creating movement off of colors in the space... 5. SOUND. Sound created by and from the architecture.”
SPATIAL RELATIONSHIP
“The distance between things onstage, especially (1) one body to another; (2) one body (or bodies) to a group of bodies; (3) the body to the architecture.”
TOPOGRAPHY
The landscape, the <i>floor pattern</i> , the <i>design</i> we create in movement through space...staging or designing for performance always involves choices about the size and shape of the space we work in.”

Viewpoints has much in common with the NVC systems of Ekman and Friesen, and with Hall’s proxemics. Viewpoints of space can be mapped to proximal relationships, and to kinesic gestures and postures. Unlike Laban’s movement analysis, viewpoints is not primarily concerned with the communicative motivations behind the motions described. Instead, viewpoints is interested in decomposing movements into salient, and readily understood sub-movements, which are amenable to conscious control and modification. Unlike Birdwhistell’s micro-kinesics, which attempted to isolate meaning in tiny moments of movement, Viewpoints is concerned with only those movements which may be apprehended by a viewer without the aid of slow-motion or frame-by-frame replay.

4. CHAPTER REVIEW

In this chapter we have discussed a range of theories, systems, and perspectives on nonverbal communication and bodily expression. Rooted in the embodied experience of being in the world, it should come as no surprise that there are many points of overlap and agreement between these systems, even when spread across multiple disciplines with distinctly different goals. One need not understand all of these systems in detail in order to begin exploring the potential of NVC for virtual worlds; however a passing familiarity with their principles can deepen a designer’s process and sharpen a theorist’s insight into how we use avatar bodies to communicate.

5. BIBLIOGRAPHY

- Argyle, M., & Dean, J. (1965). Eye-Contact, Distance and Affiliation. *Sociometry*, 28, 289 - 304.
- Badler, N., Allbeck, J., Zhao, L., & Byun, M. (2002). *Representing and Parameterizing Agent Behaviors*, Geneva.
- Bartenieff, I. (1980). *Body Movement: Coping with the Environment*. New York: Routledge.

- Birdwhistell, R. L. (1952). *Introduction to Kinesics: an annotation system for analysis of body motion and gesture*. Louisville, KY: Louisville University Press.
- Bishko, L. (1992). *Relationships between Laban Movement Analysis and Computer Animation*.
- Bogart, A., & Landau, T. (2005). *The Viewpoints Book: A Practical Guide to Viewpoints and Composition*. New York: Theatre Communications Group.
- Burgess, J. W. (1983). Developmental Trends in Proxemic Spacing Behavior Between Surrounding Companions and Strangers in Casual Groups. *Journal of Nonverbal Behavior*, 7(3), 158 - 169.
- Calero, H. H. (2005). *The Power of Non-Verbal Communication: What You Do is More Important than What You Say* (First ed.): Silver Lake Publishing.
- Cassel, J. (1998). A Framework for Gesture Generation and Interpretation. In R. Cipolla & A. Pentland (Eds.), *Computer Vision in Human-Machine Interaction* (pp. 191-215). New York: Cambridge University Press.
- Darwin, C. (1874). *The expression of the emotions in man and animals*: D. Appleton and Company.
- Dell, C. (1970). *a primer for movement description using effort-shape and supplementary concepts*. New York: Dance Notation Bureau, Inc.
- Duncan, S. (1969). Nonverbal Communication. *Psychological Bulletin*, 72(2), 118 - 137.
- Eastman, C. M., & Harper, J. (1971). A Study of Proxemic Behavior: Toward a Predictive Model. *Environment and Behavior*, 3(4), 418-437.
- Eddy, M. (2009). A brief history of somatic practices and dance: historical development of the field of somatic education and its relationship to dance. *Journal of Dance and Somatic Practices*, 1(1), 5-27.
- Ekman, P., & Friesen, W. V. (1981). The Repertoire of Nonverbal Behavior: Categories, Origins, Usage, and Coding. In T. A. Sebeok, J. Umiker-Sebeok & A. Kendon (Eds.), *Nonverbal Communication, Interaction, and Gesture*. The Hague: Mouton Publishers.
- Fast, J. (1970). *Body Language*. New York, N.Y.: Neirenberg.
- Hall, E. T. (1966). *The Hidden Dimension*: Anchor Books.
- Hall, E. T. (1968). Proxemics. *Current Anthropology*, 9(2/3), 83-108.
- Joinson, A. N. (2003). *Understanding the Psychology of Internet Behavior*. New York: Palgrave Macmillan.
- Jolly, S. (2000). Understanding Body Language: Birdwhistell's Theory of Kinesics. *Corporate Communications: An International Journal*, 5(3), 133 - 139.
- Kestenberg, J. S. (1979). *The role of movement patterns in development* (Vol. 1). New York: Dance Notation Bureau Press.
- Lamb, W., & Watson, E. (1987). *Body code: The meaning in movement*: Princeton Book Company.
- Maletic, V. (1987). *Body - Space - Expression: The Development of Rudolf Laban's Movement and Dance Concepts*. Berlin: Walter de Gruyter & Co.
- Marsella, S., Carnicke, S. M., Gratch, J., Okhmatovskaia, A., & Rizzo, A. (2006). *An exploration of Delsarte's structural acting system*, Marina del Rey, California.
- McNeill, D. (1992). *Hand and mind: what gestures reveal about thought*: University of Chicago Press.
- McNeill, D., & Duncan, S. (2000). Growth points in thinking-for-speaking. In D. McNeill (Ed.), *Language and Gesture* (pp. 162-185). Cambridge: Cambridge University Press.
- Moore, C.-L. (2005). *Movement and Making Decisions: The Body-mind Connection in the Workplace*: Dance & Movement Press.
- Moore, C.-L., & Yamamoto, K. (1988). *Beyond Words: Movement Observation and Analysis*. New York: Gordon and Breach.
- Muybridge, E. (1979). *Muybridge's Complete Human and Animal Locomotion: All 781 Plates from the 1887 Animal Locomotion* (Vol. 1-3): Dover Publications.
- Nixon, M. (2009). *Enhancing Believability: Evaluating the Application of Delsarte's Aesthetic System to the Design of Virtual Humans*. Simon Fraser University.
- Shawn, T. (1954). *Every Little Movement*. Brooklyn, New York: Dance Horizons.
- Stebbins, G. (1977). *Delsarte System of Expression*. Brooklyn: Dance Horizons.

4

BASICS OF NONVERBAL COMMUNICATION IN VIRTUAL WORLDS

By Theresa Jean Tanenbaum, Michael Nixon, and Magy Seif El-Nasr

The research space for digitally mediated nonverbal communication is quite broad, encompassing research into animation, computer mediated communication, virtual reality, virtual worlds, massively multiplayer online games, and embodied cognition. As a consequence, much of the relevant work to studies of NVC in virtual worlds lies outside the disciplinary core of virtual worlds research. In this chapter we will first look at research NVC research that doesn't directly engage with virtual worlds, but which does have significant implications for the field. We will then look at how research in virtual worlds is currently engaging with issues of nonverbal communication. We also include a discussion of the recent work that deals more broadly with social interactions in virtual worlds and massively multi-player online role playing games (MMORPGS)

1. STUDIES OF NVC IN DIGITAL ENVIRONMENTS AND VIRTUAL REALITY

NVC IN DIGITAL ENVIRONMENTS

A number of important studies have been performed that do not fit under the of Virtual Worlds research. One of the foundational papers on creating realistic nonverbal actions for actors in virtual worlds was Ken Perlin and Athomas Goldberg's paper on Improv: a scripting system and architecture for animating virtual agents (Perlin & Goldberg, 1996). Although it did not specifically engage the literature on NVC, the Improv system laid the groundwork for much of the future work in virtual world development, creating systems that created procedurally generated animations for virtual characters and directed them in autonomous behaviors.

One common thread of investigation for studies of NVC in virtual worlds was aimed at gauging human reactions to different degrees of "realism" in nonverbal cues given by virtual characters. Katherine Isbister and Clifford Nass performed preliminary work on the relationship between verbal and nonverbal cues. They asked human participants to evaluate interactions with animated virtual characters that exhibited differing levels of extroversion and introversion (Isbister & Nass, 2000). They determined that participants preferred characters where the verbal and non-verbal cues were internally consistent. Similarly, van Es et al. performed a study, in which participants were asked to interact with one of three different "talking heads", each of which exhibited different eye gaze behaviors (van Es, Heylen, van Dijk, & Nijholt, 2002). Unsurprisingly, their results indicated that participants regarded agents with more "natural" gaze behavior more positively than agents with random gaze behavior.

Another active area of inquiry has looked at social behaviors in textually mediated spaces. Walther et al. tested whether or not the medium of communication – in this case face-to-face vs text chat – impacted the creation of subjectively experienced affinity, and the communication of affect between participants, concluding that there was no measurable difference between the communication channels (Walther, Loh, & Granka, 2005). Amy Bruckman wrote about gender switching and gender identity in MUDs (Bruckman, 1993). Sherry Turkle discussed the sociological systems of MUDs, through a lens of postmodernism (Turkle, 1995). Richard Bartle developed a typology of players in Multi User Dungeons/Dimensions (MUDs) (Bartle, 1996). Julian Dibbell wrote about his time in the textual world of Lambda Moo, and most famously on the impact of a virtual sexual assault on the members of the online community (Dibbell, 1998). This body of work provides a valuable historical context for our study of Nonverbal Communication in Virtual Worlds, however, the lack of any sort of non-textual avatar body makes the vast majority of NVC frameworks inapplicable to these spaces. For this reason, we have restricted the remainder of this chapter to work that operates within graphical spaces.

VIRTUAL WORLDS VS. VIRTUAL REALITY

Virtual Reality (VR) and Virtual Worlds (VWs) have much in common, but are separated by their assumptions about embodiment and proprioception. Embodiment is a broad term that can refer to a whole range of phenomena. Tim Rohrer divides embodiment into twelve distinct dimensions, based on a survey of how embodiment has been approached within cognitive science (Rohrer, 2007). Two of Rohrer's dimensions are of particular relevance to work in VR and VWs:

- **Phenomenology:** This dimension of embodiment deals with our conscious awareness of our own bodies, and their role in mediating our experience of the world. Rohrer distinguishes between “conscious phenomenology” and the notion of the “cognitive unconscious” from cognitive psychology which deals with the autonomic reactions of the body. In Virtual Reality research, this is commonly the form of embodiment under discussion.
- **Perspective:** Embodiment can be taken to refer to a particular vantage point, or point-of-view, from which an embodied perspective is taken. Perspectives imply bodily orientations, which arise from a situatedness within the world. In Virtual Worlds, embodiment is often a function of the perspective by which an interactor orients herself to the world, often as mediated by an avatar body of some sort.

In Virtual Reality, one's sense of being embodied in the experience is often considered to overlap consistently with one's phenomenological sense of embodiment: the virtual body and the physical body of the interactor overlap. Research in Virtual Reality is often located within a laboratory environment, utilizing controlled experimental methods. Due to the infrastructure needed for most VR experiences, these experiments are done with individual interactants, or small local groups. Participants are often placed in a head mounted display that attempts to simulate the visual experience being bodily present within the virtual space. In these conditions, the participant's sense of embodiment is assumed to be focused on her own body, and it's overlapping virtual representations.

In Virtual Worlds, which are primarily mediated via a monitor, keyboard, and mouse, embodiment becomes more about the player's ability to imagine her perspective into the virtual space via an avatar⁵.

⁵Early work on avatar embodiment in virtual worlds was primarily concerned with the creation of rudimentary avatar bodies for players to inhabit, as is the case in (Bowers, Pycock, & O'Brien, 1996) and (Benford, Bowers, Fahl'en, Greenhalgh, & Snowdon, 1997). More recently, T.L. Taylor has written about the affordances of two dimensional avatars in “*The Dreamscape*” (Taylor, 2002). Taylor breaks her analysis in two categories: Social Life and Avatar Identity. Social life discusses issues around presence, communication, affiliation, socialization, and sexuality; avatar identity deals with issues around customization, personal expression, avatar autonomy, and social experimentation.

Proprioception is the perception of the physical state of the body. Up until recently, this was used primarily to refer only our perception of our own bodily state, however recent work in the neurological representation of movement in the brain, and the discovery of “mirror neurons” has been used to argue for proprioception as an aesthetic sense that includes our perception of others bodies on motion (Montero, 2006). This sense of projected embodiment helps explain how players experience the bodies of their avatars in virtual worlds.

Research within existing commercial (or open source) Virtual Worlds does not lend itself to the same experimental controls as research in VR. Instead, many of the studies of NVC in virtual worlds take the form of ethnographic studies, or virtual travelogues. Some notable exceptions to these exist, in which experimental procedures have been undertaken with and without the knowledge of the participants of the virtual world under study. In some cases, researchers have developed their own in-house virtual world, in which they can enact experimental controls by modifying the design of the world as needed. In all of these cases, one important factor is the quantity of people simultaneously interacting within the world. Unlike VR research, which tends to focus on individuals, virtual worlds research often must contend with hundreds of simultaneous interactors in an ever shifting population. In virtual worlds research, the participants and researchers pilot or “puppet” an avatar on a screen, alternating between third and first-person perspectives. The sense of embodiment felt in this case is assumed to be a virtual embodiment, in which the body of the interactor disappears, and the proprioceptive sense is mapped onto the avatar body instead.

STUDIES OF NVC IN VIRTUAL REALITY

Some of the earliest work in social signaling in digitally mediated environments takes place in virtual reality systems, within laboratory settings. While there is very little work that directly relates to Nonverbal Communication in VR, the few studies that *do* exist are very important, and establish a critical precedent for looking at NVC in mediated environments. In particular, the work of Bailenson et al. has investigated proxemic spacing behavior and mutual gaze in VR settings.

In (Bailenson, Blascovich, Beall, & Loomis, 2001) a series of controlled experiments were devised to determine the impact of realistic gaze behavior on participants interacting with a virtual agent. Participants were placed in a virtual space (via head-mounted display) and told to inspect the details of a virtual agent. Three different agent conditions were tested: a “photorealistic” agent, a flat shaded agent, and a non-human pillar. The two human agent conditions were varied for different levels of gaze behavior [Figure 4-1].

To interact with the agents, participants were allowed to move freely through a physical room, while being tracked by cameras. Their results showed that participants were more likely to respect the personal space of the agent as the realism conditions increased [Figure 4-2]. Their results also showed that female participants were more likely to regulate their proxemic behavior than male participants.

In a follow up study, Bailenson’s group refined their experimental design. In (Bailenson, Blascovich, Beall, & Loomis, 2003) two agents were used: one male and one female. The gaze conditions were simplified to a low-realism condition and a high-realism condition, and a new variable of “perception of agency” was introduced in which participants were led to believe that the virtual agent was actually an avatar under human control in some cases. This new study determined that perception of agency⁶, along with realistic gaze was actually a key factor in how participants perceived personal space around the virtual human.

⁶A later study by Shilbach et al. indicated that there is a measurable neural correlate between perception of social entailment and perception of communicative intent. Using fMRI to measure neural activity, researchers placed participants in a virtual social situation in which facial expression and gaze behavior was varied to either indicate specific communicative intent, or to shift arbitrarily (Schilbach et al., 2006). Participants interacting with the non-arbitrary conditions evidenced neural activity associated with social communication.

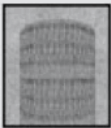








CONTROL	CYLINDER		
COND 1	EYES CLOSED		
COND 2	EYES OPEN		
COND 3	BLINKING		
COND 4	BLINKS And HEAD TURNS		
COND 5	BLINKS, HEAD TURNS, And PUPIL DILATION		

Figure 4-1: Gaze behavior conditions in (Bailenson, et al., 2001).
Reprinted by permission of MIT Press Journals.

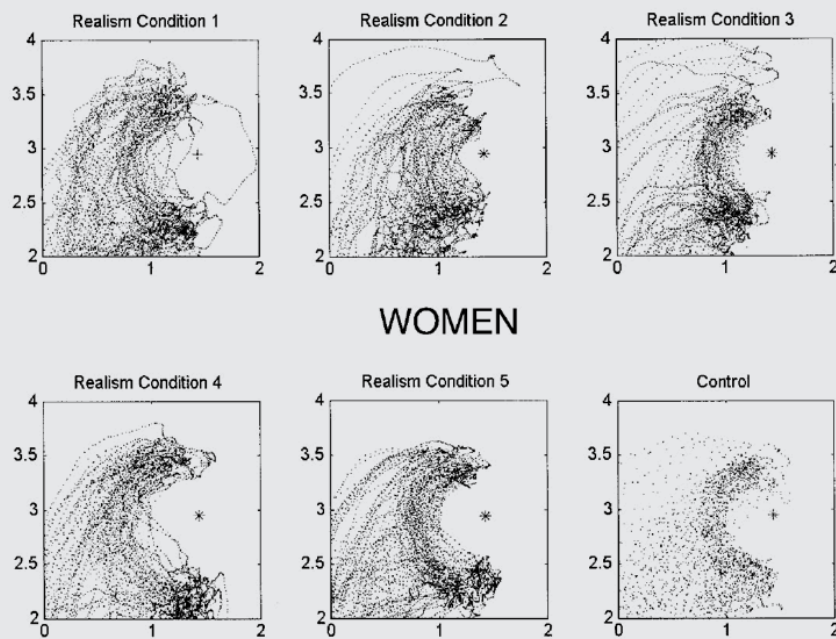


Figure 4 2: Results for Female participants interacting with virtual agents in (Bailenson, et al., 2001). In this “overhead view” dotted lines indicate the rout taken through the room by participants, while the * indicates the location of the agent.
Reprinted by permission of MIT Press Journals.

Their results indicated that within VR, proxemic spacing behaviors and gaze operate almost identically to how they operate in the physical world.

2. STUDIES OF NVC IN VIRTUAL WORLDS

In 1998 the Joint European Commission and National Science Foundation Strategy Group met to discuss the key research challenges and opportunities in information technology, including human-centered computing, online communities, and virtual environments (J. R. Brown et al., 1999). Their initial report included a call for more study of virtual worlds, a recommendation that signaled a growing interest and concern with these growing social spaces. In the last 15 years we have seen an explosion of research in how people communicate within graphically mediated virtual environments, and an ever evolving landscape of virtual worlds in which to conduct this research.

In his 2006 article on *The Demographics, Motivations, and Derived Experiences of Users of Massively Multi-User Online Graphical Environments* Nick Yee builds a case for social interactions in virtual worlds being more intense than those that occur in the physical world, due to the particular way in which they are mediated (Yee, 2006). Citing (Walther, 1996) he writes that:

“...one of the reasons why hyperpersonal interactions – interactions that are more intimate, more intense, and more salient because of the communication channel – occur in computer-mediated communication (CMC) is because participants can reallocate cognitive resources typically used to maintain socially acceptable nonverbal gestures in face-to-face interactions and focus on the structure and content of the message itself. The message itself then comes across as more personal and articulate. Indeed, in virtual worlds where we do not have to constantly worry about how we look and behave, we would be able to dedicate more cognitive resources to the message itself.” (Yee, 2006)

This is interesting, in that it suggests that the narrower communication channels in virtual worlds serve to filter out communicative “noise” from conversations. Within current virtual environments, kinesic nonverbal communication requires the same degree of active monitoring and attention as verbal communication, such that it also fits within this privileged-narrow channel. A consequence of this is that there is more opportunity for participants in a conversation to construct idealized versions of themselves, by presenting carefully curated visuals, behaviors, and utterances. The inclusion of more nuanced NVC mechanisms in virtual worlds – especially autonomic and subconscious NVC such as posture and status – has the potential to disrupt communication, or at least render it less amenable to dissemblance.

EARLY WORK ON KINESICS IN VIRTUAL WORLDS

In one of the earliest works on NVC in virtual worlds, Guye-Vuilleme et al. developed a simple button-based GUI for puppeteering avatars in a shared virtual environment (Guye-Vuilleme, Capin, Pandzic, Magnenat, & Thalmann, 1999). In their review of NVC in the social sciences, they cite (Corraze, 1980) and usefully establish three categories of information that can be conveyed:

1. Information about the affective state of the sender
2. Information about his/her identity
3. Information about the external world

To communicate this information,(Corraze, 1980) identifies three main channels:

1. The body and its moves
2. The artefacts linked to the body or to the environment
3. The distribution of the individuals in space

We see some precedence for item #2 on the second list in the later Kinesic work of Michael Argyle. In his *Bodily Communication* he describes six broad areas for which NVC is used: for expressing emotion, for communicating interpersonal attitudes, for expressing personality, for augmenting speech, as a form of social ritual or ceremony, and as a social form of persuasion (Argyle, 1975). Interestingly, Argyle's framing of NVC is not limited to just the body in space, and its movements; he also includes visual elements like hair, physique, and clothing, arguing that appearance is an important sphere for nonverbal communication (Argyle, 1975).

Guye-Vuilleme et al. go on to describe their prototype shared virtual environment VLNET (Virtual Life Network), which they modified to support kinesic NVC. Their initial experiment gave precedence to affect displays, as represented in facial expressions and postures, and to emblems in the form of specific triggered gestures. They developed a button based "control panel" interface that allowed participants to trigger and control several gestural parameters [Figure 4-3]. To evaluate this interface, an informal experiment was run to determine if participants using the VLNET framework and tools could replicate their real world relationships with each other within the digital environment. Participants were placed in the environment with no tasks to perform, and invited to interact (or not) freely. They observed that participants of varying degrees of real life intimacy exhibited varying interactional distances in the virtual world, consistent with findings in proxemics. They also observed that participants were more likely to use a wide range of gestures, while postures were often selected at the beginning of the interaction, and left alone for the duration of the exchange. This is in keeping with Ekman's framework of NVC, and appears to be a function of the participant's awareness and intentionality in regards to these different modes of communication.

Interestingly, the VLNET system was configured so that participants experienced the world from a first person perspective only. Participants expressed a desire to be able to see the body of their avatar during the study, but Guye-Vuilleme et al. feared that this would take away from the user's sense of immersion

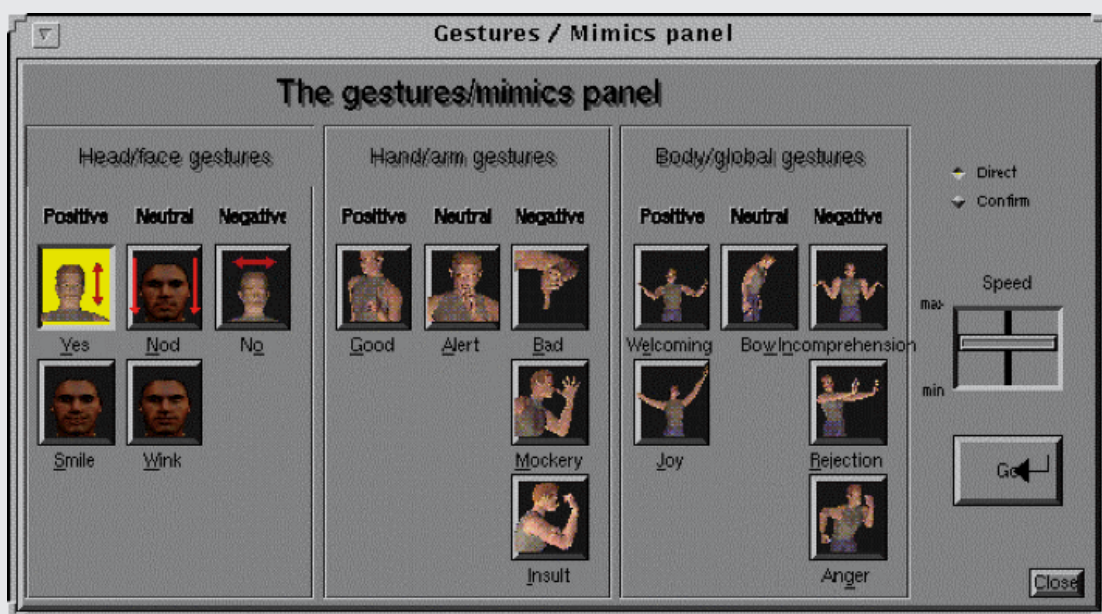


Figure 4-3: The NVC application interface from (Guye-Vuilleme et al., 1999).
With kind permission from Springer Science and Business Media.

in the world. The near ubiquity of third person views in virtual worlds today, and research into virtual embodiment and the importance of the avatar as discussed by (Benford et al., 1997) and (Bowers et al., 1996) prior to Guye-Vuilleme’s work strongly contradicts this claim. Later work, such as that of (Schroeder, 2002), (Taylor, 2002), and (Morie & Verhulsdonck, 2008) all deals with the extreme importance of the visible avatar body to users of virtual worlds.

THE IMPACT OF DESIGN AND AFFORDANCES OF VIRTUAL WORLDS

Guye-Vuilleme et al. developed a prototype virtual world to study NVC; at the same time Barbara Becker and Gloria Mark were investigating social signaling behaviors across three existing virtual worlds: Active Worlds, Onlive Traveler, and Lambda MOO (Becker & Mark, 1998). Their paper on social conventions in collaborative virtual environments is one of the first extensive comparative ethnographic studies of virtual worlds. They identified a number of social conventions in the physical world that also were present within the virtual worlds under study. These included greeting behavior; leave-taking; group formation; privacy indication; nonverbal expression; social positioning and intimacy; and sanctions of unwanted behavior. Their analysis of the frequency of use, and the effectiveness of these behaviors indicated that one of crucial factors in supporting social conventions in virtual worlds was the technological affordances of the world itself. It should come as no surprise to us that this is the case: the affordances of the interface, the design and animation of the avatars, and the physical rules governing how avatars occupy space all have profound implications on what nonverbal behaviors people will favor.

The relationship between affordances, design, and observed communication is the subject of much of the literature surrounding virtual worlds. In 2002 Cheng et al. published an article describing their experience developing and deploying virtual worlds for Microsoft (Cheng, Farnham, & Stone, 2002). Their paper reported on lessons learned from seven years of building graphical virtual environments, including Microsoft V-Chat and the Microsoft Virtual Worlds Platform. It is rare to find scholarly articles discussing the design of a commercial game or software system from the perspective of the developer⁷, and this paper provides a helpful bit of insight into the design process of virtual worlds, and in particular to role that NVC plays in this process. The article identifies nine design lessons to foster sustainable dynamic communities, which the authors classify into three areas [Table 4-1]. Their research methods consisted of a combination of qualitative and quantitative approaches including informal observations of people in the graphical environments, interviews and surveys, analysis of log data, and experimental studies. These evaluation processes fed into the ongoing revision of the virtual world platforms in an iterative design process.

Table 4-1: Design lessons for fostering sustainable, dynamic communities (Cheng et al., 2002).

GENERAL AREA	SPECIFIC DESIGN LESSONS
Individuals	“1. Provide persistent identity to encourage responsible behavior, individual accountability, and the development of lasting relationships.” “2. Support custom profile information that addresses the privacy concerns of individuals.” “3. Encourage individuals to invest in their self-representation by supporting custom end user graphical representations.”
Social Dynamics	“4. Support the ability for groups to form and then self-regulate.” “5. Frequent and repeated interactions promote cooperative behavior. Help people coordinate finding and meeting those they care about to increase the likelihood of positive interactions.” “6. Make community spaces more compelling by supporting the development of reputation and status.”
Context, Environments and User Interface	The rhythm and tempo of the movement is proportionate to the mass (including emotion) to be moved. “7. End users and world builders preferred 3D, non-abstract environments with a third person view.” “9. Different communities have different needs and require different user interfaces.”

⁷One of the most comprehensive guides to the design of virtual worlds comes from Richard Bartle, who developed the first MUD. It is recommended reading for anyone interested in the practice of world design (Bartle, 2004).

Items #7 and #8 on this list are of particular interest to any study of NVC in virtual world, in that they deal with the design and usage of avatars, gestures, and environments. Of NVC in virtual worlds, Cheng et al. write:

“We found that people did use the 3D space for non-verbal communication. They used their ability to position in the 3D environment to stand near and to look at the person with whom they were talking. Furthermore, because they were able to communicate the direction of their attention non-verbally, they were less likely to address their chat messages with user names than if they communicated only through text chat.” (Cheng et al., 2002)

Following their analysis and observations, they concluded that the use of NVC allowed people to express emotions, and communicate interest and attention, but that it often interfered with verbal, text communication.

In their longitudinal ethnography of *There.com*, Brown and Bell identified interactions with objects within the environment as a crucial component of collaborative social actions; the presence of interactive objects within the space provided a structure for social collaboration (B. Brown & Bell, 2004). More recently, in 2006 Williams et al. wrote about the social life of guilds in *World of Warcraft* (Williams et al., 2006). Using a combination of interviews and a “social network census” administered via a survey they identified a number of social norms in MMORPGs. Their analysis identifies the game mechanics of *World of Warcraft* as the “key moderator” of these social outcomes by encouraging certain types of interactions and discouraging others (Williams et al., 2006). Taken alongside the writing of Cheng et al. we find a persuasive argument for carefully integrating an analysis of the technological affordances and design parameters of a given virtual world into any model of NVC for virtual worlds. To put this in context with previous work in NVC, one of the origins of nonverbal behaviors identified by Ekman and Friesen is the physical experience of embodiment within the world (Ekman & Friesen, 1981). In the same way that our bodies shape the ways in which we can communicate non-verbally, the avatar bodies, interface mechanisms, and procedural systems (such as game rules, and the underlying logic of the simulated environment) designed into virtual worlds greatly constrain, shape, and afford particular types of nonverbal behavior that are idiosyncratic to each world.

EMPIRICAL STUDIES OF PROXEMICS IN VIRTUAL WORLDS

Yee and Bailenson (and collaborators) extended their individual earlier work into *Second Life* to study the import of NVC within an active virtual world. In their 2007 article *The Unbearable Likeness of Being Digital: The Persistence of Nonverbal Social Norms in Online Virtual Environments* they developed a script that allowed them to measure the proximal distances and gaze orientations of the 16 avatars closest to the researcher in a virtual 200 meter radius, and to track whether or not the observed avatars were talking to each-other or not (Yee, Bailenson, Urbanek, Chang, & Merget, 2007). Their measurements of interpersonal distance (IPD) and mutual gaze allowed them to calculate whether or not equilibrium theory and proxemics could be used to account for player behavior in *Second Life*. They concluded that “our social interactions in online virtual environments, such as *Second Life* are governed by the same social norms as social interactions in the physical world” (Yee et al., 2007).

In a very similar study, Friedman et al. developed a set of “social bots” that traversed *Second Life* seeking out other avatars and engaging in pre-programmed social behavior (Friedman, Steed, & Slater, 2007). These automated research instruments measured the proximal responses of other avatars in two different experiments. In the first experiment, rather than measuring group interactions, like (Yee et al., 2007), they focused on avatars in dyadic interactions that were isolated from other groups of players. Their data allowed them to claim that “users tend to keep their avatars in non-arbitrary proximity from the other avatars they

are interacting with”, however they challenged Yee et al’s claim that proxemic spacing behavior translated directly from the physical world to the virtual one. In their second experiment, they programmed their bots to initiate an interaction with a user in *Second Life*, and then move into an uncomfortable close interaction distance of 1.2 meters. The bot then observed the response of the user for 10 seconds and reported the results. Many of the subjects approached in the way retreated from the bot, which allowed them to determine that interactors in *Second Life* have an awareness of personal space (Friedman et al., 2007).

Some of the most recent work in NVC for virtual worlds has attempted to use proxemics to guide the behaviors of autonomous agents and NPCs in simulated worlds. Laga and Amaoka described an agent control system that incorporated the rules of interpersonal distance and gaze to manage the proxemic spacing of groups of virtual agents (Laga & Amaoka, 2009). Their system claims to be able to achieve more natural movement within groups that obey the rules of personal space and gaze described in proxemics.

3. CHAPTER REVIEW

There is still much work to be done in the space of nonverbal communication in virtual worlds. In this chapter we have reviewed some of the most significant research within the field, across a number of disciplines.

One thing that unifies most of the observational and experimental work in virtual worlds is an emphasis on proxemics and gaze rather than kinesics and gesture. This is unsurprising, since positioning avatars in these systems is a fundamental and transparent interaction: one which requires little additional UI affordances beyond the entry level interactions. Contemporary virtual worlds more naturally afford unconscious proxemic behaviors, whereas the literacies required for more sophisticated kinesic communication are often much more advanced. While a case might be made for the ease of gathering data about IPD and Gaze in virtual worlds, as opposed to gathering data about the use of kinesic communications, we see many opportunities to broaden studies in virtual worlds to include analysis of appearance, gesture, posture, and even facial expression. We regard the lack of this work as a significant gap in current research around NVC in virtual worlds, and an important direction for future research.

4. BIBLIOGRAPHY

- Argyle, M. (1975). *Bodily Communication* (1st ed.). London: Methuen & Co. Ltd.
- Bailenson, J. N., Blascovich, J., Beall, A. C., & Loomis, J. M. (2001). Equilibrium Theory Revisited: Mutual Gaze and Personal Space in Virtual Environments. *Presence*, 10(6), 583-598.
- Bailenson, J. N., Blascovich, J., Beall, A. C., & Loomis, J. M. (2003). Interpersonal Distance in Immersive Virtual Environments. *Personality and Social Psychology Bulletin*, 29, 819 - 833.
- Bartle, R. A. (1996). Hearts, Clubs, Diamonds, Spades: Players Who Suit MUDs. *The Journal of Virtual Environments*, 1(1).
- Bartle, R. A. (2004). *Designing Virtual Worlds*: New Riders Publishing.
- Becker, B., & Mark, G. (1998). *Social Conventions in Collaborative Virtual Environments*. Paper presented at the Collaborative Virtual Environments 1998 (CVE'98), Manchester, UK.
- Benford, S., Bowers, J., Fahl'en, L. E., Greenhalgh, C., & Snowdon, D. (1997). Embodiments, avatars, clones and agents for multi-user, multi-sensory virtual worlds. *Multimedia Systems*, 5, 93 - 104.
- Bowers, J., Pycock, J., & O'Brien, J. (1996). *Talk and Embodiment in Collaborative Virtual Environments*. Paper presented at the CHI'96, Vancouver, BC.

- Brown, B., & Bell, M. (2004, November 6 - 10). CSCW at *Play: 'There' as a Collaborative Virtual Environment*. Paper presented at the CSCW'04, Chicago, Illinois, USA.
- Brown, J. R., van Dam, A., Earnshaw, R., Encarnacao, J., Guedj, R., Preece, J., . . . Vince, J. (1999). Human-Centered Computing, Online Communities, and Virtual Environments *IEEE Computer Graphics and Applications*.
- Bruckman, A. (1993). Gender Swapping on the Internet. In P. Ludlow (Ed.), *High Noon on the Electronic Frontier: Conceptual Issues in Cyberspace* (pp. 317 - 325). Boston, MA: Massachusetts Institute of Technology.
- Cheng, L., Farnham, S., & Stone, L. (2002). Lessons Learned: Building and Deploying Shared Virtual Environments. In R. Schroeder (Ed.), *The Social Life of Avatars: Presence and Interaction in Shared Virtual Environments* (pp. 90 - 111). London: Springer-Verlag.
- Corraze, J. (1980). *Les communications non-verbales*. Paris: Presses Universitaires de France.
- Dibbell, J. (1998). *My Tiny Life: Crime and Passion in a Virtual World*. New York: Henry Hold and Company, Inc.
- Ekman, P., & Friesen, W. V. (1981). The Repertoire of Nonverbal Behavior: Categories, Origins, Usage, and Coding. In T. A. Sebeok, J. Umiker-Sebeok & A. Kendon (Eds.), *Nonverbal Communication, Interaction, and Gesture*. The Hague: Mouton Publishers.
- Friedman, D., Steed, A., & Slater, M. (2007). Spatial Social Behavior in *Second Life*. *Lecture Notes in Artificial Intelligence*, 4722, 252-263.
- Guye-Vuillieme, A., Capin, T. K., Pandzic, I. S., Magnenat, N., & Thalmann, D. (1999). Nonverbal communication interface for collaborative virtual environments. *Virtual Reality Journal*, 4, 49-59.
- Isbister, K., & Nass, C. (2000). Consistency of personality in interactive characters: verbal cues, non-verbal cues, and user characteristics. *International Journal of Human-Computer Studies*, 53, 251 - 267.
- Laga, H., & Amaoka, T. (2009). *Modeling the Spatial Behavior of Virtual Agents in Groups for Non-verbal Communication in Virtual Worlds*. Paper presented at the 3rd International Universal Communication Symposium, Tokyo, Japan.
- Montero, B. (2006). Proprioceiving Someone Else's Movement. *Philosophical Explorations*, 9(2), 149 - 161.
- Morie, J. F., & Verhulsdonck, G. (2008). *Body/Personal/Action! Emerging Non-anthropomorphic Communication and Interaction in Virtual Worlds*. Paper presented at the Advances in Computer Entertainment Technology, Yokohama, Japan.
- Perlin, K., & Goldberg, A. (1996). *Improv: a system for scripting interactive actors in virtual worlds*. Paper presented at the 23rd annual conference on Computer Graphics and Interactive Techniques.
- Rohrer, T. (2007). The Body in Space: Dimensions of Embodiment. *Body, Language and Mind*, 2.
- Schilbach, L., Wohlschlaeger, A. M., Kraemer, N. C., Newen, A., Shah, N. J., Fink, G. R., & Vogeley, K. (2006). Being with virtual others: Neural correlates of social interaction. *Neuropsychologia*, 44(5), 718-730.
- Schroeder, R. (2002). Social Interaction in Virtual Environments: Key Issues, Common Themes, and a Framework for Research. In R. Schroeder (Ed.), *The Social Life of Avatars: Presence and Interaction in Shared Virtual Environments* (pp. 1 - 18). London: Springer-Verlag.
- Taylor, T. L. (2002). Living Digitally: Embodiment in Virtual Worlds. In R. Schroeder (Ed.), *The Social Life of Avatars: Presence and Interaction in Shared Virtual Environments* (pp. 40 - 62). London: Springer-Verlag.
- Turkle, S. (1995). *Life on the screen: Identity in the age of the internet*. New York, New York: Simon and Schuster.
- van Es, I., Heylen, D., van Dijk, B., & Nijholt, A. (2002). *Gaze Behavior of Talking Faces Makes a Difference*. Paper presented at the CHI 2002, Minneapolis, USA.
- Walther, J. B. (1996). Computer-mediated communication: Impersonal, Interpersonal, and Hyperpersonal Interaction. *Communication Research*, 23, 3 - 43.
- Walther, J. B., Loh, T., & Granka, L. (2005). Let Me Count The Ways: The interchange of Verbal and Nonverbal Cues in Computer-Mediated and Face-to-Face Affinity. *Journal of Language and Social Psychology*, 24(1), 36-65.

- Williams, D., Ducheneaut, N., Xiong, L., Zhang, Y., Yee, N., & Nickell, E. (2006). From Tree House to Barracks: The Social Life of Guilds in World of Warcraft. *Games and Culture*, 1(4), 338 - 361.
- Yee, N. (2006). The Demographics, Motivations, and Derived Experiences of Users of Massively Multi-User Online Graphical Environments. *Presence*, 15(3), 309-329.
- Yee, N., Bailenson, J. N., Urbanek, M., Chang, F., & 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.

SECTION II

*IDENTITY AND COMMUNICATION
IN VIRTUAL WORLDS*

5

OUR EMPATHIC EXPERIENCE OF BELIEVABLE CHARACTERS

By Leslie Bishko

***Editor's Note:** This chapter is the first part of a larger piece on Empathy and Laban Movement Analysis for animated characters. We have broken it into three smaller chapters throughout the book, but it may be read as part of a single work. The other two parts can be found in Chapters 11 and 13. These chapters are supplemented by series of video figures, which can be found online at:*

<http://www.etc.cmu.edu/etcpres/NVCVideos/>

1.

INTRODUCTION

Historically, the concept of *empathy* has been linked with our perception of other beings as minded creatures (Stueber, 2008). Empathy is rooted in a process of embodiment, whereby we experience others by attuning our movement to theirs. In the context of virtual worlds, empathy is central to the experience of immersion in an environment as well as the process of interacting and communicating with others present in the environment. Therefore, the qualities of character movement in virtual worlds are essential to the experience of immersive being in a world.

While qualities of character movement in virtual worlds continue to advance, core believability issues still exist. The bar is continuously raised by the standards of feature-film production, which strives towards a realist aesthetic due to the proliferation of visual effects that integrates digital

LESLIE BISHKO ON HER METHODS

As an artist and animator, it is my practice to cultivate my empathic sense of movement and what it communicates. It is my art form and the aesthetic/perceptual lens of my daily life: observing and internalizing the nuances of movement.

I am not a game player – my voice is an outsider's voice, in that sense, an observer reporting on what I see, and somewhat oblivious to why things are the way they are in games.

My perspective as an observer comes from the animation process, which is the process of creating movement through my felt sense of relationship and communication in the real world. An animator creates with movement the way a painter creates with color.

Laban Movement Analysis is the glue between the continuity of my real world movement experience and the discontinuous process of creating animation frame by frame.

(continued on next page)

character animation with live action performances. In virtual worlds, the trend towards realism is complicated by the mechanisms of user-controlled character behaviors. The user has the innate expectation that the characters are an extension of her body; that they can respond to her intention, and therefore move intuitively, according to her thoughts. When this expectation is not met in a satisfactory way, the quality of a user's level of engagement in a virtual world is diminished.

This chapter explores character believability issues in depth by focusing on empathy, embodiment, and meaningful movement. Drawing on classical animation principles (Johnston & Thomas, 1981), acting theory (Hooks, 2003), and Laban Movement Analysis (Bishko, 2007; Hackney, 1998; Laban, 2011; Moore, 1988), believability issues are defined and discussed in terms of authenticity, the interplay of function (biomechanics) and expression in movement, and movement characteristics that we perceive as intentional versus behavioral.

As human beings, the development of our ability to express ourselves occurs in a reciprocal relationship with the development of functional body mechanics. The intent to reach for an object stimulates an infant's motor patterning, which evolves into crawling and walking. At the same time, our bodies are the vehicle of our expression; the motor patterning manifested by distinct physiologies makes our expression uniquely our own.

In our physical world, we cultivate these functional and expressive capabilities in relationship to gravity. Yet, in virtual worlds, we construct gravity by how we animate, capture, generate or calculate qualities of movement. Hence, the absence of gravity in virtual worlds creates the core issue of believability. Its lack lies at the root of functional movement problems that affect the appearance of neurological coordination in characters. Issues such as weight shifts, sliding feet, rapid change of orientation, physical intersections, lack of integrated body connectivity, and lack of complete "movement phrases" are common. With poor functional capability, expression is compromised. Classical animation principles address these issues, but they remain to be solved when it comes to features of real-time interaction. Where procedural solutions

I learn about patterns in communication through the constant process of observation and description. I look for my own empathic response and describe it metaphorically.

How I know this is valuable:
When people begin learning LMA, they are instantly drawn to it because it is truthful to how they express themselves in their own bodies. LMA provides access to intrinsic knowledge they already have, but never had a language for. Learning the language of movement makes movement concrete, less ephemeral, and opens our minds to its symbolism.

The methodology of Laban Movement Analysis is essentially to observe movement, describe it in terms of LMA parameters, discern patterns and salient features and make meaning from our observations. This chapter emphasizes an introduction of the terminology of LMA while I employ the methodology of observation, description and interpretation throughout, by example.

LEARNING LABAN METHOD ANALYSIS

Reading this text will provide a general, working knowledge of LMA. To really know the system, it is recommended that you embody it through movement practice. LMA is taught in dance and theater programs, within universities as introductory courses, and through several certifying bodies in an intensive format that is considered equivalent to graduate level study. See the Appendix at the end of this chapter for a list of resources for studying LMA.

exist, there are additional factors that have us perceive movement as behavioral, rather than as movement we engage with empathically through its representation of intent.

A key distinction regarding expressive movement is the difference between intentional and behavioral movement. Intentional movement communicates a character's emotions and engages our empathy, whereas behavioral movement reflects reactive or survival reflexes that project motivated action, yet feel empty or automatic. An example of this is how a character's gaze and kinetic sequencing of body movement can be coded to procedurally follow a moving target. The element of this that can't be coded is what the character thinks while tracking a target, or how he feels about his circumstances, or even why he is tracking to begin with.

Artificial intelligence solutions aim to have the potential to successfully integrate function and expression. Complex rules and conditions are embedded into run-time scenarios as a construction of character intention, which is further mediated by the end user. Barriers to believability from this perspective lie collectively in the design of the interactive experience, the effectiveness of how the AI interfaces with the movement assets, and in the qualities of the movement assets themselves.

This chapter explores these issues through animation and acting principles, and the methodology of Laban Movement Analysis (LMA), an established framework for observing, describing, and interpreting movement and what it communicates. LMA describes movement parameters through five general categories:

1. Body: posture, gesture, and patterns of coordination
2. Effort: intention manifested as qualities of movement
3. Shape: the changing form of the body in relationship to others
4. Space: the structural language of how movement forms in space
5. Phrasing: units of movement language.

LMA, as a theory of movement, offers practical and innovative solutions towards designing authentic, intentional, believable, expressive and characterized movement in virtual worlds that is perceived as meaningful through the process of empathic attunement. By using the movement framework that LMA provides, we can create character movement with deliberate attention to what makes movement authentic, intentional and believable.

This chapter will show you how and why LMA solves the issue of believability. It is the first publication of the full LMA system that is targeted for the art and industry of digital animation. My purpose is to describe to you the LMA system in a way you can apply it.

Examples of character movement from keyframe animation, interactive games, and real-time virtual worlds are discussed in LMA terms, both to introduce LMA concepts and illustrate believability factors from an LMA perspective. LMA provides a depth of understanding about movement that can be effectively employed to address the believability issues articulated in this chapter. The concepts introduced offer a framework for designing functional and expressive motion, through all aspects of production, which collectively feed into the user experience of the character's presence within a virtual world.

Movement is the vehicle of communication and interaction in virtual worlds, yet it is an underdeveloped area when we consider the levels of sophistication that are possible. This chapter proposes that LMA provides a valuable conceptual framework for furthering the effectiveness of movement communication. At the same time, it articulates what intentional movement looks like, and how we are empathically engaged by it. LMA's strength and value towards this goal is that it is an open framework for observing communication, as opposed to a closed formula that limits the interpretation of movement to literal definitions. The implication of creating virtual-world experiences through an LMA perspective is a groundbreaking approach that will make truly engaged experiences possible.

2. EMPATHY, AUTHENTICITY, AND BELIEVABILITY

In this section, my purpose is to explore the role of movement and how it creates empathic engagement. We experience this engagement with characters in virtual and interactive environments in three main ways:

- as player characters, through which we experience a videogame as a predefined character;
- as observers of characters, whom we watch and/or interact with; and
- as avatars, through which we form our own virtual-world identity and persona.

As player characters or avatars, the perception of agency (or one's own intent), in controlling character movement is not the same as observational perceptions of character movement. This section briefly discusses empathy relating to controller input, while focusing mainly on observed qualities of character movement.

EMPATHY

The word “empathy” is a translation from the German *Einfühlung*, meaning “to feel into” (“empathy,” 2011). Defined both as feeling what others feel, and projecting one's own feelings onto objects or other beings, empathy is a concept that impacts a variety of theoretical contexts, such as aesthetics, psychology, philosophy, and neuroscience (Stueber, 2008).

In neuroscience, the recent discovery of mirror neurons identifies empathy as the physiological mechanism for how we perceive others, learn through imitation, develop language, and communicate (Rizzolatti & Craighero, 2004). Mirror neurons in the pre-motor cortex of the brain fire when we take action and when we watch another's actions. Furthermore, the mirror neurons are coded to respond not just to action, but also to intentions (Morrison, 2004). In other words, observing another's action engenders a physiological response as if one had the intention and performed the action him or herself. It is also the mechanism through which we understand what we see; as we observe, we experience it ourselves, and comprehend it through our own experience.

The neuroscience model involves the perception of movement. It leads us directly to the notion that how we move as human beings, and consequently, how characters move in virtual worlds, creates engagement and communication through the mechanism of empathy. In her book *Making Connections: Total Body Integration Through Bartenieff Fundamentals*, Peggy Hackney leads readers to experience how we perceive others and differentiate ourselves from them by empathically attuning through the sensation of breath. For example, when people relate to each other intimately in a close hug, or in a relaxed state of being physically close, the rhythm of breath will naturally synchronize, or attune. It is an intimate form of connecting and relating—a state of being present to another through presence of the self (Hackney, 1998, pp. 58-60).

In *Beyond Words: Movement Observation and Analysis*, Carol-Lynne Moore describes empathic attunement through movement observation:

The easiest way to experience kinesthetic empathy is to attend to how you use your own body while watching an exciting sports event or a tense mystery program. Most involved fans find themselves muscularly participating in the event, that is, making motions like those being observed, only smaller and more subtle. These participative movements of kinesthetic empathy, drawing on imitation and movement memory, can be a valuable extension of visual perception in the understanding of human movement (Moore, 1988).

These descriptions of empathic experience point out that we communicate and perceive the meaning of others' communication through our own movement

EMPATHY AND CHARACTERIZATION

The mechanism of empathy also creates the condition in which we perceive and identify with characters. Ed Hooks, author of *Acting for Animators*, stresses that we experience characters through their emotions, not through their thoughts. Our perception of their emotions is felt empathically via how they move. “The goal of the animator is to expose emotion through the illusion of movement on screen” (Hooks, 2003). How characters move in virtual worlds has everything to do with how we perceive them as a) believable characters, which appear to have body mechanics that approximate laws of physics that are congruent with their anatomy and their environment, and b) authentic characters, whose movement style forms a specific characterization that we experience as expressive and unique. Authentic characters move in a way that is congruent with who they are, while immersed in circumstances within a narrative scene or interactive environment.

FUNCTION, EXPRESSION, AND BELIEVABILITY

The terms “function” and “expression” are important themes within the LMA system. LMA views movement as an integration of functional and expressive elements. Function can be thought of as the biomechanics of the body, or our movement potential as viewed through the laws of physics. Expression represents how our intention communicates itself through movement. Neurologically, our intent to attain objectives and express ourselves organizes the functional motor patterning of the body, while our functional abilities support and characterize how we express ourselves through movement (Bartenieff, 1980; Hackney, 1998).

Functional believability refers to character movement that has a high degree of biomechanical accuracy. Elements of believable motion include the following:

- The range of joint motion falls within normal ranges for human performance.
- Weight shifts, dynamic alignment, and distribution of weight in motion create a believable illusion of gravity. (Hooghwinkel, 2012)
- Timing of rotations of individual joints, and subtlety of degrees of rotation, adequately mimic the refinement of human motion.

Expressive believability is about creating readable intent through how the character acts; this involves designing believable gestures and facial animation. But more importantly, it means being clear about the character’s intent, and portraying that intent consistently throughout the animation. Expressive believability gets into the realm of movement stylization, caricature, and stereotype and partially depends on story, script, and the visual design of the character.

Character believability depends on the empathic perception of both functional and expressive believability. We are integrated and whole in our psycho/physical perceptual processes. Thus, we experience animation empathically, by default (Hooghwinkel, 2012). It is the qualities of the empathic engagement that we

Combining function and expression in character movement is the heart of the creative process in animation storytelling. The collective creative process of the entire production team contributes to consistently believable characters. The director relies on the teams’ creative contributions, yet functions as the integrating force that brings story, characters, and action together. For example, in the 2008 film *Kung Fu Panda*, five animal characters embody the fighting characteristics on which the tradition of Shaolin Kung Fu is based. The animals each move and fight in a stylized way that is designed to reflect the physical and instinctual characteristics of their species. This element of the animal characters also provides physical comedy, as in the scene where the smallest, and presumably the weakest, a praying mantis, is left to hold up a rope bridge that the others are fighting on.

respond to in our evaluation of the experience. Through empathy, we “know” that the movement we perceive doesn’t “feel right.” Therefore, in observing virtual characters, we have a hard-wired perceptual mechanism for detecting functional and expressive flaws in the movement.

INTENTION, BEHAVIOR, AND AUTHENTICITY

Where function and expression are believable in character movement, we can study the next layer: authenticity. An authentic character performance is one in which a character’s movement expression is congruent with the character as a whole: it reveals how a particular individual copes with circumstances within the narrative or interactive environment. When presented with well-crafted, authentic character performances, not only do we believe in the illusion of a virtual being, we believe there is an actual, specific being. We willingly suspend disbelief and forget that we are observing an animated illusion. In mainstream film or video game entertainment, creators of animated characters generally strive for this level of believability and interpretation.

Intentional movement contains the elements of function and expression, taking them to the level where character movement appears authentically motivated. Every action reflects the character’s thought, emotion, and choice of action in the moment. There is a congruence of movement style with the gestalt of the character.

Perhaps movement that is not intentional best exemplifies this. For me, the exaggerated animation style in the 2005 film *Madagascar* simply felt like too much movement. Each action was played to the fullest, having the impact of an exclamation mark. The characters felt hyperactive. I found it exhausting to watch because I was attuned empathically to function/expression, but could not relate to a felt sense of the characters’ motivation. There was functional and expressive believability, however, the degree of exaggeration created a lack of clear intention. The flaw was in the characterization; the movement didn’t match the characters. About twenty minutes into the film, the story and characters had developed, and I had become accustomed to the movement style, all of which made the rest of the film more watchable.

Related to this example, but distinctly different, is the idea of behavioral movement. I once visited an aquarium and became mesmerized by a seal swimming in great loops through her tank. Her speed was fast by my standards; yet it seemed casual and easily paced for an active seal enjoying the dynamics of her own energy in motion. After studying her a while, I gradually sensed that this looping pattern was behavior without intention; that no impulses were satisfied by it. What had first impressed me as playful, patterned motion now occurred to me as automatic, and listless, as if this was all she could do with herself in this confined space.

The concept of behavioral movement in procedural animation became clear to me in 2007, at a presentation given by Electronic Arts about their procedural character animation research and tool development (Armstrong, 2007). Through procedurally blended motion capture, a character interactively followed a target with his gaze. As the target approached the limits of visual range, the character’s position and orientation would seamlessly adjust. A sophisticated range of adjustments was demonstrated as the target challenged the character through a full set of spatial possibilities. A commendable technical achievement, the movement was beautiful to watch and seamlessly believable. However, as I attuned to the movement, I gradually became uncomfortable. At first, the character appeared to have intent: to follow and adjust to the target. But in time, I realized that I was watching procedural behavior, not characterization of intent. While the movement was stunningly, humanly realistic, the lack of character intent gave the impression that the character was a robot. Ed Hooks, author of *Acting for Animators* (2003), writes, “An action without a thought is impossible, and an action without an objective is just a mechanical thing, moving body parts.” (Hooks, 2003, p. 5)

Software engineer, dancer and Laban Movement Analyst Sandra Hooghwinkel observes that rule-based movement can have the appearance of intention by virtue of acting within a rule-space context (Hooghwinkel, 2012). For example, a character programmed to pursue proximity to other characters will appear to have the intent of physical closeness. What distinguishes it from intentional movement is whether it feels authentic. If nothing has motivated the character to pursue closeness, and if the goal of closeness is not ever satisfied by having achieved it, the movement lacks authenticity and appears as behavioral.

The following quote underscores this distinction between behavior and intention. In discussing the biological evidence of empathy found within mirror neurons, Morrison describes what triggers these neurons to fire:

Of particular importance is that these neurons fire only during the initiation or observation of actions, not just movement. For example, they fire in response to observing another monkey grasp an object, but not to observing simple opening and closing of the hand. They also do not fire to grasping by non-living objects such as a mechanical hand (Rizzolatti et al., 2001). This suggests they code for a particular relationship in motor terms between subject and object. In other words, they are coding for intention, and not just for movement. (Gallese & Goldman, 1998, in (Morrison, 2004))

Examples of intentional and behavioral animation crop up everywhere in films and interactive games. In keyframe animation, lack of phrasing, or functional movement problems can look unskilled and not believable. Where animation is working well functionally, issues of characterization (such as the over-exaggerated style in *Madagascar*) can crop up in ways that seem more behavioral.

In the case of interactive games and real-time puppeteering, so much of the action is navigating the environment. We see the same walk and run cycles without variation, and because this is most of what the character gets to do, their movement barely attracts our attention. Idle animation sequences are another example of behavioral movement. The repetition makes characters appear nervous or agitated, while actions such as looking around seem arbitrary and without stimulus [Video Figures 1–4]. The economies of game production necessitate the repeated use of cycles and movement clips. The repetition, as with the confined seal, occurs to us as behaviors without intent.⁸

Intentional movement is the mechanism for our empathic experience of a character. It is the vehicle through which we believe that the character is a conscious being, and perceive what the character is communicating. We can consider that intentional movement appears authentic, whereas behavioral movement generally seems more inauthentic. Both types of movement certainly have their place, so distinguishing them is valuable for this discussion. They are both believable on some level.

CHARACTERISTICS OF INTENTIONAL MOVEMENT:

- There is a clear objective. The movement communicates an idea, or information. We know what motivates the character.
- The movement consists of complete phrases, including preparation > action > recuperation > transition. Phrases are units of communication in movement, rhythmically organized by the breath. The Breath patterning within a movement phrase signals aliveness and emotion. (Ventrella, 2011) We engage empathically through the rhythm of breath.

⁸Ubisoft is circumventing this with procedural animation layered on top of motion clips. A tiny adjustment of the controller means the character tilts his or her head slightly, pushes a foot out slightly further, extends an arm to counterbalance, etc. See Jay White's sidebar on Procedural Pose Modification in Chapter 13, section 1.3.

- There is a direct and reciprocal relationship to environmental factors (including other characters). The mover is responding to environmental stimuli, and the mover has an impact in the environment.

CHARACTERISTICS OF BEHAVIORAL MOVEMENT:

- Movement that has no clear purpose or motivation
- Movement that appears repetitive, mechanical or automatic
- Repeated actions or repeated responses to environmental stimuli, e.g. a character in a dark space shields his eyes from a light source, and does so repeatedly. In real life, our eyes gradually adjust to the darkness, so shielding the eyes from light would diminish over time.
- Is mostly reactive to environmental stimuli, as opposed to taking action that has an impact in the environment
- Movement based on survival instincts such as the following:
 - maintaining verticality
 - self protection: protecting vulnerable body parts
 - avoiding contact
 - avoiding impact
 - avoiding or pursuing proximity (Ventrella, 2011)
- Movement created via procedural, physics-based and cyclical animation methods are generally behavioral. These methods for creating motion excel at supporting behavioral action because they are based on conditions and rules.

Story context provides a general methodology for choosing when characters need to communicate intention or behavior. Context brings focus and clarity to what needs to be communicated. The established practice in production is to foreground lead characters that have dialogue with intentional movement, while assigning behavioral movement using procedural, physics-based, and cycled animation methods to background characters. Story, character-based, or general aesthetic guidelines for procedural behaviors can prevent the behavior from attracting the viewer's attention to its lack of intent.

The challenges of overcoming barriers to intentional, authentic movement are unique to animated virtual characters. Live actor performances benefit from being inherently embodied; live actors can access qualities of authentic characterization directly through their body. In contrast, virtual character performances are constructions of bodies in motion. Intent is an illusion created through character movement. This is why authenticity is a delicate pursuit. The empathic experience of authenticity carries a sense of truthfulness. It has immediacy, presence, and innate comprehension.

3. EMPATHY, INTERACTION AND METAPHOR

My discussion in this chapter focuses on observed character movement, but interaction can't be ignored. This section briefly looks at the projection of intention in interaction. It concludes with an example of an embodied interface, which allows me to introduce the important concept of mapping and metaphor.

In videogame play, our emotions are elicited through empathic attunement with the player character, through fulfilling the character's objectives, and through the immersive qualities of the environment. Using a game controller to control a character's movement engages a kinesthetic sense that one's own intention is creating the intentional action of a character in a virtual world. Situated in the context of the game, we project our intent, not our own personality or personal objectives, onto the character.

The interface to a virtual world, both in the context of driving character activity and immersion in the world, is a metaphorical mapping of movement information to a control device via a conceptual framework. With her immersive virtual-reality piece, *Osmose* (1995), artist Char Davies created a profound example of an embodied interface mapped to an embodied conceptual framework. She uses the metaphor of scuba diving to map breath control to virtual-world embodiment:

In contrast to manually based interface techniques such as joysticks and trackballs, *Osmose* incorporates the intuitive processes of breathing and balance as the primary means of navigating within the virtual world. By breathing in, the immersant is able to float upward, by breathing out, to fall, and by subtly altering the body's centre of balance, to change direction, a method inspired by the scuba diving practice of buoyancy control.

Whereas in conventional VR, the body is often reduced to little more than a probing hand and roving eye, immersion in *Osmose* depends on the body's most essential living act, that of breath—not only to navigate, but more importantly—to attain a particular state-of-being within the virtual world. (Davies, 2012)

Davies reports that people who experience *Osmose* have had profound personal, even transformational experiences. By connecting people with their breath, and allowing breath to become the means of agency, Davies has created an interface that connects inner feelings to the outer environment.

What kinds of experiences are possible through kinesthetic mappings for motion-based control devices? They provide direct physical experience; therefore, empathy operates differently. Would something like a full-body puppeteering interface deliver something as profound as *Osmose*? I believe that *Osmose* succeeds due to metaphor. *Osmose* creates the possibility of meaning through providing an embodied interface to a virtual world. The immersant finds their own meaning in the experience.

LMA articulates the elements of embodied kinesthetic experience. In theory, when integrated with an intuitive control device, it can offer both direct and metaphoric mappings of control device motion to motion in virtual worlds. LMA movement parameters remain true under varying contexts. As a result, LMA solves the problem of literal mappings. This aspect of the system must be understood in order to embark on procedural representations of LMA.

There is . . . no simple one-to-one correspondence between a movement and what it signifies, or between a meaning and the movement and what it signifies, or between a meaning and the movement chosen to encode it (Moore, 1988).

4. SUMMARY

We need believability and authenticity to create fully embodied character performances that engage the viewer/player and communicate meaningfully. The more we attend to character movement, the more mastery we can achieve over the nuances of empathic experience. Elements of narrative and environmental circumstances will always engage us conceptually. For a believable experience, we need to focus on empathy and physical attunement with characters.

In Chapter 11 I will undertake a discussion of the Principles of Animation, and Laban Movement Analysis. We can apply these movement frameworks to create believable, authentic character movement. In Chapter 13 I will describe observations of character movement and discuss qualities of empathic engagement.

5. BIBLIOGRAPHY

- Armstrong, E. (2007, August 9). Improving Real-Time Motion. *Spor(T)*. San Diego, California, USA: ACM SIGGRAPH Conference.
- Bartenieff, I. a. (1980). *Body Movement: Coping with the Environment*. New York: Gordon and Breach Science Publishers.
- Bishko, L. (2007, December 9). *The Uses and Abuses of Cartoon Style in Animation*. Retrieved February 3, 2011, from Animation Studies Online Journal: <http://journal.animationstudies.org/leslie-bishko-the-uses-and-abuses-of-cartoon-style-in-animation/>
- Davies, C. (2012). Osmose. Retrieved May 5, 2011, from Immersence: <http://www.immersence.com/empathy>. (2011, February 3). Retrieved February 3, 2011, from Collins English Dictionary - Complete & Unabridged 10th Edition: <http://dictionary.reference.com/browse/empathy>
- Hackney, P. (1998). *Making Connections: total body integration through Bartenieff Fundamentals*. Amsterdam: Gordon and Breach.
- Hooghwinkel, S. (2012, December 31). Personal communication. Vancouver, British Columbia, Canada.
- Hooks, E. (2003). *Acting for Animators: A Complete Guide to Performance Animation* (Revised Edition ed.). Heinemann Drama.
- Johnston, O., & Thomas, F. (1981). *Disney Animation; The Illusion of Life* (1st Edition ed.). New York: Abbeville.
- Laban, R. (2011). *The Mastery of Movement* (Edition 4 ed.). (L. Ullmann, Ed.) Hampshire: Dance Books.
- Moore, C.-L. (1988). *Beyond Words: Movement Observation and Analysis*. London: Gordon and Breach.
- Morrison, J. (2004). Understanding Others by Understanding the Self: Neurobiological Models of Empathy and their Relevance to Personality Disorders. *The Canadian Child and Adolescent Psychiatry Review*, 13 (3), 68-73.
- Rizzolatti, G., & Craighero, L. (2004). The Mirror-Neuron System. *Annual Review of Neuroscience*, 27, 169-192.
- Stueber, K. (2008, August 29). Empathy. (E. N. Zalta, Editor) Retrieved May 15, 2011, from The Stanford Encyclopedia of Philosophy: <http://plato.stanford.edu/archives/fall2008/entries/empathy/>
- Ventrella, J. (2011, January 20). Personal communication. Vancouver, British Columbia, Canada.

6. LABAN MOVEMENT ANALYSIS RESOURCES

CERTIFICATION PROGRAMS IN LABAN MOVEMENT ANALYSIS

EUROLAB (Germany)

<http://www.laban-eurolab.org/>

Integrated Movement Studies (Utah, California)

<http://www.imsmovement.com/>

Laban/Bartenieff Institute of Movement Studies

(New York, Maryland, Massachusetts, Tennessee, Koolskamp (Belgium), Toronto, Edinburgh)

<http://www.limsonline.org/>

Laban International

<http://www.labaninternational.org/>

Columbia College Chicago

The Department of Dance/Movement Therapy & Counseling

<http://www.colum.edu/Admissions/Graduate/programs/dance-movement-therapy-and-counseling/graduate-laban-certificate-in-movement-analysis.php>

http://www.colum.edu/Academics/DMTC/movement_pattern_analysis_consultant_certificate/

INTRODUCTORY WORKSHOPS: PREREQUISITES FOR LMA CERTIFICATION

United States

Integrated Movement Studies

<http://www.imsmovement.com/index.php/workshops-classes/>

Laban/Bartenieff Institute of Movement Studies

<http://www.limsonline.org/introductory-workshops>

Moving On Center

<http://www.movingoncenter.org/Workshops.htm>

Canada

Nadine Saxton

<http://www.nadinesaxton.com/workshops/labem-movement-analysis/>

Helen Walkley

<http://www.helenwalkley.com/teaching.html>

United Kingdom

Moving Forth

<http://movingforth.org/training/labam-movement-analysisbartenieff-fundamentals/>

Italy

Academia Dell'Arte

<http://www.adalife.it/2012/10/18/professional-laban-training-at-ada/>

Germany/Austria

EUROLAB

http://laban-eurolab.org/index.php?option=com_content&view=category&layout=blog&id=7&Itemid=35&lang=en

UNIVERSITY/COLLEGE PROGRAMS THAT OFFER LMA CLASSES

Source: <http://www.movementhasmeaning.com>

United States

Brigham Young University (Provo, Utah)

<http://cfacweb.byu.edu/departments/dance>

SUNY Brockport (Brockport, New York)

<http://www.brockport.edu/dance/>

SUNY Potsdam (Potsdam, New York)

<http://www.potsdam.edu/academics/AAS/Theatre/index.cfm>

University of Utah (Salt Lake City, Utah)

<http://www.dance.utah.edu/>

Utah Valley University (Orem, Utah)

<http://www.uvu.edu/dance/>

Ohio State University (Columbus, Ohio)

<http://dance.osu.edu/>

Hobart William Smith College (Geneva, New York)

<http://www.hws.edu/academics/dance/>

University of Virginia (Charlottesville, Virginia)
<http://www.virginia.edu/drama/danceminor.htm>

University of Maryland (College Park, Maryland)
<http://tdps.umd.edu/>

University of California, Irvine (Irvine, California)
<http://dance.arts.uci.edu/>

Glendale Community College (Glendale, Arizona)
<http://www.gccaz.edu/performingarts/12049.htm>

University of Wisconsin (Madison, Wisconsin)
<http://www.dance.wisc.edu/dance/default.asp>

Towson University
<http://www.towson.edu/theatre/>

Canada

Simon Fraser University (Vancouver, British Columbia)
<http://cgi.sfu.ca/~scahome/?q=dance>

York University (Toronto, Ontario)
<http://dance.finearts.yorku.ca>

United Kingdom

Trinity Laban Conservatoire of Music & Dance (London, England)
<http://www.trinitylaban.ac.uk>

The Laban Guild for Movement and Dance
 (non-university professional courses, workshops and ongoing community classes)
<http://www.labanguild.org.uk>

Queen Margaret University (Edinburgh, Scotland)
<http://www.qmu.ac.uk/mcpa/default.htm>

Italy

Accademia Dell'Arte
<http://www.adalife.it/2012/10/18/professional-laban-training-at-ada/>

RELATED PRACTICES IN DEVELOPMENTAL MOVEMENT PATTERNING & PSYCHOLOGY

Amazing Babies Moving
<http://www.amazingbabiesmoving.com>

Body Mind Centering
<http://www.bmc-nc.com>

Bobath Centre
<http://www.bobath.org.uk>

Kestenbergh Movement Profile
<http://www.kestenberghmovementprofile.org/index.html>

NOTATION

Dance Notation Bureau
<http://www.dancenotation.org>

International Council of Kinetography Laban
<http://www.ickl.org>

ONLINE DISCUSSION

CMAPlus Listserve

<https://listserv.cc.denison.edu/sympa/info/cmalist/>

GENERAL LABAN RESOURCES

Motus Humanus

<http://www.motushumanus.org>

Laban Analyses

<http://www.laban-analyses.org/>

Movement Has Meaning

<http://movementhasmeaning.com>

The LMA Effort Bank

<http://www.lmaeffortbank.com>

DNB Theory Bulletin Board

<http://dnbtheorybb.blogspot.ca>

The Laban Project

<http://www.labanproject.com>

LABAN PRACTITIONER DIRECTORIES

Laban/Bartenieff Institute for Movement Studies

<http://www.limsonline.org/services-certified-movement-analysts>

Movement Has Meaning

<http://movementhasmeaning.com/locate-a-practitioner/#all>

Laban British Columbia

<http://labanbc.wordpress.com/>

ACTING RESOURCES

Ed Hooks

<http://www.edhooks.com/>

Keith Johnstone

<http://www.keithjohnstone.com>

6

VIRTUAL GAZE: THE COMMUNICATIVE ENERGY BETWEEN AVATAR FACES

By Jeffrey Ventrella

1.

INTRODUCTION

Gazing (making eye contact, fixating on another person's eyes, and optionally aiming your face and body towards that person) is an important component of body language. Since *gaze* serves many well-established communicative purposes in society, it can be encoded into a succinct representation. This allows the equivalent of the physical act to be invoked remotely and virtually, without requiring a physical human head and eyes. I refer to this as *virtual gaze*.

Gaze is so powerful that even the *idea* of it, as uttered using commands in text-based virtual worlds, can create visceral responses. Michele White, who studies the ways in which technology renders and regulates users, points out that in text-based virtual worlds, the commands that allow a player to "watch", "scope", "peep", "gawk", etc. can be disturbing to the gazee, especially if she is being continually "looked at" by other players (White 2001). In animated virtual worlds, gaze becomes spatial, sensory, and dynamic, and so it has even more expressive and communicative effect. In this paper, let us use the term *virtual gaze* specifically in reference to human-like avatars in 3D virtual worlds that can rotate their eyeballs, heads, and/or bodies, so as to aim towards other avatars, as a form of nonverbal communication.

JEFFREY VENTRELLA ON HIS METHODS

Since I am more of a developer/designer than a researcher, I would describe my method as a kind of design research. What I am concerned with are the problems of how we (inventors, designers, and programmers) solve problems of engineering tools for creativity, and media for affective communication.

Here is an interesting phenomenon I have noticed about us developers of virtual worlds: many of us (not all) are indifferent to the livelihoods of the people who populate our worlds - we are more interested in building them than living in them. This has some positive, and some negative consequences. On the positive side, we can be more agnostic and objective about some things. On the negative side, we may be out of touch with what drives people to use our virtual worlds and to spend so much of their lives in them, what they want, and why.

(continued on next page)

The procedural animation technique of making a virtual character aim its body, face and/or eyes at a position in a scene, has been explored since the early days in the history of virtual human craft. Badler and others have researched eye contact and other behaviors for virtual agents (Lee, et al 2002) (Chopra-Khullar, et al. 2001). In many gaze systems, including *SmartBody* (Thiebaut, et al, 2009), gaze is manifested not only in the rotation of an avatar's head, but in the neck, and several joints of the spine. The various ways to look at a gaze target are then adjustable to a fine degree. "Gaze", in this case, need not be a term that applies to the eyes or head only. One can represent gaze with the whole body, the upper body, the head, the eyes, or any combination, to create many subtle variations of attention or annotation. The avatars in *There.com* employed multiple levels of gaze. In Figure 6-1, two *There.com* avatars are gazing at each other in a "chat prop" (a location in the world where special UI is made available for chatting, social signaling, and camera behaviors that respond to body language). These avatars are gazing with their whole bodies.

In some cases (in my case) being in-world can be uncomfortable – even painful, because the limitations and compromises that were made are difficult to ignore - they are bad memories. Co-founding a virtual world company means remembering what dreams had been dashed, and what features were canceled, often for the sake of achieving business goals, rather than implementing good design. In this case, studying people's uses and behaviors of a virtual world that I helped make would be problematic, as I cannot fully engage in the experiential world of the users, their ambitions, and their livelihoods. To me, the man behind the curtain is within full view, warts and all. For this reason, the papers I have written for this book, as well as the chapters of *Virtual Body Language*, where some of the material comes from, are told from the standpoint of a designer/engineer.



Figure 6-1: Avatar body/head/eye gaze in *There.com*

Lance and Marcella (2010) have researched the problem of gaze, attempting to map the gaze shifting of a virtual human to particular emotional states, in order to build better models for virtual gaze. Morency (2008) has built systems to detect user head rotation and eye tracking to provide inputs for engaging with embodied conversational agents (ECAs). Gratch, et al. (2006), Bailenson (2005), and others experiment with “virtual rapport” – the effects of positive interactions between humans and avatars with behaviors, such as mimicking, eye contact, nodding, etc.

These experiments shed light on how humans interact nonverbally. However, there are conceptual problems when considering these techniques for practical uses. Having avatars generate nonverbal cues that are not necessarily in line with the user’s original body language puts us in murky territory, especially with regards to authenticity. Steptoe et al. (2010) study how avatars reveal whether users are lying by using eye contact, blinking, and pupil size. These features are read from the user’s faces. All very interesting research, but when commentators and bloggers start referring to an “avatar that lies” (Ani, 2010) they are being careless with language. Avatars are incapable of lying because they are illusions caused by changing pixel colors on a computer screen – determined by running software. My conclusion is this: the only way an avatar can be said to reveal its user’s true emotional self is if the user is speaking through voice chat, and wearing a full-body motion-capture suit, with every minute nonverbal motion (including pupil size) being projected onto the avatar. Furthermore, the avatar must be viewable at a high pixel resolution and at a high animation frame-rate. In this case, one could argue that we are no longer talking about “avatars”: avatars are *semi-autonomous* by nature. Any other variation of control (that is, almost all kinds of avatar control) are mediated to some extent, and are therefore subject to sending artificial nonverbal signals, whether intended or not. In this case (i.e., in most cases) we cannot speak unambiguously of *lying* or *truth-telling*.

Let’s assume, then, for the purposes of this chapter, that you are not so concerned with whether your avatar is expressing your true emotions or intentions. Let’s assume that your avatar has some degree of autonomy, and that you are interested in controlling your avatar in order to communicate to others in creative ways. There is a touch of *theatre*: you may want to project nonverbal behavior that does not necessarily correspond with what you are feeling at any given moment. “Expressing” and “communicating” are the operative words. This is not the same as being monitored by a lie-detector.

As cinematic language makes its way increasingly into computer games and virtual worlds, the true power of gaze will reach its potential. But, considering the sheer power of virtual gaze, and the availability of software techniques to enable it, I believe that virtual gaze is underutilized in social virtual worlds – it should be more advanced than it currently is. In this paper, I explore some possible reasons why such an important component of natural body language is not well-represented in virtual worlds, and I also propose a framework for thinking about the nature of virtual gaze, and the technical and user-interaction problems with implementing it.

2. CONCEPTS

“Puppeteering” refers to the act of controlling an inanimate object or a virtual character in real-time, to create the illusion of life. It could be your hand, a sock, a ragdoll, or a high-tech marionette. It could also be a virtual character in a game or virtual world. One can easily imagine the term “puppeteering” to refer to moving a body part (like what happens when you pull a marionette string to lift the puppet’s hand). But rotating an avatar’s eyes and head to look at another avatar is special: it has a particular effect that extends out beyond the body of the avatar. It projects an implied line of sight onto the thing being gazed at (*the gaze*). The effect is a psychological “energy” vector that connects *gazer* with *gaze*. You could think of this vector in one of two ways: (1) as originating from the gaze (photons are bouncing off of the face of the gaze and landing on the retinas of the gazer), or (2) as originating from the gazer (the gazer sends a nonverbal signal to the gaze). This direction (gazer to gaze) is the one I am most interested in: it is a form of silent communication.



Figure 6-2: Gaze vectors originating from people's eyes, and landing on what they are looking at.



Figure 6-3: Two gaze vectors fused as a result of mutual eye-contact.

Eye contact established between two people can have a powerful effect. Sometimes gaze accompanies feelings of love; sometimes it accompanies feelings of being threatened or overpowered. Sometimes a split-second glance can be the telltale signal that clinches the answer to a convoluted emotional question – influenced by the context in which the gaze act is done. This “energy” I refer to is illustrated in Figure 6-2. It shows imaginary beams of light originating from people’s eyes, and landing on the things they are looking at.

The purpose of showing these imaginary beams is to make a point: the space between people’s faces is continually being charged with psychological energy, especially when they are using natural language. If you’ve spent as much time as I have watching and thinking about this psychological energy (several years developing avatar gaze algorithms and writing about it), you can certainly *feel* its power – especially in a room with several people in conversation. Let’s keep these images in mind as we explore the ways in which gaze can manifest in virtual worlds.

SEMIOTIC SACCADDES

The rapid changes in eye rotation are referred to as *saccadic* eye movement, or simply, *saccades*. As the eyes take snapshots of the environment – quickly jumping from focal point to focal point, the brain builds up a stable model of the environment, even though the “images” landing on the retina are jumping around like shots in a wild music video. Many birds don’t have the same eye-orbiting talents that we have, and so they have to use their heads to take these snapshots for visual fixation—a behavior called “head bobbing”. These bird-saccades are timed with their struts, famously demonstrated by chickens and enthusiastic funk musicians. This strut allows the bird’s head to be stationary for a brief moment between thrusts, so that the brain can take a picture. It’s not unlike *spotting*: the head-shifting technique used by ballet dancers when they are rapidly spinning in place.

Saccades evolved for building stable models of the world via snapshots. Saccades are important for us primates: since we have high foveal acuity (a high density of photoreceptor cells in the region of the fovea) we aim precisely at points in the environment to resolve details. But the evolution of saccades might have been intertwined with social adaptation as well. In the context of this chapter, what is interesting about saccadic behavior is not just its utility for taking in reality, but how this behavior has become a part of natural language. It’s not just used for input; it is used for output as well. The eyes of many social mammals have evolved distinct visual features—with clarity and utility reaching a high level in humans. Consider the following sentence that you might read in a novel: “Immediately after Mary told Bob what had happened, she made a split-second glance over to Frank—silently alluding to his involvement”. Recent scientific research is helping to validate the notion that variations in saccades and fixation can be used for communication and annotation, as well as for establishing joint attention (Müller et al. 2009).

THE PUPPETEERING PROBLEM

Designing user-interaction schemes for making avatars walk, fly, or strike dance moves in 3D virtual worlds and games is already difficult: but when you consider the problem of how to generate nonverbal communication, it becomes much more complex. I am talking about controlling an avatar using any technology other than fully immersive virtual reality (VR) with total-body motion-capture (in this case there is no “puppeteering” since the user effectively IS the avatar). Full-on VR with total-body motion-capture is great for research, education, and art. But it is bloody expensive, and it does not permit casual, asynchronous usage. I am referring instead to typical avatar-based worlds using common, everyday input devices, such as keyboards, mice, touch screens, Wii devices, and the Kinect, all of which – to varying degrees – allow some subset of human movement to be translated to a virtual body for natural body

language expression. Because of the necessary mediation of these technologies, there are mappings that need to be made from physical body to virtual body. The puppeteering problem also refers to text input and voice input as a vehicle for controlling an avatar's nonverbal behavior verbally. I include all of these forms of input because I want to look at the puppeteering problem in the most general way.

Controlling avatar gaze vectors for nonverbal effect is problematic for several reasons. One reason is that virtual worlds are typically experienced in third-person, which is a form of virtual astral projection: you do not occupy the same space as your avatar and so it is not always straightforward to specify directions relative to your body. This problem is alleviated somewhat with first-person view, in which your view onto the world is synonymous with your avatar's gaze. (One could argue that when you cannot see your own avatar setting its gaze, you lose some cinematic immersion). Another problem with virtual gaze is how to specify *what* your avatar is looking at, and whether and how to set *smooth-tracking* so as to keep the gaze fixed on the gazee even if it is moving. These user-interaction problems may be partially responsible for the lack of gaze behavior in avatars in virtual worlds: it's just a hard design problem to solve. This is especially true if you want to build an avatar control system that allows the high-frequency signaling of saccades.

Another theory as to why avatars lack gaze behavior (or any nonverbal communication) is because virtual worlds evolved largely out of computer game technology. Computer games have perfected the art of blowing things up, leveling up, navigation, and, more recently, building stuff (i.e., *Second Life*). Expressing yourself within a virtual body is a rather new occupation, in the scope of the computer game history. But virtual worlds may finally be starting to grow out of this adolescent stage.

3. TECHNIQUES

For developing avatars for social virtual worlds, it is best to allow eye rotation to be independent of head rotation. So, an avatar could throw a shifty side-glance with no head motion, or it could gaze with the entire head – for an effect that everyone in the room notices. The speed at which an avatar rotates the eyes and head to look at something could be adjustable—a parameter roughly associated with “alertness”, or “attention” (or perhaps “caffeine”). So, if an avatar were tracking a humming bird darting through the air, a “low-alertness” setting would create a rather drunken meandering of the head and eyes in attempt to follow the bird's trajectory, while a high alertness setting would result in quick head/eye motions, keeping a tight aim on the bird the whole time. The same kind of tight alertness response is used in first-person shooters (the camera is very *tight*: responsive to mouse movement) so that the player can take aim quickly.

JUMPING TO A MOVING TARGET AND TRACKING IT

Your visual system uses saccades as you watch a crowd of people walking by, darting your focus among the many faces. If an attractive prospect walks by and catches your eye, you switch from saccades to smooth-tracking (or *smooth pursuit*). This kind of eye behavior is what happens when you watch a distant bird arc across the sky. Smooth tracking uses a very different eye-brain control system than saccadic motion.

In virtual worlds and 3D games, one way to create smooth tracking is to establish a virtual link between the gazer's eyes and the mobile object to be pursued. In a prototype I developed for There.com, I had a mode where circles could be overlaid onto the heads of all the avatars in the scene (including the user's). This is illustrated in Figure 6-4.



Figure 6-4: Circles overlaid on avatar heads serving as selectable gaze targets

The circles provided affordance to the user: they were selectable regions where the mouse cursor could be clicked-on to trigger a gaze shift, causing your avatar to rotate its head (and/or eyeballs) to fixate on that avatar's head. Once chosen, this gaze shift would stay fixed even as both avatars moved about in the world. In this illustration, the user's avatar is shown at left (the black woman). Selecting one's own avatar head circle cancels the gaze that may have been established on another avatar.

We explored some variations of the user interface, including having head circles appear only when the user's mouse cursor hovers over the avatars' heads. We did not test the effects of these visual overlays, however. The implications of this kind of interaction, and the various visual affordances, would make for an interesting study for the design of social virtual worlds.

A thought experiment is illustrated in Figure 6-5. The user of the dog avatar rolls the mouse cursor over the cat at the left. Once the circle over the cat avatar has been selected, the dog avatar will begin to *smooth-track* that cat. (One can imagine this as the initial act leading up to a chase scene).

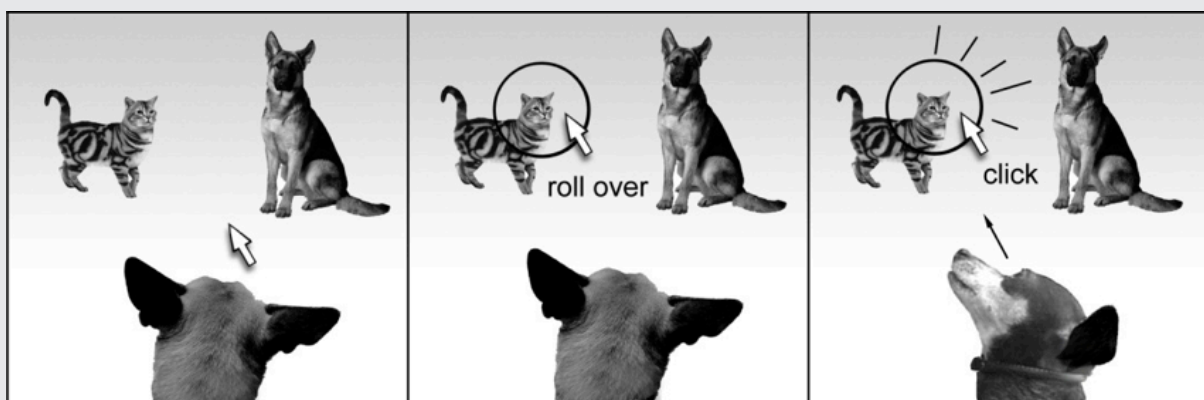


Figure 6-5: Circles overlaid on avatar heads serving as selectable gaze targets

Figure 6-6 shows another example. Imagine yourself sitting on a park bench reading a book. A person walks by and your gaze shifts to the person's face. You continue to track that person for a while, and then your gaze returns to the book. Two salient events have occurred: a gaze-shift to the person, and a gaze-shift back to the book.

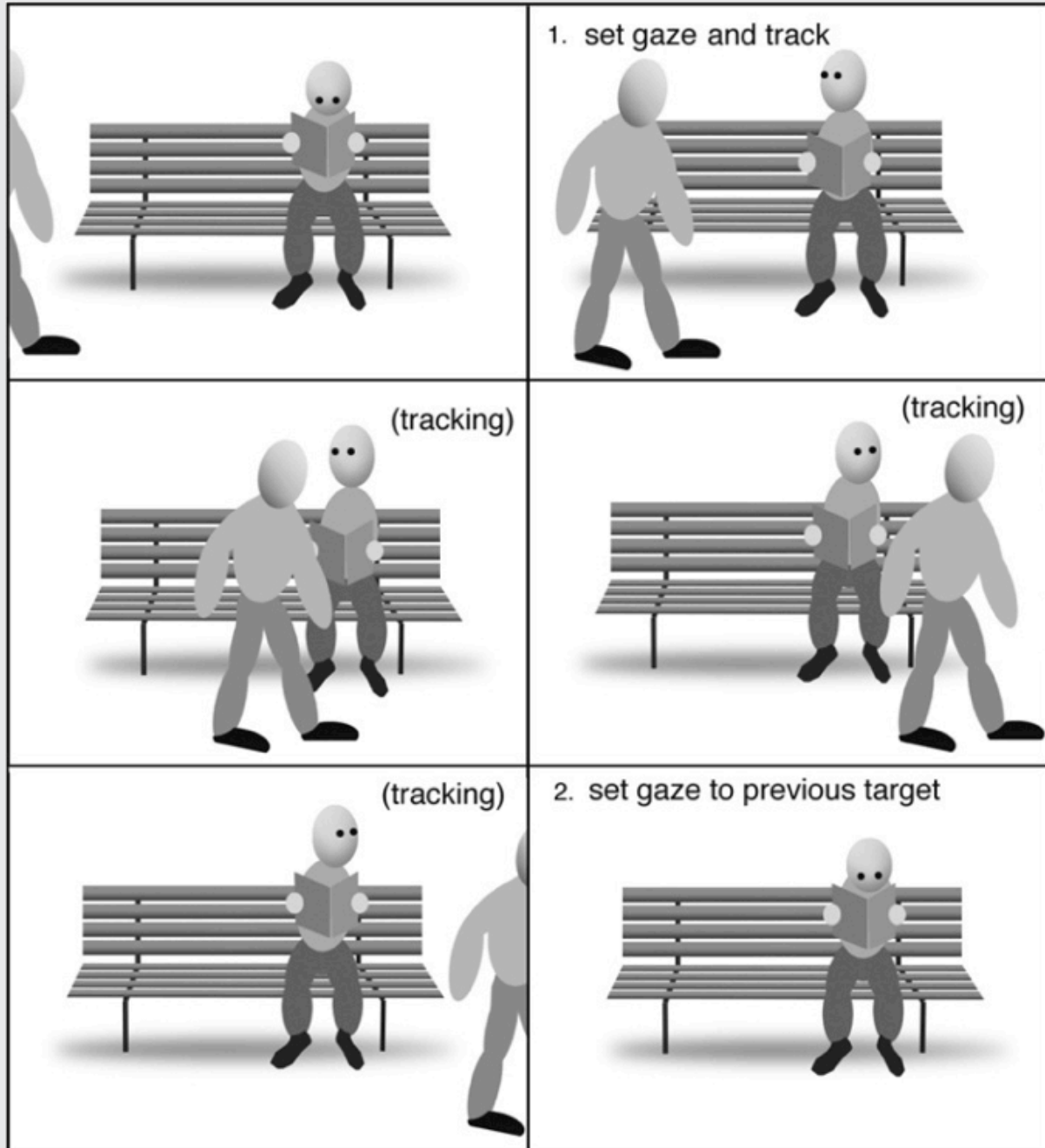


Figure 6-6: Setting your avatar gaze to a passerby's face, tracking it, and then setting it back to the previous target

In a virtual world, these two events could be specified by two messages sent up to the server that specify the identifier of a gaze target (the passerby), followed by a “time-out”, or else a specific message to set the gaze back to the book, or some other, new gaze target. Everything that happens in-between those two events does not require any internet messaging, as long as the first message specifies to “track” the gaze target wherever it moves.

All the tracking can be done on the client-side, and that is because the location of the two faces (the *gazer* and the *gazee*) are cached on the local client; they are instantiated in every user’s view of the world whose avatar happens to be logged-in at that location, and can potentially see the gazing avatar. All the expensive 3D vector math can be computed locally. If I happen to be controlling the avatar that just walked by, I will see that you have set your gaze on me and tracked me for a while. And the reason I can see this is that your avatar (as instantiated on my client) received the gaze message, and *my* client started running the tracking behavior.

It is not only for technical reasons that a behavior like eye-tracking should happen automatically in an avatar: there really is no need for the user to continually control this tracking behavior as long as the high-level goal is to “keep looking at the passerby”. Now, what about the decision to look at the passerby in the first place? Should that be automated? Like most channels of virtual body language, that would be served best by offering multiple levels of control, each useful at different times and for different purposes. The hardest level of gaze to design, and the easiest to use, is automatic gaze: the avatar system decides when the avatar should shift its gaze, and to whom or what. Ideally, a large set of controls should be tweakable as part of the avatar customization interface, to modulate these automatic behaviors—sort of like “personality filters”.

On the opposite extreme, a manual gaze system would be useful for times when the user wants to puppeteer gaze using a short leash: “I want to look at you...*now*”.

A virtual world in which users could choose other avatars to gaze at with ease, and to switch that gaze as easily as clicking on the heads of nearby avatars, would become charged with nonverbal energy. The idea of clicking on the heads of avatars is synonymous with directing your attention to various elements on a web site. By clicking on an avatar head, you are saying that you want to look at that avatar (that is, *you* the user, as well as *you* the avatar). The gaze of you the user is not broadcast into the world, but your avatar’s gaze is. In some cases your avatar’s gaze may represent your physical gaze, and in some cases it would not. The point is that having this level of puppetry would enrich the semiotic landscape.

If the gazee-avatar reciprocates to the gazer-avatar’s gaze, then you have established a natural, wordless social link between avatars (and by implication, between the physical users as well). This could be all that’s needed (as in two people who simply want to acknowledge each other using eye-contact). Or it could be used as a queue to start a verbal conversation. By allowing this behavior to be operated manually, the social utility of gaze would be entirely up to the users, and they might in fact develop their own nonverbal conventions, forming their own semiosis, without the avatar system imposing a code of behavior. The downside is of course that it requires extra controls, and more attention from the user.

4. THE SOCIAL COORDINATE SYSTEM

Human saccades shoot psychical beams with a complex rhythm and symmetry, like visual music playing over the dynamics of verbal communication. Smooth-tracking eyes perform violin sweeps and clarinet glissandos. A virtual world where avatars cannot look at each other is a world without psychical energy; it has no *musical soundtrack*, as demonstrated by the avatars in the virtual world Kataspaces [Figure 6-7].



Figure 6-7: Setting your avatar gaze to a passerby's face, tracking it, and then setting it back to the previous target

In most virtual worlds it is not easy for users to set the gaze of avatars at will. In *Second Life*, well-crafted portraits of fashionable residents are often depicted looking off into empty space, which reinforces a kind of persona we often see expressed in these worlds: aloofness and *serene ennui*. Might this kind of persona ultimately be a symptom of a design deficit rather than a conscious artistic decision?

When a character animator is composing an animation in a standalone application, the geometry is defined in terms of the character's local coordinate system. For instance, the character's x-axis extends from the left to the right of the body. In most systems, the y-axis extends vertically: toe-to-head, and the z-axis extends from back-to-front. A joint rotation is typically specified in relation to the joint's *parent*. An elbow is rotated *off* the shoulder. An animator may use built-in tools for *inverse-kinematics* (the technique of adjusting joint rotations to cause an end-effector to move to a particular location – such as calculating wrist, elbow, and shoulder rotations to cause an avatar's hand to touch the shoulder of another avatar) or *forward dynamics* (simulating Newtonian physics for momentum, forces of gravity, friction, etc.) to create special effects—and indeed these procedural animation systems require the temporary use of a coordinate system in a higher frame of reference. But, once the animation is completed, and a file is exported, all joint rotations are normalized to the local coordinate system of the character. It is essentially a record of body movement without a home—without an environmental context—floating in Einsteinian relativistic space.

When this floating animation file is imported into a virtual world and starts running on an avatar in realtime, procedural techniques take charge and help ground the animation in the context of the world. Inverse kinematics in particular is used to modify the leg joints of the animation and adjust them to conform to an uneven terrain as the avatar ambulates through the world. This set of events could never be anticipated during the initial act of animating. Forward dynamics can also be applied to flexible parts of the avatar (hair, tails, etc.) causing them to sag naturally with gravity or shift from wind or collisions with other objects. The same goes with enabling avatar heads to swivel so as to face each other, or for avatars to hold hands, or to pet a dog, as illustrated in Figure 6-8.

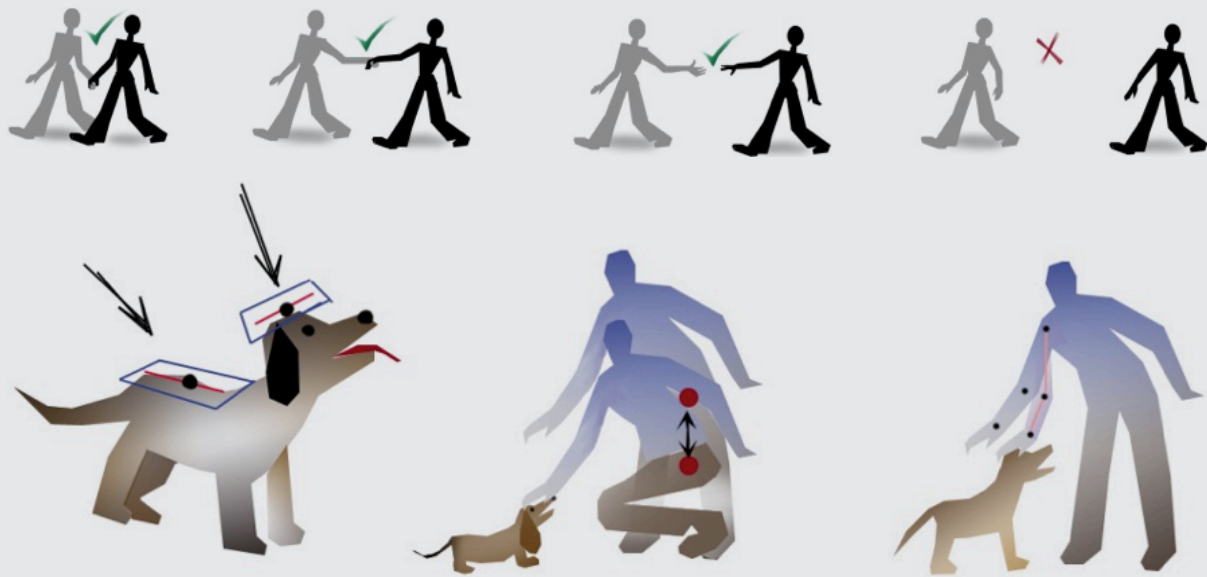


Figure 6-8: (top) Avatars holding hands; (bottom) an avatar petting a dog

These activities *ground* an avatar’s motion in the world from moment to moment. We human users (the meat puppets who control avatar puppets) function in the same way: neuroscientists have discovered cells in the brain that they call “place cells” and “grid cells”. These cells ground our sense of where we are in the world; they are linked to the hippocampus, and help us incorporate our body schema with the environment (O’Keefe and Dostrovsky 1971)(Hafting et al. 2005). Place cells and grid cells are part of a broad system for dynamic representation of self-location. They can be thought of as the brain’s way of coordinating personal space with the global coordinate system. Mirror neurons have their own way of connecting up personal spaces. According to Blakeslee and Blakeslee, they create *shared manifolds of space*: “Watch a fast pass in hockey, listen to a piano duet, or watch two people dance the tango. Mirror neurons help people coordinate joint actions swiftly and accurately, providing a kind of ‘we-centric’ space for doing things together” (2007).

The local coordinate system—by itself—is *lonely*. The global coordinate system—the frame of reference which all avatars occupy, and which allows all the various lonely coordinate systems of the world to “transform” to each other—is *social*. The mathematics of social gaze might be described as a *we-centric* connective glue that infuses the social into the personal. Effective virtual body language requires seamless translation between the lonely coordinate system and the social coordinate system. Mathematically-speaking, that means having software interfaces to transform the local geometry of the body to the global geometry of the world. Inverse-kinematics is one of many techniques to make this happen, and it is also how avatar heads swivel to look at each other in the global coordinate system; in the *social coordinate system*. Instead of manipulating an arm bone or a thigh bone, we are manipulating a gaze vector. This *gaze vector* is in global space, as illustrated in Figure 6-9.

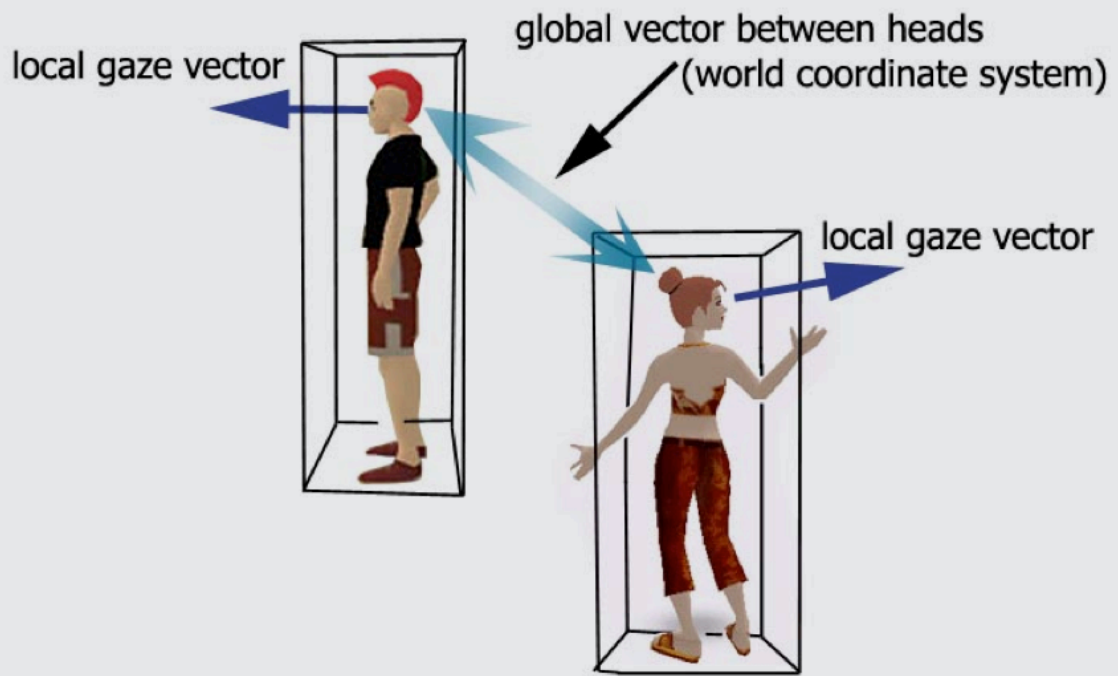


Figure 6-9: Local gaze vectors must be transformed to the global (social) coordinate system

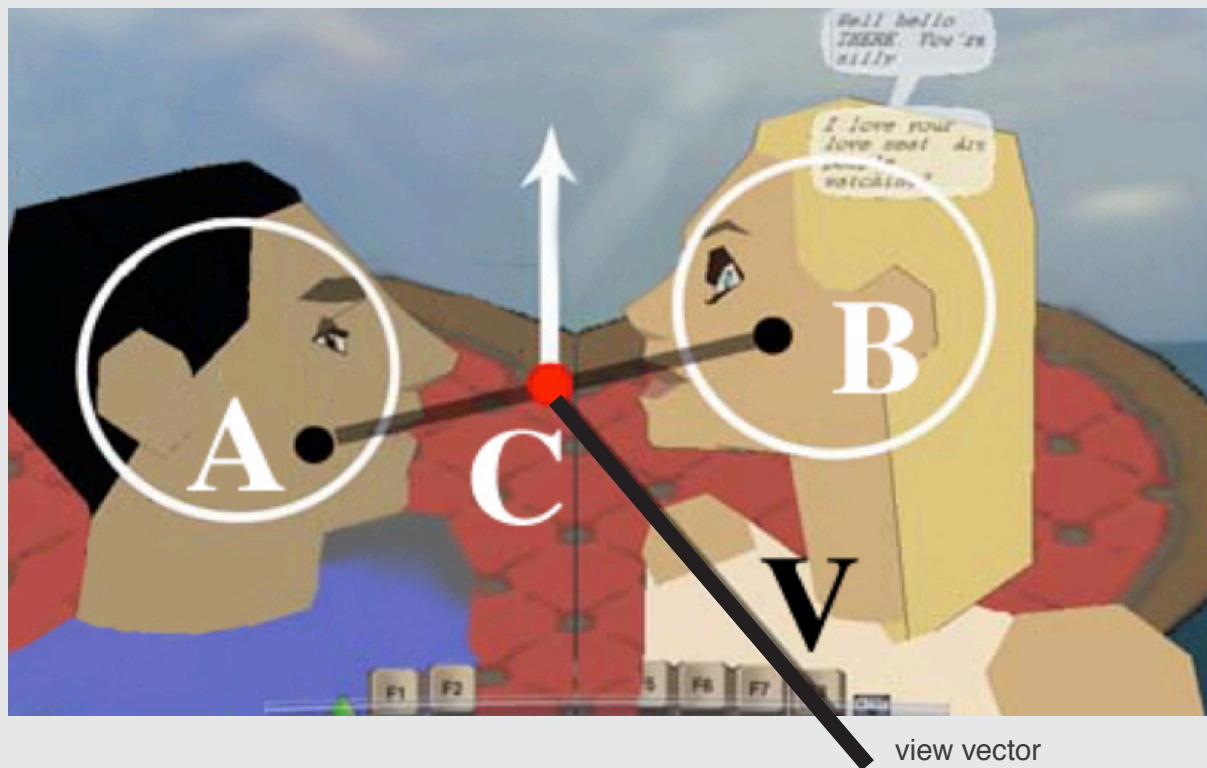


Figure 6-10: The intimacam

INTIMACAM

Speaking of mathematics, there is a fairly simple bit of math that can be brought to bear on avatar body language, with strong emotional effect. I call it “intimacam” – it is a camera behavior that responds to two avatars that are gazing at each other in close proximity. A particular chat prop that was designed by game designer Chuck Clanton and myself was called the Loveseat (Clanton, 2003). This is illustrated in Figure 6-10.

When a user’s avatar was placed into the Loveseat, we provided various UI elements for invoking body language such as eye gaze, head gaze, full body gaze, and facial expressions. If the avatar faces were both aimed at each other (indicating that there may be some romantic interest) the camera would situate itself to aim perpendicular to the mutual gaze vector of the avatars, causing one avatar head to appear at left, and the other avatar head to appear at right. As the avatar faces drew closer, the camera likewise drew closer.

Refer to the global coordinate skeletal head joint positions of the two avatars as A and B. The midpoint between A and B is called C. Take the cross product of the vector AB with the global up direction to determine a perpendicular vector V. That vector V is used to determine the camera’s orientation and distance to C. The length of V is thus proportional to the length of AB. A few tweekers were thrown in to take into account the camera’s field of view.

The cinematic effect we strove to achieve was one of *increasing intimacy*, as the three elements (left avatar head, right avatar head, and camera/voyeur) drew closer. When (or if) the *kiss* happened, the camera would be moved-in to its absolute closest: the two avatar heads would be taking up the entire view. Similar techniques like this were prototyped, influenced by Chuck’s ideas on using cinematic techniques in games, as described in (Isbister, 2006). Camera close-ups are being used increasingly in 3D computer games that involve complex characters with social interactions, such as Rockstar Games’ *LA Noire*. In this game, the ability to read eye contact to determine if a character is lying is an important aspect of gameplay.

5. CLOSING THOUGHTS

Although much research has been applied to the problems and techniques for virtual gaze, popular virtual worlds have not taken full advantage of this research. Users do not have the amount of gaze control that one should expect in a socially-focused virtual world, where simulated embodiment and proxemics constitute a large proportion of social currency. Take *Second Life* for example. In this virtual world, it is not straightforward for a user to aim his or her avatar head or eyes at another avatar. Changes in the “look at target” often happen unpredictably, based on automated algorithms and invisible – often mysterious – timing mechanisms. There is no clear, direct, discoverable way to puppeteer your avatar to look at an arbitrary location in the world, or an arbitrary avatar’s face.

Virtual gaze is a tricky bit of body language indeed. On the one hand, it is a cool feature for a virtual world to puppeteer your avatar automatically – making it look up when his/her name has been uttered in a chat, or to make it smile when you type a :) into your chat window. But, these behaviors steal autonomy away from you: while your avatar performs your body language for you, it sometimes may not be your original intention. Consider the avatars in *Blue Mars*: they will sometimes turn around and look directly at the camera – at the user..you! This behavior apparently creeps-out some people (Au, 2010). And no surprise: when you – the user – are yanked into the world (via a gaze vector “self-puppeteered” by your own avatar), your sense of immersion gets scrambled. The cinematic effect is intense, but what did the *Blue Mars* people have in mind? It is possible that they were just playing with a cool effect, without taking into full consideration the media effects. Figure 6-11 shows an image from a blog post by Ironyca (2010): an avatar gazes back at its user.



Figure 6-11: I think my avatar might have a crush on me!

My conclusion regarding gaze is this: as in real life, I should *always* have the option to look at whomever I want, whenever I want. If I want my avatar to look at me (for whatever strange psychological reason), that should be an option. I should also be able to turn that off (“off” seems like a good default to me). I should be able to override any of my avatar’s “autonomic nervous system” behaviors to generate deliberate purposeful body language. Should all my avatar’s body language be deliberate? Should every action be the result of conscious control on my part? Of course not: avatars are semi-autonomous puppets – by nature, by design. The issue is just *how* autonomous, and *under which situations* they should be automatic vs. user-puppeteered. That is the big question. It makes sense to leave in place the lowest-level autonomic systems, such as blinking, and the “moving hold” technique of character animation to avoid absolute robotic stillness, as well as various techniques such as Perlin noise. This subject of balancing autonomic avatar behavior with deliberate puppeteering is covered in detail in *Virtual Body Language* (Ventrella, 2011).

This chapter presents the topic of avatar gaze as a powerful form of nonverbal communication, which has not yet seen its full potential in virtual worlds. The reasons are partly cultural (the ancestor medium of computer games is traditionally indifferent to self-expression), and partly technical (puppeteering gaze is not straightforward, given the standard ergonomics of virtual world interfaces, especially with third-person views).

Once these human-computer-interface tools have been worked out – along with the normal evolution of user behavior that ultimately modifies design choices – this channel of nonverbal communication will become very expressive. In real life, my wife only needs to make one quick glance at me, and I know that it is time to take out the trash. Or...depending on the timing, or the situation...it may mean something else entirely: something that is more fun than taking out the trash. This simple bit of body language is powerful indeed, and once the puppeteering problem has been worked out, it will enliven virtual worlds, further validating them as a communication medium of embodiment.

6. BIBLIOGRAPHY

- Ani. Avatars with real eye movements can spot liars in the virtual world. 2010.
 Web site: <http://science.gaeatimes.com/2010/04/09/avatars-with-real-eye-movements-can-spot-liars-in-the-virtual-world-10358/>
- Au, Hamlet. Blue World Notes: Why Your Blue Mars Avatar Seems to Look You in the Eye.
 Blog site: http://blog.bluemars.com/2010_11_01_archive.html. 2010
- Bailenson, J., et al. 2005. "The independent and interactive effects of embodied-agent appearance and behavior on self-report, cognitive, and behavioral markers of copresence in immersive virtual environments". *Presence: Teleoperators and Virtual Environments*. Volume 14, Issue 4. MIT Press.
- Blakeslee, S., and Blakeslee, M. 2007. *The Body Has a Mind of its Own*. Random House.
- Chopra-Khullar, Sonu., Badler, N. Where to look? Automating attending behaviors of virtual human characters." /Autonomous Agents and Multi-agent Systems/ 4(1/2), 2001, pp. 9-23.
- Clanton, C., Ventrella, J. 2003. Avatar-centric Communication in There. Presentation at the People, Computers and Design seminar, Stanford University.
<http://hci.stanford.edu/courses/cs547/abstracts/02-03/030404-clanton.html>.
- Gratch, J., Okhmatovskaia, A., Lamothe, F., Marsella, S., Morales, M., van der Werf, R. J., and Louis-Philippe Morency. 2006 Virtual Rapport, 6th International Conference on Intelligent Virtual Agents (IVA'06).
- Hafting T., Fyhn M., Molden S., Moser M. B., Moser E. I. 2005. "Microstructure of a spatial map in the entorhinal cortex". *Nature*. Aug 11, 2005.
- Ironyca. WoW meets Blue Mars – about the Avatar Gaze. 2010.
 Blog: <http://ironyca.wordpress.com/2010/09/26/wow-meets-blue-mars-about-the-avatar-gaze/>
- Isbister, K. 2006. *Better Game Characters by Design: A Psychological Approach*. Elsevier.
- Lance, B., and Marcella, S. 2010. "Glances, glares, and glowering: how should a virtual human express emotion through gaze?" *Autonomous Agents and Multi-Agent Systems*, Volume 20, number 1. Springer.
- Lee, S. P. Badler, J., Badler, N. Eyes Alive. SIGGRAPH 2002: 637-644
- Morency, L., P., Whitehill, J. and Movellan, J. 2008. Generalized Adaptive View-based Appearance Model: Integrated Framework for Monocular Head Pose Estimation, 8th International Conference on Automatic Face and Gesture Recognition (FG 2008).
- Müller, R., Pannasch, S., Velichkovsky, B. M. 2009. "Comparing eye movements for perception and communication: Changes in visual fixation durations and saccadic amplitudes". *Perception* 38 ECVF Abstract Supplement, page 23.
- O'Keefe, J., and Dostrovsky, J. 1971. "The hippocampus as a spatial map. Preliminary evidence from unit activity in the freely-moving rat" in *Brain Research* Volume 34, pages 171-175.
- Steptoe, W., et al. Lie tracking: social presence, truth and deception in avatar-mediated telecommunication. CHI '10 Proceedings of the 28th international conference on Human factors in computing systems. 2010.
- Thiebaut, M., Lance, B., Marsella, S. 2009. Real-Time Expressive Gaze Animation for Virtual Humans. International Conference on Autonomous Agents and Multiagent Systems. (www.ifaamas.org).
- Ventrella, J. Virtual Body Language. ETC Press, 2011.
- White, M. "Visual Pleasure in Textual Places: Gazing in multi-user object-oriented worlds". *Information, Communication & Society*. Volume 2, Issue 4. December, 2001. Pages 496-520.

7

AVATAR APPEARANCE AS PRIMA FACIE NON-VERBAL COMMUNICATION

By Jacquelyn Ford Morie

The Gods will not speak to us face to face until we ourselves have a face.

-C. S. Lewis

1. INTRODUCTION

Who are we when we are online in virtual worlds? That is a fundamental question – one that begs to know what constitutes our identity in the beginnings of the 21st Century. Other questions follow intuitively: How do we communicate in these new spaces of social experience? What are the new norms for virtual interaction?

What differentiates virtual worlds from other forms of computer-mediated communication is that in these spaces we possess a form of embodiment. We have a representation that stands for our self in these worlds, and with which we navigate space, interact with objects and communicate with others. This representation – known as an avatar– presents our *self* to others in the virtual world. We can make it move, we can take it hither and yon, we can shape it and dress it and give it a voice. We can put on a very different appearance, changing it like a new outfit from our wardrobe to suit certain situations or moods.

Yet, these avatars are not as complete or exquisitely functioned as are our physical bodies. They cannot yet transmit the richness of physical expression we as humans have employed for millennia to share what we are thinking and feeling with another, often whether we want to or not. They give us only a means of partial communication, allowing us to use our real voice, but otherwise limiting us to the impoverished capabilities of avatar facial expressions, behaviors, and appearance. And yet, in this last aspect – appearance– we do have a myriad of options to explore that transcend the limitations of physical reality.

There may be a day in the future when we can more fully connect our muscle and bone body, our senses and neuronal circuits, our unknown and psychological selves, including dreams and wishes, to our avataric representations and behaviors. But until we are finally connected via the proverbial “jack” that integrates our physiology with our avatar body, enabling us to more directly transfer non-verbal interaction to our avatars, we must find other means to enable understanding. Our avatar’s appearance, as a malleable form, can become a transcendent means for compensation. We can design it, change it, and have it represent a secretly held part of our psyche. We can hide in it, revel in it, and express with it. It can become all the facets of the self we live, want to live, or are afraid to live. It is a multifaceted projection of our self, and can even embody parts of our psyche we may have never known before virtual worlds.

Many facets of avatars beg research and explication. In this chapter I explore one form of Non-Verbal Communication (NVC) that avatars allow us: the ability of communication by means of our avatar's constructed appearance. The avatar as our projected self is what we let others see in a virtual world. Just as we are aware that we present ourselves differently in various situations in everyday life, so that we conform to specific expectations of a social situation, we are just as aware that an avatar is a means for us to present a face to our compatriots in the virtual world. What others experience of us, through our avatars, includes our appearance, how we move, how we chat and communicate via voice, where we live and how we define our interests and social groups. But many of these aspects of our avatars, especially communication and movement, do not have as rich a palette of expression as we have in the physical world. Because of this, I contend that it is the appearance of the avatar that provides our largest channel of non-verbal communication.

TRADITIONAL COMMUNICATION: VC AND NVC

To set a common stage for understanding I first present the modes of communication we have developed over millennia in our every days lives, and compare these to what is available to us as avatars in virtual worlds.

The definition of the word *communication* I use is from Shannon and Weaver's (1949: 95), which includes "all of the procedures by which one mind may affect another." Broadly, the term communication is divided into two large subsets: Verbal (VC) and Non-Verbal (NVC). Verbal Communication includes any modality that uses symbolic interaction. This includes spoken language; signs that stand for or represent verbal meaning, such as writing; signs that we see as a command or request (no smoking, no turn on red); and even sign language. According to Adam Kendon, noted for developing the semiotics of gesture, NVC is by contrast, "... generally considered to refer to communication as it is effected through behavior whose communicative significance cannot be achieved in any other way" (Kendon, 1981:3).

Non-verbal communication also encompasses Non-verbal behaviors (NVBs), as well as one's outward appearance. Some behaviors fit neatly into the realm of NVC, such as a facial expression of disgust, or even more innate physiological signals such as the startle response.⁹ Many NVBs are learned in childhood or from one's cultural milieu. Most of the emphasis here will be on the visual aspects of NVC, but some attention to NVBs will also be included.

There is an open research question concerning whether verbal and nonverbal communication can, in fact, be separated cleanly from one another (See, cf. Jones & LeBaron, 2002; Kendon, 1977), but I will assume a line of demarcation between them for argument's sake.

Various estimates place the contribution of non-verbal language at anywhere from 65-90 percent of human communication channels. For instance, Birdwhistell and other early researchers of non-verbal communication note that, despite our predilection for language and words, the real contribution of verbal communication seems to be quite small. His findings resulted in an oft-cited rule that 7% of communication used words; 38% was about tonality, and the remaining 55% involved our physiology (Birdwhistell, 1970). Philpott, stated that only about 31% of what is expressed during communication is due to verbal communication (Philpot, 1983: 158). The remaining amount encompasses those things whose "communicative significance cannot be achieved in any other way."

⁹The startle response is an unconsciously elicited physiological response to a surprising occurrence.

If this estimate is correct, then the higher percentage attributed to NVC must encompass deeper and more expansive modalities than one might think. How exactly do we communicate this NV percentage? Since sociologist Erving Goffman's influential book, *The Presentation of the Self in Everyday Life* was published (1973), researchers in diverse disciplines have defined a variety of specific non-verbal communication cues or code systems. These code systems include kinesics (posture, gesture, stance and bodily movement), oculosics (eye behavior), physical appearance (including body, face, hair and clothes), proxemics (our comfort with interpersonal distances), haptics (the role touch plays), objectics (everything we surround ourselves with, including personal artifacts), chronemics (temporal elements), olfactrics (messages odors convey) and vocalics (other attributes of speech besides actual words). (Kendon, 1977; Burgoon and Hoobler, 2002)

VC and NVC are not cleanly separated within our brain functions, with many areas dependent on or triggering other areas, according to the latest psychological and neurological findings (Haxby and Gobbini, 2011; Nestor et al., 2011). Likewise the various code systems of NVC listed above are also highly intertwined. This chapter is primarily concerned with two NVC code systems: physical appearance and personal objectics. Physical appearance includes those characteristics such as our genetic physical makeup (height, weight, eye color, shape of mouth, etc), as well as the clothes we wear, and the overall style we present via hair, nails, shoes, and grooming. Objectics can augment the message of personal appearance by helping to reinforce status and power through clothes and other items with which we surround ourselves – the Armani suit, the Gucci handbag, the hunting rifle or the sport utility vehicle we drive – all serve to send a message about who we are, what we like and our relative social status. Unlike in the physical world, where we are typically limited by the high monetary cost and difficulty transcending social structures that contribute to our presentation of self, in the virtual world these aspects can be a pure fabrication. In the virtual world, my avatar can drive a Ferrari, even if I cannot afford anything like that in the actual world.

Our general physical appearance has as much to do with conscious and unconscious decisions as it does with our genetic makeup. No matter what form of visage we are given in life by virtue of our heredity, we also tend to dress a certain way, usually appropriate to a situation; we wear our hair according to predominate styles for our age group and culture, and we hold our posture and distance from others per codified social norms. Goffman sees appearance as *crafted* by a person (say, a middle aged man walking down a public street) to convey intentional attributes, in this example perhaps: “sobriety, innocent intent, suitable aliveness to the situation, and general social competence.” (Goffman, 1981:88) In crafting an appearance a person “externalizes a presumed inward state and ... renders himself easy to assess” for any onlookers who might otherwise judge him negatively. (ibid.: 89) In most circumstances a person will perform honestly, with nothing to hide, but situations will exist where some masking of true feelings will take place (Goffman, 1973: 64).

Goffman enumerates two primary parts of what he calls performance: the “front” and the “personal front.” The “front” comprises elements of an environment – physical setting, furniture and accoutrements. The “personal front” includes those elements most closely identified with the performer, such as race, or clothes. He further divides the personal front into appearance and manner. He defines appearance as the “stimuli which function at the time to tell us of the performer’s social statuses.” Manner, by contrast, refers to the clues (or NVBs), that speak to the “interaction role the performer will expect to play in the coming interaction.” (ibid: 24)

I will henceforth use the term appearance to refer to both physical appearance (as defined by Goffman) combined with objectics, which closely resembles Goffman’s “front.” Some attention will be given to Goffman’s manner, as it affects initial impressions vis a vis avatar appearance.

FIRST IMPRESSIONS

In the physical world, when we see another person, his or her appearance evokes an immediate reaction. That primary gut-reaction to appearance informs us if we might want any form of interaction with this person. Do we like the way they look? Are they too unlike us? Do they seem approachable? Does it look like I can gain any advantage in trying to meet them? Like wise we are ourselves presenting cues that allow them to answer the same questions about us.

Our first perceptions are far from arbitrary; they have an evolutionary adaptive function (Zebrowitz, 2004). First impressions are made across several areas, but generally include a quick assessment of social status, social warmth, honesty, general health, physical fitness and intelligence. We tend to hold more attractive people in higher esteem, in what is called the Halo Effect (Dion, 1972; Cialdini, 2001). If they look good, they must have positive traits, like intelligence. If they do not conform to our ideas of attractiveness, then we will view them in a more negative light, according to concepts formulated from our own experiences (Kelley, 1955). These personal constructs modulate the immediate responses to a person's general appearance. (If someone looks like my mother and I had a good relationship with her, then her overall appearance may influence me differently than it might someone else.)

Gibson's ecological approach to perception (1979) asserts that appearance provides clues – adaptive information – about social interaction affordances. In his words: “perceiving is for doing” (Gibson 1979). In essence, these clues are invitations for the range of socially acceptable actions. Expanding Gibson's work to the broader arena of Social Cognition, Fiske and Taylor look at a range of areas that contribute to how we judge others, from attention and memory to affect and motivation (Fiske and Taylor, 1991). People tend to make acceptable judgments based on their perceptions, and by relating what they see to their own experiences and most importantly, to their own goals (Fiske, 1992). We look at others, especially in groups, and categorize them via characteristics that may have potential benefit for us, or with which we self-identify (Rosch, 1978; Neisser, 1987). We may look for a leader in the group, who may be the one that most fully embodies the prototype of that group (Abrams and Hogg, 1990). We may form opinions about whether or not we conform well enough to this group to invest in it. We may look at the group through the lenses of stereotypes, or overgeneralizations, which are all part of the quick assessment, streamlined, adaptive survival behaviors (Zebrowitz, 1996; 2003).

We are aware, even if it is a subconscious awareness, that we are being judged in the same ways. Goffman affirmed that humans carefully attend to the way they look in any given situation to best affect those with whom they interact. He even defined it in theatrical terms, using a dramaturgical perspective for how one chooses one's appearance with props to “act” or convey a coherent message about the self. Each situation is a little play, in the overall continuous theater of our lives. We play the role we need to, and switch between nuances of self for each that we may not even be consciously aware of. Work, play, parenting, job interviews, and other facets of life all demand slightly different versions of our self. We construct these as patterns of communication and as rituals of social exchange, as Catherine Bell describes:

The limited and highly patterned nature of these interactions serves the purpose of creating a self that can be constructed only with the cooperative help of some and the contrasting foil provided by others. In effect, Goffman suggests, one constructs one identity, or “face,” as a type of sacred object constituted by the ritual of social exchange. The social construction of self-images and their relations with other self-images generates a total “ritual order,” he argues, that is a system of communication that deals not with facts but with understandings and interpretations... (Bell, 1997: 141).

Goffman also notes that people will underplay those parts of themselves that “are incompatible with an idealized version of himself and his products” and quotes William James as saying that a person “has as many different social selves as there are distinct groups of persona about whose opinion he cares” (Goffman, 1973: 48).

Just as we require multiple social faces in the physical world, individuals in today’s digitally connected world need to develop and maintain multiple representations and go fluidly from one to another. A person may belong to multiple social networks, and each one might require some form of avatar representation, leading to many avatars in use concurrently, with each look adjusted to the purpose of the application in which it is used. One’s Linked-In image may be professional. On FaceBook that image may indicate a more leisured persona. In games a player’s look can reveal valuable cues about expertise and level. In Virtual Worlds, like games, we see a full avatar representation, one that is broadly customizable by the user.

With the development of avataric representations in all forms of networked social and game media, we are developing, by need, a newly emerging and complex layer of non-verbal cues that are rapidly becoming part of our repertoire for communication. But we are doing so within the confines of a software structure that defines our world and the affordances that shape what we can or cannot do therein. It is a different from the physical world. Virtually, we have less of some aspects of NVC we have come to know over millennia of being human, and in a few ways, we can be seen to have more. This situation means that our forms of communication are evolving within the structures extant today.

WHAT’S MISSING IN THE VIRTUAL WORLD?

Creating one’s avatar can provide a way to compensate or correct for perceived or imagined real-world shortcomings we may have due to genetics, injury, or age. Being able to modify one’s avatar to reflect an idealized self is a compelling method to correct discrepancies between the self we feel inside and our current and actual physical self. But this process is not yet “perfect”.

Jonathan Gratch, virtual humans and emotions researcher at USC’s Institute for Creative Technologies, notes that while virtual worlds such as *Second Life*[™] seek to create a portal whereby people can establish real emotional relationships through media, these technological systems strip out much of the subtlety of human interpersonal communication.

Even though researchers such as Gratch are diligently working on the problem, it will still be some time before we have truly emotionally resonant media.... Where we can both exhibit and perceive emotional signals of any complexity. P 375 though progress is being made especially in osculetics, kinesics and speech prosody (vocalics). Such progress adds to the persuasiveness of a virtual representation, it is still very much in the research domain (Gratch, 2010: 370, 375).

What is currently available to us (and our avatars) is still very limited.

In terms of behaviors, an action or expression taken in a VW (with the exception of idle behaviors and animation overrides) must be deliberate. Body codes that can be enacted naturally or with little thought in the physical world must be performed rhetorically in the virtual one (Verhulsdonck and Morie, 2009). We have some control over how we look, but less over how we exhibit behaviors or facial expressions. Yet, such behaviors can impart key information about cultural, power, status, sexually proclivity and the like. To show some of those same behaviors in the virtual world, I must specifically choose which actions I want to perform. Any context-specific action though, such as laughing at a joke, or showing displeasure, or reacting in surprise must be chosen from one’s inventory or gesture library. This state of affairs exists because the affordances we are given in a virtual world are the results of design decisions made by someone else.

ROLE OF THE DESIGNER/CODER

T. L. Taylor notes that programmers and designers of virtual worlds who shape the underlying software are key in deciding what opportunities a user has for making his or her avatar. She states that the code on which the virtual world is built “then acts as the material upon which an experience of embodiment is built.” (Taylor 2004: 261) In other words, each and every avatar has programmatic underpinnings that shape and limit it. Limitations might have to do with bandwidth or data storage concerns, the game engine that is used, and constraints designed to make life easier for the designers (Taylor describes one world where everyone must be the same size to facilitate coding “parameters of interaction” like the height of doors and seats. (ibid. :264)).

There are the decisions the designers make that affect our abilities in the virtual world, and there are just some things that programmers can’t get software to do, such as following our normal modes of communication, because it is a non-trivial task. There is much work to be done, some of which includes difficult research topics to really improve the state of communication affordances in virtual worlds.

Smiljana Antonijević, in a fascinating 2008 study of avatar communication in *Second Life* done from an ethnographic perspective, points to both missing cues as well as imposed ones. She notes there is little meaningful distinction between what the user generates as opposed to what the software itself imposes. She enumerates four primary non-verbal categories of communication in virtual worlds related primarily to proxemics and kinesics. These are *Predefined, User-defined, Blended and Missing*. Predefined are those provided by the software system itself, User-defined actions have been set up or bought by a user. Blended refers to a combination of the two and missing is self-explanatory, but can change as new functionality is introduced via the software. She found that the user-generated category is closest to what we expect from NVC in the physical world. She states that blended and system-defined tend towards stereotypical and do a fairly poor job of simulating our normal non-verbal behavioral actions (Antonijević, 2008). For instance, when we are communicating with text chat in *Second Life*, our hands start typing on a keyboard. Likewise, when we are “away from keyboard” or not paying attention to our avatar for a period of time, idle behaviors are generated by *Second Life* that keep the representation from going dead or still. If we ignore the avatar too long we drop forward in a stance that tells everyone we are not there, which is a very different type of signaling than we have in our normal communication with other people.

In spite of this gap between our physical world and virtual world NVC modalities, there are things we can control and chief among these is our appearance. Even here, the tools we have been provided will influence what we are able to create. Taylor says how influential the avatar customization interface is thusly: “It contains within it the explicit imaginations about how participants not only will, but should, be constructing identities and inhabiting that space.” (Taylor, 2004: 265). And, it turns out, the appearance we create is a really big deal.

2. WHY APPEARANCE IS SO IMPORTANT

Appearance matters. We know this implicitly, and we see its importance in social and ecological theories. But why should this be so? What are the biological correlates that create and maintain this importance? We evolved to make near instantaneous judgments when encountering an “other” to maximize our chances for survival in the world. We are quick to determine our reaction via immediate perceptions, most importantly sight, which has the largest carries the most information (up to 83% in some estimates) of any of our senses. (Pease and Pease, 2004). Survival is important and thus quick determination of whether someone poses a threat or not can be critical to living or dying. We react this way whether we are consciously aware of it or not. The old adage about not judging a book by its cover is contrary to our

evolutionary adaptations to survival. In virtual worlds, though, the stakes are not yet that high. We will not fail to survive and pass on our genes if we guess wrong in the VW. Yet we are still wired to judge quickly, and we do so in the virtual world just as we do in the physical one.

INITIAL MEETINGS

A popular social encounter theory, Uncertainty Reduction (URT), was developed by communications theorists Charles Berger and Richard Calabrese in the mid 1970s. It was inspired, in part by Shannon and Weaver's 1949 work in information theory, which noted that the motivation for communication behavior is the need to reduce uncertainty.

“Especially in initial encounters, there exists a high degree of uncertainty given that a number of possible alternatives exist in the situation (Shannon & Weaver, 1949). But individuals can use communication to reduce this uncertainty. Berger and Calabrese (1975) maintained that “communication behavior is one vehicle through which such predictions and explanations are themselves formulated” (: 101). Individuals have the ability to decrease uncertainty by establishing predictable patterns of interaction. Because of this, reducing uncertainty can help foster the development of relationships.”

Reducing uncertainty is the first step towards developing a relationship. This is true in the virtual as well as the physical world. We try to find common ground when we meet a new person through a series of disclosures. An initial meeting typically progresses through stages (Berger and Calabrese name three), which take the encounter from the initial visual impression into verbal communication. Each of these steps toward relationship development helps to reduce the uncertainty about the other. Berger and Calabrese, while focused primarily on more verbal stages, also state that:

As nonverbal affiliative expressiveness increases, uncertainty levels will decrease in an initial interaction situation. In addition, decreases in uncertainty level will cause increases in nonverbal affiliative expressiveness.

Such theories have generally held up over the past 30 years as an approach to understanding another person in an initial encounter. While Berger and Calabrese look to the combination of verbal and non-verbal behaviors as complementary and connected processes in URT, there is no doubt that the use of visual ascertainment (does this person look like me? How are they dressed? Does their demeanor seem suitable for this occasion?) is a huge initial step in the process of reducing uncertainty about one's self, a potential partner, and the relationship that might exist between them.

THE IMPORTANCE OF THE FACE

Facial recognition is one of our primary cues to social affordances as faces are the most expressive part of our bodies. Darwin noted that humans have more facial musculature than most of the apes, with many of these muscles finer and capable of more nuanced expression (Eckman 2006). We humans can form a vast range of general and idiosyncratic expressions. While these expressions are connected to a smaller range of emotional expressions, they vary little from culture to culture but more from person to person. How and why we see the face as we do has a long evolutionary path, as knowing whether a person was good, bad, or physically fit enough to be a mate were key pieces of information for doing well in the physical world (Zebrowitz, 2008).

Extensive work has been done on facial behaviors, or expressions (Eckman and others). It may be the most well known of the NVB our physical body engages in. The impression we derive from a face takes less than 100 milliseconds to form in our brain (Willis and Todorov, 2006). Research is also robust in the effects of how the face as a whole looks to us, its basic appearance and whether we perceive it as pleasing or not. Why is this? According to Zebrowitz and Mantepare (2008) “Appearance matters because some facial qualities are so useful in guiding adaptive behavior that even a trace of these qualities can create an impression.” It is highly likely, according to these same researchers, that a basic facial structure triggers some universal mechanism. They note that the stronger clues concern those facial elements that indicate the individual is not fit (e.g. exhibits signs of disease or ill health). Neurologically, it seems our brains are well wired to recognize faces, and may respond more strongly (in a negative response) to maladaptive ones. Our brains have significant areas devoted to seeing static and moving facial cues, and these are connected to regions that perceive/recognize social meaning and emotional valence.

It is obvious why we tend to make generalizations about people we see. We are trying to find something familiar, something that says “this person is not a threat,” “this person is okay” or “maybe even someone I want to be friends with.” The quickest way to do this is to have a template (stereotype) quickly available that covers a host of similarities we consider safe. This first precept is “Does this person look like me?” The most salient part of the body on which we make this determination is the face. Does this person look like anyone I have good template for, or a bad one?

Facial perception, then, is of key importance to the appearance we project in the physical world. Ecological theory states the importance not only of static cues, but also those that incorporate movement, which is the natural state of being alive. The facial expressions defined by Eckman, et al. are parts of a continuum that is more properly studied in valid living and behavioral contexts.

As important as the face and facial expressions are in the physical world, their importance in the virtual one is still open for discussion. One might logically assume that they are or should be as important. But we do not have either the full range of facial expressions available to us, or the number of virtual facial muscles that led Darwin to believe we could express more nuanced expressions than even our nearest animal relatives – the primates. Not only do most virtual worlds not provide access to a reasonable number of parameters to create meaningful expressions, the way we must go about using any expression via our avatar is by selecting that expression – a conscious decision, unlike our more intuitive modes of expression in everyday life (Verhulsdonck and Morie, 2009).

In spite of this, and the lack of nuanced expressions, our virtual faces are read as faces by our facial recognition brain areas. However, because our brain is keenly attuned to them, we do not know how we are responding to a face that may be lacking. We should be aware, in terms of NVC of facial expressions, the virtual world is a pale substitute for the physical reality, especially in this regard, and that the limitations inherent to the virtual may skew our ultimate perceptions.

Even if we had ways to create a more accurate range of human expression with our avatar (which is likely as virtual worlds become more sophisticated) these would still need to be explicitly chosen. As yet, there is no direct link between what the human is thinking or feeling and what their avatar can express. And yet, if the behavioral range of facial actions is lacking in the virtual world, that ability we have for adjusting the general *appearance* of our avatars face is quite broad even now, and can allow a person to have an avatar that transcends perceived or actual physical limitations.

BODIES, CLOTHING AND EMBELLISHMENTS

After faces, clothes are perhaps the next most important cue in humans generating first impressions of other human. Every culture in the world engages in embellishing the body with some form of decoration, from clothing to jewelry to tattoos. (Entwistle, 2000)

Clothes protect, adorn and mediate our body in both physical and social contexts. Seldom neutral, they are ordered and constrained within both society and situations. As humans we are born with bodies, but we seldom engage in social situations without some form of covering or embellishment. Joann Twigg says: “Clothes mediate between the naked body and the social world, the self and society, presenting a means whereby social expectations in relation to age act upon and are made manifest in the body” (Twigg, 2007: 285).

Dress conveys both propriety and our sexuality and is the means by which we conform to social standards and codes, including moral ones. Entwistle maintains fashion is a huge part of how we form/create our identity through its “articulation of the body.” She also notes how we react with unease and anxiety when we fall outside the standards required within specific social spaces. Twigg describes “how individuals feel vulnerable and embarrassed if their dress lets them down, through laddered tights, drooping hems, and other failures of appearance” (ibid.: 295) Virtual world mores reflect our everyday ones, with nudity being forbidden in most areas, and ratings of PG and Mature setting standards even for the type of outfits allowed in specific spaces. Why should this system of propriety be in place in a graphic world? Quentin Bell, in his 1976 work, *On Human Finery*, suggests that all humans possess a “sartorial conscience”, and notes that humans try to conform to at least some modicum of proper dress in most situations so that they circumvent social censure (Bell, 1976: 18-19). We can understand a reasonable transference of social mores of a sexual or revealing nature, but another aspect of sartorial conscience has little concern for virtual residents. For example, virtual clothes are not subject to the vicissitudes of everyday wear and never become shabby with continued use. We don't have to worry about being caught with a stain on our jacket or tie, or a rip in the seat of our pants. Likewise, stockings don't get runs (ladders) unless they are designed that way. In this way, the virtual world is a more perfect incarnation of fashion.

Tseëlon (1992) contrasts Goffman's dramaturgical approach towards our appearance with popular ideas from “Impression Management,” a social psychology perspective that equates conscious management of one's appearance as more insincere than spontaneous actions. In this regard, getting dressed for an occasion would always seem to belie a certain amount of insincerity since it is not spontaneous. Tseëlon argued that if putting on clothes were related to insincerity, there would be no reason for people to care what they looked like when by themselves or with close family. Just because we think consciously about how we appear to others and modify that look to expectations, it does not mean it is necessarily deceptive. From his study done with forty British women, he concludes, siding more with Goffman, that appearance is less about manipulating our look for deceptive purposes, but about presenting who we believe our self to be. Thus, in spite of a theatrical approach to our appearance, it will nevertheless tend to indicate where we are along several continuums, from social status (Breward, 2000), to gender (Entwistle, 2007). (See Figure 7-1.)

Because we give out both conscious and unconscious messages by what we wear, we can use dressing to explore different (and maybe deeply hidden) facets of ourselves. The virtual world provides much more flexibility to self-actualize foreground characteristics or internal feelings, or even try out something new, which tends to reflect something within us of which we have little conscious awareness. As Suler notes in his article *Identity Management in Cyberspace*, “How we decide to present ourselves in cyberspace isn't always a purely conscious choice. Some aspects of identity are hidden below the surface. Covert wishes and inclinations leak out in roundabout or disguised ways without our even knowing it” (Suler, 2002: 457).



Figure 7-1: Avatars in various forms of dress.

3. ON LOOKING AT ONE'S SELF

Today we live in a physical world where image is vastly important. Not only do we see images all around us, we are also aware that we present an image of our own to others. We are constantly reminded that we need to work at making the best presentation of our self that we can, from the color of our hair, to the type of deodorant we wear. Buy this, we are told, and it will make you more acceptable, better, or more loveable. It will improve your image. Because of the ocean of images in which we exist, we are acutely aware of what we present to others and how that might make them judge us, love us, and understand us. What we present is not about who we really are.

There appears to be less of this coercion in the virtual world. In spite of social mores and a rampant commercialism and software limitations, it is still more accepted to “wear” the self-representation you want rather than one foisted on us by immersion in relentless media. In the virtual world, communities of animal, robot and other non-human representations are accepted and thriving. It may be that much of what we present in the virtual world is informed by a desire to “play dress up” in ways that are more open or socially acceptable than people usually find in everyday life (Fron, et al. 2007).

Not only do we have more options for our appearance in the virtual world stage, we also have more options for observation, especially of our self, than we do in the physical world. (Irani, Hayes and Dourish, 2008) Most virtual worlds are presented to a user with an over the shoulder or view from behind that avatar. We can often set the distance at which we want the camera to follow us, and sometimes even the focal length of that camera, or how much of the field around us is in view. Whatever your settings, your avatar is almost constantly in view of your visual system.¹⁰

We also have control of our location of looking, in the guise of a virtual camera that can be controlled, moved around, zoomed in etc. This means we can turn the camera around so that we see our front side, and some VWs have a setting that clicks you immediately into this mode. We also can use that camera to

¹⁰We can also use a function that allows one to “look through one’s eyes”, or a first person view. It is only occasionally used as it is actually somewhat harder to navigate vis this viewpoint.

view others surreptitiously, sometime far from where our avatar is located in virtual space. This leads to a certain amount of voyeurism (camera hopping, looking, checking out profile information, moving your camera to private areas, etc.) that is usually not obvious to the other person or people. Regardless, if we do it to others, we are aware that they may use these same actions to view us.

In the physical world a first meeting has observed codes, conventions and customs, some of which are based on survival mechanisms and some of which are culturally inculcated. The ritualistic meeting in a virtual world can therefore be quite different. People may check out a new person in one of several ways: simply looking at the avatar within the scene they are immersed in, zoom the camera up to and around them and/or click on them to check out their profile information (which can be considered an extended part of the avatar's appearance). In my experience this non-verbal behavior typically precedes the effort to start talking to someone. This looking can be considered another aspect of non-verbal communication that is provided to others as an extension of one's appearance. This has not been given much study, however. There are questions associated with these new modes of initial assessment of the other, such as its idiosyncratic nature, unequal cues, mutual interpretability and social acceptability of those acts among people in that world.

What is particularly of interest here, though, is the ways in which these different modalities of looking – those that draw upon the conventions of the virtual world, those that draw upon the conventions of the graphical user interface, and those that draw upon the conventions of a software engine – are intertwined. They cannot easily be separated. Interaction in the virtual world extends beyond the confines of the simulation window. There are direct correlations to how we have thus far learned to look at ourselves in our original lives. We have some objectification, some projection, and some feedback loop between the two. Because we have more visual awareness of the self we present in the virtual world, we may in fact have a much more rapid feedback loop in play.

We first come to the realization that our self can be an objectified when we are children. This occurs, according to Freudian psychiatrist Lacan and others, when we first become aware that the image we are seeing in the mirror is us. This precipitates a fundamental change in our relationship with, not only the world around us, but also with our inner self. According to Lacan, this event may actually be the start of our perception of Ego (Lacan, 1953). Before this realization we do not recognize, as Maurice Merleau-Ponty notes, “*that there can be a viewpoint taken*” on our self. “Through the acquisition of the specular image, the child notices that he is visible, for himself and for others” (Merleau-Ponty 1964: 136).

Knowing that others can view him as he views himself in a mirror, sets in motions new thoughts about what others may think of this person they see. Before this awareness, the focus for a young child is on his own thoughts and how those around him relate to *him*. The mirror expands and shatters the singular internal viewpoint of self forever. Merleau-Ponty notes that with this recognition comes a growing realization that others can not only look at him, but in some sense, also judge him as well. From this point on our minds will wonder what others think of us. So we dress and present ourselves in ways that will stave off bad judgments and elicit good ones.

With our avatars within a virtual world, we have a new means for viewing ourselves, unlike any other in history, because the virtual world accommodates new forms of viewing. How others viewed us was limited to our imagination, our assumptions of what people might think. In the virtual realm we may not know exactly what they think, but we are surer of the appearance of the self we are presenting. We don't need a mirror because our self is most always in view.

So now, as important as how others see and react to our appearance in the virtual world is how we see ourselves. It is a wholly new form of Lacanian mirror recognition. This is a topic that begs continued research and appraisal, but as Kathy Cleland notes in her dissertation:

Responses to our image avatars may also change over time. The uncertainty and existential shock provoked by encountering an unfamiliar view of the self may give way to feelings of fascination, delight and engagement as the image becomes more familiar; or the avatar image may continue to provoke feelings of unease and alienation, particularly if it does not match our own mental self-image or inexperienced as an uncanny and disconnected other. (Cleland, 2008: 87)

In comparison, the previous technological revolution in how we see ourselves, photography, generated not only wonder, but consternation. According to Roland Barthes, people typically react poorly to seeing themselves in a photograph, as that frozen slice of space-time never quite matches what we imagine our self to be.

“The photographic image is frequently a source of anxiety and disappointment rather than a reassuring affirmation of the subject’s idealised self-image. Even when the mediating subjectivity of the photographer is removed, the ‘objective’ mechanical image generated by the inhuman eye of the camera is no better. As Barthes points out, “the Photomat always turns you into a criminal type” (12). (Barthes, 1993: 12).

In comparison, control of the image is in our hands with avatars. We see what others see, which amounts to more control. This control is part of the attraction of creating and maintaining one’s avatar – a vicarious pleasure in making the avatar be what we want to present or feel inside. We are both subject AND object and we are fully aware that we are an object, not only because we most often see ourselves in 3rd person view... but also because... we are able to choose what our appearance is in more varied ways than in the physical world. We cannot escape this “digital mirror” in current virtual worlds. It pervades and is the substance of our interactions in the world. Yet, as much as people hate to see themselves in a photograph we are more likely to be *enamored* of the look of our avatar (as evidenced by quotes in Kristen Shomaker’s 1000 Avatars books, 2011a, 2011b).

Why is this? An avatar look differs significantly from the sterile lens that captured us on the camera’s film. We have more direct control and can alter the avatar’s manifestation to match our internal representation. We can digitally primp and preen and decide exactly how we want to appear. We are aware of how we will appear (at least in a visual sense) because we too, are seeing ourselves in 3rd person. This is a phenomenological change over previous forms of picturing ourselves. And, we never look like a criminal, unless that is what we are aiming for.

While our appearance is the strongest means we possess to communicate non-verbally to others, it should be emphasized that, at the same time, it is also concurrently communicating non-verbally to our self too. The study that is cited as the *classic* one in this regard was done by Nick Yee and Jeremy Bailenson (2007). In a phenomenon they termed the Proteus Effect, people whose avatars possessed particular traits in the virtual world were found to carry over some of the effects of those traits to the physical world. In another study, Fox and Bailenson found that participants who watched an avatar that bore their photographic resemblance exercise vigorously in a virtual environment, would do more exercise over the subsequent 24 hours than if they had watched their avatar do no exercise, or watch an avatar that did not resemble them do exercise (Fox and Bailenson, 2009).

In seeing our avatar, we are changed. These studies were an amazing first step at understanding how avatars affect us, even though they were not done within the parameters of a fully socially enabled virtual world. In the future we have more to investigate, such as what might be the effect of long-term avatar habitation, reasons for what we wear, how avatars achieve status in their domains and much more. We will next look at some initial studies that have been done with avatars.

4. SURVEYS AND HINTS OF MEANING

There have been several studies that cover various aspects of the relationship between a person and an avatar they inhabit, even if that habitation is for the briefest of time. Most of these have been done looking only at the avatar and their single user, or at most an avatar with one other, as in Bailensen's work describe earlier. An exhaustive search of studies done with avatars found few that focused on avatars in social settings. The studies that have been concluded are presented here, rather succinctly, as efforts informing a realm of knowledge that needs much more research.

A 2010 survey done by Wagner James Au (Au, 2010) and reported in his long time *Second Life* blog, New World Notes, showed that of 850 respondents, 22.8% had avatars that were human and resembled their physical world self; 20.9% had human avatars that fit to some real life style or stereotype. Another 12.6% were child avatars (adults that take on the VW appearance of a child). Mythological or whimsical themed human-like avatars (e. g. fairies, vampires) weighed in at 14.4%. Fantastic non-humanoids accounted for 5.2% (e. g. dragons).

The category of tinies (diminutive avatars with their own culture, living spaces, furniture, clothes, etc.) weighed in at 6.7%, 8.7% were cat-like Nekos (with some part of the appearance resembling a cat) and 4.9% were furrries (various anthropomorphic animal costumes). There were also a few – 3.3% - that fit in none of these categories (an example of which might be a fruit-filled Jello mold avatar, or a simple geometric shape).

In a study by Conrad et al. (2010) students were directed to use an avatar for a particular class-related task in the VW. Of 208 surveys received from 283 participants 89.9% chose an avatar that 1) matched their own gender, 2) resembled their physical self in some way (as was possible within the framework of a limited number of choices given to them), and 3) gave their avatar a name that was the same of similar to their real world one.

Conrad relates this to the avatar being an extension of the person's self within the virtual world. Even though the relationship may not be a strong one (especially in the case of students who must use it for a class purpose), he notes "This vague relationship legitimizes and cements that association; in this way, the user adopts as part of themselves a subjectivity which has been constructed externally, a self-image not of the user's own creation."

Conrad also states "This stress upon the user's absolute and essential control of the avatar, in denying the existential influence which the avatar's form and performance may have upon the user's own subjectivity, may of course suggest a telling anxiety in relation to that notion" (Conrad et al., 2010: 7).

Now if this amount of coupling happens in a person who has just been introduced to the concept of avatar representation, these extrinsic characteristics, how much more cognitive connection occurs for long time avatars – those who have used the virtual world affordances to bond even more tightly with their avatar?

Martey and Consalvo (2010) induced a role-playing activity with 211 subjects to induce a quick form of such bonds, offering them free themed costumes during their experience. They found that social pressures trumped all else (e.g. the experimenter's expectations and encouragement, or personal tendencies), that is, if the group members agreed or refused to wear the costumes, others went along.

Messinger et al.'s study (2009) looked at the relationship between people and how their avatar looked. They found that people tended to manifest aspects of their real selves in their avatars, using both self-verification (similarity) and self-enhancement (betterment) as operational motives. Most users give their avatars characteristics they would like to have (e.g. more curves, a muscular physique, wild color hair, longer legs, less body fat etc.) so that their virtual self is seen as more attractive in some ways. They also found that, to some extent, most people felt less inhibited when performing with their avatar, even though they tend to use behaviors they normally would. Finally, this study found that if a person has created their avatar to be more attractive, they would act more extroverted in the virtual world, with the largest difference being seen by those who were low extroverts in the physical world.

One interesting outcome of this study was that it tended to disprove the assumption that people try to make their avatars look as much like their real-world selves as possible. Most people said their avatar was a mix of "similar and unrecognizable features" as compared to their real appearance (Messinger et al., 2009: 11).

Neustaedter and Fedorovskaya (2009) conducted a survey of twenty-two participants (half male and half female) of the virtual world to see how they both constructed and related to their avatar appearance. They identified four types of avatars: *Realistics*, *Ideals*, *Fantasies* and *Roleplayers*. Realistics need a high congruence between their virtual and real self appearances; Ideals could handle some disconnect between the two, but often used their avatar to "overcome perceived inadequacies" in the physical world. Fantasies keep their virtual and real selves separate, using their avatar as a costume or a masquerade they don to play an ongoing persona. Unlike Fantasies, who tended to keep a consistent (unchanging) appearance, Roleplayers, Neustaedter's final category, tended to put on an avatar to "fulfill different identity needs." For the most part, however, the researchers found that though the potential for multiple identities is certainly available, most people do not take advantage of it, opting instead for a more consistent identity (ibid.: 7).

While not a formal survey, per se, Kristen Shoemaker's recent book, *1000 Avatars* (<http://1000avatars.wordpress.com/>), contains snapshots she took of 1000 avatars in the virtual world SL (fascinatingly enough, all photographed from behind). A count of these avatars by type reveals that only 6.5% of the participating avatars are non-human. In her second book, done several months later with photographs of another 1000 avatars, just over 22% are non-human. What caused this increase is a matter for speculation. It could be that after seeing the results of the first book, people felt more trusting of the photographer's intent, or wanted to be included as part of this project. Word of mouth could have reached to more fringe audiences that might not have received or responded to the first posting for volunteers, increasing the percentage of uniquely designed non-human avatars.

Finally, a recent study done by University of Florida researchers (Black et al., 2009) looked at which characteristics a user might change about an avatar (using *Second Life* avatar creation tools), given the opportunity, under four circumstances: a story in which they were either a *hero* or a *villain*, or instructions to create their *ideal* or their *actual* self. These four categories were assigned randomly to each of the 102 undergraduate participants (13 male, 89 female), none of whom had experience with *Second Life*. The subjects were then introduced to the avatar tools and were allowed technical help by the researchers, but no aesthetic help. The resulting avatars were analyzed along 12 aspects, from relatively invariant physical world characteristics such as skin tone, musculature, and gender, to ones that we can change easily in the real world like hair color, length and style, accessories and clothing.

They found that in all categories people created avatars that were visually similar to their physical selves, determined, in part, by an initial photograph taken when the participant arrived. In general, subjects significantly changed characteristics that were changeable in real life, and not the more enduring ones like skin tone and body structure. The one exception was gender specific characteristics, which were enhanced.

Less change was made to the hero and villain avatars than might have been expected. For the hero significant changes were made to four aspects: eye color, gender characteristics, general appearance and clothing. For the villain, only clothing, general appearance and hair style were changed with any significance. For the ideal avatar, however, 7 characteristics showed significant change: accessories, clothing, eye color, gender characteristics, general appearance, hair length and hair style. For the actual self category hair color was also modified along with the seven parameters changed in the ideal situation.

This study was valuable to show how someone might approach their first avatar creation. The story vignettes that were presented did little to encourage much change from the starting avatar. The situations that addressed their ideal and actual selves elicited that largest change, perhaps showing that there is more investment in how one portrays themselves as themselves, rather than within some fictional and arbitrary role.

These studies suggest that the most common way people approach making an initial avatar (at least as evidenced in these studies) is to make it look like a slightly idealized version of themselves. This most likely does not indicate so much a change from a person's self image, but rather that the self-image a person has may be idealized from an internal stance to begin with.

5. CREATING AN AVATAR

For those unfamiliar with the process of creating an avatar, this section will exemplify the process using the fairly rich tools in *Second Life*[™] as the chief example. *Second Life* has much the same functionality of any social or game-based virtual world, allowing one to choose from a set of starting avatars. However in games, a starting avatar may indicate status level or capabilities. In *Second Life*, the avatar's look does not do this and serves more as a springboard for Taylor's "explicit imaginations." Sliders are provided that encompass over 200 parameters that can be adjusted, more than in most games or virtual worlds. There is inherent limitation in these parameters, of course. They assume a humanoid skeletal framework on which the 3D geometry sits. This is parameterized to fit within standard body shapes, heights, head size, etc. This tends to set limits on the variety of human-avatar types one can design using these sliders.

As noted, virtual worlds typically provide a range of starting avatars, and tools to customize those avatars after the initial choice. Some few permit no modifications after sign up, but that is rare. Why do the companies allow this? Is it to get people more engaged, giving them a way to form some sort of bond with their avatar? Does it facilitate change and novelty so participants don't get bored? In any case, avatar customization is an accepted part of most virtual worlds.

STARTING OUT

Default avatars have been offered by Linden Lab since they started in 2003 and these choices have gotten progressively more sophisticated over the years. (See Figure 7-2 and Figure 7-3.) The original starting avatar choices offered by Linden Lab tended towards human or humanoid representations. Animals, vampires, robots, and even vehicles ("vehitars") are now given their own tabs in the pantheon of Linden Lab avatar selections.



Figure 7-2: Starting avatars offered by Linden Lab' *Second Life* world circa 2007. The first two were the original default avatars.

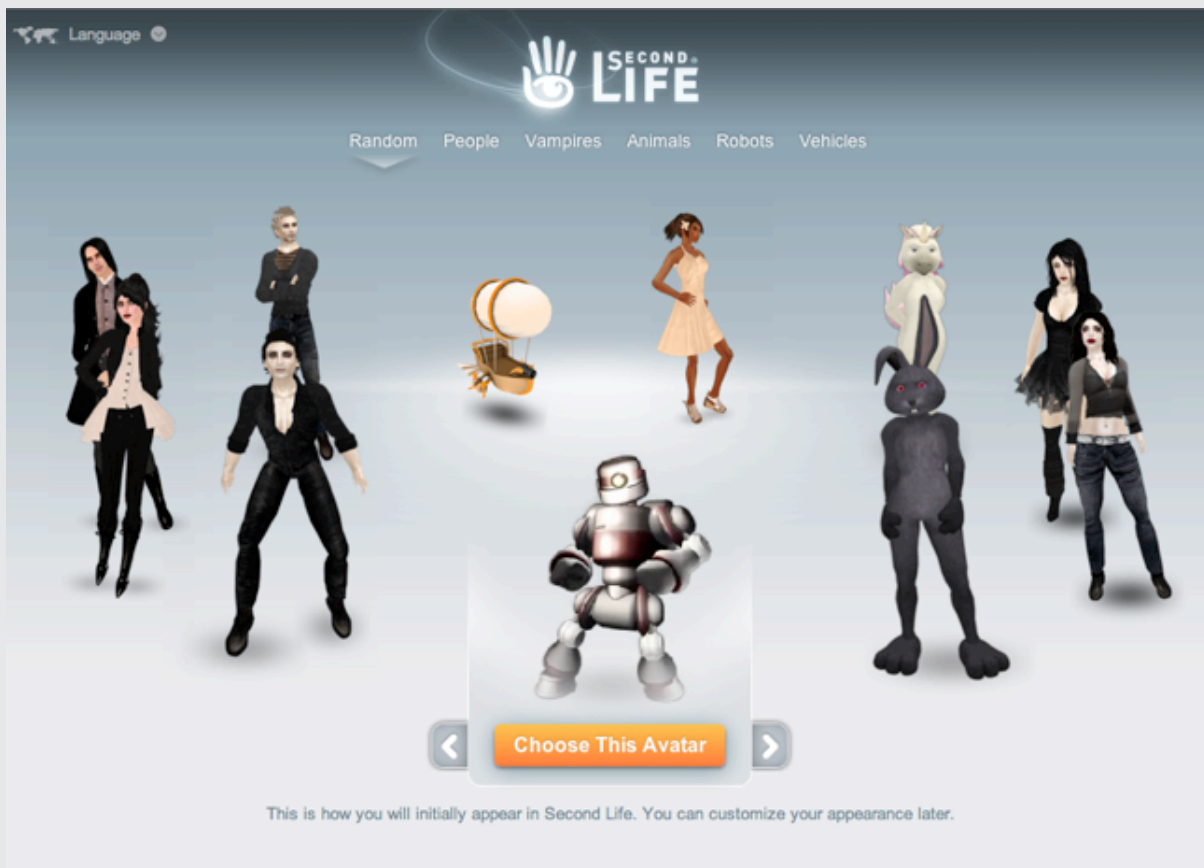


Figure 7-3: June 20, 2011 – Starting avatar choices on the *Second Life* site.

Once “in-world” one can choose to alter one’s appearance, which brings up an interactive menu of various parameters that can be changed, from facial shape to size of hands, to color of lips. These parameters are changed by means of sliders, and cumulatively offer about 200 configurable areas. (See Figure 7-4.)



Figure 7-4: Changing avatar appearance with sliders

Linden Lab also offers downloadable templates for users to be able to customize their skin (as well as to make clothes), which allows for more creative expression than the in-world sliders. One can paint a detailed skin for one’s face, upper body and lower body (3 separate templates) that can then be uploaded into the world as a texture to be wrapped onto the geometry of the avatars. (See Figures Figure 7-5 and Figure 7-6.)



Figure 7-5: Two variations on the same basic head (skin) template. While the cheek tattoo was placed within this 2D image, separate tattoo layers also can be used.



Figure 7-6: The two facial textures above mapped onto SL avatar head.

FOR SALE

If one does not want to remain with a starting avatar (which may carry the stigma of being a “noobie”), or go to the extra effort to customize one, a thriving market for “designer avatars” offered for sale has emerged. On June 4, 2011, the *Second Life* Marketplace listed over 17,000 complete avatars for sale, divided among Human female (2409), Human Male (1307), Human child (445), Sci Fi & Fantasy (4420), Furrries (2482) Animal (1411), and less than 1000 each for anime, monster and robot avatars. The “other” category numbered 1600 (lumped together) and included things like stick figures, cartoon characters, toys, and even vegetables and foods.

Avatar components (separating shape and skins, as shown in the figures above) are offered in even greater numbers. One can find over 10,000 avatar shapes offered, nearly 15,000 avatar skins, and another 4822 skin and shape combinations.¹¹ A complete avatar body (skin and shape) can be found on the SL Marketplace from one of the merchants or stores offering virtual goods for sale. The price for a complete well-designed avatar ranges between \$1000L and \$5000L – which might sound pricey if it was real world dollars, but at the average exchange rate its value in real life (IRL) is between \$4.00 and \$16.00.

If you decide to be part of a group in the virtual world, the style of your avatar may need to conform to group norms. For example, you can join a Steampunk group, or one that is devoted to role playing, such as the Goreans. Another category often found in *Second Life* is the tiny – a diminutive character that often inhabits its own specially built place with small furniture and houses, and other accouterments sized specifically for them. They have their own groups (as so many other “themed” avatars types). Groups that have formed around each category may range from one or two to hundreds. Tinies have over 500 groups and Furrries over 400, as noted in a recent search of the *Second Life* site for these groups.

BEING UNIQUE

If one does not want to fit into a group, there are many ways to create an avatar that is unique. As previously noted, one can create a custom avatar by playing with sliders within the program, and accoutering your form (tattoos, clothing, jewelry, hair, other appurtenances, etc.). You can transcend the inherent limitations via clever tricks, such as reducing the skeleton to the smallest size, folding it up and/or making portions invisible to portray a smaller character than a standard human one, such as a quadruped, a tiny, a butterfly or a most abstract form. This is actually the design procedure for making a tiny, mentioned previously. (See Figure 7-7)

A creator can even place transparent prims (basic 3D shapes) around an arm or a leg, to cause that appendage to disappear. A genie could thus be made to float above the ground, and a curling smoke texture (even animated!) could be placed on the prims to look like the genie’s body ends in a wisp of smoke. An entire default body shape could be masked in this way and overlaid with replacement shapes designed to create a fantastic character, like a flying spaghetti monster, or a tiny wispy fairy light. The range of personalized avatar that can thus be made is literally endless.

¹¹<https://marketplace.secondlife.com/products>



Figure 7-7: An avatar in two phases of transforming from a humanoid to a hedgehog shape.

Why choose something like this? What does it mean if I want to assume the persona of a hedgehog?¹² It might mean that I wish to be perceived as fun, or funny. I might want to try a different means of locomotion (rolling instead of walking or running). I may want to be cute or identify with some aspects of hedgehogness. Rather than being adopted to create a response in others, the reason may be more inwardly directed. For example, it may be a means to “walk in another’s shoes” for a time. In a stunning example of this, Micha Cárdenas took on the persona of a dragon by living as this avatar for 365 hours to emphasize the current requirement that a transgender candidate must live as their target gender for a period of time before any surgery can take place (Cárdenas, et al., 2009).

Many people, faced with the choice of creating an avatar, go with what they know best – what they have bonded to all their life – their own body and their own face. I, myself, have made hundreds of avatars for people, and over 90% of first timers request that the avatar resemble them as much as possible given the *Second Life* tools. Recalling the Neustader findings, the Realistics are the largest group. For these Realistics, having their own face would most likely be an important option.

¹²It should be stated that I purchased this avatar, made by Daryth Kennedy, and did not create it myself.

Mapping a photographic image of your own face onto the avatar's form does seem to be the ultimate personalization. Interestingly enough, Linden Lab does not offer a service to do this (though they do offer a do-it-yourself tutorial originally posted in 2005).¹³

One company, Cyberextruder, offered facial photo mapping as a service from approximately 2007 to 2010. From their website during this time, they advertised a service that took approximately 2 hours at \$150 per hour. They also set up services in-world on Avatar Island in *Second Life*.¹⁴ Ultimately though the business model proved unsustainable and currently their primary product offering is supports facial recognition for security.

Currently, Second Skin Labs offers what they call a Portrait Service “for your residual digital self image” where they show you how to take a series of take a high quality photos of yourself that they process and return it as an avatar skin for around \$400.¹⁵ The results are fairly convincing, to the extent they can be with a geometric structure that may or may not be customizable to truly reflect a person's facial shape. Matching the photo texture map to the shape obviously takes a bit more work.

It would seem, however, that is there really was a large demand for this photographically congruent avatar appearance, that many more companies would be offering it. This seems to underscore the fascination with creating a more idealized form with our digital self.

6. DISCUSSION

In relative terms we are still at the beginning of the age of virtual worlds. Even those who have inhabited graphical avatars from the earliest days (e.g. Habitat, 1985) have been in a virtual form for less than 30 years. For *Second Life* that time period is less than ten years.

In the experiments described above people did not live as their character for any length of time. Just as in the physical world we age and change, it is normal for one to alter one's appearance in the virtual world. First of all, the avatar we start with is rarely the one we inhabit for long, as with such an avatar we are perceived as being a newcomer to the world. Neustader calls this the “Social stigma of being a default avatar. Continuing to wear a starting avatar shape signals to everyone that you are a newbie and not all that invested in the world or your self within it (Neustader and Fedorovskaya, 2009).

In the culture of *Second Life* there is strong peer pressure to refine one's appearance beyond the basics, to express your commitment to the whole world. (There may be a community of avatars who all still relish their newbie bodies, but I have not found it!). Newbie-looking avatars are often taken under the wing of someone more entrenched in the world. An experienced avatar might take a newbie shopping, offer them better clothes, or help them find a hair style that appeals to them. This helps a newcomer learn the possibilities for appearance, and as one gets more acclimated to living digitally, allows significant change. As well, the software systems of virtual worlds gets updated and improved, which can lead to more sophisticated looks. There is also expanding creativity on the part of skin, shape and clothes designers, who are always perfecting their craft.

¹³<http://forums-archive.secondlife.com/109/e1/246059/1.html>

¹⁴www.cyberextruder.com

¹⁵<http://www.secondskinlabs.com/Portraits/aboutportraits.html>

In late 2011 Ars Avataria (run by Harper Bresford) launched a request for people to post pictures of how their *Second Life* avatar had changed over time.¹⁶ A quote from the Ars Avataria site states looking at the evolved avatar is “a very visual way to consider what’s essential in someone’s expression of themselves as an avatar, and what can be changed along the way.”

Looking at a series of temporal snapshots of an avatar, it is easy to see that most people have evolved their digital selves (some substantially) over time. Unlike physical life, we can improve visibly with age, as there are no effects from the vicissitudes and ravages of time to our virtual body. In the vast majority, however, it can be seen that there are central core elements that remains, so recognition is maintained. (See Figure 7-8.)

What does this new ability to manipulate our selves through avatars, and what that is doing to us needs to be explored in the coming years, as it is here to stay.



Figure 7-8: The author’s continually evolving avatar, ChingALing Bling, in *Second Life*

It is also true that people may have many avatars, either in different types of worlds or for different purposes in one. We don’t yet know the long term effects inhabiting multiple avatars might have on us. Sherry Turkle, in *Life on the Screen*, her tome on identity in the Internet age, argues that today, multiple identities are a basis for how we develop a sense of self (Turkle, 1995).

In its virtual reality, we self-fashion and self-create. What kinds of personae do we make? What relation do these have to what we have traditionally thought of as the “whole” person? Are they experienced as an expanded self or as separate from the self? Do our real-life selves learn lessons from our virtual personae? Are these virtual personae fragments of a coherent real-life personality? (Turkle: 180).

This idea that there is no single ‘fixed’ self or identity is one of the key tenets of postmodern theories of the self. Stuart Hall summarizes this idea of the fragmented postmodern self as follows:

We can no longer conceive of the ‘individual’ in terms of a whole, stable and completed Ego or autonomous, rational ‘self’. The ‘self’ is conceptualised as more fragmented and incomplete, composed of multiple ‘selves’ or identities in relation to the different social worlds we inhabit, something with a history, ‘produced’, in process. (Hall, 1996: 226).

¹⁶<http://www.flickr.com/photos/ruthlatour/6319678609/in/pool-1756997@N20/>

Avatars definitely represent some part of ourselves, and who we are becoming in post modern times. Some people may think that avatar use is a passing fad, but the numbers belie this assumption. Kzero, a British research firm focusing on use of social technologies including virtual worlds, have followed the increasing use of avatars by all ages in all virtual worlds over the past several years. Their latest statistics report that there are nearly 1.8 billion people playing in virtual worlds via an avatar form. (KZero, 2011) Even if we decrease that number by admitting that many people have more than one avatar, the numbers are still staggering. Even more astonishing is that just over 1 billion of these avatar inhabitants are between the ages of 5 and 15. Children are growing up immersed in this new paradigm. They will never know a life where they could not put on new avatars just as easily as we out on outfits.

Avatars should be considered a disruptive technology that both enables and demands new forms of communication. What these forms are, whether they are based on our appearance, internal software expressions, or being connected via brain signals, only by more research and use will we uncover, define, and give names to them.

The virtual body is our means to experience not only new digital realms, but also to interact with others in those spaces. We present the persona we wish others to see, within the means afforded by the virtual world. The correspondences to the physical world have been laid out above. The virtual has every potential to affect us in similar, profound ways.

As virtual reality philosopher Michael Heim comments "...when we put on our avatar, we also put off the habitual self. ... We shed our form like a changeling. We lay aside the illusory fixity of being a hard ego encapsulated in a shell of flesh (Heim, 1999: 12).

Experimentation with our numerous virtual social selves can be seen as part of a process of modern personal growth and development. We suspect, but do not yet know for sure, that virtual worlds and their affordances provide an impetus for corresponding changes in the individuals psyche and subjective understanding of themselves. Where past generations have been more or less aware of our home, work and play modes of being as being somewhat separate identities, today's youth are extremely aware that they have different parts to play and these each may require distinct mindsets. My Habbo Hotel avatar exists in a very different world than my Club Penguin one! These are the postmodern extensions to the self of old, and we cannot be sure what this means for our future relationship to the world and to others around us.

7. CONCLUSION

Avatars are simulations, but they are ones that can externalize our inner feelings in a powerful, visual way. It remains to be seen how avatars will fit into our lives in the coming years. The evidence is emerging that there may be something very powerful at work as they become a common part of who we are. Davey Winder, in *Being Virtual: Who you really are online*, relates stories that reveal how being an avatar in a virtual world has affected his life. After incurring devastating health issues and the subsequent breakup of his marriage, his life was at a nadir. But his online experiences were able to provide the means for a more positive outlook. "The realization that the identity I was developing online was someone I rather liked pre-empted a remarkable change in my real world circumstances" (Winder, 2008: 29). This kind of change could help many people come to better terms with themselves. There may be staggering potential packed in those 3D representations.

We have come a long way along the digital road of NVC, from the early days of emoticons to the full body appearance of one's avatar in a virtual world. What does our increasing participation in virtual worlds portend for our future selves? Could the disruptive technology of avatars become the vessels of how we

communicate our accumulated self to subsequent generations? Ray Kurtzweil embraces a concept called the singularity – that mythical point in time where we will be able to save our brains – our essence – into robots or their future ilk (Kurtzweil, 2005). I can foresee a day, rather, when we have figured out how to download our memories, thoughts and experiences into an avatar form that can live on after us. This may not be our consciousness living on in full, but I believe, rather than a singularity in our future, we will reach a multiplicity. This is when our multiple avatars representations have become so thoroughly us, and we them, that our essence remains in their crucibles after our deaths.

8. BIBLIOGRAPHY

- Abrams, D., & Hogg, M. A. (Ed.). (1990). *Social identity theory: Constructive and critical advances*. New York: Springer-Verlag.
- Anderson, P. A. (2008) *Nonverbal Communication: Forms and Functions Second Edition*. Waveland Press, Inc. Long Grove, IL.
- Antonijevic, S. (2008) From Text to Gesture Online: A microethnographic analysis of nonverbal communication in the *Second Life* virtual environment. *Information, Communication & Society* 11 (2): 221-238
- Au, J. W. (2010) *Second Life* Survey Results: Majority Choose Non-Realistic Avatar Types in New World Notes. Available online at <http://nwn.blogs.com/nwn/2010/05/second-life-avatar-survey-results-majority-choose-nonrealistic-avatar-types.html> (accessed December 20, 2010).
- Barthes, R. (1993) *Camera Lucida: Reflections on Photography*. London: Vintage.
- Bell, C. (1997). *Ritual Perspectives and Dimensions*. Oxford University Press, New York.
- Bell, Q. (1976) *On Human Finery*. Hogarth Press, London.
- Berger, C.R., & Calabrese, R.J. (1975). Some explorations in initial interaction and beyond: Toward a developmental theory of interpersonal communication. *Human Communication Research*, 1, 99-112.
- Birdwhistle, R. (1970) *Kinesics and Context: Essays on Body Motion Communications*. University of Pennsylvania Press, Philadelphia, PA.
- Black, E. W., Ferdig, R. E., DiPetro, J. C., Liu, F., and Whalen, B. (2009) Visual Analysis of the Creation of Avatars. *International Journal of Gaming and Computer-Mediated Simulation*. 1 (1) IGI Global: 90-107.
- Breward, C. (2000) Cutures, Identities, Histories: Fashioning a Cultural Approach to Dress. In White, N. and Griffiths, I. (eds). *The Fashion Business: Theory, Practice, Image*. Berg, Oxford: 23-36.
- Burgoon, J. K., & Hoobler, G. (2002). Nonverbal signals. In M. L. Knapp & J. Daly (Ed.), *Handbook of interpersonal communication* (pp. 240-299). Thousand Oaks, CA: Sage.
- Cárdenas, M., Head, C., Margolis, T., and Greco, K., (2009) Becoming Dragon: a mixed reality, durational performance in *Second Life*. In Proceedings of the SPIE Vol. 7238: The Engineering Reality of Virtual Reality 2009, Ian E. McDowall; Margaret Dolinsky, Editors.
- Cialdini, R. B. (2001) *Influence: Science and Practice*. Boston, MA: Allyn and Bacon.
- Cleland, K. (2008) Image Avatars: Self-Other Encounters in a Mediated World. Unpublished PhD Thesis University of Technology, Sydney.
- Conrad, M., Neale, J. and Charles, A. (2010) This is my Body: the Uses and Effects of the Avatar in the Virtual World. *International Journal for Infonomics (IJI)*, 3 (4): 4.
- Dion, K. K., Berscheid, E., and Walster E. (1972) What is beautiful is good. *Journal of Personality and Social Psychology*, 24: 285-290.
- Eckman, P. (2006) *Darwin and Facial Expression: A Century of Research in Review*. Malor Books.
- Entwistle, J. (2000) *The Fashioned Body: Fashion, Dress and Modern Social Theory*. Polity, Cambridge.
- Fiske, S. T., & Taylor, S. E. (1991) *Social Cognition (Second Edition)*. McGraw-Hill, NY.

- Fiske, S. T. (1992) Thinking is for doing: Portraits of social cognition from Daguerreotype to laserphoto. *Journal of Personality and Social Psychology*, 63 (6): 877-889.
- Fron, J., Fullerton, T., Morie, J. & Pearce, C. (aka Ludica) (2007). "Playing Dress-Up: Costume, roleplay and imagination." *Philosophy of Computer Games Online Proceedings*, January 2007.
- Fox, J. and Bailenson, J. (2009) Virtual Self-Modeling: The Effects of Vicarious Reinforcement and Identification on Exercise Behavior. *Media Psychology*, 12: 1-25.
- Goffman, E. (1973) *The Presentation of the Self in Everyday Life*. New York: Overlook Press.
- Goffman, E. (1981) *Forms of Talk*. University of Pennsylvania Press, Philadelphia.
- Gratch, J. (2010) Emotionally resonant media: Advances in sensing, understanding and influencing human emotion through interactive media in *Handbook of Emotions and the Mass Media (Chapter 21)*, Katrin Döveling, Christian von Scheve, and Elly A. Konijn, Eds. Routledge, NY: 385-402.
- Hall, S. (1996). The Meaning of New Times. In *Stuart Hall: Critical Dialogues in Cultural Studies*. D. Morley & K.-H. Chen (Eds.), Routledge, NY.
- Haxby, J. V. and Gobbini, M. I. (2011) Distributed Neural Systems for Face Perception, In *Oxford Book of Face Perception*, Andrew J. Calder, Gilliam Rhodes and Mark Johnson, eds. Oxford University Press: 93 -110.
- Heim, M. (1999). Transmogrification. Retrieved 5 July 2005, from <http://www.mheim.com/files.transmog.pdf>
- Irani, L., Hayes, G. R., Dourish, P. (2008) Situated Practices of Looking: Visual Practice in an Online World. (Proceedings of CSCW 2008, San Diego, CA. ACM Press.
- Jones, S., & LeBaron, C. (2002). Research on the Relationship Between Verbal and Nonverbal Communication: Emerging Integrations: Emerging Integrations. *Journal of Communication*, 52, 3, 499-521.
- Kelley, G. (1955) *The psychology of personal constructs*. Vol. I, II. Norton, New York. (2nd printing: 1991, Routledge, London, New York)
- Kendon, A. (1977) *Studies in the Behavior of Social Interaction* Bloomington: Indiana University Press.
- Kendon, A., editor (1981) *Nonverbal Communication, Interaction and Gesture: Selections from Semiotica (Vol.41, Approaches to Semiotics)*. Mouton and Co., The Hague.
- Kurtzweil, R. (2005) *The Singularity Is Near: When Humans Transcend Biology*. The Viking Press.
- Kzero (2011) *Virtual Worlds: Industry and User Data: Universe Chart for Q4, 2011; Slide 7*. Available online at <http://www.slideshare.net/nicmitham/kzero-universe-q4-2011>. Accessed Jan. 20, 2012.
- Lacan, J. (1953) Some Reflections on the Ego, In *Écrits: The First Complete Edition in English*, transl. by Bruce Fink, 2006, W.W. Norton & Co. New York.
- Martey, R. M. and Consalvo, M. (2011) Performing the looking-glass self: Avatar appearance and group identity in *Second Life*. *Popular Communication*, 9 (3): 165-180.
- Merleau-Ponty, M. (1964) *The Primacy of Perception: and other essays on phenomenological psychology, the philosophy of art, history, and politics*. Evanston, Ill.: Northwestern University Press.
- Messinger, P. R., Stroulia, E., Lyons, K., Bone, M., Niu, R. H., Smirnov, K., and Perelgut, S. (2009). Virtual world - past, present, and future: New directions in social computing. *Decision Support Systems*. 47: 204- 228.
- Neisser, U. (1987). Introduction. In U. Neisser (Ed.), *Concepts and conceptual development: Ecological and intellectual factors in categorization*. Cambridge University Press. Cambridge: 1-10.
- Nestor, A., Plaut, D. C., and Behrmann, M. (2011) Unraveling the distributed neural code of facial identity through spatiotemporal pattern analysis. *Proceedings of the National Academy of Sciences (PNAS)* June 14, 2011 108 (24) 9998-10003.
- Neustaedter, C. and Fedorovskaya, E. (2009) Presenting Identity in a Virtual World through Avatar Appearances, In *Proceedings of Graphics Interface*, ACM Press.
- Pease B., Pease A. (2004). *The Definitive Book of Body Language*. New York, NY: Bantam Books.

- Philpott, J.S. (1983) The Relative Contribution to Meaning of Verbal and Nonverbal Channels of Communication: A Meta-analysis. Unpublished master's thesis, University of Nebraska, Lincoln.
- Rosch, E. (1978). Principles of categorization. In E. Rosch and B. B. Lloyd (Eds.), *Cognition and categorization*. Lawrence Erlbaum Associates Hillsdale, NJ: 27-48.
- Schomaker, K. (2011a) *1000 Avatars Vol. 1*. Self-Published. Available at <http://www.blurb.com/books/2289173>
- Schomaker, K. (2011b) *1000 Avatars Vol. 2*. Self-published. Available at <http://www.blurb.com/books/2828775>
- Shannon, C.E. & Weaver, W. (1949). *The mathematical theory of communication*. University of Illinois press, Urbana, IL
- Suler, J. R. (2002) Identity Management in Cyberspace. *Journal of Applied Psychoanalytic Studies*, 4 (4): 455-459.
- Taylor, T. L. (2004) The Social Design of Virtual Worlds: Constructing the User and Community through Code. In M. Consalvo and Kate O'Riordan (eds) *Internet Research Annual Volume 1: Selected Papers from the Association of Researchers Conferences 2000-2002*. New York: Peter Lang.
- Tseëlon, E. (1992) Self-Presentation through Appearance: A Manipulative vs. A Dramaturgical Approach. *Symbolic Interaction* 15 (4): 501-513.
- Turkle, S. (1995). *Life on the screen: identity in the age of the Internet*. New York: Simon & Schuster.
- Twigg, J. (2007) Clothes and Age: A critical review. *Ageing and Society* 27:285–305
Cambridge University Press.
- Verhulsdonck, G. and Morie, J.F. (2009) Virtual Chironomia: Developing Standards for Non-verbal Communication in Virtual Worlds. *Journal of Virtual Worlds Research* 2(3): Technology, Economy, and Standards. Available online at <http://journals.tdl.org/jvwr/article/viewArticle/657>
- Willis J, and Todorov A (2006) First impressions: Making up your mind after 100 ms exposure to a face. *Psychological Science* 17: 592–598.
- Winder, D. (2008) *Being Virtual: Who you really are online*. John Wiley and Sons. UK.
- Yee, N. & Bailenson, J. (2007) The Proteus Effect: The effect of transformed self-representation on behavior. *Human Communication Research*, 33 (3), 271-290.
- Zebrowitz, L. A. (1996). Physical appearance as a basis of stereotyping. In C. N. Macrae, C. Stangor, and M. Hewstone (Eds.), *Stereotypes and Stereotyping*. (pp. 79-120) New York: The Guilford Press.
- Zebrowitz, L.A. (2003). Overgeneralization effects in perceiving nonverbal behavior: Evolutionary and ecological origins. *Journal of Nonverbal Behavior*, 27, 133-138.
- Zebrowitz, L.A. (2004). The origins of first impressions. *Journal of Cultural and Evolutionary Psychology*, 2, 93-108.
- Zebrowitz, L. A. and Montepare, J. M. (2008) Social Psychological Face Perception: Why Appearance Matters. *Social and Personality Psychology Compass*. 2(3): 1497-1517.
<http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2811283/>

8

TIMETRAVELLER™: FIRST NATIONS NONVERBAL COMMUNICATION IN SECOND LIFE

By Elizabeth LaPensée and Jason Edward Lewis

INTRODUCTION

In order for Indigenous knowledge to survive in a post-colonial context, Indigenous peoples must both know and understand Western culture (Mihe-suah, 2003). This understanding includes how to use Western technology. For our current purposes, we are particularly interested in how the ability to manipulate digital technology can be used to reify Indigenous knowledge in virtual spaces that have the promise to be radically intercultural. Our approach to materializing Indigenous knowledge involves sharing Indigenous stories, whether traditional or contemporary, with both Indigenous and non-Indigenous peoples via digital networks. Such sharing requires that Indigenous storytelling practices, including nonverbal communication components, be adapted for new contexts.

Contemporary Art curator Candice Hopkins (Métis/Tlingit) describes Indigenous stories as continually changing, individualized and communal, original and replicated, authored and authorless (Hopkins, 2006). Hopkins employs Victor Masayesva's explication of Indigenous aesthetic—measured by the artwork's ability to subvert colonization—as a way to frame Indigenous creation in cyberspace. Drawing from Dana Claxton (2005), we can also see the futility of distinguishing between “traditional” and “contemporary” styles, given Indigenous art's situation within a context of constant change, hybridity, and timelessness. The stories told by Indigenous new media artists such as Skawennati (Fraguito, 2010), Archer Pechawis (2011), Cheryl L'Hirondelle (2011), and Ahasiw Maskegon-Iskwew (2003) incorporate these characteristics as they seek to occupy, transform, appropriate and reimagine cyberspace. Virtual worlds like Linden Lab's *Second Life*, which allow users to constantly generate new environments and perform within them, offer an ideal space for Indigenous representation in a digital context.

Accounting for nonverbal communication is essential to diverse and robust cultural representation in virtual worlds. Examining user-generated assets such as animations, clothing, and objects, as well as the placement and use of avatars can help inform intercultural communication in virtual spaces. Specifically, considering the transference of First Nations traditional and contemporary nonverbal communication into virtual worlds can be useful in future design of and research of First Nations representation in a virtual context.

Our case study for considering the role of nonverbal communication in virtual worlds focuses on *TimeTraveller*™, a machinima series and Alternate Reality Game (ARG) in *Second Life* created by Mohawk artist Skawennati. *TimeTraveller*™ is a project by Aboriginal Territories in Cyberspace (AbTeC)—a research network of artists, academics and technologists centrally concerned with Indigenous representation in digital media. AbTeC investigates innovative methods for Indigenous peoples to participate in networked culture to tell our stories, and in so doing, strengthen our communities and participate actively in shaping cyberspace. Jason Edward Lewis and Skawennati co-direct AbTeC and Elisabeth LaPensée is a Research

Assistant who writes for the *TimeTraveller*™ ARG. Although we have not directly participated in the creation of nonverbal communication assets in *TimeTraveller*™, we have in-depth knowledge of the project from personal communication with Skawennati over years of development.

We first look broadly at nonverbal communication in culture and nonverbal communication in virtual worlds. We then discuss the background, motivation, process, assets, and challenges of *TimeTraveller*™. The ongoing development of *TimeTraveller*™ offers insight into the possibilities and importance of culturally based nonverbal communication in virtual worlds.

1. NONVERBAL COMMUNICATION

NONVERBAL COMMUNICATION IN CULTURE

Definitions of culture are as numerous as the number of different cultures themselves. All definitions of which we are aware, however, have a central claim that culture is learned—rather than biologically inherited—through a process that involves assigning symbolic meanings (Jandt, 2010). Cultures are distinguished by how and to what they assign these meanings. Intercultural communication both concerns the internal communication of an individual culture and cross-communication between different cultures. The integral parts of intercultural communication include perception (e.g. beliefs, values, attitudes, worldview, and social organization), verbal processes (e.g. verbal language and patterns of thought), and nonverbal processes (e.g. nonverbal behavior) (Samovar & Porter, 1991).

Nonverbal communication most generally refers to wordless communication, including gesture, body language, facial expression, intonation of speech, and clothing (Innocent & Haines, 2007). Communication scholars Samovar and Porter (1991) divide nonverbal communication into four categories: (1) kinesics, (2) proxemics, (3) paralanguage, and (4) chronemics. Kinesics refers to body movements (or body language) made during communication, such as facial expressions, eye contact, hand gestures, and touch. Proxemics refers to the use of space during communication, including the range from architecture and furniture to the distance between communicators. Paralanguage includes all sounds made with voices that are not words, such as laughter, tone, and pacing. Chronemics relates to how time is used in communication, such as perceptions of past, present, and future as well as the literal passing of time during communication.

The intercultural communication research of anthropologist Hall (1959), later developed by social psychologist and anthropologist Hofstede (1983), identifies a difference in nonverbal communication between “low context” and “high context” cultures. Low context cultures rely on individual value orientation, line logic, direct verbal interaction, and individualistic nonverbal style. High context cultures function on group value orientation, spiral logic, indirect verbal interaction, and contextual nonverbal style. Intentions and meanings in high context cultures are interpreted within a larger shared knowledge.

Indigenous cultures tend to be high context cultures, whereas Western cultures tend to be low context cultures (Hall, 1976). As with any culture, nonverbal communication is intrinsically linked with all modes of communication. Native American historian Donald L. Fixico suggests that looking at nonverbal communication in the oral storytelling tradition informs our understanding of Indigenous communities and their thought processes (Fixico, 2003). Indigenous languages are, mostly, verb-rich process and/or action-oriented languages that rely heavily on nonverbal communication (Little Bear, 2000) as well as high context environments. Further, words and hand signals usually describe “happenings” rather than individual objects (Little Bear, 2000).

It is important to note that some nonverbal communication is shared among Indigenous peoples, but often each Indigenous tribe/nation/band has its own unique set of nonverbal communications. For example, Cherokee, Navajo, and Hopi, among others, make minimal eye contact to express that they are listening fully (Chiang, 1993). While Kiowa point with their lips (Kirch, 1979), Mohawk point with their chins.

As anthropologist Claude Lévi-Strauss has noted and advocated against, the difference in contextual importance combined with colonialist dynamics have contributed to the devaluing of Indigenous cultures by Western cultures (Samovar & Porter, 1991). This devaluing privileges colonialist perspectives over Indigenous perspectives (Smith, 1999), while the discrediting of oral and nonverbal communication traditions requires that Indigenous peoples actively work to counteract misrepresentations that arise due to the exclusive valorization of written or recorded communication methods (Chamberlin, 2000). Self-representation by Indigenous individuals of nonverbal communication in virtual worlds is one method to encourage intercultural communication with diverse communities. It is also a political act designed to reassert the value of such in the real world.

NONVERBAL COMMUNICATION IN VIRTUAL WORLDS

Virtual worlds, which act as persistent, avatar-based social spaces, afford users with opportunities for intercultural communication and self-expression (Thomas & Brown, 2009). Users share the same space, see physical representations of each other, and communicate and act in the shared space through both verbal and nonverbal means. Nonverbal communication in real world spaces is learned as part of one's cultural literacy in a deep and continuous way, such that participants are often not aware of their performance of it. In virtual worlds, however, nonverbal communication by players through their avatars requires *intentional acts* on the part of the user (Verhulsdonck & Morie, 2009). More precisely, nonverbal communication in virtual worlds is "stylized through animation, appearance, and performance of the player's avatar" (Innocent & Haines, 2007). Avatars and objects may have altered appearances, be embedded within the space, react to input and interaction, and be created, combined, and modified (Innocent & Haines, 2007). These actions all result from users choosing to design the world in which they are participating.

Most virtual worlds are capable of supporting all four of Samovar and Porter (1991)'s forms of nonverbal communication, as well as an additional category, appearance, added by Kujanpää and Manninen (2003). Kinesics (body language) is realized through animations with particular emphasis on facial animations (Kujanpää & Manninen, 2003). Proxemics (use of space) can change from moment-to-moment based on an individual avatars' placement in space but can also be statically fixed for objects such as furniture or buildings. Paralanguage (sounds) are integrated on an individual avatar basis. Chronemics (use of time) often reflect the users' perception of time acted through avatars (Halbert & Ghosh, 2008). Appearance involves customizing the physical aspects of avatars (e.g. hair color and style, skin tone, eye color) and wearing or using items (e.g. clothing and accessories).

Second Life, in particular, supports all of the above forms of nonverbal communication as well as text and voice chat (Ryzmaszewski, et al, 2006). Support includes preset assets provided to every participant as well as enabling users to create their own assets using internal or third party tools.

Second Life contains numerous user-generated assets that can be used for nonverbal communication. Thousands of skins, shapes, hairstyles, tattoos, jewelry, clothes, props and animations are available for sale by independent creators. Photoshop templates for basic clothing are freely downloadable from *Second Life's* website, as are dozens of tutorials for learning how to create assets on your own. The world supports the creation of more complex non-verbal communication assets. For example, researchers interested in integrating haptic (touch) interactions in *Second Life* created a haptic-jacket system as an add-on to the communication channel (Hossain, et al, 2010). The haptic-jacket allows avatars to activate animations such as "encouraging pat" and "comforting hug." Similarly, a wearable humanoid robot recognizes nine

emotions from text chat in *Second Life* and haptically augments the user's emotionally immersive experience (Tsetserukou & Neviarouskaya, 2010). The hope of such projects is to enhance user interaction and immersion in virtual worlds communication.

Communication scholar Smiljana Antonijevic, who spent six months conducting an ethnographic research study of nonverbal communication in *Second Life*, found a significant difference between predefined and user-defined nonverbal communication (Antonijevic, 2008). The use of predefined nonverbal communication—meaning nonverbal acts generated by the system—was often not related to avatar physical appearance or any co-occurring textual discourse. Notably, nonverbal communication assets made by users were tied closely to an avatar's physical characteristics (e.g. skin tone, height, shape) as well as the specific local context.

By virtue of its global reach, *Second Life* provides rich opportunities for intercultural encounters and exposure to unique cultural perspectives (U-Mackey, 2011). The challenges are many in designing effective nonverbal communications for exchanges where one participant might be logging in from China, another from Egypt, and a third from the Cherokee Nation (Oklahoma). The *TimeTraveller™* project provides an interesting case study for understanding those challenges, devising design approaches appropriate for surmounting those challenges, and generally considering how culturally diverse assets for nonverbal communication can be effectively integrated into the host environment.

2. TIMETRAVELLER™ CASE STUDY

BACKGROUND

Skawennati is an artist whose work addresses history, the future, and change. She has been working in new media since 1996, beginning with the pioneering online exhibition and chat space, *CyberPowWow* (1997). *Imagining Indians in the 25th Century* (Fragnito, 2001); *Thanksgiving Address: Greetings to the Technological World* (with Jason Edward Lewis, 2002); and *80 Minutes, 80 Movies, 80s Music* (2002-present), have been widely exhibited. Her awards include imagineNATIVE's 2009 Best New Media Award and a 2011 Eiteljorg Fellowship for Native American Fine Art.

Skawennati initiated *TimeTraveller™* in 2008. It is a multi-platform project featuring a website (www.TimeTravellerTM.com), a machinima series, and an Alternate Reality Game. The story revolves around Hunter, a young Mohawk man living in the 22nd Century. Hunter possesses an impressive range of traditional skills, but finds difficulty fitting in to an "overcrowded, hyperconsumerist, mediated world" (Fragnito, 2011). With the help of his edutainment system, *TimeTraveller™*, he embarks on a technologically-enabled vision quest that takes him back in time to historical conflicts that have involved First Nations. Along the way, he meets Karahkwenhawi, a young Mohawk woman from our (2011) present, whose unique perspective on Aboriginal issues deeply affects him. Together, they discover the complexity of history, place, culture, and their place within it.

The *TimeTraveller™* machinima series is based in *Second Life*. *Second Life* is a popular platform for producing machinima, or 'machine cinema', a form of animated film created using game engines or virtual worlds (Marino, 2004). Machinima has been used for education (Carroll, 2005), art (Picard, 2007), activism (Jones, 2007), and entertainment. The *TimeTraveller™* Alternative Reality Game (ARG)—like all ARGs, a story-driven game that utilizes the real world and virtual world as the play space—serves to extend the machinima's storyline. Together, the machinima episodes, the ARG, and the website encourage players to learn First Nations history while offering a shared experience for a diverse community of players consisting of Indigenous and non-Indigenous youth, artists, *Second Life* enthusiasts, and history buffs.

The *TimeTraveller*[™] story begins at TimeTravellerTM.com (Fragnito, 2009). The website is presented as if it was made by an “edutainment” company in the 22nd Century to promote its latest product, “TimeTraveller[™],” and appears in our present by means of a rift in space-time. Resembling a pair of futuristic sunglasses, the TimeTraveller[™] device creates perfect 3D renderings of real-life locations and their inhabitants from any time in history. Through the “patented” History HUD (Heads Up Display), users can interact with these historical events in a fully-immersive experience that feels like one has actually travelled back in time.

The website also hosts the machinima episodes. In the future context, Hunter has won a contest to identify the most extreme and exciting user of the technology, and, as the winner, he is starring in a reality TV series that follows him on the adventures he undertakes with the TimeTraveller[™] device. The machinima episodes are presented as installments in this series.

SECOND LIFE

Skawennati was drawn to *Second Life* in part because it promised a rich set of tools for creating and supporting nonverbal modes of communication. She initially evaluated *Second Life* as a production platform due to its similarities with *The Palace*, the technology she used to host four iterations of the *CyberPowWow* online gallery from 1996 to 2004. *The Palace* was a virtual 2D chat space that allowed users to create avatars and backgrounds using 2D photos and art (Time Warner Interactive, 1995). *Second Life* incorporates the shared chat functionality of *The Palace* while extending the virtual environment into the third dimension. Skawennati saw that *Second Life* support for creating and customizing 3D avatars, building and scripting objects, and leveraging a hybrid economy (real money exchangeable for virtual money and vice versa) to exchange objects, materials, and land (Rymaszewski, et al, 2008) was a natural evolution/progression of her work in virtual worlds.

Given that much of the *TimeTraveller*[™] story is based on historical events, the series requires very distinct sets, props, wardrobes, hairstyles and actor-avatars (skins and shapes) that look like they come from a specific era. In 2008, when a user opened a standard *Second Life* account, she received a small library of assets including several choices of avatars, some basic gestures (which have not changed very much), some sounds (like a kissing sound), a few textures, and a list of vendors that gave away free assets to assist with the fact that, when she started out (then, as now), she was penniless. The user discovers, however, that the world is filled with thousands of user-generated assets that can be purchased.

However, Skawennati and her team soon found that it was quite difficult to find existing assets that were (1) appropriate for Indigenous representation and (2) would work within the context of the *TimeTraveller*[™] storyline. After a series of exhaustive but fruitless searches of the whole range of both built-in and user-generated assets, they decided that the only way to proceed was to custom-create the appropriate assets. As a result, the team produced a whole series of assets for use in nonverbal communication, either as part of the machinima episodes or in the ARG.

The resulting nonverbal communication assets are influenced by the *TimeTraveller*[™] story’s particular emphasis on portraying Indigenous peoples as technologically adept to an advanced degree. Further, the design of the assets stem from the individual perspective and choices of Skawennati, an urban-based, reserve-adjacent raised, half-Italian Mohawk who is quite conscious of speaking from that context. Consequently, the assets are not intended to encompass or represent all Indigenous peoples.

However, looking closely at the creation and use of the nonverbal communication assets for *TimeTraveller*[™] can be quite useful in informing the design of assets for the representation of other Indigenous peoples, both in *Second Life* and in other virtual worlds. Even more generally, the project’s cultural design process has implications for the importance of nonverbal communication for enriching and broadening the diversity of cultural representation in virtual environments of all kinds.

PROCESS

TimeTraveller™ originated as the brother piece to *Imagining Indians in the 25th Century* (Fraguito, 2000). *Imagining Indians* draws on two traditionally “feminine” pastimes: paper dolls and journaling. In contrast, Skawennati envisioned *TimeTraveller*™ as a first-person shooter, which is archetypically considered “masculine”. Skawennati chose *Second Life* in part because it offered the first-person point of view. Secondly, it offered the built-in physics of teleporting and flying, which was exactly what her time-travelling, jet-packing characters required.

Prior to working in *Second Life*, Skawennati had positive experiences with taking existing technologies and customizing them. She found *Second Life* to be highly customizable. She and the team spent time self-teaching using hands-on experience as well as community and company support. The process of making assets and creating machinima was initially slow due to the newness of making and shooting in *Second Life*, particularly given that there existed very few resources on how to do it. The team learned by trial and error. The workflow sped up in later episodes when challenges were resolved.

The workflow developed by Skawennati’s team incorporated a unique development process for each class of nonverbal communication asset. Animations are made using QAvimator (open-source animation software developed for use with *Second Life*). Clothing is made in Photoshop (image editing software) using the Linden Lab’s templates. Textures are made from photographs of real materials, and then post-processed in Photoshop. Simple objects (e.g. walls, floors, benches) are built in *Second Life* using the provided tools. Complex objects (e.g. the *TimeTraveller*™ glasses) are built in and exported from Maya (commercial 3D modeling and animation software).

3. NONVERBAL COMMUNICATION IN TIMETRAVELLER™

Most of the user-generated nonverbal communication in *TimeTraveller*™ can be divided into two categories: (1) appearance and (2) performance. “Appearance” includes assets such as skin tone, clothing, hairstyles, and objects. “Performance,” or kinesics, includes body language such as gestures, facial expressions, touch, gaze, and posture (Allmendinger, 2010). For example, the team made assets such as whispering animations for the crowd in the futuristic powwow scene of Episode 04. Nonverbal communication in *TimeTraveller*™ is not limited to these categories, however. The machinima series also uses paralanguage, proxemics, and chronemics. Paralanguage (sound) appears frequently in the actors’ voiceovers, including emphasis, volume, and intonation. Autonomous forms of paralanguage such as laughter and sighing accompany default in-world animations. Chronemics (use of time) is incorporated in voiceovers and the representation of time in the story. For example, voiceover actors use pauses during speech or take deep breaths. Proxemics (space) is used in the placement of avatars in relation to each other and the sets. An example is the face-off scene in Episode 03, where the team could not use existing assets to get the Canadian soldier avatar and Warrior avatar to stand extremely close without touching. They ended up creating poses, which are like animations but static, so that the avatars would not overlap each other when they breathed.

For the purposes of the case study, we will focus on the categories of appearance and performance. These two categories describe the majority of user-generated nonverbal communication assets for First Nations representation both in *TimeTraveller*™ and in *Second Life* in general.



Figure 8-1: Hunter's distressed Iroquois Confederacy t-shirt.



Figure 8-2: Traditional ribbon shirt.

APPEARANCE

Clothing provides ongoing individual self-expression for avatars in the *TimeTraveller*[™] machinima and ARG. For Hunter (the main character), Skawennati designed a collection of “distressed” t-shirts, all black and each featuring a First Nations symbol that would have differing levels of recognition to different viewers. In Episode 01, he wears one with an Iroquois Confederacy (Haudenosaunee) symbol, supporting his stated nationality (Mohawk), but also possibly indicating his rejection of Canada’s jurisdiction over him, his people, and their territory [Figure 8-1]. In Episode 02, his shirt features a twice-bisected circle representing the Four Directions. Many viewers know this symbol as a representation of sacred Anishnaabe teachings. Skawennati’s rendition of the symbol is very modern, and followers of her practice might recognize it from her 2000 work, *Imagining Indians in the 25th Century*, where a character that visits the future as an Olympic athlete wears it and it is displayed as the symbol that has been adopted as Canada’s flag. His t-shirt in Episode 03 depicts a stylized turtle that refers to Turtle Island. The decision to make the symbols “distressed”—or worn-looking—was to convey two things: the venerability of the symbols, and Hunter’s status as an outsider, not fitting in to the society around him.

Pinky, a Mohawk activist who appears in Episode 03 during the Oka Crisis of 1990, wears t-shirts that feature 1980s bands, indicating both her age and her unsterotypical—for a Native woman— taste in music. The Mohawk Warrior motif, which has been adopted by many Indigenous groups as a symbol of resistance, is seen on t-shirts worn by a warrior in Episode 03, the four-year old version of Karahkwenhawi in Episode 03, and a punk band drummer in the future powwow scene in Episode 04. Ribbon shirts, which are worn by many First Nations and Native Americans at special occasions including ceremonies and powwows, adorn various avatars [Figure 8-2].



Figure 8-4: Long braided hair for male avatars



Figure 8-5: Hunter wearing the TimeTraveller™ glasses

Jingle dresses were originally worn by Anishinaabe women for healing dances and during the last few decades have been adapted to powwow culture. The “jingles” are metal cones attached to the fabric and are unique in both their appearance and in the sound that they make. They appear in an array of colors in the futuristic powwow scene in Episode 04 [Figure 8 3]. The haute couture “Ovoid” gowns from the same episode were designed by Skawennati and her team to reflect a flourishing trend in the contemporary Aboriginal world to integrate traditional textiles, styles and symbols into high fashion. The ovoid, which is one of the basic “building blocks” of West Coast imagery/design, is used to show how the future powwow embraces many Nations, including those that historically did not participate in powwows.

Hair selection proved to be especially challenging. Braided hairstyles are almost non-existent, probably because it is not possible to make them “flexi” (the in-world building tool that gives otherwise rigid objects flexibility). Hair assets were purchased for the most part. However, the team made long braided hair for one male avatar that needed a traditional hairstyle [Figure 8-4], which, unfortunately, was quite rigid.

Objects were also important for conveying the Native context of the locations and situations of the scenes. In Episode 01, Hunter’s weapons wall includes a tomahawk and bow and arrow (both purchased). In Episode 03, a shell and a custom-made sage bundle are used in a ceremony. The team made a wampum belt—traditionally woven with beads, sinew and leather and used as a mnemonic device to keep account of treaties and contracts. The TimeTraveller™ glasses worn by Hunter and later used by Karahkwenhawi in Episode 04 serve to show futuristic advanced technology in First Nations hands [Figure 8 5]. Karahkwenhawi, who lives in our present, also carries an iPhone. Hunter, by comparison, flies using a jetpack in his 2121 world.

PERFORMANCE

Performance is most noticeable in full body movement, which was primarily customized in Episode 04. Karahkwenhawi is seen filming with her iPhone in a church. After she activates the TimeTraveller™ glasses, she joins a futuristic powwow where avatars are drumming and singing while competitors jingle dance in their dresses.

Karahkwenhawi has several hand gestures, including waving her hands in front of her face, answering the iPhone, putting on the glasses, and pressing the button on the glasses. Hunter shares the gestures required for the glasses but also uses specialized gestures including taking off his jetpack. The animations fight against stereotypes that First Nations people are not technology-capable.

In Episode 03, hand gestures from other avatars include gesturing to the flag on the treatment center, gesturing to the trees while explaining, and waving hands at the television. Hand gestures that acknowledge the subject of communication are especially important in First Nations nonverbal communication. Specialized hand positions also had to be made for holding the wampum belt and smudge shell [Figure 8-6]. The positions are meaningful since ceremonies have particular protocols.



Figure 8-6: The smudge shell and bundle of sage used in Episode 03.

Head gestures were made mainly to replace the default *Second Life* animations that were too exaggerated for *TimeTraveller*™. In Episode 03, a female warrior shakes her head. The team also made three variations on what they referred to as a “non-spastic nod of agreement.” Most notably, they made a head gesture that enables avatars to point with their chins—a specifically Mohawk form of nonverbal communication. Touch animations vary in terms of purpose and type. In Episode 02, an avatar elbows another to “shut up.” In Episode 03, touch is used to show relationships. For example, Pinky puts a hand on Hunter’s shoulder to disengage him from the face-off with the soldier. Her simple gesture shows her sense of control of the situation, unafraid and able to guide the young warrior to make the right choice. Lance Thomas pats the young hothead on the back, to tell him he was proud of him. Mavis McCumber, an elder, kisses her male counterpart on the cheek, expressing her affection for him. Finally, Hunter picks up the four year-old Karahwehawi and playfully tosses her in the air, demonstrating the fact that he had gained her trust, and that they had become friends.

APPROPRIATE ASSETS

The *TimeTraveller*™ machinima series and ARG required creating original nonverbal communication assets for three reasons: (1) the assets needed were simply not available in *Second Life*, and/or (2) the default assets were unsatisfactory, and/or (3) many of the user-generated assets were unsatisfactory.

Skawennati found that many of the default assets were either problematic or unusable. The gestures and facial expressions, for instance, were too extreme. Partially, the dramatic animations did not fit the tone of the machinima series. More generally, proper representation of Indigenous nonverbal communication required a full range of subtle gestures and expressions. For example, the default laugh is a belly laugh; there are no little chuckles or giggles. An angry or quizzical face had to be on or off, and even when the team's avatar operators turned them on and quickly off again, they still usually lasted too long.

Similarly, *Second Life* residents also make a lot of animations to accompany the virtual world's sex industry. Thus, one can find a plethora of physically intimate animations, yet not one animation for a kiss on the cheek—they all were designed as direct mouth-to-mouth kisses.

Skawennati also found *Second Life*'s default hair to be hideously designed (Fragnito, 2011), which is why, she suspects, there is so much user-generated hair available for purchase. Unfortunately, even with all that hair around, the team could not find long braided hair suitable for a Native male in the 1800s.

Whether default or user-generated, most assets were not appropriate for representing Indigenous cultures. In terms of cultural appropriateness, Skawennati felt that the representations of Indigenous people and cultural artifacts in *Second Life* are largely romanticized depictions based on pan-Indian stereotypes from the 1800s. Her experience observing and talking to other participants in-world points to three reasons that users create Indigenous avatars: (1) Western role-playing sims, (2) exotic eroticism, (3) and solidarity with Indigenous peoples. Western role-playing sims (simulations) invite players to role play in settings reminiscent of Western films with props and clothing such as feathered guns and moccasins. Exotic eroticism is found throughout *Second Life*, and it has been adapted to Indigenous representations through the creation of “animal hide” bikinis or outfits of feathers, skillfully placed to cover genitalia. Lastly, some ‘Indigenous’ assets are made by and for users who want to experience or express their support, understanding of, or desire to participate in Indigenous spirituality, beliefs, and politics. Clothing and accessories in this category include sage bundles, drums, flags, and political t-shirts. Animations include dances and spiritual practices.

It was rare to find any assets that seemed to be made by or for Indigenous people when Skawennati first joined *Second Life* in 2007. Skawennati's first encounter with Indigenous representations in *Second Life* was with two female avatars dressed in sexualized regalia. After, she found more clothing, but at the time, *Second Life*'s building limitations as well as the misinterpretations of Indigenous design by users resulted in poor quality assets, such as feather accessories for hair. When Skawennati started the machinima series, only light-colored skins were available for avatars. She settled on the darkest skin she could find—one identified as “Latino”—for Hunter. Because no traditional Mohawk hairstyles were available, she had to give Hunter the punk-inflected dreadhawk—though she later came to regard it as a better fit for his personality and era. In general, hair that accurately depicted both historic and contemporary Indigenous styles was the most difficult asset to find. Eventually, more clothing became available, but most of it was overtly erotic.

By the time Skawennati shot Episode 02 in 2010, the situation had improved, albeit marginally. Since Episode 02 takes place in the 1800's, the team was able to find and use a number of user-generated assets, including a bow and arrows, moccasins, raccoon hides, a pouch, a necklace, turquoise belt buckles, and feathered rifles. However, many clothing assets had to be mixed and matched for better historical accuracy—a loin cloth worn with a European-style cotton shirt, for example. For later episodes, the team was unable to locate useable ribbon shirts, fancy dance dresses, and jingle dresses.

Objects were also difficult to locate. The team found examples of shells and sage bundles available for sale from users, but they were all designed poorly. *TimeTraveller*[™] required objects that were realistic in size compared to the size of avatars. All of the existing options were oversized. However, the team was able to use some objects made by users, such as guns with a “Native” option (meaning they could be worn with feathers attached) that were used in Episode 02 of the machinima.

One of Skawennati’s main goals with the project was to integrate First Nations cultural imagery (historical, contemporary, and futuristic) with imagery of high-tech equipment and processes (Fraginito, 2011). Similar in concept to the “Native” versioning of guns, Skawennati integrates First Nations imagery with imagery of advanced technology. In particular, she uses screens, from large-scale home displays to tiny mobile displays, as a motif engaging screen culture and invoking issues of remediatization (Bolter & Grusin, 2000). She included objects such as an 1800’s Panorama, a big screen television in the future, a 1970’s/1980’s era television, an iPhone in the present, and a gigantic, anti-gravity revolving screen at a powwow in the future.

CHALLENGES

Even though *Second Life* is currently the ideal virtual world for *TimeTraveller*[™], the team ran into numerous disadvantages when using nonverbal communication. Avatars in *Second Life* have notable limitations: (1) they are mainly in an on or off state, (2) in-world customization is limited, and (3) default characteristics and animations are problematic.

The most challenging limitation is the fact that avatars are either in an on or off state; for example, they either sit or stand without a continuing flow of animation. Using custom animations with precise timing when shooting machinima or live performances with avatars provides a way of overcoming this challenge but doing so requires a well-organized team working in the same space to allow them to effectively communicate directly with each other out-of-world, in other words, in real space.

Secondly, in-world customization has limitations. Avatars cannot be made child-sized using *Second Life*’s controls in-world (meaning within the *Second Life* application). On the other end of the spectrum, it is also hard to make an avatar appear old in body structure. Making the necessary modifications to avatar size and structure must be done using third party software, which in turns requires separate skill sets and an often cumbersome and error-prone process for importing the subsequent models back into *Second Life*.

When Skawennati first began *TimeTraveller*[™], the default walk animation was jerky. Users responded to this problem by creating Animation Overrides (AO), which are now widely available. However, AOs are problematic due to the fact that custom animations often will not work while an AO is activated.

The team’s biggest challenge was making unblinking eyes for a dead character. In fact, the team is unable to make facial expressions in *Second Life*, which is a major concern for nonverbal communication. Many of *Second Life*’s default facial expression animations are overacted. To resolve this, during the editing phase, shots are cut short to minimize the length of the animation.

However, as *Second Life* improves with software updates and users create new solutions, so do the possibilities of overcoming current challenges in customizing and using nonverbal communication.

4. FUTURE WORK

Development on the *TimeTraveller*™ machinima series and ARG are ongoing. Four episodes of the machinima are complete and another six are planned. Episode 05, in which Karahkwenhawi finds herself at the deathbed of the Blessed Kateri Tekakwitha, is currently in the research phase. Episode 06 will feature a recreation of part of Alcatraz Island in 1969, and Episode 07 part of the thriving metropolis of Tenochtitlan in 1490. Episodes 08 through 10 explore Karahkwenhawi's and our Future. The ARG is currently in the development stage as a *Second Life* game component. The gameplay includes a scavenger hunt that draws participants to elements of the storyline.

Future plans for the development of *TimeTraveller*™'s nonverbal communication assets include distribution to other creators, the creation of new assets, and additional research into technical and design workflows to better accommodate their creation. Currently, only the team's avatars can access the current assets. AbTeC plans to release select assets to other users through the ARG as a part of a reward mechanism to encourage participation. Ultimately, most of the assets will be made available to the general *Second Life* public through AbTeC's proposed in-world store.

Each additional episode has substantially new settings, and a constant flow of new characters. Skawennati sees each episode as an opportunity to experiment with making new assets, such as Aztec regalia and feather headdresses for Episode 06. Over time, as her team becomes more facile within the environment, she is planning for a growth in the complexity and sophistication of her assets. Already, one can see a huge jump in ambition between the first episode and Episode 04—which incorporates a detailed, to-scale replica of the real-life Saint Francis-Xavier Mission in Kahnawake Mohawk Territory, as well as the massive, Olympics-sized powwow stadium.

Once all of the assets from *TimeTraveller*™ are finished and made accessible to other *Second Life* users, AbTeC will conduct additional research to observe and analyze the acquisition and use of the *TimeTraveller*™ assets. Possible research questions include: (1) What types of avatars are using the assets (do they self-identify as Indigenous or not)?, (2) How do they use the assets?, (3) Where do they use the assets?, and (4) What is their interpretation of the assets? Answers to questions such as these will help to determine their impact as First Nations representation in virtual worlds. Virtual world builders could greatly improve the cultural reach of their environments by soliciting feedback from different cultural groups on what sort of assets should be part of the default libraries.

5. CONCLUSIONS

In *TimeTraveller*™ specifically, the most customizable forms of nonverbal communication assets include (1) appearance and (2) performance. Foremost, the process of making *TimeTraveller*™ resulted in a growing collection of nonverbal communication assets for First Nations representation in *Second Life*.

The process of making the assets in a self-determined context has implications for designing Indigenous assets in other current and future virtual worlds. The designer needs to operate from a specific context—which Nation, which tribe, and when—as well as have some familiarity with the actual texture of life in Indigenous communities. Indigenous people are particularly well-placed to understand and honor such contextual information, but it is possible for non-Indigenous designers to make contributions if they ground themselves in traditional forms of Indigenous nonverbal communication. The designer needs to understand how the available tools have been built with a particular set of representations in mind, and she must consciously develop ways of working around those limitations rather than accepting them and the

homogenizing results. The virtual nature of the exercise does not mean no consequences exist in the real world for these choices; the constructed nature of the exercise creates a heavy burden on the designer to take responsibility for the cultural significance of everything she makes. Designers should put pressure on the builders of virtual worlds—regardless of the cultural backgrounds of the developers—to better facilitate culturally grounded representation and self-representation by (1) expanding standard asset libraries in ways that are useful for Indigenous representation (e.g. appropriate skin tones), and (2) thinking deeply about how their customization tools can accommodate a wider range of representation and action. Skawennati's *TimeTraveller™* project—representing First Nations peoples in a future context in a way that acknowledges individual nations and expresses that First Nations peoples are indeed technologically advanced—provides a strong example of what can be done in such spaces, but only with much customization and fighting against the grain of the tools. We contend that all virtual world builders should make an effort to solicit the feedback from different cultural groups.

6. BIBLIOGRAPHY

- Allmendinger, K. (2010). Social presence in synchronous virtual learning situations: The role of nonverbal signals displayed by avatars. *Educational Psychology Review*, 22(1), pp. 41-56. Retrieved May 5, 2011, from <http://www.eric.ed.gov/>
- Antonijevic, S. (2008). A microethnographic analysis of nonverbal communication in the 'Second Life' virtual environment. Paper presented at *10th Biennale Conference of the European Association of Social Anthropologists*. Ljubljana, Slovenia (August, 2008).
- Battiste, M., & Henderson, J. Y. (2000). *Protecting Indigenous knowledge and heritage : A global challenge*. Saskatoon: Purich Publishing.
- Bolter, J. D. & Grusin, R. (2000). *Remediation: Understand new media*. Cambridge, MA: MIT Press.
- Carroll, J. & Cameron, D. (2005). Machinima: Digital performance and emergent authorship. Proceedings of Digital Games Research Association Conference: *Changing Views, Worlds in Play*. Vancouver, British Columbia. Retrieved January 5, 2011, from www.digra.org/dl/db/06276.32151.pdf
- Chamberlin, J. E. (2000). From hand to mouth: the postcolonial politics of oral and written traditions. In M. A. Battiste (Ed.), *Reclaiming Indigenous voice and vision* (pp. 124-141). Vancouver, BC: UBC Press.
- Chiang, L. H. (1993). Beyond the language: Native Americans' nonverbal communication. *Proceedings of the Annual Meeting of the Midwest Association of Teachers of Educational Psychology*. Anderson, Indiana. Retrieved May 5, 2011, from <http://www.eric.ed.gov/>
- Claxton, D. (2005). Re:wind. In Townsend, M. A., Claxton, D., Loft, S., Art Gallery of Hamilton, Walter Phillips Gallery, & Indigenous Media Arts Group (Eds.), *Transference, tradition, technology : Native new media exploring visual & digital culture* (pp. 15-41). Banff, Alta.: Walter Phillips Gallery Editions.
- Denzin, N. K., Lincoln, Y. S., & Smith, L. T. (2008). *Handbook of critical and Indigenous methodologies*. Los Angeles: Sage.
- Fixico, D. L. (2003). Oral tradition and traditional knowledge. In *The American Indian mind in a linear world : American Indian studies and traditional knowledge* (pp. 21-39). New York, NY: Routledge.
- Fragnito, S. (1997). CyberPowWow. Retrieved March 4, 2011, from <http://www.cyberpowwow.net/>
- Fragnito, S. (2000). Imagining Indians in the 25th Century. Retrieved March 4, 2011, from <http://www.imaginingindians.net/>
- Fragnito, S. (2009). *TimeTraveller™*. Retrieved March 14, 2011, from www.timetravellertm.com
- Fragnito, S. (2010). *Skawennati.net*. Retrieved August 11, 2011, from www.skawennati.net/
- Fragnito, S. (2011, March 5). Email interview.

- Halbert, D. & Ghosh, S. (2008). Real spaces: *The proxemics, chronemics, and haptics of virtual worlds and social networking*. Paper presented at *User-Generated Content, Social Networking & Virtual Worlds*. Nashville, Tennessee (November 14-15, 2008).
- Hall, E. T. (1959). *The silent language*. New York, NY: Doubleday.
- Hall, E. T. (1976). *Beyond culture*. New York, NY: Doubleday.
- Hofstede, G. (1983). The cultural relativity of organizational practices and theories. *Journal of International Business Studies*, 14.
- Hopkins, C. (2006). Making things our own: the indigenous aesthetic in digital storytelling. *MIT Press Journals*, 39(4).
- Hossain, S.K.A., Rahman, A.S.M.M., & El Saddik, A. (2010). Interpersonal haptic communication in *Second Life*. *Proceedings of 2010 IEEE International Symposium on Haptic Audio-Visual Environments and Games*. Retrieved April 4, 2011, from <http://ieeexplore.ieee.org/xpl/mostRecentIssue.jsp?punumber=5614820>
- Innocent, T., & Haines, S. (2007). Nonverbal communication in multiplayer game worlds. *Proceedings of the 4th Australasian Conference on Interactive Entertainment*, RMIT University, Melbourne, Australia.
- Jandt, F. E. (2010). An introduction to intercultural communication: *Identities in a global community* (6th ed.). Los Angeles: Sage.
- Jones, R. (2007). Saving worlds with videogame activism. In R. Ferdig (Ed.), *Handbook of Research on Effective Electronic Gaming in Education*. Retrieved January 5, 2011, from www.waffler.org/wp-content/uploads/2009/05/saving-worlds-with-videogameactivism.pdf
- Kelland, M., Morris, D., & Lloyd, D. (2005). *Machinima*. Boston, MA: Thomson / Course Technology.
- Kirch, M. S. (1979). Non-verbal communication across cultures. *The Modern Language Journal*, 63(8), pp. 416-423. Retrieved April 4, 2011, from <http://www.jstor.org/stable/326027>
- Kovach, M. (2009). *Indigenous methodologies: Characteristics, conversations and contexts*. Toronto: University of Toronto Press.
- Kujanpää, T. & Manninen, T. (2003). Supporting visual elements of non-verbal communication in computer game avatars. *Proceedings of Digital Games Research Association Conference: Level Up*. Utrecht, Netherlands. Retrieved April 4, 2011, from www.digra.org/dl/db/05163.24426.pdf
- L'Hirondelle, C. (2011). *Cheryl L'Hirondelle*. Retrieved August 11, 2011, from <http://www.cheryllhirondelle.com/>
- Little Bear, L. (2000). Jagged worldviews colliding. In M. A. Battiste (Ed.), *Reclaiming indigenous voice and vision* (pp. 77-85). Vancouver, BC: UBC Press.
- Marino, P. (2004). *3D game-based filmmaking: The art of machinima*. Scottsdale, Ariz.: Paraglyph Press.
- Maskegon-Iskwew, A. (2003). *Drumbeats to Drumbytes: Ahasiw Maskegon-Iskwew*. Retrieved August 11, 2011, from <http://drumbytes.org/about/ahasiw.php>
- McAdam, S. (2009). *Cultural teachings: First nations protocols and methodologies*. Saskatoon: Saskatchewan Indian Cultural Centre.
- Mihesuah, D. A. (2003). In the trenches of academia. In *Indigenous American Women: Decolonization, empowerment, activism* (pp. 21-38). Lincoln, NB: University of Nebraska Press.
- Pechawis, A. (2011). *Archer Pechawis | Performance & Media Artist*. Retrieved August 11, 2011, from <http://apxo.net/home.html>
- Picard, M. (2007). Machinima: Video game as an art form? *Loading...*, 1(1). Retrieved January 5, 2011, from journals.sfu.ca/loading/index.php/loading/article/view/17
- Rymaszewski, M., Au, W. J., Ondrejka, C., Platel, R., Van Gorden, S., Cezanne, J., et al. (2008). *Second Life: The official guide* (2nd ed.). Sybex.
- Samovar, L. A. & Porter, R. E. (1991). *Intercultural communication: a reader*. California: Wadsworth, Inc.
- Sebeok, T. A. (1979). *The sign & its masters*. Austin: University of Texas Press.
- Smith, L. T. (2011). *Decolonizing methodologies: Research and indigenous peoples* (2nd ed.). London: Zed Books Ltd.

- Tsetserukou, D. & Neviarouskaya, A. (2010). World's first wearable humanoid robot that augments our emotions. *Proceedings of the 1st Augmented Human International Conference*. France. Retrieved August 11, 2011, from <http://portal.acm.org/citation.cfm?id=1785463>
- Time Warner Interactive (1995). The Palace. Retrieved March 4, 2011, from <http://www.thepalace.com/>
- Thomas, D. & Brown, J. S. (2009). Why virtual worlds can matter. *International Journal of Learning and Media*, 1(1), pp. 37-49.
- U-Mackey, A. (2011). *Second Life* as a tool to develop intercultural communicative competence. Paper presented at *International Association for Language Learning Technology*. Auckland, New Zealand (June 21-25, 2011).
- Verhulsdonck, G. & Morie, J. F. (2009). Virtual chironomia: developing non-verbal communication standards in virtual worlds. *Journal of Virtual Worlds Research*, 2(3). Retrieved April 4, 2011, from http://usc.academia.edu/JacquelynFordMorie/Papers/425975/Virtual_Chironomia_Developing_Standards_for_Non-verbal_Communication_in_Virtual_Worlds
- Weber, A., Rufer-Bach, K., & Platel, R. (2008). *Creating your world: The official guide to advanced content creation for Second Life*. Indianapolis, IN: Wiley Publishing, Inc.

SECTION III

*VIRTUAL PERFORMANCE
AND THEATER*

9

LESSONS FROM THE ARTS: WHAT THE PERFORMING ARTS LITERATURE CAN TEACH US ABOUT CREATING EXPRESSIVE CHARACTER MOVEMENT

By Michael Neff

The performing arts literature offers arguably the most comprehensive analysis of expressive character movement available. While it is not always straightforward to apply material written for human performers to virtual characters, this material provides excellent guidance on the movement repertoire we need to imbue synthetic characters with in order to give them rich personalities and strong emotional expression. This chapter reviews material from actor training, animation and movement theory in order to provide a broad grounding in the movement principles distilled in the arts. It also discusses some of the challenges in using this material and shows examples of how it can be successfully applied for procedural animation. The chapter aims to provide a clear understanding of the kinds of movement characters in virtual worlds will likely require in order to convey rich personalities and emotional nuance.¹⁷

¹⁷Acknowledgements: Partial support for this work has been provided through NSF grant IIS 0845529. You may contact the author at: neff@cs.ucdavis.edu

MICHAEL NEFF ON HIS METHODS

This work is motivated by the goal of creating computer tools that allow the easy generation of expressive characters for virtual worlds. In support of this goal, the chapter provides a broad overview of movement insights gained from the arts. This knowledge was gained from three different ways of understanding expressive movement: secondary research in the arts literature, embodied knowledge from movement training and applied work representing movement in computer animation tools. It is argued that this understanding of movement provides a basis for understanding the types of movement necessary for virtual characters.

The literature based research has been approached by completing a broad survey of writings on expressive movement, including research on actor training, mime, clown, animation and movement

(continued on next page)

No field has studied character movement more intently than the performing arts as expressive motion is critical for creating the rich characters that have been produced on stage for centuries, and more recently, in film and animation. One of the great challenges in building character systems for virtual worlds is generating such movement that can express a character's unique personality and change appropriately with his/her changing mood. As this challenge is also at the heart of the work of actors and animators, we can learn from the substantial literature outlining the principles they have uncovered.

This chapter has three goals. First and foremost, it will summarize the key movement properties discussed in the arts literature, with a focus on material most useful for virtual characters. Emphasis will be placed on understanding which movement aspects support the expression of personality and emotion, with a brief discussion of action selection. These movement qualities provide the palette that we use to design virtual characters and it is thus crucial to have a clear understanding of them. The second goal of the chapter is to show how these ideas have been incorporated into computational models for animation. This will be done through "implementations" sections, placed alongside a description of the qualities. I will discuss my own work in this area, as I know it best, but also try to link to the significant amount of important work done by other researchers. This will necessarily be incomplete and I apologize if I fail to mention worthwhile work for reasons of space. Finally, the chapter will conclude with a discussion of the challenges inherent in applying this material computationally and discuss potential future uses of the material.

A comprehensive summary of the key properties of expressive movement provides a map that outlines the qualities any character animation system should seek to support. This chapter argues that this material offers significant value no matter the motion representation used, as it can inform the design of procedural algorithms for motion control, provide guidance for building a motion capture library working with a team of actors, or be used as a measuring stick to gauge the completeness of learning based approaches.

theory. Through a synthesis process, movement properties that are repeated across the literature are identified and categorized. These "common properties" provide a good basis for understanding movement. This survey focuses on "functional" or "objective" descriptions of movement, i.e. what the body is actually doing, and supplements this by possible meanings implied by those movements. These functional descriptions lend themselves more readily to computational representation.

The literature based research is tested in two ways: against embodied experience and for its ability to support computational models. Practical movement work provides the quickest and most effective way to explore movement ideas. We can try them out in our body and observe them in others and through this gain a deeper understanding of how they function and what they mean. I've experimented with most of the properties discussed in the chapter through my own movement practice, particularly in my training to become a Certified Laban Movement Analyst. For virtual worlds, movement must ultimately be mapped to a computational representation. Performing this mapping process both reveals whether the property is clearly understood and indicates which movement properties lend themselves most easily to computational use and where future research challenges lie. I've spent several years developing computational models of movement based on these insights.

1. SIDEBAR: A BRIEF INTRODUCTION TO COMPUTER ANIMATION

It is helpful to introduce some basic animation concepts and terminology in order to be able to discuss how ideas from the arts have been applied computationally. Characters are generally represented computationally by a skeleton, a set of fixed length bones connected by joints. Movement is specified by changing the angles of these joints over time. Joints may rotate in up to three dimensions (x, y, z) and each of these rotations is referred to as a *Degree of Freedom* (DOF). The visual appearance of the character is normally defined by a 3D mesh which is bound to the skeleton and deformed by its movement.

Most researchers that have sought to leverage off material from the arts have done so by building *procedural models*, i.e. they have written software that encodes particular movement ideas. A common representation is to divide movement into a series of *poses* specified at particular points in time, known as *keyframes* in traditional animation. These poses may be specified by providing the value of all the joint angles in the skeleton or by providing the position of key body parts, such as the hands.

Interpolation is the process of moving from one pose to the next, something known as *inbetweening* in traditional animation. There is a wide range of velocity profiles that can be used in making these transitions and this is referred to as the *motion envelope*.

Forward Kinematics (FK) is the process of solving for the character pose based on a specified set of joint angles. *Inverse Kinematics* (IK) is the process of solving for a set of joint angles that will satisfy a world space constraint, such as the position of the character's hand. You can think of FK as controlling the character through joint angles and IK as controlling the character by specifying positions for body parts.

2. KEY SOURCES

This work draws on three main fields that have studied expressive movement in the context of creating convincing characters: traditional animation, actor training and movement theory. The paper will synthesize findings across these fields rather than adopting a single framework.

For traditional animation, we rely on the work of Thomas and Johnston (1981) and Lasseter (1987) that describes the principles of traditional animation developed at the Disney studio in the 1930's. These are: Squash and Stretch (deforming a character to show it's mass and rigidity/fluid nature), Timing (spacing actions to reveal a sense of weight and personality), Anticipation (preparing the audience for what is about to happen), Staging (arranging action so that it is easily read by an audience), Follow Through and Overlapping Action (different parts of the body should start and stop at different times; one action should flow into, and overlap with, the next), Straight Ahead Action and Pose-To-Pose Action (drawing either every frame in a sequence or first drawing key moments and then filling in the in-betweens), Slow In and Out (adjusting the spacing between key poses), Arcs (using curved paths through space to give more natural movement), Exaggeration (emphasizing the key points to be communicated), Secondary Action (actions that result from other actions), and Appeal (creating design and motion that an audience will enjoy).

For actor training, we survey the work of Stanislavski, Meyerhold and Grotowski, along with more applied texts on actor training (e.g.(Alberts, 1997)). Probably no one had a greater influence on the art of acting in the last hundred years than Constantin Stanislavski, director of the Moscow Art Theatre and inventor of the "Stanislavski System", as the root of "Method Acting" in North America. Stanislavski wrote a trilogy

on acting technique: *An Actor Prepares* (Stanislavski, 1936), *Building a Character* (Stanislavski, 1949) and *Creating a Role* (Stanislavski, 1961). The first book focuses on the inner life of the character, the second deals with external representation (physical movement) and the third discusses how an actor can create a role. He argued that actors must *live* on stage. They must not merely try to mechanically reproduce a set of actions, but there must instead be emotional truth to what they do (Moore, 1984). We need to also imbue our virtual characters with this quality. While considering physical training important, he emphasized first focusing on methods to help an actor achieve psychological realism.

Vsevolod Meyerhold developed an exercise based form of actor training known as *biomechanics* (Law & Gordon, 1996). Sergei Eisenstein was his student during this time and developed a related method known as *Expressive Movement* (Law & Gordon, 1996). They argued that what an actor felt inside was not important if it was not communicated to the audience and counter to Stanislavski's inward focus, sought to increase an actor's physical vocabulary (Law & Gordon, 1996). They proposed that a highly trained actor could achieve emotional involvement through his movements (Moore, 1984).

Polish director Jerzy Grotowski formed the Theatre Laboratory in Poland with the goal of understanding the nature of acting. He drew widely on Soviet and Western theatre traditions, especially Meyerhold, as well as training methods used in the Orient (Grotowski, 1968e). Grotowski developed a training regime which focused on physical and vocal exercises, with the idea that actors approach movement ideas through their body (Tarver & Bligh, 1999). By repeatedly participating in a set of exercises, normally in silence, the actor gains a greater understanding of the nature of movement by experiencing it through his or her body. He argued that man does not behave "naturally" in heightened emotional moments, so one must go beyond naturalism to reveal the deeper truth (Grotowski, 1968e).

The movement theorists include: Francois Delsarte, a 19th century French theoretician whose work was influential in actor training and the development of American modern dance (Shawn, 1963); Barba and Savarese, who developed the field of *theatre anthropology* to study the art of the performer; and Rudolf Laban and his collaborators who developed Laban Movement Analysis (LMA). LMA offers a systematic study of expressive aspects of movement, divided into Body, Effort, Shape and Space. Body includes structural aspects of movement related to the anatomy of the body and how motion passes through the body. Effort refers to a person's attitude of indulging or resisting four movement qualities: Weight, Space, Time and Flow. Shape describes the process of change between poses as well as particular poses a character assumes. Space relates a character's movement to external pulls in the environment.

The LMA system aims to describe movement without defining a particular meaning for it. In other words, it aims to describe what the mover is doing, rather than applying an interpretation as to what emotion that movement suggests or what impression it should make on an observer. Delsarte's work on the other hand often maps movements to specific meanings. Some of these mappings will be included in the discussion below, but without careful experimental validation, are best viewed as an initial, potential interpretation rather than an immutable truth.

Motion Qualities

This section will review the key aspects of movement discussed in the literature. After a brief discussion of general movement principles, the discussion will be organized around four aspects of movement I term: Shape, the poses adopted by a character; Transitions, all the transients aspects of movement as a character moves from pose to pose; *Timing; and Phrasing*, how motion qualities are layered and combined.

To begin, it is worth noting that the motion we want from our animated characters may be different from the motion we experience in daily life. Barba argues that, "[t]he way we use our bodies in daily life is substantially different from the way we use them in performance." (Barba, 1991b, p.9), suggesting that "[w]hile daily behaviour is based on functionality, on economy of power, on the relationship between the

energy used and the result obtained, in the performer's extra-daily behaviour each action, no matter how small, is based on waste, on excess." (Barba, 1991a, p.55). The purpose of stage movement is to infect the audience with emotion (Eisenstein & Tretyakov, 1996) and it is generally when the attributes of an actor's movement are out of the ordinary, that they will have the greatest significance for the audience (Alberts, 1997).

Driven by the need to clearly communicate with an audience, performance movement is based on two basic principles: simplification (Lawson, 1957; Thomas & Johnston, 1981; Lasseter 1987; Barba, 1991b; Moore, 1984) and exaggeration (Thomas & Johnston, 1981; Lasseter 1987; Barba, 1991a). These two properties work together to clarify the meaning of a character's movement in the spectator's mind.

Simplification, also known as the "Virtue of Omission" (Barba, 1991b), works to bring focus to certain elements of a character's movement by eliminating extraneous movements (Barba, 1991b). Traditional animation principles suggest having a character only do one thing at a time so that the action reads clearly (they also suggest having one action bleed into the next for fluidity) (Thomas & Johnston, 1981; Lasseter 1987). In her summary of Stanislavski's teachings, Moore suggests that "[e]very form of expression must be simple and clear", with an emphasis on precision (Moore, 1984, p.54). All this brings a focus on communication, and arguably clarity, to performance motion that is often lacking in the movements of daily life.

Once a movement has been simplified, it is exaggerated to ensure that its meaning is conveyed to the audience. Thomas nicely summarizes the interplay of simplification and exaggeration as follows:

As artists, we need to find the essence of the emotion and the individual who is experiencing it. When these subtle differences have been found, we must emphasize them, build them up and at the same time, eliminate everything else that might appear contradictory or confusing. (Thomas, 1987b, P.6)

Movements should never be vague; audiences demand motions which they can follow (LeCoq, 2002).

Implementations

Applying these ideas to virtual characters suggests that verisimilitude of daily life may not be the ultimate goal. Rather, movement more focused on clear expression may be more effective in virtual worlds and more consistent with other performance contexts.

SHAPE

The shape category gathers properties that refer both to a pose at an instant in time and how these poses change.

EXTENSION

Extent or extension refers to how far an action or gesture takes place from a character's body. It can be thought of as how much space a character is using while completing an action. Laban (1988) refers to the area around a person's body as the *Kinesphere* and defines three regions within it: the near region, anything within about ten inches of the character's body, is the area for personal, intimate or perhaps nervous actions; the mid area is about two feet from the person's body and this is where daily activities take place, such as shaking hands; the area of far extent has the person extended to full reach. It is used for dramatic, extreme movements and in general is used more on stage than in daily life.

Delsarte suggests that excitement, explosive anger, strong and violent emotions that are aggressive all act to expand action (or increase extent) (Shawn, 1963). Thought, meditation, concentration, fear, suspicion and repulsion contract a body's movements. Normal emotions and gestures are in-between. He also suggests that slow movements that emphasize vastness and grandeur are aided by full extension. On a related note, Stanislavski suggests that an actor must be decisive in his big movements (Stanislavski, 1949).

Implementations

Several computational models have implemented control over extent, such as the EMOTE model based on Laban Movement Analysis (LMA) (Chi, Costa, Zhao, & Badler, 2000), Hartmann et al.'s (Hartmann et al., 2002, 2006) model based on social psychology (it is one dimension of their six dimensions of expression) and my previous work on motion editing (Neff & Fiume, 2003). All of these models scale the desired position of the hands relative to the body. They also include control over the swivel angle, defined by a rotation along the axis running from the shoulder to the wrist. This controls how close a character's elbows are to his sides and impacts his overall size.

BALANCE

Balance adjustments can move a character from being stably balanced on both feet, to a single foot, to the very edge of balance and are fundamental to expressive movement. Indeed, Barba claims that the "dance of balance" is revealed in the fundamental principles of all performance forms (Barba, 1991b):

The characteristic most common to actors and dancers from different cultures and times is the abandonment of daily balance in favour of a 'precarious' or 'extra-daily' balance. Extra-daily balance demands a greater physical effort -- it is this extra effort which dilates the body's tensions in such a way that the performer seems to be alive even before he begins to express. (Barba, 1991c, P.34)

Such balance adjustments can work to intensify a motion. "A change of balance results in a series of specific organic tensions which engage and emphasize the performer's material presence, but at a stage which precedes intentional, individualised expression." (Barba, 1991c, P.35) Many performance traditions will impose physical constraints to make the balance task more difficult, for example binding feet or having a ballerina work on toe. Balance adjustments can activate chains of muscles that connect multiple segments of a person's body.

When we stand in daily life, we are never still, but rather are constantly making small adjustments, shifting our weight to the toes, heels, right side, left side, etc. These movements should be modeled and amplified in performance (Barba, 1991b). A proper balance adjustment can even give a static pose a sense of motion:

The performer's dynamic balance, based on the body's tensions, is a balance in action: it generates the sensation of movement in the spectator even when there is only immobility. (Barba & Savarese, 1991, p.40)

Being off balance also reflects on the action that has just been completed (Tarver & Bligh, 1999).

Improved control of balance is part of the actor training of Meyerhold, Grotowski and Stanislavski (Law & Gordon, 1996; Grotowski, 1968b; Stanislavski, 1936). Delsarte suggests that balance can create a mood of security and control; imbalance suggests insecurity, indecision, fear and worry (Shawn, 1963).

Implementations

Computational models of balance have often focused on static balance in which a character's center of mass must project to the support polygon defined by the perimeter of its feet (Wooten, 1998). We (Neff & Fiume, 2004, 2006) presented a kinematic model based originally on Wooten's controller (Wooten, 1998) and later extended to handle more complex adjustments (Neff & Kim, 2009). Wooten (Wooten, 1998) and van Welbergen et al. (Welbergen, Reidsma, Ruttkay, & Zwiers, 2010) have applied this to physically simulated characters. Other approaches (Tak, Song, & Ko, 2000; Tak & Ko, 2005; Shin, Kovar, & Gleicher, 2003) rely on the Zero Moment Point, which acts as a measure of dynamic balance and can be used to correct the balance of more rapid movements.

POSTURE AND POSE

Posture is one of the clearest indicators of both a character's overall personality and emotional state in a particular scene. Alberts (1997) suggests that posture is a combination of two components: the degree of tension displayed in the body and the overall body position. Body position includes standing, leaning, kneeling, sitting and lying down. Alberts proposes the following posture scale: hunched, stooped, slumped, drooped, slouched, sagging, tired, relaxed, straight, upright, uptight, erect and over-erect (at attention). This range runs from a tired old man to a rigidly erect army officer.

Shawn argues that one of the main contributions of Delsarte was realizing that the torso is the main instrument of emotional expression (Shawn, 1963). It was for this reason that modern dance moved away from the stiff and fixed spine that is traditional in ballet.

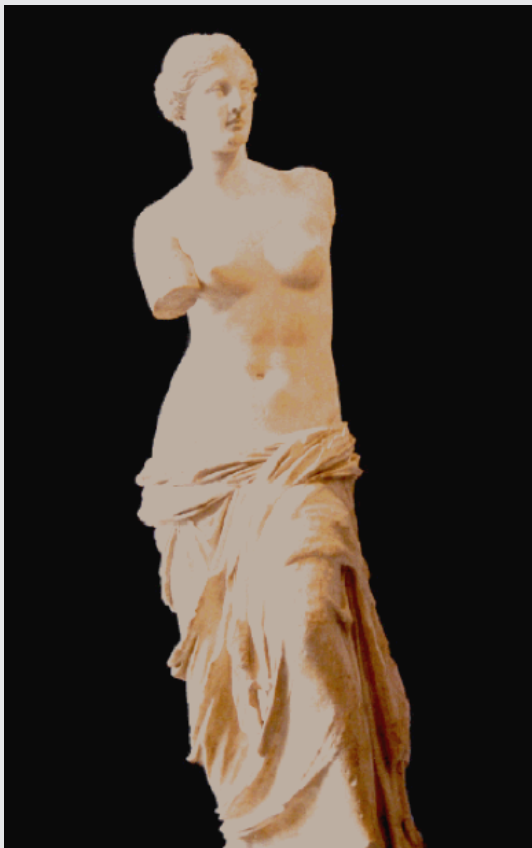


Figure 9-1: Venus de Milo

Delsarte suggested that the part of the torso that a person habitually holds forward is a strong indicator of what kind of person they are (Shawn, 1963). If they hold their chest high, this indicates self-respect and pride. If their abdomen is protruding, this indicates animality, sensuality and lack of bodily pride. A normal, balanced carriage will have the middle zone of the abdomen carried forward and the chest and abdomen withdrawn. This triad can be augmented by considering people who carry their head forward, normally indicating a mental or academic disposition.

The shape of the spine in the coronal plane is also important. The "S" curve or "Beauty Line" involves the legs, torso and neck in making a large S curve with the entire body. It is a key pose in Indian dance, where it is called *tribhangi*, meaning three arches. It was also prominent in ancient Greek sculpture, the Venus de Milo offering a clear example [Figure 9 1], and was taken up by Florentine sculptors in the 14th century (Barba & Savarese, 1991). This also serves to illustrate the importance of the interplay between the legs, torso and neck.

Laban suggests that there are three principal components of trunk movement: rotational movement about the length of the spine; “pincer-like” curling from one or both ends of the trunk and “bulge-like” shifting of the central area of the trunk out of its regular position (Laban, 1988).

The torso can be expanded, contracted or relaxed. According to Delsarte, expansion indicates different degrees of excitement, vehemence and power of the will. Contraction indicated different degrees of timidity, pain, effort or convulsion of the will. Relaxation indicates different degrees of surrender, indolence, intoxication, prostration and intensity of the will (Shawn, 1963).

The legs of a character have a functional role to play in supporting it, which limits their expressive range. Nonetheless, stance is very important (Shawn, 1963). The main sources of variation in the legs are the width of the stance, the bend in either knee, whether the legs are turned out, in or straight and whether there is a twist in the pelvis. Stanislavski argues that the legs should be turned slightly out; rising on one’s toes suggests flight; feet and toes modulate jerkiness and give a quality of smoothness and gracefulness to motion (Stanislavski, 1949).

The arms and hands are used to create a wide range of motions. Delsarte refers to the shoulder, elbow and wrist joints as thermometers because he feels they indicate *how much* rather than *what kind* of expression (Shawn, 1963). Raised shoulders act to strengthen any action, the more they are raised, the stronger the action. He argues no intense emotion is possible without the elevation of the shoulders, or forward contraction in the case of fear or backward pull to show aggression or defiance. “The elbow approaches the body by reason of humility, and moves outward, away from the body, to express pride, arrogance, assertion of the will.” (Shawn, 1963, p.41) Elbows within the column of the body denote a lack of self-respect. Finally, a not rigid, but strong wrist indicates a strong, healthy condition. A limp wrist indicates weakness or a devitalized condition. A sharply bent wrist indicates a crippling influence.

Wrist movement can also add definition to a gesture (Lawson, 1957). Stanislavski argues that the arms should neither hang in-front nor behind the body, but at the actor’s side. Elbows should turn outward, not inward, but this cannot be excessive (Stanislavski, 1949).

Poses can be either symmetric or asymmetric. Lasseter (1987) recommends avoiding perfect symmetry (twinning) as this generally looks unnatural. This is particularly true with regards to how the pose reads in the 2D image of the animation. Grotowski argues “if something is symmetric it is not organic!” (Grotowski, 1968b, p.194). Asymmetric poses generate tension through opposition, whereas symmetric poses lack opposition and give balance (Barba, 1991a).

Within the LMA Shape category, *Shape Qualities* describe how a pose can change along the cardinal axes and capture an overall movement tendency. Vertical movement can be either *Rising* if the pose stretches upwards, or *Sinking* for downward movement. Forward movements are *Advancing* and backward movements *Retreating*. Horizontal movements that go sideways out from the body are *Spreading* whereas those that move inwards and cover the chest are *enclosing*. *Gathering* and *Scattering* are related but slightly more general concepts. Motions in which a person wraps their arms around space and brings it to them (think of a hug) are *gathering*. On the other hand, if the character starts with her hands near her middle and thrusts them out and away from her body to the sides (picture spreading seeds onto a field), this is a *scattering* gesture. Delsarte observed this same pattern. Generally speaking, scattering actions suggest openness, sharing and an external focus. Gathering actions suggest the character is closed, coveting or tormented and internally focused. Open and closed static postures have similar connotations. Posture changes are done relative to a character’s body attitude, the habitual default body configuration the character holds.

Closely tied to posture, a commonly referenced (Law & Gordon, 1996; Tarver & Bligh, 1999) principle of effective movement is *full body engagement*, which indicates that the whole body should be engaged in all movements. Grotowski (1968b, p.193) writes:

Our whole body must adapt to every movement, however small....If we pick up a piece of ice from the ground, our whole body must react to this movement and to the cold. Not only the fingertips, not only the whole hand, but the whole body must reveal the coldness of this little piece of ice.

Full body engagement acts to clarify and emphasize a movement.

The literature offers little guidance on how to actually achieve full body engagement. One potential starting point is the LMA Body concept of *Patterns of Total Body Connectivity* (PoTBC). Emerging out of work developed by Laban's student/collaborator Irmgard Bartenieff, as well as work by Karl and Berta Bobath, Bonnie Bainbridge Cohen and Peggy Hackney, PoTBC provide an ordered set of patterns for organizing overall body activity that follow the developmental stages of humans (Hackney, 1998). The first pattern is the expansion and contraction of *Breath*. This can be divided into three dimensions of torso movement, lengthening and shortening in the vertical dimension, bulging and hollowing in the sagittal and widening and narrow in the horizontal. The next pattern, *Core-Distal* involves movements of the six limbs (arms, legs, head and tail) from the core out distally, and back in towards core; connecting core and distal. The *Head-Tail* pattern explores connectivity through the spine. *Upper-Lower* involves coordinated up and down movements of the arms and legs and can generate the first stages of crawling. *Body-Half* movements divide the body along the sagittal plane and one half stabilizes while the other half is mobile. Finally, the *Cross-Lateral* pattern features diagonal connections from one arm through to the opposite leg. It is the pattern behind walking.

Another organizing principle for full body movement is the idea of *Posture-Gesture Mergers* (PGM), developed by another of Laban's student/collaborators, Warren Lamb (1965), (Lamb & Watson, 1979). Gesture is defined as some movement of a part of the body (hands, arms, head, etc.). Posture is defined as a movement of the entire body. A posture-gesture merger occurs when the quality of the movement is the same for both the posture and the gesture and synchronized in time. Here "Quality" is defined in terms of the Effort and Shape Qualities of Laban Movement Analysis. The qualities need not overlap for the entire gesture, but there could be a partial PGM for a portion of the gesture. Lamb considers PGM a fundamental aspect of human movement and the particular patterning of PGMs to be a unique part of a person's movement style.

Implementations

When compared to the attention paid to arm gestures, there has been relatively little work on developing effective models of posture and torso movement. The expressive importance of the arm swivel angle has been recognized by several researchers, for example, in the EMOTE model (Chi, Costa, Zhao, & Badler, 2000), in our previous work (Neff & Fiume, 2004, 2006) and the work of Hartmann et al. (Hartmann, Mancini, & Pelachaud, 2006). This is an important control for any character system.

The EMOTE model (Chi, Costa, Zhao, & Badler, 2000) includes an implementation of Shape Qualities which deform the torso. They define a maximum orientation for each DOF in the spine, neck and collarbones for each movement quality (e.g. Rising). User parameters control the interpolation between these values and a neutral pose.

In previous work (Neff & Fiume, 2004, 2006), we developed a posture model that allowed a particular shape to be specified for the torso (e.g. a hunched posture, or the S beauty curve seen in the Venus de Milo) in combination with a desired balance point. In satisfying reach constraints on the arm(s), the system could change the degree of the shape, but its form would be maintained. This work illustrated the expressive importance of posture variation and different types of torso shaping, illustrating that a wide range of poses may look “natural”, but they will communicate different messages.

Data-based approaches to inverse kinematics (e.g. Rose III, Sloan, & Cohen, 2001; Grochow, Martin, Hertzmann, & Popović, 2004) work off either a motion clip or complete motion library and then find output poses that are both similar to these samples and satisfy end effector constraints. These methods offer an approach for generating realistic, full body poses. For character design, this shifts the problem from needing to determine the correct parameters to use in a model to deciding on the correct input motion to use to build the data-based algorithm in order to generate appropriate poses for a particular character.

Breath is an important component of torso change that has a strong expressive impact. This has seen limited attention in virtual character work, but hopefully this will change with the recent development of CG models for breathing (e.g. Zordan, Celly, Chiu, & DiLorenzo, 2004; Kider Jr, Pollock, & Safonova, 2011). Breath can support and coordinate with movement of the limbs. For example, the concept of Shape Flow Support in LMA (also called Breath Support) describes how the fluid nature of the torso can be engaged through breath to combine with the movement of the limbs in integrated, full body movements. The full set of PoTBC patterns offer a potential framework for generating full body motion.

SHAPE OVER TIME

Disney animators considered *squash* and *stretch* to be the most important movement principle (Thomas & Johnston, 1981; Thomas, 1987a). Only the most rigid objects do not deform when they are moved (Lasseter 1987) and as Thomas observes “...all living flesh is subtle and stretches or bulges or sags or becomes taut in reaction to the forces working on it” (Thomas, 1987a, p.23). Deforming their characters and objects as they animated became essential in order to give them a sense of life. Lasseter (1987) points out that hinged objects such as Luxo can squash and stretch without deforming. This of course applies to the human skeleton as well, for example, consider the volume changes in gathering and scattering movement or extension changes. Squash and stretch is also particularly important for facial animation (Lasseter 1987).

Recoil is one of the most frequently cited movement properties (Eisenstein & Tretyakov, 1996; Barba, 1991a; Lasseter 1987; Thomas & Johnston, 1981; Laban, 1988; Eisenstein, 1996; Taylor, 1999; Shawn, 1963). In its most basic form, recoil involves first making a movement in the opposite direction of the intended movement, followed by the intended movement itself (Eisenstein, 1996). It creates a negative space into which the motion can travel. Recoil serves to underscore and accentuate a movement (Eisenstein & Tretyakov, 1996; Eisenstein, 1996) and is one form of the traditional animation principle of Anticipation. Recoil also allows the momentum of an action to be built, as with the backswing before a punch. This also relates to phrasing, discussed below.

The path a movement takes in space can also have an important expressive impact. Gathering vs. Scattering movements offered an example of this. Within the LMA Shape category, Modes of Shape Change describe different ways a person can interact with their surroundings. *Shape Flow* movements are self-focused and connect with breath and the inner fluid nature of the body (think of a tremor or a dance where the person’s focus is on feeling the connections in their own body). *Directional* movements connect the character to the environment, either with straight, linear movements (*Spoking*), such as pointing at an object, or rotational movements (*Arcing*), such as waving. *Carving* describes three-dimensional movements that cut through space in voluminous way, for instance to describe the shape of a large ball.

The Space component of LMA further explores how a person connects to their environment. *Central Spatial Tension* involves radial pulls from the character's center out to the periphery of the Kinesphere. In *Peripheral Spatial Tension*, the movements of the hands or feet maintain a fixed distance from the body, for example envision someone running her hands along the edge of an imaginary cylinder centered at the her body. *Transverse Spatial Tension* involves movements that travel between the center and the periphery, but without a radial orientation, for example grabbing an object in front of you and swiping it back, past your side.

Spatial pulls may also be dimensional, aligning with the horizontal, vertical or sagittal (forward-back) axis. Combining two pulls gives movements in a plane and combining three pulls gives full, three dimensional movement. Ordered patterns of these spatial pulls are developed in Laban's concept of Space Harmony, defining sequences of spatial pulls known as scales which a person moves through, but that lies beyond the scope of this article.

Directionality can give movement a clear sense of focus. Describing one of his actors, Grotowski says "All his movements have a well-defined direction that is followed by all the extremities and, on closer observation, even by all the muscles" (Grotowski, 1968b,p.186). This leads to a decisive and clear movement piece.

Implementations

There has been limited implementation work for much of this area. There are numerous deformation techniques that allow a character mesh to be varied over time, but outside of muscles and clothing, less work that explores what types of variations are effective for virtual characters. Recoil is closely related to the traditional animation concept of *anticipation*. A simple way to implement this in pose based approaches is to adjust the interpolation function to first have a character move away from a target pose before reversing directions and moving towards it (Chi, Costa, Zhao, & Badler, 2000; Neff & Fiume, 2005). IK provides a way to specify a motion path through space, but I am unaware of work that uses this to organize the movement of the full body.

TRANSITIONS

Transitions include all the transient aspects of movement as a character moves from pose to pose.

MOVEMENT FLOW

Laban (1988) suggest that movement can flow continuously, be intermittently interrupted, yielding a trembling kind of movement, or stopped yielding a pose. It is worth distinguishing between a motion that is paused (Laban, 1988; Tarver & Bligh, 1999) or suspended (Alberts, 1997) and one that is stopped. When a motion is paused, there is no perceptible loss of intensity nor break in intention. When a motion is stopped, the energy of that motion has been lost and the actor's focus is no longer on the completion of a motion. It cannot be seamlessly continued with the same intensity (Tarver & Bligh, 1999). Motions can also be either complete or incomplete (Tarver & Bligh, 1999). Many actions in daily life are incomplete. They are interrupted before they reach their natural conclusion. Being able to break an action before completion can have a powerful impact as it can be a strong indicator of a character's internal mental process.

Stanislavski related flow to how energy moved through the body, "...external plasticity is based on our inner sense of the movement of energy" (Stanislavski, 1949, p.67). A smooth and regular flow of energy gives a smooth, measured and elastic step. Energy in jerks creates an uneven, choppy gait. Stanislavski told his actors that flow must be controlled and to think of a bead of mercury that they are consciously moving through their veins. Creating an endless, unbroken line with this bead will give smooth, flowing movement (Stanislavski, 1949). This idea relates closely to the chains of connectivity underlying Patterns

of Total Body Connectivity. It should be noted that developing a sense of flow is very closely related to the use of successions discussed with phrasing.

Implementations

I am aware of little work that explicitly addresses these concepts, although models for succession and physics-based tension change (discussed later) capture some aspects of flow. Disrupted flow still needs a precise, computationally definition. The ability to pause or stop a movement midstream is a useful feature for designers of virtual worlds to include.

MOTION ENVELOPE

The *motion envelope* describes the speed profile of a movement over a transition; its patterning of acceleration and deceleration. For example, some movements will start slowly and end quickly while other movements will do the opposite. Disney animators found it effective to have the bulk of footage near extreme poses and less footage in between in order to emphasize these poses (Thomas & Johnston, 1981; Lasseter 1987) and referred to this as *slow in, slow out*. In computer animation, an ease-in, ease-out curve provides the same effect, transitioning most quickly during the middle of the movement and slowing at the beginning and end. Animators will often go beyond this to create more varied control of timing, for instance using just an ease-in curve or just an ease-out curve to transition between poses.

Implementations

Control of the motion envelope is widely supported, most often by controlling the tangents of splines that are used to control the interpolation between poses. The double interpolant method first proposed by Steketee and Badler (Steketee & Badler, 1985) decouples the path through space from the timing of the motion. This is very useful for allowing expressive variation in timing without altering a motion's path. Tools such as tension, continuity and bias splines (Kochanek & Bartels, 1984) offer additional flexibility by allowing control over the shape of the interpolating curve. These techniques are key components in the EMOTE model (Chi, Costa, Zhao, & Badler, 2000) and Hartmann et al.'s style model (Hartmann, Mancini, & Pelachaud, 2006), and we provide comparable controls in (Neff & Fiume, 2005).

EFFORT

The Effort category of LMA is probably the one most frequently applied to character animation. It describes a person's inner attitude towards four qualities that are either indulged in or resisted, giving a bipolar scale for each.

Weight Effort (Strong - Light) refers to the person's attitude towards the use of force. If they are moving very forcefully (e.g. stomping up stairs), this is *Strong Weight*. If they are moving with gentleness, carefully regulating the amount of force they use, this is *Light Weight*. A physically heavy person may move in a Light way and vice versa.

Time Effort (Sudden - Sustained) describes changes in timing compared to surrounding movements, rather than overall speed. If a movement is very rapid compared to the surrounding movements, this is *Sudden Time*. If a movement is prolonged, this is *Sustained Time*. A person might have a sense of urgency or wish to linger.

Space Effort (Direct - Indirect) relates to how a person focuses their attention. If they are focusing on a particular point, this is *Direct Space Effort*. If they are attuning to the entire space, this is *Indirect Space Effort*. Movements of the rest of the body will generally reflect this focus, for example, providing a clear sense of directionality towards a single point of focus, such as staring, pointing and marching towards a misbehaving child spotted across a room, or moving in a way that attunes to the global space, such as maintaining awareness of a large group of children at play across a field.

Flow Effort (Bound - Free) describes how controlled someone's movements are. *Bound* movements are tightly controlled, precise, possibly stiff, whereas *Free* movements are loose and uncontrolled.

A movement may contain any subset of the qualities (e.g. there might be no aspect of Time present) and it is rare for all of these qualities to be present at once. Pairs of qualities combine to form an Effort State and triples form Effort Drives. For example, the Action Drive combines Weight, Time and Space, with evocative names given to each combination of poles. A *punch* is combination of Strong Weight, Sudden Time and Direct Space, whereas a float is a combination of Light Weight, Sustained Time and Indirect Space.

Implementations

Researchers have long sought to develop a computational representation for Effort because it encapsulates an important range of expressive variation. The EMOTE model (Chi, Costa, Zhao, & Badler, 2000) provides a procedural implementation based on interpolating key poses to generate motion. It adjusts the interpolation space that is used to move between poses (end effector, joint angle, or elbow), as well as the parameters of the interpolation functions, both to vary the path through space and adjust the timing. It also adds flourishes that add extra wrist bend or sinusoidally deflect the elbow to evoke particular Effort qualities.

Torresani et al. (Torresani, Hackney, & Bregler, 2007) take a data driven approach and capture motion of trained dancers performing different Effort constellations. From this, they learn an interpolation function that allows them to vary between the various Effort qualities.

TENSION AND RELAXATION

The interplay of tension and relaxation is another widely cited movement property (Dorcy, 1961; Laban, 1988; Shawn, 1963; Lawson, 1957; Barba, 1991a), closely related to the Flow Effort (Free to Bound) dimension in LMA. Tension and relaxation naturally interleave: there must first be relaxation in order for there to be tension and tension is followed again by relaxation. There is a consequent ebb and flow of energy that accompanies changes in tension (Shawn, 1963). This also relates to the preparation, action, and recuperation pattern of phrasing discussed below.

Tension changes can take place through the entire body or a tiny part. They can occur suddenly or gradually and there can be spasmodic changes back and forth (Lawson, 1957). A rise in tension can serve to accent a movement (Laban, 1988). For example, physical or emotional pain can be shown by spasmodic contractions of muscles, followed by relaxation as the pain eases (Shawn, 1963).

Stanislavski stresses the importance of an actor being relaxed and avoiding tension in his body (Stanislavski, 1936; Stanislavski, 1949; Moore, 1984) arguing "*You cannot ... have any conception of the evil that results from muscular spasm and physical contraction.*" (Stanislavski, 1936, p.91) An actor must learn to identify the sources of tension in his body and relax them (Moore, 1984). Stiff arms and legs give the body a wooden quality, looking like a mannequin. "The resulting impression is that the actor's soul is likely to be as wooden as his arms. If you add to this a stiff back, which bends only at the waist and at right angles, you have a complete picture of a stick. What emotions can a stick reflect" (Stanislavski, 1936, p.102).

Implementations

We provided tension control using a physical simulation approach in which a variant of a proportional derivative controller is used as a simple model of muscle (Neff & Fiume, 2002). This is a first order approximation of real muscle, representing it as a spring and damper. The gain on the spring term in a PD controller regulates the amount of tension in the motion. Reducing the tension is useful for making more floppy motion, for example as a wrist follows the movement of an arm, and pendular movement when

an arm drops to a character's side. Tension control also determines how forces transfer through the body, allowing a character to look very rigid, or relaxed and loose.

WEIGHT

It is critical that characters have a sense of weight (Thomas & Johnston, 1981; Thomas, 1987a, Appia, 1962 (original 1921), fifth edition 1982; Laban, 1988). This creates the physicality necessary for realism by providing the sense that the character is actually inhabiting and interacting with his/her environment. Thomas and Johnson suggest that it is an inability to correctly capture a sense of weight that makes cartoon characters lose credibility when viewed next to live action (Thomas & Johnston, 1981;).

Laban relates weight to the use of muscular energy or force to either move a weight or to react to a resistive force (Laban, 1988). Weight can come either from the weight of a body part to be moved or from an object being moved. Resistance can be internal, coming from the antagonistic actions of the character's own muscles and reflecting internal conflict, or external, coming from other objects or people. Resistance may involve strong, normal or weak muscular tension.

Implementations

An accurate sense of weight is probably the biggest win for simulation based approaches to character animation. There has been a great deal of research in this area over the past 25 years. We are starting to see these approaches applied for virtual worlds, for instance in the work of van Welbergen et al. (van Welbergen, Reidsma, Ruttkay, & Zwiers, 2010).

PHRASING

Phrasing deals with issues ranging from how joint activations are ordered in a particular movement to how an entire motion sequence is put together.

JOINT ACTIVATION ORDER

Successions deal with how a movement passes through the body. Rarely will every limb involved in a motion start and stop at the same time. Delsarte defined two types of successions: true or normal successions and reverse succession (Shawn, 1963). In a normal succession, a movement starts at the base of a character's torso and spreads out to the extremities. In a reverse succession, the movement starts at the extremities and moves in towards the centre of the character. Shawn claims that the conscious use of successions was fundamental to the development of modern dance (Shawn, 1963).

Successions are part of the "Follow Through and Overlapping Action" principle of Disney animation (Lasseter 1987). Thomas and Johnson write:

Our most startling observation from films of people in motion was that almost all actions start with the hips; and, ordinarily, there is a drop -- as if gravity were being used to get things going. From this move, there is usually a turn or tilt or a wind up, followed by a whiplash type of action as the rest of the body starts to follow through....Any person starting to move from a still, standing position, whether to start walking or pick something up, always began the move with the hips. (1981, p.72)

This also emphasizes the role of weight shifts discussed earlier and introduces the phrasing concept of initiation, where a particular body part begins a movement. Grotowski similarly argues that "The driving impulse, however, stems from the loins. Every live impulse begins in this region, even if invisible from outside." (Grotowski, 1968b, p.191), again reflecting the outward flow of movement consistent with a

normal succession. Stanislavski also gives examples of exercises designed to teach the use of successions in order to increase fluidity (Stanislavski, 1949).

The Body component of LMA defines three forms of movement sequencing (Hackney & Meaden, 2009). In *simultaneous* movements, all body parts start and stop at the same time. In successive motions, the movement travels from one body part to the next adjacent body part (e.g. collar bone to shoulder to elbow to wrist). Finally, in sequential ordering, movement jumps from body part to body part (e.g. left elbow to right elbow, or wrist to shoulder).

Implementations

Control based on IK provides simultaneous movement because all joint angles in the kinematic chain are updated at the same time. It is for this reason that animators will often avoid using IK in favor of the greater control over phrasing provided by FK (Benseghir, 2007). Successions have been explicitly represented in my work on motion editing (Neff & Fiume, 2003) and are part of the Hartmann et al. style model (Hartmann, Mancini, & Pelachaud, 2006). They can be simply implemented by offsetting the start and end time of each joint in a motion, moving from the root out to the extremities. Adding flourishes, such as an accompanying flexion and extension of the wrist, can aid the effect (Chi, Costa, Zhao, & Badler, 2000). Coleman et al. (Coleman, Bibliowicz, Singh, & Gleicher, 2008) introduce “staggered poses”, a representation for this offsetting of the motion of different joints, along with related editing commands. Adding successions to motions can increase a sense of flow.

PHRASES AND PHRASING

A *phrase* is defined as “any movement with a through line containing a beginning and an end” (Hackney & Meaden, 2009). It could thus be viewed as representing a “unit” of movement which may contain multiple actions. The phrase definition places no particular limit on duration. A phrase could be a sub-second flick or a kilometre long run if it is done with a single intent. The changes between phrases are marked by changes in intent or focus and can be somewhat difficult to identify. Individuals may have a preference for a particular phrase length and use this often in their movement (cf. tempo-rhythm).

Phrasing can be viewed as the process of patterning over time. It is the “manner of execution or the way in which energy is distributed in the execution of a movement or series of movements.” (Yvonne Rainer cited in Maletic, 2000).

A movement phrase can be broken into a series of phases (Hackney, 1998). These include, in order:

1. *Inner preparation*, the internal thought process that precedes the movement when the brain formulates a motor plan.
2. A *moment of initiation*, when the mover begins the motion. Action pathways are setup during this phase. The part of the body that initiates movement should be noted.
3. *Main action/Exertion* includes the primary action of the movement.
4. *Follow-through* carries the momentum of the body onwards once the main action has been completed.
5. *Recuperation* allows the mover to recover from the exertion of the phrase. Phrases are often preceded and followed by momentary pauses.

Emphatic phrases have a point of emphasis during the phrase, for example a Sudden wrist flick in the middle of a movement. Emphasis is generally notated as being at the beginning, middle or end, although more complex, multi-emphasis phrases are possible. Beginning emphasis could occur during initiation, middle during the main action and end emphasis either at the end of the action or in the follow-through phase, if it's present. Non-emphatic phrases are even throughout, such as a move in Tai Chi.

Maletic (2000) adds three additional types of phrasing. *Accented Phrasing* consists of a series of accents which together form a sequence. These involve exertions of energy followed by long or short periods of stillness. *Vibrating Phrasing* consists of a series of “sudden, repetitive movements”. *Resilient Phrasing* “creates several rebounding, resilient movements together forming an entity”.

Phrasing often involves changes in Effort qualities. Effort can be used to show emphasis in three ways.

Loading adds an additional effort dynamic to the movement. The *intensity* can be changed by increasing the “volume” of the effort qualities. Finally, in a *change of effort*, the effort qualities present are changed, often by reversing the factors (e.g. from Light Weight to Strong Weight). Accents can be added to movements through a combination of Weight and Time Effort (Hackney & Meaden, 2009). A Light accent uses Sudden Time and Light Weight. A Strong accent uses Sudden Time and Strong Weight.

Maletic also observes that phrases can be layered in time and across different body parts. In *consecutive* phrasing, one phrase completes and then the next begins. *Simultaneous* phrasing occurs when different parts of the body execute different phrasing patterns (e.g. middle emphasis or beginning emphasis) at the same time. In *overlapping* phrasing, one part of the body starts a new phrase while another part of the body is completing the previous phrase. This corresponds to the animation principle of *overlapping action*. As Disney once instructed his animators, “Things don’t come to a stop all at once, guys; first there’s one part, and then another” (cited on p.59, Thomas & Johnston, 1981).

Implementations

Detailed computational models of phrasing are lacking. This is in part because there has been limited work focused on the composition of detailed movement sequences. LifeForms (e.g. (Bruderlin, Teo, & Calvert, 1994)) is a notable exception as a system that foregrounds the composition process. In gesture work, the preparation, stroke, retraction model has become common (e.g. Hartmann, Mancini, & Pelachaud, 2002), and this shows clear parallels to the phases of a movement phrase.

COMPOSING MOVEMENT SEQUENCES

Actions need to be given adequate time if the audience is going to follow the inner thoughts of the character. Delsarte argued that there is a sequence of perception, recognition and action that underlies movements (Taylor, 1999). Allowing time for these different phases can make the communication with the audience more clear. This relates to the more general form of the *anticipation* principle in traditional animation. It is important to give the audience hints as to what is coming so that they are prepared for it and will perceive it when it arrives (Lasseter 1987; Thomas & Johnston, 1981).

Lasseter (1987) argues that it is important to spend the correct amount of time on anticipation for an action, the action itself and then the reaction to the action (Lasseter 1987). Timing can be used to indicate if a character is nervous, lethargic, excited or relaxed (Lasseter 1987). In order to give a sense of life, it is also important to show character’s thinking (Thomas & Johnston, 1981) and time must be allotted to this in the animation.

Secondary action can be added to a movement sequence in order to strengthen and clarify the meaning of a sequence. This can include small actions like a character wiping a tear, shaking his head or putting on glasses (Thomas & Johnston, 1981), as well as physical reactions to a movement, such as the flow of long hair.

REACTIONS AND CHARACTER INTERACTION

For characters to appear real, it is vital that they react to both other characters and objects in the virtual world. Much actor training focuses on preparing performers to react fluidly and spontaneously to other actors and their environment. Grotowski (1968d, p.225) describes the process of acting as follows: “Something stimulates you and you react. That is the whole secret.” The main principle of Grotowski’s work is *via negativa*; a process of eliminating the blocks that are inhibiting the actor (Barba, 1968) so that an actor’s body can be free and completely responsive to stimuli. According to LeCoq (2002, p.10), “reaction creates action”. “There is no action without reaction” (LeCoq, 2002, p.89). Movement is often a reaction to various stimuli: those that are internal, from another character or from the environment. Our characters must both sense these stimuli and react to them.

Contact is Grotowski’s term for an awareness of other people, their actions and their moods. He argues that theatre is composed of elements of human contact; “give and take”. These are what define the score for the actor. “Take other people, confront them with oneself, one’s own experiences and thoughts, and give a reply” (Grotowski, 1968a, p.212). Contact between characters must be maintained during a performance. As one actor makes small changes to the set performance, the other actor should make adjustments as well and hence small changes to the performance should emerge from that contact (Grotowski, 1968d). Contact is important for maintaining believability. Stanislavski also emphasized the importance of an actor relating his behaviour to the other characters on stage. He must honestly absorb what they say and do and react to them (Moore, 1984). The inner and outer adjustments people make to one another Stanislavski terms *adaptation* (Stanislavski, 1936). In the animation setting, it is similarly important that characters are sensitive to each other’s behaviour and fit their actions to what the other characters are doing. They must have *contact* and *adaptation*.

Characters must also adapt to overcome physical obstacles (Moore, 1984). A character must perceive and interact with objects. He must treat them as if they are what he wants the audience to perceive them to be. It is important to assign meanings to objects and then show those meanings in movement. For instance, a drink might be wine or poison. The character’s movements should reflect which it is. (Moore, 1984)

Regulating behavior (e.g. gaze changes) by which a speaker maintains or gives up the floor helps determine turn taking in conversations and is another important aspect of interpersonal interactions (Alberts, 1997). It is often dropped in stage movement or animation as the interactions are preset and so there is no need for it. Adding regulating behaviour to a piece of performance motion, however, can help increase realism. It is even more important for spontaneous character/avatar interactions in which the conversational turn must be actively managed.

If two characters are placed in the same scene, a number of factors become important, such as:

How each character moves relative to the other character (e.g. do they mirror postures?), what changes they make in stance and posture, how they position themselves in space and their relative orientation. These not only speak volumes about the relationship between the characters, but also comment on the personality and inner state of each character individually.

Following Alberts, the amount of space around the body that a person considers to be private or personal varies a great deal from culture to culture (Alberts, 1997). In North America, personal space extends outwards from the body roughly three feet in all directions, which is more than most other cultures. The intimate zone extends to about eighteen inches. The social zone is about four to ten feet. Four to six feet is safe for informal conversation, six to ten feet is more suited for formal interactions. The kind of greeting someone makes when passing a friend on the street is a function of the distance at which they pass. If they

are more than about twelve feet apart, a wave is fine. Ten to twelve feet normally requires them to verbalize. Passing within six feet, they should generally enter into a conversation.

Eye gaze is an important factor in indicating both a character's interests and personality. The interpretation ultimately made of a character's gaze direction will depend very much on the overall context of the scene.

Implementations

Work on these ideas is occurring at two interacting levels: agent architectures that determine how a character should behave in a given interaction and lower level movement models that determine the details of a character's motion. Agent architectures are beyond the scope of this chapter, but they must deal with perception issues, especially when agents are interacting with a human, and decision making to plan appropriate responses. At the motion level, characters are beginning to adjust to their interlocutors through mimicry. (For an example of work on turn taking, see (Cassell, Torres, & Prevost, 1999)).

ACTION SELECTION

The problem of action selection – deciding what movements a character should make – is not the main focus of this chapter. However, since it is one of the main tasks in animation, some discussion is warranted. Moore goes so far as to suggest “[t]he creative process of an actor's work is choice of actions...” (Moore, 1984, p.56).

The rule of simplification applies to action selection and restraint is important (Stanislavski, 1949) as everything that happens on stage must have a purpose (Stanislavski, 1936). Complex human psychological life should be expressed through simple gesture (Moore, 1984). Stanislavski writes “Unrestrained movements, natural though they may be to the actor himself, only blur the design of his part, make his performance unclear, monotonous and uncontrolled” (Stanislavski, 1949, p.69).

The actions performed must relate to the inner life of the character, or they will lack meaning (Stanislavski, 1949). Stanislavski's work allows an actor to develop the proper inner process, from which the outer expression will flow (Moore, 1984). Every psychological aim should be expressed physically and every action should have a psychological aim (Moore, 1984). Furthermore, gestures and movements must have concrete justifications. Otherwise, we are left with beautiful gestures with the “emotions of a fairy dance” (Grotowski, 1968c).

“Frequently physical immobility is the direct result of inner intensity, and it is these inner activities that are far more important artistically.” (Stanislavski, 1936, p.34) Stillness can denote inner emotional weight.

Action choice determines the type of character which is created. Different characters might share an objective, but they will choose different actions to achieve it. There should be a continuous line through a character's actions and these should build towards the superobjective; the overall goal of the character (Moore, 1984).

One of Stanislavski's fundamental principles is that a performer should never try to act a feeling.

Fix this for all times in your memories: On the stage there cannot be, under any circumstances, action which is directed immediately at the arousing of a feeling for its own sake. To ignore this rule results only in the most disgusting artificiality. When you are choosing some bit of action leave feeling and spiritual content alone. Never seek to be jealous, or to make love, or to suffer, for its own sake. All such feelings are the result of

something that has gone before. Of the thing that goes before you should think hard as you can. As for the result, it will produce itself. The false acting of passions, or of types, or the mere use of conventional gestures, -- these are all frequent faults in our profession. But you must keep away from these unrealities. You must not copy passions or copy types. You must live in the passions and in the types. Your acting of them must grow out of your living in them.” (Stanislavski, 1936, p.38, original emphasis)

Actions must arise out of a context. Trying to act an emotion without connecting it to the context of the story will make for empty gestures.

Grotowski makes a similar argument, saying that actors should not illustrate words, suggesting that a bad actor who is asked to act bored will try to show boredom. His actions and gestures illustrate the word. A man who is actually bored will be very active. Perhaps he will read a book, lose interest and put it down. Then he'll look for some food, but nothing tastes good today. Then maybe he'll try to have a nap, but be unsatisfied with this as well (Grotowski, 1968d). It is good to remember that behaviour is composed of small, logical, concrete actions (Moore, 1984).

Emotions should be made specific to a given character. Again, we can turn to Stanislavski:

Most actors do not penetrate the nature of the feelings they portray. For them love is a big and generalized experience.

They try immediately to ‘embrace the unembraceable.’ They forget that great experiences are made up of a number of separate episodes and moments. These must be known, studied, absorbed, fulfilled in their entirety. Unless an actor does this he is destined to become the victim of stereotype. (Stanislavski, 1949, p.272)

Emotions are shown through a number of movements that are customized to be meaningful for a given character in a given situation. Emotional truth lies in finding these correct actions and performing them appropriately.

When creating a particular character, it is important to avoid easy stereotypes. Stanislavski's description of an old man is illustrative of the dangers here. “[The joints of an old man] rasp and squeak like rusty iron. This lessens the breadth of his gestures, it reduces the angles of flexibility of his torso, his head. He is obliged to break up his larger motions into a series of smaller ones and each has to be prepared before he makes it” (Stanislavski, 1949, p.29). A young man might turn his hips 50-60 degrees while an old man moves 20 degrees and slowly. The tempo and rhythm of an old person's motions are slow and flaccid. Yet Stanislavski criticizes an actor that directly implements this: “You are keeping constantly to the same slow rhythm and pace as you walk to an exaggerated caution in your gestures. Old people are not like that” (Stanislavski, 1949, P.30). Old people will change their rhythm, have periods of preparation followed by periods of speed. At times they are limbered up and there is momentum to their movements. They prepare actions much more than the young. For example, when sitting down they will feel for the chair and prepare to sit, sit, pause and then finally lean back. Once seated, the difficult part is done. The person can show more vigor, but while seated and with a limited joint range (Stanislavski, 1949). There is a varied texture to the movement. The old man does not perform all motions in the same way. Also, notice that both the sequence of actions performed and the manner change when an old man sits down versus a young. The old man adds additional actions to reach for the chair seat and make sure he is positioned correctly. A structural change must be made to the motion in order to take these differences into account. Both how the motion is performed and what motions are performed must be changed when the character is changed.

Clichés should be avoided. Grotowski suggests that when a character says “What a beautiful day”, he should not always say it with a happy intonation nor “Today I am a little sad” with a sad intonation. These are clichés. Strive instead for what the character’s deeper intention is, what lies below (Grotowski, 1968d). To determine such deeper intentions, it is helpful to analyze the character’s larger life, to understand her goals beyond a particular scene and then use her actions in a particular moment to provide insight into the larger internal thoughts and motivations of the character.

Implementations

Action selection is as much, or more, a problem of AI as it is of animation. The AI work is beyond the scope of this chapter, but there is some particularly relevant work drawing from the arts literature that is worth mentioning. Part of Delsarte’s analysis is a set of suggestions on how particular poses convey emotions or other connotations. Some researchers have suggested that this may be particularly useful in virtual agent research as it provides a defined set of behaviors for a character to perform in order to achieve a given impact, which is what the agent designer is normally after. Marsella et al. (Marsella, Carnicke, Gratch, Okhmatovskaia, & Rizzo, 2006) explore Delsarte’s “Attitudes of the Hand”, which considers the “gesture cube”, an imaginary box placed in front of the agent, and suggests meanings associated with having the hand posed on each face of the cube. They perform perceptual experiments in which an agent moves to one of these faces and ask users to rate the impression, confirming that some, but not all, of Delsarte’s mappings appear to hold for virtual agents. Nixon et al. (Nixon, Pasquier, & El-Nasr, 2010) introduce DelsArtMap, a mapping between nine emotions and character poses, that allows a user to specify emotional intensity values and yields an output pose. The poses are based on Delsarte’s division of body part positions into three categories: excentric, normal or concentric. For the head, these can be thought of as away, neutral or towards and are applied to both the twist and up-down rotation, yielding nine combinations. Their system includes poses for the head, torso, legs and arms. In future work, they plan to experimentally validate these mappings.

Another pose-driven approach to animation, focused on supporting choreography, DanceForms (Calvert, Wilke, Ryman, & Fox, 2005; Wilke, Calvert, Ryman, & Fox, 2005) provides a way to go from a structured representation of movement to 3D animation. The system takes Labanotation as input, a method for notating dance that is comparable to a score for music and specifies poses at points in time, along with additional information about the nature of the movement. DanceForms resolves ambiguities in the notation and generates output 3D animation. While this work is aimed largely at supporting composition, it demonstrates the great difficulty of generating movement from an external specification. Even with a very detailed description language like Labanotation, ambiguity remains that must be resolved in order to generate appropriate motion. Indeed, the authors argue that “a recurring issue ... is the need for a unique, unambiguous way to represent human movement”[p.6]. Having such a precise movement language would be of enormous value.

3. SIDEBAR: SIMPLE TECHNIQUES FOR IMPROVING CHARACTER MOTION

I describe here a few simple computational methods that are derived from material in the arts, can easily be added to a computational model and, in my experience, improve the quality of animation.

The arms do not start at the shoulders. Arm movements should include the collarbones, and ideally, one should also mobilize the torso and whole body as part of a gesture (cf. Patterns of Total Body Connectivity and Posture Gesture Merger). A simple way to add collarbone movement is to map the angle of the collarbones to the height of the hand. Simple IK can be applied from the shoulder to the wrist and the

additional mapping will automatically raise the collarbones as the character raises his arm. Beyond this basic biomechanical connection, the collarbones are very important for showing emotional changes and opening and closing the torso.

The torso has a strong expressive impact on character motion. A simple, but surprisingly effective method for engaging the torso is to map the curvature of the spine to the motion of the arm so that the torso varies in unison with arm movement. Different types of curvature will read differently. For instance, a forward C curve in the sagittal plane can be used to create an old or tired character. A large C curve in the coronal plane (vertical plane in LMA terms) can be used to create a more happy go lucky character that sways from side to side. An S curve in the same plane can create a more sensual feel. For a more detailed discussion of how this can be implemented as a motion editing tool and ways to automatically extract these relationships, see (Neff & Kim, 2009; Neff, Albrecht, & Seidel, 2007).

The swivel angle can be defined as the rotation around an axis that runs from the collarbone to the wrist. It moves the elbow close to the torso or out from the character's side. As Delsarte observed, this has a strong impact on the perception of personality and emotional state. Allowing direct control of this parameter provides a useful expressive handle. Packages like IKAN (Tolani, Goswami, & Badler, 2000) implement this and are freely available.

Our work (Neff & Kim, 2009) also presents algorithms that provide efficient control over the lower body through specifying intuitive parameters: control over the balance point, pelvic twist, foot positions and knee bends. We find these parameters that have the most expressive impact on the pose of the lower body. Gaze behavior can contribute significantly to making a character appear as though he is thinking. IK applied to the head and neck provides a simple way to implement this. Looking up and to the left is often associated with thinking.

Providing direct control over the motion envelope has become a standard technique in animation tools through variation of spline tangents and should be supported in any virtual character application. It is important to allow control over where the velocity peak occurs in the motion and to support anticipation and overshoot effects.

Controlling the succession in a movement can be implemented by first solving for a full body pose and then offsetting the timing of each joint in the pose. For example, the shoulder could start a couple of frames after the collarbones, the elbow a frame after that, etc.

One of the main advantages of physically simulating a conversational character is that it provides an effective method for adding pendular arm movement to the character's motion when they drop their arm to their side (e.g. (Neff & Fiume, 2002)). This enhances the naturalness of the motion. Simulation also provides a way to capture the impact of momentum on a character's movement.

4. LOST IN TRANSLATION: CHALLENGES IN APPLYING ARTS MATERIAL COMPUTATIONALLY

Character animation for virtual worlds must generate the correct movements to express personality and mood, a goal often shared with the arts. Nonetheless, it is a challenge to directly apply lessons from the performing arts literature to automatic animation production for a number of reasons. First, much of the literature is meant for embodied performers who can interpret this material within their own physical

bodies and against their substantial experience as movers. Actors develop much of their knowledge of movement through physical exercises, as is routinely emphasized (e.g. (Stanislavski, 1949; Moore, 1984; Grotowski, 1968b). These exercises serve a range of purposes, from providing a way for an actor to investigate movement (Barba, 1968), to developing concentration and an ability to notice and react to external stimuli (Barba, 1968), to developing a sense of rhythm.

Actors can also tap their emotional memories. Only a living actor has the muscle memory of actions performed during an emotional crisis, illustrating the psycho-physical connection between the internal mental experiences of actors and their outward movements that is lacking in virtual characters. For example, Stanislavski considered *emotional memory* to be a very important source for an actor (Moore, 1984), suggesting that an actor must execute the correct physical action to relive the emotion required for a scene. Grotowski also argued for the importance of drawing on specific, real, intimate experiences (Grotowski, 1968d). Our ability to recall emotions and find through our body correct actions for them is likely one reason why animators often work out scenes in front of a mirror or by videotaping themselves. They need to be able to see what their body decided to do (or what movement they subconsciously triggered).

A second challenge is that motion is very complex, occupying a high dimensional space, with somewhere between 40 and over 100 degrees of freedom required to specify a pose of the human body, depending on the amount of detail required. While these DOFs are not all independent, it remains a challenging specification task. Moreover, this level of detail is entirely different than what is required in the arts when describing ideas for performers that can interpret these against their own experience. Actors do not need an external representation of movement, whereas computer models need to generate an explicit and precise representation. Applying the arts material thus requires the establishment of a mapping between the literature and concrete, computational representations such as procedural algorithms or motion libraries. This can be a challenge as movement experts may talk in analogies or at a more general level of detail than the very fine-grained representation required computationally. The arts properties are not precisely defined, in a computer science sense, and determining definitions is a substantial challenge. Indeed, defining this mapping is equivalent to defining the complete, unambiguous language for movement that Calvert (Calvert, Wilke, Ryman, & Fox, 2005) argues is missing.

Even when movement qualities can be given a precise definition, it is difficult to know the correct movements to perform in order to generate a particular impression with an audience. This decision is highly dependent on context, which may include the role of the character, his environment, his current emotions, who he is talking to, events earlier in the conversation, etc. The mappings are generally not simple and not fixed. Earlier in the chapter, I made an effort to focus on the most concrete and easy to apply ideas from the literature, but it is a mistake to blindly apply these properties in a prescriptive way. They are not recipes that can be followed to give a specific result, but abstract principles that give insight to the nature of movement. As Shawn says,

We have to extract [the laws of Delsarte], make simple statements of them for study purposes; but in actuality, in the final use, a vast number of these laws are operating simultaneously and mutually modifying and affecting each other to produce complex yet seemingly simple results. (Shawn, 1963)

There is a level of artist interpretation that is fundamental to these principles: “[N]o two artists using the same principle would produce forms of expression identical with each other, or with anyone else.” (Shawn, 1963, p.26). LeCoq (2002) argues that there is generally not one meaning for a movement nor vice versa.

Finally, perhaps the biggest challenge is that these motion qualities are not easily separable nor necessarily orthogonal. As Shawn suggests above, many qualities are combined in a single piece of movement. The qualities interact with each other in generating a final impression, and related descriptive properties

may alter the same aspects of movement. There is also a many to one relationship between the terms and movement. While trained observers may agree that a particular movement contains a particular movement quality, for instance Strong Weight, there is a vast array of different movements that could also contain this quality. Many more details need to be specified to move from a description, such as “Strong Weight”, to a complete movement that realizes it.

5. STEPS FORWARD

Despite the challenges, the arts literature still provides very important guidance for creating effective virtual worlds. It provides a road map that illustrates which movement properties are important and must be included in any motion representation. It is a valuable resource for anyone developing character movement, regardless of the domain, and indeed, can support a range of applications. It indicates useful ways to edit motion capture, provides guidance on the range of motions that should be included in a motion library and also indicates the type of AI problems that will need to be solved to generate effective character motion. There are deep questions here, in terms of how movement should be varied to adjust to context, how to generate a particular character with a unique movement signature and how to decide the right actions to use in a given situation.

A great deal of progress has been made on showing how ideas from the arts can be implemented to improve virtual character motion. I think there are three areas particularly ripe for further exploration. Full body engagement is still lacking from most virtual character systems. It is common for the arms to be controlled independently of the rest of the body and this leads to unnecessarily stiff and robotic movement. Phrasing has only been lightly explored. It is a complicated topic, involving issues in animation, AI and motion perception, but better models of phrasing will likely be important in developing personal movement profiles for characters. Finally, virtual worlds need to start making better use of deformations for breath and other variations of the character’s mesh, ranging from muscle movement to dynamic effects. There has been significant research on this in the character animation community, but it has been slow to migrate to virtual worlds.

While it remains an open question as to whether procedural approaches are the best way to leverage off the arts material, or if learning approaches applied to motion data will provide greater strides, the literature offers value to either approach. It defines what matters in movement and can therefore be used as a measuring stick to evaluate various techniques. What set of movement qualities does a technique cover and what set does it leave out? How complete is any given model?

Perceptual experiments will likely gain importance as this work develops. They allow us to both test claims made in the literature and validate particular implementations of these ideas. Marsella’s (Marsella, Carnicke, Gratch, Okhmatovskaia, & Rizzo, 2006) pioneering work on testing a hand orientation mapping proposed by Delsarte is inspirational in this regard. Such efforts will need to be continued and refined to better understand the role that various components of motion play in generating a particular impression. It will also be important to be able to compare the output of various animation systems, both to ensure consistency and to better refine movement models by learning best practices.

6. BIBLIOGRAPHY

- Alberts, D. (1997). *The expressive body: Physical characterization for the actor*. Portsmouth, N.H.: Heinemann.
- Appia, A. (1962 (original 1921), fifth edition 1982). *The work of living art: A theory of the theatre*. Coral Gables, Florida: University of Miami Press. (Translated by H.D. Albright)
- Barba, E. (1968). *Actor's training (1959-1962)*. In J. Grotowski (Ed.), *Towards a poor theatre*. New York: Routledge, A Theatre Arts Book.
- Barba, E. (1991a). Dilated body. In E. Barba & N. Savarese (Eds.), *A dictionary of theatre anthropology: The secret art of the performer*. London: Routledge.
- Barba, E. (1991b). Theatre anthropogoy. In E. Barba & N. Savarese (Eds.), *A dictionary of theatre anthropology: The secret art of the performer*. London: Routledge.
- Barba, E. (1991c). Theatre anthropology: First hypothesis. In E. Barba & N. Savarese (Eds.), *A dictionary of theatre anthropology: The secret art of the performer*. London: Routledge.
- Barba, E., & Savarese, N. (1991). *A dictionary of theatre anthropology: The secret art of the performer*. London: Routledge.
- Benseghir, D. (2007). Keynote address, *Intelligent Virtual Agents 07*.
- Bruderlin, A., Teo, C. G., & Calvert, T. (1994). Procedural movement for articulated figure animation. *Computers & Graphics*, 18 (4), 453-462.
- Calvert, T., Wilke, W., Ryman, R., & Fox, I. (2005). Applications of computers to dance. *Computer Graphics and Applications, IEEE*, 25 (2), 6-12.
- Cassell, J., Torres, O., & Prevost, S. (1999). Turn taking vs. discourse structure: How best to model multimodal conversation. *Machine conversations*, 143-154.
- Chi, D. M., Costa, M., Zhao, L., & Badler, N. I. (2000). The EMOTE model for effort and shape. In *Proc. SIGGRAPH 2000* (pp. 173-182).
- Coleman, P., Bibliowicz, J., Singh, K., & Gleicher, M. (2008, July). Staggered poses: A character motion representation for detail-preserving editing of pose and coordinated timing. In *2008 ACM SIGGRAPH/ Eurographics Symposium on Computer Animation*.
- Dorcy, J. (1961). *The mime*. Robert Speller and Sons, Publishers, Inc. (Translated by Robert Speeler, Jr. and Pierre de Fontnouvelle)
- Eisenstein, S. (1996). On recoil movement. In A. Law & M. Gordon (Eds.), *Meyerhold, Eisenstein and biomechanics: Actor training in revolutionary Russia*. Jefferson, North Carolina: McFarland and Company, Inc., Publishers.
- Eisenstein, S., & Tretyakov, S. (1996). Expressive movement. In A. Law & M. Gordon (Eds.), *Meyerhold, Eisenstein and biomechanics: Actor training in revolutionary Russia*. Jefferson, North Carolina: McFarland and Company, Inc., Publishers.
- Gallaher, P. E. (1992). Individual differences in nonverbal behavior: Dimensions of style. *Journal of Personality and Social Psychology*, 63 (1), 133-145.
- Grochow, K., Martin, S. L., Hertzmann, A., & Popović, Z. (2004). Style-based inverse kinematics. *ACM Transactions on Graphics*, 23 (3), 522-531.
- Grotowski, J. (1968a). *The actor's technique*. In E. Barba (Ed.), *Towards a poor theatre*. New York: Routledge, A Theatre Arts Book.
- Grotowski, J. (1968b). *Actor's training (1966)*. In E. Barba (Ed.), *Towards a poor theatre*. New York: Routledge, A Theatre Arts Book.
- Grotowski, J. (1968c). *American encounter*. In E. Barba (Ed.), *Towards a poor theatre*. New York: Routledge, A Theatre Arts Book.
- Grotowski, J. (1968d). *Skara speech*. In E. Barba (Ed.), *Towards a poor theatre*. New York: Routledge, A Theatre Arts Book.
- Grotowski, J. (1968e). *Towards a poor theatre*. In E. Barba (Ed.), *Towards a poor theatre*. New York: Routledge, A Theatre Arts Book.

- Hackney, P. (1998). *Making connections: Total body integration through Bartenieff fundamentals*. Routledge.
- Hackney, P., & Meaden, J. (2009). Berkeley laban movement analysis certification program. *Integrated Movement Studies*.
- Hartmann, B., Mancini, M., & Pelachaud, C. (2002). Formational parameters and adaptive prototype installation for MPEG-4 compliant gesture synthesis. In *Proc. Computer Animation 2002* (p. 111-119).
- Hartmann, B., Mancini, M., & Pelachaud, C. (2006). Implementing expressive gesture synthesis for embodied conversational agents. In *Proc. Gesture Workshop 2005* (Vol. 3881, pp. 45-55). Berlin; Heidelberg: Springer.
- Kider Jr, J., Pollock, K., & Safonova, A. (2011). A data-driven appearance model for human fatigue. In *Proceedings of the 2011 ACM Siggraph/Eurographics Symposium on Computer Animation* (pp. 119-128).
- Kochanek, D. H. U., & Bartels, R. H. (1984). Interpolating splines with local tension, continuity, and bias control. *Computer Graphics (Proceedings of SIGGRAPH 84)*, 18 (3), 33--41. (Held in Minneapolis, Minnesota).
- Laban, R. (1988). *The mastery of movement* (fourth ed.). London: Northcote House. (Revised by Lisa Ullman).
- Lamb, W. (1965). *Posture and gesture: an introduction to the study of physical behavior*. London: Duckworth.
- Lamb, W., & Watson, E. (1979). *Body code: The meaning in movement*. Londres.
- Lasseter, J. (1987). Principles of traditional animation applied to 3d computer animation. *Proceedings of SIGGRAPH 87*, 21 (4), 35-44.
- Law, A., & Gordon, M. (1996). *Meyerhold, Eisenstein and biomechanics: Actor training in revolutionary Russia*. Jefferson, North Carolina: McFarland and Company, Inc., Publishers.
- Lawson, J. (1957). *Mime: The theory and practice of expressive gesture with a description of its historical development*. London: Sir Isaac Pitman and Sons Ltd. (Drawings by Peter Revitt)
- LeCoq, J. (2002). *The moving body: Teaching creative theatre*. New York: Theatre Arts Book, Routledge. (Translated by David Bradby)
- Maletic, V. (2000). The identification of phrasing types in movement and dance.
- Marsella, S., Carnicke, S., Gratch, J., Okhmatovskaia, A., & Rizzo, A. (2006). An exploration of Delsarte's structural acting system. In *Intelligent virtual agents* (pp. 80-92).
- Moore, S. (1984). *The Stanislavski system: The professional training of an actor*. Penguin Books.
- Neff, M., Albrecht, I., & Seidel, H.-P. (2007, September). Layered performance animation with correlation maps. *Computer Graphics Forum (Eurographics)*, 26 (3), 675-684.
- Neff, M., & Fiume, E. (2002). Modeling tension and relaxation for computer animation. In *Proc. ACM SIGGRAPH Symposium on Computer Animation 2002* (pp. 81-88).
- Neff, M., & Fiume, E. (2003). Aesthetic edits for character animation. In *Proc. ACM SIGGRAPH/Eurographics Symposium on Computer Animation 2003* (p. 239-244).
- Neff, M., & Fiume, E. (2004). Methods for exploring expressive stance. In *Proc. ACM SIGGRAPH/Eurographics Symposium on Computer Animation 2004* (p. 49-58).
- Neff, M., & Fiume, E. (2005). AER: Aesthetic Exploration and Refinement for expressive character animation. In *Proc. ACM SIGGRAPH/Eurographics Symposium on Computer Animation 2005* (pp. 161-170).
- Neff, M., & Fiume, E. (2006). Methods for exploring expressive stance. *Graphical Models*, 68 (2), 133-157.
- Neff, M., & Kim, Y. (2009). Interactive editing of motion style using drives and correlations. In *Proceedings of the 2009 ACM SIGGRAPH/Eurographics Symposium on Computer Animation* (pp. 103-112).
- Nixon, M., Pasquier, P., & El-Nasr, M. (2010). DelsArtMap: Applying Delsarte's aesthetic system to virtual agents. In *Intelligent Virtual Agents* (pp. 139-145).

- Rose III, C. F., Sloan, P.-P. J., & Cohen, M. F. (2001). Artist-directed inverse-kinematics using radial basis function interpolation. *Computer Graphics Forum*, 20 (3), 239-250.
- Shawn, T. (1963). *Every little movement: A book about Francois Delsarte* (Second Revised ed.). New York: Dance Horizons, Inc.
- Shin, H., Kovar, L., & Gleicher, M. (2003). Physical touch-up of human motions. In *Proceedings of the 11th Pacific Conference on Computer Graphics and Applications* (Vol. 194).
- Stanislavski, C. (1936). *An actor prepares*. Theatre Arts, Inc.
(Translated by Elizabeth Reynolds Hapgood)
- Stanislavski, C. (1949). *Building a character*. Theatre Arts Books.
(Translated by Elizabeth Reynolds Hapgood)
- Stanislavski, C. (1961). *Creating a role* (H. I. Popper, Ed.). Theatre Arts Books.
(Translated by Elizabeth Reynolds Hapgood)
- Steketeer, S. N., & Badler, N. I. (1985). Parametric keyframe interpolation incorporating kinetic adjustment and phasing control. In *Computer graphics (proceedings of SIGGRAPH 85)* (Vol. 19, pp. 255-262).
- Tak, S., & Ko, H. (2005). A physically-based motion retargeting filter. *ACM Transactions on Graphics (TOG)*, 24 (1), 98-117.
- Tak, S., Song, O. young, & Ko, H.-S. (2000). Motion balance filtering. *Computer Graphics Forum*, 19 (3), 437-446.
- Tarver, J., & Bligh, K. (1999). *The physical art of the performer [Workshop: A 30 hr. intensive introduction to Laban and Grotowski]*. (Workshop: A 30 hr. intensive introduction to Laban and Grotowski held at the Nightwood Theatre studio, Toronto.)
- Taylor, G. (1999). *Francois delsarte: A codification of nineteenth-century acting*. Theatre Research International, 24 (1), 71-81.
- Thomas, F. (1987a). Can classic Disney animation be duplicated on the computer? In B. Kroyer & P. Bergeron (Eds.), *3d character animation by computer* (Vol. 5). ACM Siggraph. (Originally appeared in *Computer Pictures*, July/Aug 1984)
- Thomas, F. (1987b). The future of character animation by computer. In B. Kroyer & P. Bergeron (Eds.), *3d character animation by computer* (Vol. 5). ACM Siggraph.
- Thomas, F., & Johnston, O. (1981). *The illusion of life: Disney animation*. New York: Abbeville Press.
- Tolani, D., Goswami, A., & Badler, N. (2000). Real-time inverse kinematics techniques for anthropomorphic limbs. *Graphical models*, 62 (5), 353-388.
- Torresani, L., Hackney, P., & Bregler, C. (2007). Learning motion style synthesis from perceptual observations. *Advances in Neural Information Processing Systems*, 19, 1393.
- Welbergen, H. van, Reidsma, D., Ruttkay, Z., & Zwiers, J. (2010). Elckerlyc-A BL realizer for continuous, multimodal interaction with a virtual human. *Journal of Multimodal User Interfaces*, 3 (4), 271-284.
- Wilke, L., Calvert, T., Ryman, R., & Fox, I. (2005). From dance notation to human animation: The labandancer project. *Computer Animation and Virtual Worlds*, 16 (3-4), 201-211.
- Wooten, W. L. (1998). *Simulation of leaping, tumbling, landing, and balancing humans*. Ph.D. dissertation., Georgia Institute of Technology.
- Zordan, V., Celly, B., Chiu, B., & DiLorenzo, P. (2004). Breathe easy: model and control of simulated respiration for animation. In *Proceedings of the 2004 ACM SIGGRAPH/Eurographics Symposium on Computer Animation* (pp. 29-37).

10

THEATER AS VIRTUAL REALITY

By Jim R. Parker

For millennia, people have participated in and have attended theatrical performances. Early examples would have taken place around a fire and were likely to have portrayed the glories of the day's hunt, a battle against a neighboring village, or the complex and fascinating relationships between the gods. Before formal rules of theatre were established these performances would likely have been interactive. Audience members would have participated in the actual events being described, and could jump up and contribute (or be dragged in) to the presentation; children could ask questions, which could be answered immediately as an extension of the performance; audience members could shout and sing, in response to the prompting of the 'cast'; and so on. At these times an entire village could be engaged simultaneously. One imagines that there would be food and drink, perhaps music and dancing. It would be a key social activity for the village, as well as a way to transmit history and tradition.

As the centuries passed and theatre became more of a formal entertainment, rules and conventions developed for its conduct. These rules were culturally derived, so that Egyptian theatre differed from Greek or Chinese, but in most cases the interaction between player and audience gradually disappeared. The audience was assigned a passive role which is largely the situation today. Audiences sit quietly in a darkened theatre and observe the proceedings on stage. As film and television became prominent the distance between the audience and the performers grew; in fact, the performers in many instances can't see or hear the audience at all. The distance is now one of space and time, as film preserves a performance for the ages with perfect replication. With the relatively recent development of the video game technologies and virtual reality this distance has been reduced, and new opportunities for communication have been enabled.

The term *virtual reality* was first used by French director Antonin Artaud in a theatrical context (Artaud, 1958). He described the virtual reality (*la realite virtuelle*) of the theater as a place where the characters, objects, and images of theatre form a purely fictitious and illusory world. In a very natural way, a theatrical performance represents a kind of virtual reality by nearly any definition. UNESCO, for instance, defines virtual reality as¹⁸:

an immersive and interactive simulation of either reality-based or imaginary images and scenes

This definition is clearly also a description of theatre.

The role of the director of a play is to engage the audience (immersion) in an experience that is a narrative based on reality or fiction presented in real time on a stage by people pretending to be participants in the experience (simulation) (Geigel, 2004). So, virtual reality would seem to have a natural and historical connection to theatre, even if the technology of virtual reality is not commonly used as a part of modern stagecraft. The next steps towards a truly virtual theatre have been made, so that not only can a theatrical production be thought of as a virtual reality but it can be conducted within a non-real space, one made tangible by the computer.

¹⁸www.unesco.org/education/educprog/lwfi/doc/portfolio/definitions.htm

1. VIRTUAL THEATRE

Virtual reality was originally developed in a practical way as an adjunct to *telepresence*. The idea was to allow a human to operate machinery from a distance by using telecommunications technology and local (to the operator) interfaces and displays. An example would be the remotely operated vehicles (submarines) that were used to help repair the oil well leak left from the Deepwater Horizon which sank on 22 April 2010 (Newman, 2010). The well was a mile beneath the surface of the sea, and the pressure at the site would have been about 2350 psi (a ton per square inch) and the temperature was near freezing. It makes sense to use robotic devices to work in such dangerous places, and providing a natural interface to the operator permits him/her to work more effectively. The operator senses the remote site through robotic data collection devices and these data need to be presented as the normal human senses of sight, hearing, and touch. In the best situation, the operator would feel as if they were actually at the remote location. This feeling is called presence. The remote senses define a *virtual space*, which is that volume that can be detected by the sensors and the ability of those sensors and the display to present a view to the operator.

In what is sometimes called *virtual theatre*, the remote location is imaginary. It is a simulated theatre consisting of visualized polygons, existing only in the computer's memories and storage devices. The sensations that are returned to the operator are simulated too, or represent a real sound or image from one of the other participants. In virtual theatre a play can be performed live on a computer graphic stage by real actors, whose voices are transmitted to the participants through computer networks. The actors are not in the virtual space, of course, but are represented by a graphical rendering of the actor, which could look like anything at all. This representation, or *avatar*, could look like the real actor, or like some other real actor; it could look like an animal or an animated object like a lamp; or it could be an abstract shape or lighting effect. The important thing is that the story can be told in this virtual space, using music, voice, and animate objects. It could be told on a stage, but could just as easily be in an open space, in the sky, or under water.

JIM PARKER ON HIS METHODS

Non-verbal communication is a key aspect of theatre, and has been from the beginning. Playwrights weave nonverbal cues into their plays, and actors take those and expand them as they see the need to clarify and expand on the messages being sent to the audience. A computer game is both a virtual world and, in many cases, a theatrical performance. Theatre has a virtual aspect that has not been lost on dramaturges of the past, and some techniques of dramaturgy can be and have been drawn into game design.

Computers have been used by theatrical designers for some time. The ability of a designer to change a design quickly has become standard practice. The use of virtual spaces for design, where the work is done in three dimensions from the beginning, is less common, but serves the theatre better in many ways. Being able to actually walk through a set design like an actor on a real stage presents some clear advantages. Use of a virtual world for design adds a simulation aspect to the basic graphical capability of computers that has been used in the past.

A final step is the use of virtual worlds, not just for design, but for performance. Many thousands of people can be connected together in spaces like *Second Life* and *World of Warcraft*, and can see and hear the same things at the same time. Since people are getting more and more of their entertainment online, it makes sense to offer live theatre online too. This provides many advantages, including the fact that the designs not become the reality - the designs live in the same space where the performance will take place, and are real in that space. Changes to the set now take place in minutes.

(continued on next page)

In a virtual theatre there is in fact no 'operator' per se, but there are normally two classes of participant. In the first class are the performers who take their traditional place on stage, if such exists, as those who transmit the narrative. Their view of the virtual space is that of the more traditional operator in telepresence; they, along with the technical support people, control the operation of the play. Their motions and speech are sent to the virtual space, and the view from that space is sent back to them. The second class of participant is the audience. In a traditional performance their role is passive. They observe the performance without actively participating. However, they can not only see the play being performed by their motions and speech are also transmitted to the virtual space. There is in fact no practical difference between a performer and an audience member except the roles that have been assigned to them.

What does this do to live theatre? Nothing except to add a new option, and to increase the value of computers for designers. It does open new venues and offer new lines of research for people working in computer mediated arts, and it creates new audiences where old ones are declining in number. Live virtual theatre provides one strong motivation for the development of higher resolution virtual spaces, ways to more thoroughly use available bandwidth, ways to capture face and gesture information from actors, ways to create dialogs and narrative dynamically, to allow for more complete audience involvement. It could involve nothing less than a rethinking of how theatre works at all levels, and a concomitant effect on game design.

This is true in a traditional theatre too, of course. Audience members do not jump on stage and engage the performers because there is a strong social convention against it. Interestingly, in the virtual performances that have been described no audience member interfered with the show in any way. The social convention seems to apply even though the participants are effectively anonymous and in disguise.

Mounting a performance in a virtual space requires the same steps as does a performance in a real theatre, but requires that they be done differently. There are also new aspects to virtual performance that are not obvious at the outset, especially to one who is trained in traditional design or stagecraft. In many cases physical laws are negotiable in virtual spaces, so gravity, for example, is not a constraint. On the other hand, motion and facial expressions of actors are constrained, sometimes severely, and interactions with props and sets are restricted as well. The unconventional properties of the space need to be understood by the director and stage manager, and a set of protocols need to be established so that the performance can best take advantage of the possibilities.

Second Life is an online virtual world that allows free access to anyone who can connect through the Internet. It allows voice communication, so actors can deliver lines in a traditional manner, and has facilities for building objects and costumes. There have already been performances in *Second Life* (SL). Theatres and musical venues exist and are patronized by the tens of thousands of denizens of the world. This world is typical of the spaces available online, and will be the subject of most of the remaining discussions, although others will be mentioned.

PERFORMANCE SPACES

It is a strange fact that, while many theatres exist in *Second Life*, they are mostly of a conventional thrust stage or proscenium type. There is no need for a stage at all. The audience can teleport to the place where they can observe, so no front of house would seem to be required except for pre-show social intercourse. The performance itself can take place in a quite natural looking location without need of a stage; in fact, a stage would be restrictive. If the play calls for a street, then the performance should take place in a street. If more viewing room is needed, then perhaps the audience could be arranged along a transparent wall. So why are stages so prevalent in virtual performances? An informal poll shows that designers and directors are simply more familiar with the techniques used in a proscenium stage than they are in virtual reality.



Figure 10-1: Sample performance spaces in *Second Life*. They all tend to have audiences seated facing a stage, which forms a frame around and constrains the actions of performers. Such stages are not necessary in virtual spaces, but are comfortable for the designer and director.



Figure 10-2: The Cheap Theatre venue, designed to be used for virtual performances. (Top) Traditional layout with ramped audience and stage. (Bottom) 2D Street layout designed for a performance of *Waiting for Godot*.

A lot of their training is in techniques that accommodate the constraints associated with traditional spaces. A significant minority of designers are intrigued by the opportunities offered. The idea of having a spherical theatre with no gravity offers some interesting possibilities.

Figure 10-1 shows a sample of typical performance venues in *Second Life*. All have seats and a stage of some type as a focus. On the other hand, Figure 10-2 shows the *Cheap Theatre* venue configured as a more or less traditional theatre for a collection of short plays, and as a street, for a performance of *Waiting for Godot*. All areas of *Second Life* are reconfigurable, so any novelty is in how the spaces are actually used. This space resides on the smallest parcel of virtual land that can be leased from Linden Labs, the operators of *Second Life*. The traditional theatre configuration has a stage and seats of a sort but no walls except when needed, and no roof. The *Godot* set has no stage or seats but offers a space where the audience can ‘hover’ while watching the play. The curved wall behind them can be made transparent if a large group shows up or can be removed completely in a few seconds. Or, in fact, the entire theatre construction can be removed for open space performances.

The thought that a venue can be tailored to fit a production is a foreign one to stage designers. Yes, the stage can be set and decorated, but the structure of the theatre itself can’t normally be altered. This may be one of the key advantages in virtual spaces: to make the space suitable to the show.

MIS EN SCENE

The French phrase for “put in scene” or “place on stage”, *mise en scene* is the term used to describe the design aspects of a theatre production. It’s an ill defined term, but can refer to set design, costume, props and lighting. The related term stagecraft refers to more technical matters, including constructing sets, hanging and focusing lights, makeup, prop construction, and other such issues. Stagecraft is the implementation, whereas *mise-en-scene* is the design. In virtual theatre, the two are really one.

While design takes place in the mind, an early stage in communication of a design is to create a drawing or model. Many designers use computer based tools these days in addition to or instead of pencil and paper. The next step would normally be to realize the design as an object: a set, prop, or costume item. In virtual reality the computer design can often actually be used within the computer generated space as the object itself. Both the computer design tools and the virtual space naturally manipulate objects that are native to computer representation, and so should be able to share them. Such an object is really an artifact of 3D graphics. Polygons and solid objects are used to build graphical constructions within a simulated 3D volume, and these can be rendered from any particular viewpoint: that of the audience members or the actors.

This blurring of the design representation with the actual object is the second important new feature of virtual theatre (the first being the lack of default physical laws). This can be both an advantage and a disadvantage to stage managers. On the one side, “Hey, no set painting”! On the other side, objects will only do what the designer says they can. An actor can only pick up a coffee cup if the cup and the character have been designed to do so, and if the right programs have been written to allow it. This can be painful, as we’ll see.

SET DESIGN

Set design frequently starts with pencil sketches or renderings using simple drawing software. Many designers carry a sketch book with them at all times in case inspiration strikes, or on the chance they may see something while out and about that they can use. Of course the specific process depends on the designer and the tools at their disposal. Set designers do sometimes use software, such as computer aided design (CAD) programs. Vectorworks is typical of these programs, which are typically built for architects

and engineers. When using such a program, objects are built from lines and then are collected into bigger collections, which are then coloured. The objects are not 3D in nature, but shapes can be extruded to create a 3D effect (Figure 10-3, left). Theatre designers find it awkward, but there's very little software specifically designed for the theatre.

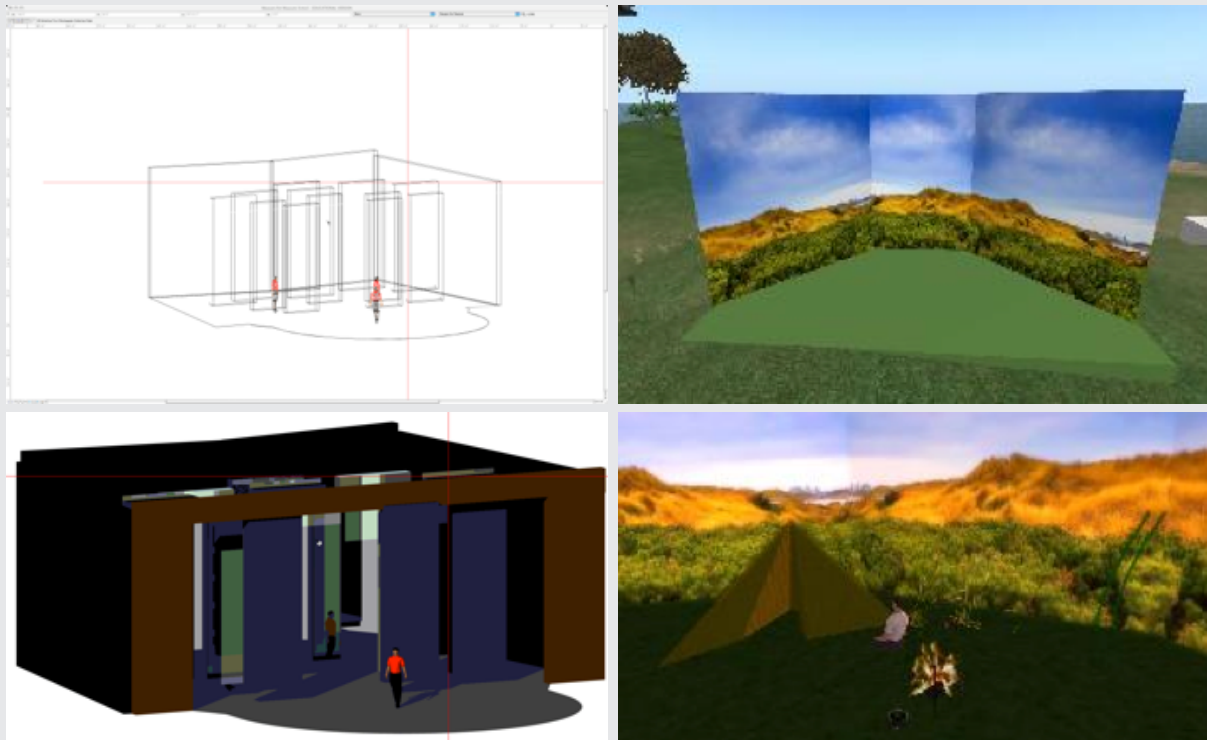


Figure 10-3: Set design, old and new. (Left) A professional designer uses Vectorworks to design a set of the play *Measure for Measure*. (Right) A set intended to represent a prairie scene was built in *Second Life* in an hour. This set can now be the venue for the actual production, whereas the other now needs to be built and painted.

On the other hand, in *Second Life* all constructed objects are built in three dimensions from the outset. In this solid modeling scheme the primitives are not polygons, but are cubes, spheres, and prisms. These 'prims' (primitive objects) can be hollowed out, twisted, sheared, cut, and combined in dozens of ways. The surface of the objects can be given textures (uploaded images). Objects can be coloured, made transparent and even glow.

Because these objects already reside within the virtual space they can be used immediately within that space. For virtual theatre, the value of this can't be over stated: the designed object is a working 3D model that can be viewed from any point in 3D space. In fact, the designer can walk through and around the set as it is being built. It can be used immediately for blocking and for rehearsals. A complete set can be constructed in a matter of hours, used, and rebuilt if necessary to accommodate perceived problems. In fact, sets can be modified during rehearsals as problems are discovered or to explore a new idea. They can be saved and restored in a moment. This gives directors unprecedented freedom to experiment.

Things that are essentially two dimensional have to be built from 3D prims. A simple wall is an example; it is constructed by starting with a cube (a prim) which is then shrunk in one dimension until it is flat and stretched in the other two dimensions until it is the correct size. It can be placed where needed and given

a texture, which is simply some image that was uploaded. The sets in Figure 10-2 and Figure 10-3 were created in that way. In the first case two images of derelict houses were created using drawing tools like Paint and Photoshop, uploaded, and mapped onto plane surfaces. In Figure 10-3 an image of a prairie scene was found on the Internet and applied to three surfaces arranged as a tri-fold backdrop.

There are problems with sets of this type, and the biggest one is the main problem with online spaces in general - resolution. Objects rendered in *Second Life* are typical of all other spaces, and are of relatively low graphical quality. The user/player must use their computer to render all objects in the field of view many times each second. Obviously, high resolution objects require more time to be rendered than lower resolution objects, so a compromise is made between graphical quality and animation quality (frame rate). This compromise is temporary; improvements in graphics software and in home computer hardware technology will gradually permit higher resolutions and frame rates until real-time photorealistic sets and objects are possible.

COSTUME

Costume design uses methods similar to those of set design, and in virtual space also yields a useable prototype. In *Second Life*, a sketch can be made on a simple 2D template using standard drawing tools. The templates are specific to the virtual world, and resemble a pattern such as is used in clothing construction in the real world. The base pattern is simply a diagram showing the boundaries of the clothing item. Figure 10-4 shows the pattern for a shirt, and has the front and back of the body and of a sleeve. The assumption here is that the two sleeves will be the same. The image is uploaded to the space, in this case *Second Life* again, and is used as clothing. The system is geared to understand images of this type, and they can be used as textures on a wearable item. Avatar shapes are defined by the system designers, and so mapping a shirt texture onto the torso of an avatar is a well defined operation. It is allowed to map trousers onto a torso, but the results are not attractive.



Figure 10-4: Building a costume using a graphical template as a texture.

Standard templates exist that can be mapped onto standard *Second Life* avatars. (Left) A basic template for a shirt. (Centre) An avatar wearing the template to show how it maps onto the body. (Right) A T-shirt design drawn on the template yields a fair approximation to a shirt when worn.

Using this scheme, it is possible to achieve very rapid trials and modifications of costumes. Iterations can be accomplished in moments that used to require days; the costume can be viewed in 3D on a moving body within a few minutes of conception. And the actor never complains about the fit. However, the clothing is really a rigid shell over the avatar shape, much the way a texture gives the appearance of a house when applied to a plane surface. Real clothing seems to have softness; it can move and flow as the actor moves, and the actor can adjust the clothing on stage as a part of the performance. What would *Streetcar Named Desire* be without the scene where Stanley rips at his shirt in anguish and cries upstairs to his wife: “Stella!”? The Greek gowns in *Antigone* need to flow freely about the legs. These things are hard to do with a graphical texture, so there is a second way to clothe an actor: essentially, with an object created in-world.

Building clothing this way involves using the same prims that were used for making sets. A skirt or kilt, for instance, would likely begin as a cone or cylinder. A texture is pasted onto the surface, and then the object properties are modified so that the object becomes soft and floppy. Finally this object is moved into position on the avatar’s body using the standard positioning operations and attached to the avatar. Now the object (kilt) moves with the avatar, but has the flowing characteristic that is associated with kilts.

Whichever method is used it should become familiar to a 21st century costume designer. It’s simply the final step that is missing, the production step which is also the most time consuming and expensive. Not only are the virtual designs useable on the virtual stage, but as many copies as needed can be created for no extra effort. An army can be costumed for the same effort as a single soldier.

PROPS AND OBJECT INTERACTIONS

Props in most systems are constructed in the same manner as are sets and costumes. They have an additional characteristic that sets don’t have, though: they can be manipulated. That is, a prop can be moved, reoriented, and reconfigured. A coffee cup, for example, is a prop; it can be moved from one place to another, usually by an actor, it can be rotated, and it can be broken (reconfigured). None of these things can happen ‘naturally’ in a virtual space because the normal laws of physics don’t apply by default. That’s been an advantage until this point in the discussion, but now becomes a burden.

Having an actor pick up a coffee cup and drink from it involves a gesture, but also an interaction with a prop. In virtual spaces this can be the most difficult aspect of a performance. There is no standard or natural interaction with the props. They seem to be physical objects, but are in fact merely constructs of polygons, shading, and lighting. In order to interact with them, the interaction must be located (from the script), carefully defined as a set of attachments to other objects and/or changes in position and orientation, and then implemented.

Implementing an interaction is done using software that has to be written specifically for each object and each interaction. The coordinates of the polygons for an object are specified relative to some origin within the object itself, either a vertex or the center of mass. Attaching an object to another simply means that the graphics system has to recalculate the vertex coordinates of the object’s polygons relative to some other origin, say the hand of the avatar. In confusing terminology, the program that codes for the interaction is usually called a *script*, and is written in a scripting language. In *Second Life* the *Linden Scripting Language (LSL)* allows a programmer to, for example, attach a cup to an avatar’s hand, execute a short animation for the drinking gesture, and then detach the cup and attach it to a table or saucer. Every interaction must be implemented at this level of detail, and no accidental interactions are even possible.

Object-object interactions are often more complex than this. Consider an actor who brushes against a desk, knocking a vase over, whereupon it falls to the floor and smashes. The first interaction is actor-desk. This is a collision, which may or may not be detected automatically by the virtual reality system. If not, then either a programmer must do it or, more likely, the collision will be made into a *cue*. Cues are events in a

performance that are handled by a technician. The most obvious example of a cue is a sound cue, where a particular piece of music or a sound effect is played at a precise moment in the performance. The sound of a car horn honking outside to attract one of the actors to the window is a sound effect cue. Music that leads to a scene change is another. It is someone's job (the sound technician) to play these sounds, which have been pre-recorded of course, at the moment they are needed. Lighting cues are also very common. Some are subtle, involving slow fades from one light to another as an actor moves across the stage, and others are rapid and precise, such as lightning or the switching on and off of set lighting. Again, the lighting cues are controlled by the script. At a particular point in the performance a word is spoken that the lighting technician uses to mark the moment when the lighting cue is to be invoked.

Object interactions are not generally thought of as cues in theatre. They are natural events that are described in the script as stage directions or imposed by the director. When a desk is bumped, a vase will fall over. When a ball is thrown, it will hit the ground and bounce. In virtual spaces this is not always the case. Video games often have complex physics simulations that can deal with this, but spaces like *World of Warcraft (WoW)* and *Second Life* do not. So it is that object interactions form a new type of cue. It will be someone's job, the interaction technician, to instigate pre-programmed object actions that would normally be a natural consequence of the way the real world works. The actor bumps against the desk, and the interaction technician makes the desk move just a little and causes the vase to fall over and off of the desk onto the floor - this could possibly be two cues, actually, one for the desk motion and one for the vase.

Having the vase smash is another cue. When it hits the ground the technician can run a script that causes the vase to break into pre-defined parts, and each part would follow its own path across the room. Perhaps some of the parts would break again, but these actions could be a part of the initial cue. The vase would break in exactly the same way every time, which is quite convenient but not very realistic. The cue would be programmed to work the same way each time, and it would make little sense to turn a play into an exercise in software engineering so development and debugging would be limited. Therefore, problems are possible. Not only is there a chance of a programming error showing up during a production, but inadvertently starting the cue could have odd effects. The vase could fall through the floor, for example, or could shatter a foot above it. These matters should show up during rehearsals, but might not always do so. What do we do?

Object interaction cues should have aborts. One abort should return the object to its initial location and form, one should place it in its final location, and others should be programmable events attached to the keyboard. In this way if something unexpected happens, it may be possible to recover, perhaps not perfectly, but with some possibility of continuing the performance. Ad-libs are much more difficult in virtual theatre at this time. If the vase breaks prematurely then other pre-programmed cues that involve the vase in the future might simply not function, and will be hard to recover from. The interactions and gestures will simply not be available. Returning objects to their previous state after a mishap means that things simply appear out of nowhere or fade into existence, and that's bad - but at least the rest of the play can carry on.

GESTURES

The term *gesture* is used to describe any physical action by an actor, whether involving motion or simply body pose. The gestures of an actor are typically very expressive and communicate a great deal of the message of a given scene. Actors also allow their performance to evolve, learning from one performance to the next what seems to work and what does not. As a result, while the feedback from the audience might not have much impact on the performance today, it may well modify the one given tomorrow.

The representation of the actor on the virtual stage is their avatar, and avatars are not as flexible as human bodies or as expressive. Like sets and props, avatars consist of polygons. The motion of an avatar is under human control; usually the arrow keys on the computer keyboard are used to control the basic direction of motion of an avatar. This allows for a basic walking motion, but not jumping or skipping, arm motions,

or hand gestures. This is a matter of resolution again, although a different type. It should be possible to implement motions of individual joints, and assign keys to joints so that avatars could be moved in complex ways by virtuoso performances on the keyboard, but no system does this yet. It would be puppetry essentially, the control of a human surrogate through strings or other remote manipulation.¹⁹ Puppetry is low resolution but is a known theatre form having well discussed limitations and implementations. Both *WoW* and *Second Life* provide a library of stock animations that can be invoked from the keyboard, but these provide a rather limited repertoire for performance purposes.

The more complex poses and motions can be accomplished by creating an animation and then using a set of animation cues. An animation in this context is “a set of instructions that cause an avatar to engage in a sequence of motions (Linden, 2007)”. Like sets, animations for specific gestures can be created in advance of the performance and then treated as cues. These gestures are painstakingly constructed from small individual motions of joints that are then combined and ‘played back’ in real time on demand. Animation tools that can do this abound, the most comprehensive being connected with expensive modeling tools like Maya and 3D Studio Max. One example of a free tool for this is *Avimator* (Webb, 2007).

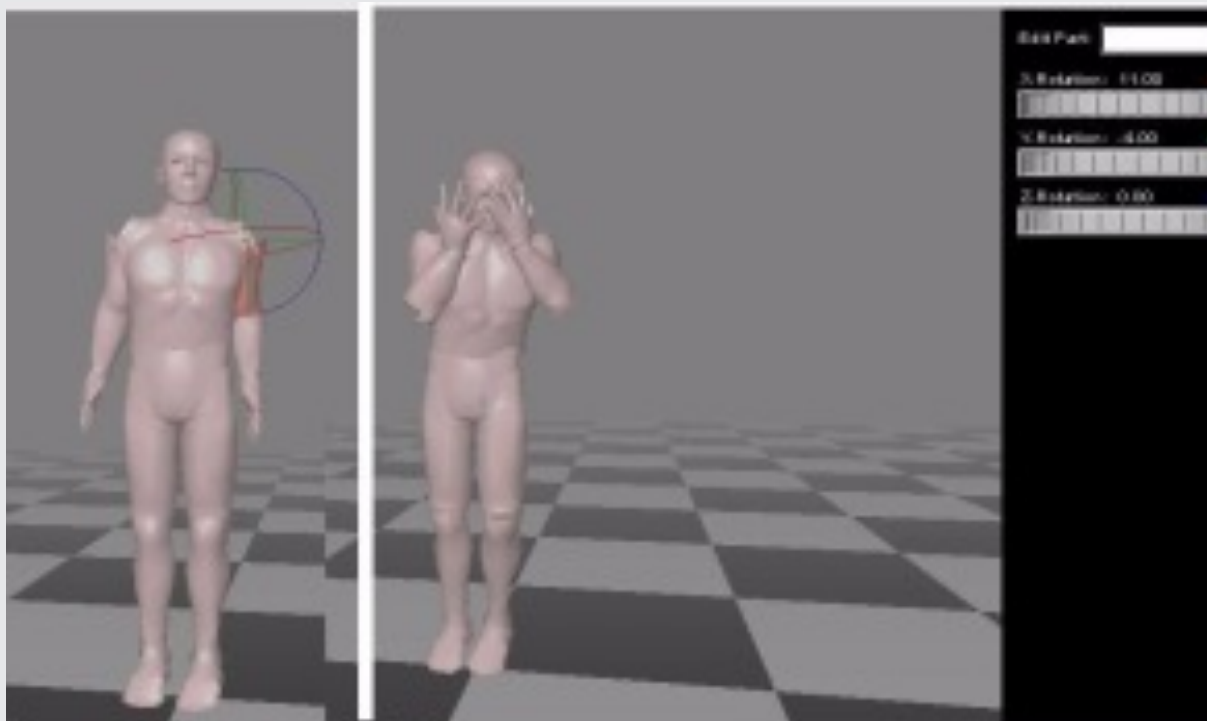


Figure 10-5: Using Avimator to create a gesture. Each joint in the body can be moved through its normal angular range in each dimension using the dials at the right. The bicep is selected in the left image, the final pose in the right.

¹⁹For more on puppeteering avatars, see Chapters 13, 14, and 16 of this book.

Avimator presented the designer with an image of a human figure. Any joint Individual (elbow, ankle, knee) can be selected with the mouse and moved through a range of angles. Building a gesture animation involves knowing the start and end positions, in terms of basic body pose, and moving each joint a small amount in each frame. The process can be quite precise, although it takes about an hour to complete 2-3 seconds of animation. The result is stored in one of 3-4 standard formats, (Weber, 2008) which can be uploaded to the virtual world server and stored for later use. The client software can play the animation on cue once it has been uploaded, and repeat it as often as needed.

My experience has shown that a page of script typically has three unrepeated gestures, and if they average 1.5 seconds each it amounts to between 1.5-2 hours of effort to animate. *Hamlet* has 153 pages more or less, which corresponds to 230-306 person-hours. This effort can easily exceed that for building and painting actual sets, and animating gestures may well be the greatest expense concerned with small virtual theatre productions. Consider that, at \$15/hr for animation, *Hamlet* could cost \$4590. Many theatre groups would either avoid the animations or avoid *Hamlet* at this rate, but there is another option.

Motion capture systems are getting less expensive each year, and are becoming a viable option to the use of prerendered animations. Such systems determine the pose of an actor's body in real time and transmit the coordinates of each joint to a computer, which in turn can send them to the virtual world server so that the avatar position can be updated correspondingly. This means that the avatar effectively moves in the same way as does the actor. Motion capture has been used extensively in animation and in video games, but is just getting to be of interest in virtual theatre because of the difficulty in having the avatar respond to live joint coordinates in stead of the downloaded pre-built animations. Custom programming is needed for this.

Motion capture systems tend to be of one of three types: mechanical, optical, or magnetic. Mechanical ones, such as the Gypsy series [Figure 10-6, left], are built into suits worn by the actors. Each joint has a sensor that records joint angles as voltages. Such systems tend to be very fast and quite accurate. They do impede the motion of the wearer, and some people find them uncomfortable. They do, however, allow multiple actors to interact on a stage. Optical systems [Figure 10-6, right] use cameras and computer vision software to track actors and recognize their pose. Actors usually wear coloured targets (markers) that the vision software can recognize more easily, and these do not impede the actors' motion. On the other hand, the markers have to be visible to be helpful, and so occlusions can be a problem. Many actors in the same space can confuse the software. Magnetic motion capture systems use sensors placed on the actors that measure a magnetic fields generated by a common source. The field has a distinct directional character, and the distance to each sensor can be measured quickly by using the strength of the field at each sensor. There are no occlusion problems, but magnetic field strength drops of very quickly with distance, and metal or electronics in the area of the stage can cause interference.

Unless gymnastics are required by the play, a mechanical system would likely be the best choice for live performance. The price of such a system is under \$8000 per actor and dropping, making this a viable option for many theatre companies. The animation needed for *Hamlet* alone costs a little over half the price of such a system, and it could be used to build the animations much more quickly than *Avimator*: the software that drives these suits can create BVH files too. This means that the suit would pay for itself quickly, and could be put to practical use immediately building animations.



Figure 10-6: (Left) Gypsy 6 motion capture suit uses mechanical and inertial devices to capture joint positions and send them by wireless link to a computer. (Right) Optical motion capture systems instrument the actor with small targets, often balls of a known colour

FACIAL EXPRESSIONS

Facial expressions are essential to a good performance on stage or film. Expressions convey emotion not conveyed by words, and in fact can speak the truth in situations where the words and expressions disagree. In person such a disagreement can convey sarcasm; if speaking to someone off stage, the expression can speak the truth while the words do not. Facial expressions in virtual spaces are treated as gestures, and have the same pitfalls, the most significant one being resolution again. It is difficult to be very emotive with a low-resolution graphical face. Or is it?

Based on Paul Eckman's work on the Facial Action Coding System (Eckman, 1978), Ken Perlin (1997) whipped up a Java applet that could express emotions on a simple two dimensional graphic face [Figure 10-7a]. It would appear that basic emotions at least can be conveyed using software, and so perhaps they could be made into gestural cues.

Facial animations are a subclass of gestures, and in *Second Life* and *WoW* at least they are treated in almost exactly the same way. There are a few facial animations that are provided by the system and that can be started using a keyboard or mouse sequence. Those are listed in the documentation with the other gestures, but have special names: 'expression_' followed by the name of the expression; 'expression anger' would do the obvious thing. These operate on a generic avatar face by manipulating the underlying structure, and therefore have much the same visual effect on all of the avatar faces. Unfortunately, these are the only facial animations supported, and no way is provided to create significantly new ones.



Figure 10-7: (a) Ken Perlin's 2D face can convey basic emotions in a low resolution image. Here we see the normal face and 'surprise'. (b) *Second Life* avatars showing a range of expressions that can be animated simply.

Subtle differences in expressions can be introduced by combining the built-in animations with alterations in the face structure. Details of avatar faces are specified by specifying a set of parameters on a scale of 0-100. Lip thickness, cheekbone height, chin width – these are three of dozens of facial parameters that the user has control over. Combining a stock emotion animation with modifications of facial structure parameters can give new expressions, or new emphasis to standard ones. For example, starting from a standard face, a slight smile can be created by narrowing the eyes and lips, adding lower lip, turning the corners of the mouth upwards a little, and making a slight underbite (See Figure 10-7b, 'Slight Smile'). It is also sometimes possible to combine two existing animations too, and while this creates unpredictable results, the animations can be observed, manually classified, and saved for future use.

Capturing the expression on a face using a camera is a variation on motion capture, and has been possible for some years. It is called *performance capture* by film director James Cameron (2009), indicating how important facial expression is for performance specifically and communication in general. It requires much higher resolution than does traditional motion capture, as facial expressions involve motions of a few millimeters, and so performance capture uses more expensive cameras and more sophisticated software. It is possible to find systems that use a webcam and 2D morphing methods for as little as a thousand dollars, and given the current state of the art in *WoW* and *Second Life* this would probably be fine. Motion picture quality systems cost ten times that, and do not operate in real time, working instead on recorded video. Still, technology moves quickly, and a real-time high resolution alternative that is affordable by smaller theatre companies should be available in a few years.

A novel idea provided by a *Second Life* object called *EmoHeart* (Neviarouskaya, 2010) is the use of text chat to control expression. No commands are used. Instead, the text is parsed for emotional content and appropriate animations are played automatically. Sadly there is no facility yet to listen to voice chat and accomplish the same task, which would be a more useful object (Vilhjalmsson, 2004). However, it would be possible to handle basic expression tasks by feeding parts of the script into this system. The archaic English of Shakespeare would likely cause some amusing errors.

LIP SYNC

An even more detailed form of gesture is that of lips while speaking. Humans expect to see a person's lips move in a way that is consistent with what they are saying. Each sound in a language (phoneme) has a mouth shape that goes with it, and we're very good at seeing that. Watching a movie where the sound is delayed can be very irritating for that reason. Virtual worlds often implement a form of lip sync when the avatars are speaking aloud. *There.com* had it before *Second Life*, but that world is now defunct. Both *WoW* and *Second Life* have a form of lip sync, but it is primitive. The mouth graphics are morphed in as a function of sound energy and duration, but do not reflect the actual speech phonemes being spoken.

In a normal theatre this is not an issue. The actors move their lips while speaking, and the audience generally ignores this – unless it is wrong, of course. Animators do lip sync in various ways, none of which are really applicable to live virtual theatre because they don't operate in real time. What exists is functional (just), and the cost for making it correct is huge, while the penalty for not doing so is relatively small. Progress in this area will depend on other applications, generally for the film industry.

SOUND

Although virtual online performances have been possible for over a decade, it has only been possible to hear the lines spoken for less than half that time²⁰. Online spaces now use Voice Over Internet Protocol (VOIP) to allow voice chat between participants. This means that actors can speak their lines as they normally would and expect that the audience can hear. The actors can hear the audience too, if the audience permits it. Before the advent of VOIP the plays used text chat to send the script to the audience as it was needed. This is not unlike the way silent films presented lines as written text in what would now be thought of as text boxes between action sequences. The text would appear in a small region of the screen dedicated to text chat messages, and would be typed in real time as the play progressed. Sometimes ambient sound is played during the event, as that can be downloaded in advance. Figure 10-8 shows how this would appear to the audience. Note the text chat area in the lower left part of the screen; note also that the character on the stage is holding its arms out as if it were typing on a keyboard. This is typical behaviour in *Second Life* when entering a text message, and means that the actor controlling the character on the stage is also entering the text.

Since its advent in 2006, voice chat in virtual spaces has become very simple to use. Usually one mouse click can start the voice channel, assuming that a microphone is already plugged in to the computer. From then on any sound that the mic can pick up will be sent to the whole world. Actors should each have their own microphone so that the sound technician can mix the levels properly, but they can sit at a table behind their mic or walk about with a wireless mic if that makes them feel more comfortable. Many actors prefer the latter because they are accustomed to moving about a stage while delivering lines; this would be the way sound would have to be done if the actors use live motion capture. Using table mics brings back the days of radio theatre, where the actors could not be seen, and spoke into microphones in a studio, live to broadcast. Many actors in those days stood before the mic, as that body posture gave them better voice control. Whichever scheme is used, the signals from multiple microphones are gathered at a mixer where levels are adjusted before sending the signal to the computer.

²⁰OZ Virtual” and “OnLive!” used a kind of voice chat as early as 1997, but its nature and utility is difficult to determine.



Figure 10-8: A screen capture from a silent performance that used text chat.

Modern USB mixers can also be connected to the *line input* of the sound card or even to the *microphone input*. One problem with using the line input is that various computers deal with it in different ways, and its use can result in frustration in trying to get it to work properly. The computer connected to the mixer is running a client for the virtual space and sends the audio to the theatre. One disadvantage of this scheme is that the sound will appear to originate at one spot in the theatre, no matter who is speaking, and that's true of any system that uses one mixer as opposed to individual computers with individual microphones for actors. Voice chat has a delay due to latency of between 1/2 to 1 second; that is, when you speak into the microphone it can take 1 second to hear the sound in the virtual space. It is therefore not possible to allow the actors to hear their voices as transmitted back from the virtual stage. The delay confounds the delivery of their lines. They hear what they just said a second ago, and they cannot focus on what's to be said next. Thus, a direct connection of the actor's headphones to *Second Life* through a single computer is not the best choice.

Figure 10-9 shows one audio configuration that is known to be effective in productions in *Second Life*. Each actor has their own microphone which are all connected to a common audio mixer. Each also has their own computer for moving their avatar, and this is connected to the server but with all sound turned off. The mixer feeds sound into a computer that is connected to *Second Life*, and that in turn sends the sound down one voice chat channel to the stage. A second computer, with a second avatar within the same virtual space, is used to hear what is happening on the stage. What this avatar 'hears' is fed into the second computer, and the sound technician controls the sound as heard in the performance space, after the delay. The short delay is acceptable for setting levels, and gives better volume levels than other options. The actors either monitor sound simply by hearing the other actors in the room, or listen through a headphone connected to the studio mixer.

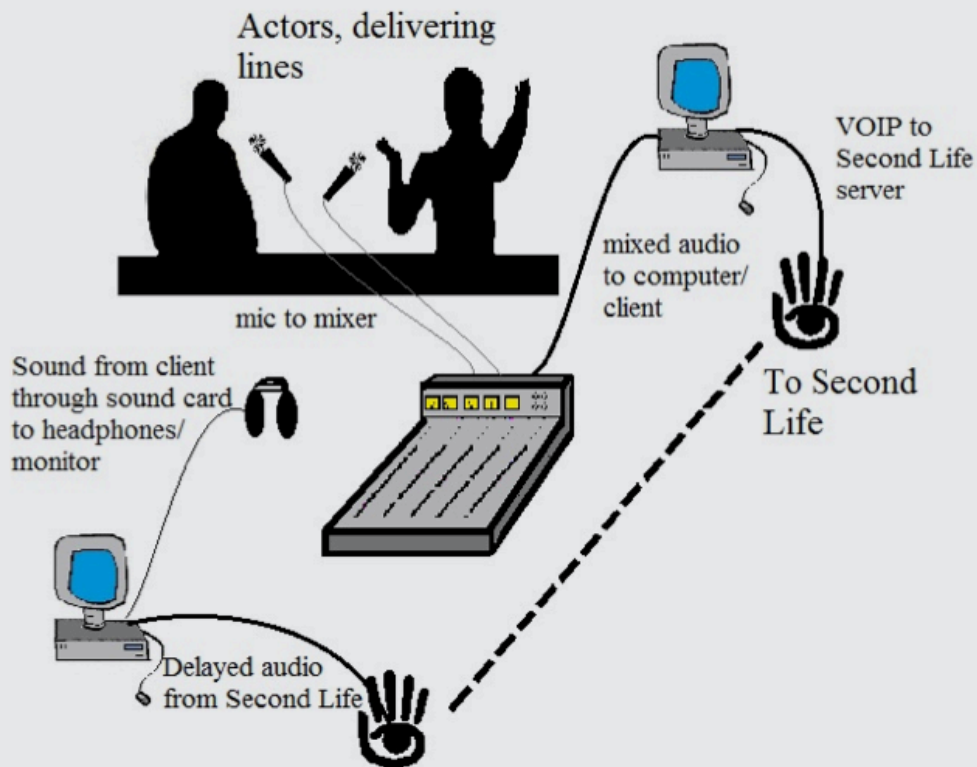


Figure 10-9: A functional audio scheme that has been used successfully in virtual theatre productions. The sound technician hears the sound after the delay introduced by the Internet connection; the actors hear the sound directly from the studio.

Sound effects are stored on the sound technician's computer and are played when needed by a standard media player program, which means that the audio will be sent directly to the virtual world server. They could also be played back on a more traditional MP3 or disk player and run through the mixer in the same way as are the microphone inputs.

There is a second way to bring audio into a performance. Streaming audio works with some spaces, and uses inexpensive free software to send a continuous audio signal to the space and hence to the user's computers. *Shoutcast* is a well known vendor of these services, and can easily feed audio into *Second Life* using the well-known sound player *WinAmp* as an audio source. One problem with streaming audio is that the latency is excessive, at least from a typical PC installation. Once the stream has started it feeds continuously, but there is a delay of up to 12 seconds from the start which makes synchronization of actor speech with avatars impossible. This discussion is predicated on the idea that the actors are present in the same room during the performance. This is the typical setup for a performance in a theatre, but is not really necessary for a virtual performance. In a *distributed* virtual performance the actors are in different places, and meet only in the virtual world where the stage lives. Each actor can see the stage and hear from the perspective of their avatar. If the latency is relatively low this works fine, especially for scripts that do not involve a rapid give and take. If actors are expected to interrupt each other, timing is difficult to get right in this case. Telephone or other dedicated audio links can be used to connect the actors directly; they can use the fast audio links, while the audience and sound technician use the voice chat. Distributed performances work best if each actor has their own sound technician who sets their levels and handles the sound cues for that actor's character. This minimizes the delay between an action and its corresponding sound. The director and stage manager must assign the sound cues to locations during rehearsals, and some experimenting may be needed to find the best arrangement.

2. MOUNTING A VIRTUAL PERFORMANCE

The steps in mounting a performance in virtual space generally are much the same as those for a performance in a real space. The key differences are in the details. The specific virtual space selected for the show may help decide the methods by which the show will be put on – distributed or in a common space, the layout of the audio system, and so on. The director, production manager, and designers are engaged, actors are auditioned and selected. There will be meetings at which the visual and auditory tone for the play will be decided.

SCRIPT ANALYSIS

A production begins with a script. The production manager must normally scan each word in a script to identify all *assets* and *cues*. Assets include scenery and set items, props, and costumes. Designers work with the assets to create a consistent visual presentation. The list of required assets is given to the designers, who will then construct samples to show the director. In a traditional performance the designers will show the director drawings, but in virtual spaces the actual items can be created, at least in rough form, and the director can see them in three dimensions in the context in which they will appear. Changes can be suggested and in many cases be implemented immediately for approval.

Normal theatre cues involve lighting and sound. The cues are indicators to the technical staff that an action is to take place. For example, when the script indicates that an actor switches a light on, this becomes a cue to the lighting technician to turn a new light on. Actors almost never really control anything on the stage; everything is done by technical people behind the scenes. Cues in traditional theatre tend to concern either lights or sounds, but can include mechanical or physical cues (E.g. things falling or being set on fire). Cues in virtual theatre are, as we have seen, more complex and numerous. The script is instrumented by the production manager with the cues that are found, as directed by the nature of the show and the plan of the director. Lighting cues and sound cues are written into a script as annotations, and copies of these scripts will be used by the technical staff during the performance. A sound cue is normally connected with a particular word being spoken, for example, and this is indicated in writing so that errors are minimized and the performance is the same each night.

In a virtual performance the gesture cues and the interaction cues have to be found and noted as well. There are now nearly double the types of cues as there were, and this probably translates into triple the number of individual cues in the script. A special cue sheet can be used to keep track of these, and the cue sheet can be an interactive document running on the computers used by the technicians. Table 10-1 shows a page from a typical cue sheet, in this instance showing a part of the grave digger scene from *Hamlet*. The text of the script is seen on the left side of the page; on the right are columns for cues, indicated by numbers. The number can be searched in a cue list for each cue type where a detailed description of that cue can be found. In the script segment, gesture cue 39 is a digging action of the grave digger. This cue is re-used a few times, since he digs a fair bit during the scene. Details of the cue – length, file name, creator, format, and so on are listed in the gesture cue list. There's a list for each cue type.

The precise timing of the cues is indicated by coloured words in the script text. Orange words indicate a gesture cue, for example, and when the orange word is spoken the gesture cue must be instigated. No more than one cue of any type is allowed per line of script. It is sometimes necessary to end a line of text prematurely to allow a cue to be added: the line "One that was a woman, sir, buy rest her soul she's dead" has a line break after "but" to allow for a second gesture cue.

Table 10-1: A section of the cue sheet for the grave digger scene in Hamlet. Each cue type is a different colour, and the coloured word in the script indicates when the cue is to be instigated. The sound cues could be blue, interaction cues green, lighting cues purple. In the software that displays the cues, the script and cues page down on demand, a line at a time or the page can be set to scroll at a particular speed. Software that allows the cues to be played when the text in the live cue sheet is clicked on, with a backup cue system quickly available, would be a valuable contribution to the technology. This would mean that all cues, lighting, sound, etc. could be run by a single technician instead of by four.

CHARACTER	TEXT	GESTURE	INTERACTION	SOUND	LIGHT	NOTES
Hamlet	Whose grave's this, sirrah?					
First Clown	Mine , sir.	41				41 Puts aside shovel
Hamlet	I think it be thine, indeed; for thou liest in't .. What man dost thou dig it for?	39		45		39 Digging 45 Digging sound
First Clown	For no man, sir.		52			52 Dirt flies from shovel
Hamlet	What woman, then?	39				
First Clown	For none, neither.					
Hamlet	Who is to be buried in't?					
First Clown	One that was a woman, sir; but, rest her soul, she's dead.	41 42				42 crosses self
Hamlet	How absolute the knave is! How long hast though been a grave-maker?	39				
First Clown	Since that day that young Hamlet was born; he that is mad, and sent into England.		52			
Hamlet	Ay, marry, why was he sent into England?					
First Clown	Why, because he was mad: he shall recover wits there; or, if he do not , it's no great matter there.	41				
Hamlet	Why?		53			53 grabs jar
First Clown	'Twill, a not be seen in him there; there the man are as mad as he.					

READ-THROUGH

After a significant portion of the pre-production are complete the actors will gather with the director and other creative staff and read through the script. Especially in virtual theatre with its accelerated production schedule the first read through may well be cold; that is, the actors will not have read it before. Two or three readings will be needed to identify parts of the script that have to be modified or bits of performance that need to be done in specific ways. A read-through is done in traditional theatre and film, and so is a familiar process to all of the participants.

REHEARSALS

The initial rehearsals will involve the actors reading the script through their microphones. At first no avatar motion is performed. This allows the actors to get a feel for the play and for the environment, and to think more about delivery and actions. Then actors and technicians work together on the gestures, possibly editing the animations for expressiveness and timing. Cue timing may need to be adjusted too, based on how the lines are delivered. A first draft of the set can be built for the first rehearsals and is adjusted according to the needs of the director and actors, sometimes on a daily basis.

One of the advantages of virtual theatre is the degree to which there is mutual feedback. Not only do the actors adjust to the sets, sounds, and stage directions, but the reverse is easy to accomplish; sets can be changed on very short notice to accommodate a new idea or to solve a problem. Recordings of the rehearsals are easy to make, and can be used for immediate feedback. Also, parts of the creative staff may be in distant locations, but they can still observe the rehearsals and make comments and changes.

At some point in the rehearsal sequence the costumes become available. These do not normally change anything, but the costumers can adjust them on the avatars for each actor when the actors are not present. One day the actor has no costume, the next day the avatar is simply wearing it. Costume changes usually involve one or two mouse clicks, and these must be set up by the costumers to make them fast and easy to accomplish.

Once all of the assets and cues are complete there are a few rehearsals called 'cue-to-cue'. Here, parts of the script that have no cues in them are simply skipped, and the lines are delivered from one cue to the next to allow the technical staff the opportunity to set sound levels, fix timings, and practice the cues sequences with the actual actors. Cue-to-cue rehearsals in virtual space may, in fact, be just like a dress rehearsal because there are many more cues than in a real-world performance. Still, it is essential to practice.

PERFORMANCE

Publicity for a virtual performance can be tricky. Online performances can be attended by nearly anyone on Earth without significant expense or travel time, which would appear to be a huge advantage. However, there's a lot of chaff on the Internet and attracting the attention of a target audience is not easy. The virtual world will almost certainly offer some ability to publicize events. Most worlds have user groups within them that allow Email communication between members, so a group attached to the theatre can be informed about new plays. And, of course, the theatre group's web site can be used and previous attendees can be placed on a notification list. Time zones must be specified in announcements, because the distribution is world wide.

On the day of the performance someone should be observing the venue to answer questions of anyone who shows up. A performance on the Internet is by definition an international one, and people from anywhere can show up at anytime. Posters should be up and the curtain time should be clearly indicated.

The venue itself does not require seating, although it is traditional. Attendees can stand or hover in space. There should be a ‘front of house’ or lobby in which attendees can mingle and speak with each other. Of course the audience can probably fly in or teleport to the venue, so an official entry is not really needed, but a place to be social is still important. It’s also important to have a marshalling area for the actors and stage crew. It’s true that teleportation is a common feature of virtual spaces, but it is not instantaneous and a small area is needed near the stage to change costumes, get props, and so on. This backstage area needs to be accessible to the actors both from the stage and from outside where the audience cannot see.

The latency that will be observed during the performance is to some extent a function of the number of attendees. Each avatar in the theatre requires computer time to support and bandwidth to update. Areas within the virtual world are handled by specific servers, and it is important to have the theatre venue located in a place (I.E server) that can support large numbers of avatars simultaneously. The location where the actors perform in the real world needs uninterrupted high speed Internet access too, so that the audio, motions, and cues are not delayed excessively.

Many virtual spaces have a currency and admission can be charged, although most theatres now depend on donations. A donation box can be placed at strategic locations about the venue. Currency has a variable value in the real world, and can be bought with actual currency; sometimes it can be sold, too, and that means virtual money is worth real money. Otherwise it can only be used to purchase things within the virtual space. *Second Life* has a currency called a *Linden* (L\$), and its conversion rate varies around 231 L\$ per U.S. dollar. Admission fees can be used to offset the expenses of the production, which includes the fees for uploading animations and textures, the purchase of prop items and costumes, and fees to programmers within the virtual world who will write scripts on request.

3. OTHER VIRTUAL SPACES FOR THEATRE

Second Life is a common host for online performances because of its size and features. There are other spaces that can be used, and each has advantages and disadvantages. The space that is chosen for a particular show would depend on the resources available and the needs of the script.

THIRD ROCK GRID AND OPENSIMULATOR

(<http://3drockgrid.com> / <http://opensimulator.org>)

3RG, as it has come to be known, is really a variation on *Second Life* that is based on open source software. It offers the same facilities as *Second Life* and has the same advantages and disadvantages technically speaking. It is reached through an open source client named *Hippo*, or the *Hippo OpenSim Viewer* which has access not only to *3RG* but a host of other ‘grids’, as they are called, that have characteristics like those of *Second Life*.

The Open Simulator project has programmers and designers working on an open source server for virtual worlds. This means that any group could build their own virtual space with rather little effort. The software can be downloaded and installed on a computer that has access to a web server and a high speed Internet link, and then the group can have their own world with no limitations imposed by a vendor. It would cost thousands of dollars per year to have a private space in *Second Life* with sufficient resources to mount plays to significant audiences. However, given a computer it will cost almost nothing to build a custom virtual space for the same purpose using OpenSimulator. The Hippo viewer can connect to Open Simulator worlds by specifying their Internet location, and the viewer is free for download too.

These custom worlds are great for developing new features – motion capture for avatar control, for example. They are also ideal for rehearsals. *Second Life*, on the other hand, has hundreds of thousands of users, each one a potential audience member. It has clubs and theatres that already exist. For the moment at least, gaining an audience is much easier on an existing grid, and *Second Life* has the biggest.

ACTIVE WORLDS

(<http://www.activeworlds.com>)

Active Worlds is the oldest existing 3D virtual space, having its origins in about 1995. It is described as a *universe* consisting of multiple named *worlds* each having variable themes and capabilities. It has, of course, evolved from its original form into a system that looks a lot like *Second Life*. So, new objects can be constructed, textures can be used, sounds can be played (wav, midi, and mp3). Objects can be created outside of the space, though, and there is a file format with the suffix RWX that is simple text that defines in-world objects. This means that the file can be edited directly if desired or drawing tools can be used to create them. Gestures are pre-defined, but animations are permitted – simply not gesture animations. Animations in *Active Worlds* are more like GIF animation, consisting of images displayed in a particular sequence.

Voice chat has been a feature of some of the worlds since 2006, meaning that *Active Worlds* had it before *Second Life*. It's open source, too, and that means that adapting it for use with motion capture and other inputs should be possible. Theatre should be as straightforward in *Active Worlds* as in *Second Life*, but does not seem to be a significant feature yet.

WORLD OF WARCRAFT

(<http://us.battle.net>)

World of Warcraft (WoW) is an online role playing game that takes place in a fictional world called Azeroth. It is a fantasy world that includes magic and trolls and dragons, and involves its players in various quests, most of which have a significant collaborative component. It would seem to be a good place to put on a play, especially something from Shakespeare. It does have significant disadvantages as a venue, though. It is a game and does not allow players to create structures or objects. Sets and props, in fact stagecraft generally, would seem to be impossible. One must make do with found spaces and pre-existing objects.

Little work has been published on the use of WoW for theatre, but Balzer (2009) describes an effort to mount a performance of a custom written play *An Ode to 10-mans* (Kozma, 2009). Among the good advice to be found in his thesis is that the scene designer should be replaced by a scenographer, whose job is to locate existing locales for the show. A costume coordinator is needed instead of a costume designer, because costume construction is very limited. Blacksmiths in the game can build only specific, generic armour for the avatars. Props too are limited to that that can be found in the space, weapons for the most part. It is no surprise that the play *An Ode to 10-mans* concerned a quest within the game – in this way the playwright made the best use of the materials that could actually be obtained for use within the play.

Balzer discusses the attitude of the denizens of WoW towards theatre as being quite variable, from more or less accepting to aggressively negative. Discussions with players has revealed that different servers actually have sets of players with differing attitudes, and that the server on which the play was to be performed would have to be selected with care.

OPEN WONDERLAND

(<http://openwonderland.org/>)

Open Wonderland is not so much a single world as a tool for building your own virtual worlds. These can be coded from scratch in Java, the underpinning of the system, or can be adapted from an existing example world, but in either case programming skill is essential. The fact that it is Java based means that it can be adapted to include almost any feature desired, but not in a point-and-click fashion. The world created is reachable on the web, just like the others, and so could be used to build a theatre space.

Wonderland was released in May of 2010 by Sun Labs, and so consists of quite recent technology. The original purpose was to construct business simulations and meeting spaces of a high quality, but obviously such spaces have many of the characteristics one would want of a theatre. The new owners of Sun decided recently to stop development of this project, and so it is now in the public domain as a community-supported open-source project.

4. THE POTENTIAL OF VIRTUAL THEATRE

In 2008, while live artistic performances showed an overall decline in attendance, 47 million Americans chose to view theatrical, musical, or dance performances at least weekly using the Internet (Trescott, 2009). North Americans are using New Media and the Web and the more and more, and are attending live performances less and less. The only categories where attendance has increased since 2002 are for musical plays and some forms of dance. In all other instances - plays, classical and jazz concerts, ballet, and opera - attendance has declined, in some cases by 12% (NEA, 2009). There is a need to look technically and creatively at ways to offer new online forms for the next decade.

Some critics do not believe that this is actually a form of theatre or that it is, indeed, live. It is certainly live – the delay from actor to audience is no more than is typical between the stage and the far seats of a large concert hall. One could demand physical proximity from live theatre, but that is relative. And virtual theatre is not a recorded performance. Gesture may be animated, but all cues have a virtual nature about them. Sound cues are certainly not live by any real definition, even in traditional theatre. Good arguments can certainly be made that virtual online theatre is live theatre in a very real sense.

In the 1970's, Brazilian theatre practitioner Augusto Boal began his work on what he termed invisible theatre. In his work, *Theatre of the Oppressed*, he describes invisible theatre as such:

[Invisible theatre] consists of the presentation of a scene in an environment other than a theater, before people who are not spectators. The place can be a restaurant, a sidewalk, a market, a train, a line of people, etc. The people who witness the scene are those who are there by chance. During the spectacle, these people must not have the slightest idea that it is a spectacle, for this would make them *spectators* (Boal, 1979).

Boal goes on to describe a scene in an actual restaurant, with actors at some tables and unknowing audience members at others, going about their day. The actors begin to complain about the prices of the food, and try to negotiate. One refuses to pay; another offers to work for food. While this was scripted for the actors, nobody else understood that it was a play, and the chance of a confrontation that could turn to violence or panic is obviously possible in these productions. Imagine a situation where an actor simulated some form of attack on another actor. A Good Samaritan could easily become involved, or the police could misunderstand, resulting in injuries or a court appearance. In *Second Life* these performances, while still a surprise to the audience, have no such consequences in the real world, and so no injury or damage can occur.

Consider the thoughts and actions of an audience member who, after being invited to a play, shows up to find no stage, no 'actors', and only a bunch of people milling about. Someone speaks; another complains – both actors, in fact. Audience reactions are merged into the script, and the show becomes part scripted part improvisation. By the time most of the audience realizes what is happening, many of them have already been a part of the show!

The University of Calgary Drama department has experience drawn from four plays produced in *Second Life* (Parker, 2010; Parker, 2011) The plays selected took advantage of the special nature of virtual spaces. The first, *Guppies* (Martini, 2001), is an exchange between fish. The set was a large aquarium within which our fish characters swam about. The guppy characters could have simply been designed as fish, because convincing costumes were possible, but it was ultimately decided to have the actors appear to be riding the guppies. This sort of design choice is normally not possible. Other productions involved a genie, living and speaking chocolate treats, and a giant talking black widow spider. Such things are difficult to present effectively on a live stage, but present merely minor problems as virtual reality. Other companies have performed various works of Shakespeare, *Oedipus Rex*, *Moulin Rouge*, and a host of other shows. Each has a set of challenges that have much to teach all directors about this new medium. It is unfortunate that the process was not documented in these cases. In fact, nobody really knows what the first performance in *Second Life* was.

Virtual theatre is still in an initial state, and will profit from higher resolution graphics, better audio, and most of all by better avatar control systems. Work is ongoing in many places on the use of motion capture so that actor movements can be translated into avatar motions in real time. Facial animation is also a crucial technology for future success - the facial animation for the film *Avatar* by Image Metrics (Waxman, 2006) is an example of what can be done, and of how quickly such technologies can change. The existing system for facial animation works well only for recorded video, and is not useable in real time as would be needed for theatre work. This will change. Most of all, what is needed are playwrights who are willing to create for this medium, and directors who want to put on these plays and develop new methods for virtual spaces.

These techniques have the potential to change the way live in-person theatre is conducted. Virtual theatre is also a new extension of an old art form, and is an attractive entertainment form with virtual spaces, which is a place many people come both to play and to work. It shows the potential to widen the audience of live theatre, to interest a younger audience, and to simplify the act of attending theatre. The process of mounting the performance can be more democratic than traditional theatre, allowing feedback from many sources, and encouraging a rapid prototyping system of development. It is not a replacement for real-world theatre in any sense, just as motion pictures were not a replacement for live theatre. It is a new medium having roots in the old ones, and offering new advantages and audiences.

5. BIBLIOGRAPHY

- Artaud, A. (1958). *The Theatre and its Double*, (Caroline Richards, Trans.). New York: Grove Press. (Original work published in 1938)
- Balzer, T. (2009). Theatre, */quit: The Potential of Community in World of Warcraft*. Unpublished master's thesis, College of Fine Arts, Ohio University, Athens Ohio, USA.
<http://www.timbalzer.com/storage/wow/Balzer%20Tim%20-%20Theatre%20Slashquit.pdf>
- Boal, A. (1993). *Theatre of the Oppressed*. (Charles A. & Maria-Odilia Leal McBride, Trans.). London: Pluto Press (Original work published in 1979)
- Cameron, J. (2009). *Avatar: Motion Capture Mirrors Emotions* [video], Interview with Discovery Channel, Dec 24, 2009. Retrieved April 29, 2011 from <http://www.youtube.com/watch?v=1wK1Ixr-UmM>
- Ekman, P. and Friesen, W. (1978). *Facial Action Coding System: A Technique for the Measurement of Facial Movement*. Palo Alto: Consulting Psychologists Press.

- Geigel, J. and Schweppe, M. (2004, October 10-16, 2004). *Theatrical storytelling in a virtual space*, Paper presented at the 1st ACM workshop on Story Representation, Mechanism and Context, New York, NY.
- Kozma, R. (2009). *An Ode to 10-mans*, Retrieved April 10, 2011 from <http://www.timbalzer.com/storage/wow/An%20Ode%20to%2010-mans.pdf>
- Linden Research, (2007). *Animation*, Retrieved March 2011 from <http://wiki.secondlife.com/wiki/Animation>
- Martini, C. (2001). *Guppies*. Festival of Ten, SUNY Brockport.
- National Endowment for the Arts. (2009). *2008 Survey of Public Participation in the Arts*, Retrieved March 21, 2011 from <http://www.nea.gov/research/2008-SPPA.pdf>
- Neviarouskaya, A., Prendinger, H., and Ishizuka, M. (2010). *EmoHeart: Conveying Emotions in Second Life Based on Affect Sensing from Text*, in *Advances in Human-Computer Interaction*, Volume 2010, January.
- Newman, P. (2010). *Remotely Operated Vehicles Involved in the Deepwater Horizon Response*, Douglas-Westwood, Canterbury, UK. Retrieved March 4, 2011 from <http://www.dw-1.com/files/files/583-Deepwater%20Horizon%20ROV%20%20AUV%20PN.pdf>
- Parker, J.R. and Martini, C. (2010). *Producing Live Theatre in Virtual Spaces: Real Performances/Alternative Landscape*, Arts in Society Conference, Sydney Australia, 22-25 July.
- Parker, J.R. and Martini, C. (2011) *Puppetry of the Pixel: Producing Live Theatre in Virtual Spaces*, CCNC 2011: 3rd IEEE International Workshop on Digital Entertainment, Networked Virtual Environments, and Creative Technology, Las Vegas, NV. Jan 9-11, 2011.
- Perlin, K. (1997 August 3-9, 1997). *Layered Compositing of Facial Expression*, Presented at ACM SIGGRAPH 97 (Technical Sketch), Los Angeles, CA.
- Trescott, J. (2009). *Web audience grows for arts groups*, Washington Post, December 10, 2009.
- Waxman, S. (2006). *Cyberface: New Technology That Captures the Soul*, New York Times, October 15, 2006.
- Vilhjalmsson, H., *Animating Conversation in Online Games*, in M. Rauterberg (ed.), *Entertainment Computing ICEC 2004*, Lecture Notes in Computer Science 3166, pp. 139-150, Springer, 2004.
- Webb, R. (2007). *A Beginners Guide to Avimator*, Retrieved Feb 26, 2011 from http://www.avimator.com/downloads/Avimator_Beginners_Guide.pdf
- Weber, A., Rufer-Bach, K. and Platel, R. (2008). *Creating Your World - The Official Guide to Advanced Content Creation for Second Life*, New York: Wiley Publishing.

11

ANIMATION PRINCIPLES AND LABAN MOVEMENT ANALYSIS: MOVEMENT FRAMEWORKS FOR CREATING EMPATHIC CHARACTER PERFORMANCES

By Leslie Bishko

Editor's Note: This chapter is the second part of a larger piece on Empathy and Laban Movement Analysis for animated characters. We have broken it into three smaller chapters throughout the book, but it may be read as part of a single work. The other two parts can be found in Chapters 5 and 13. These chapters are supplemented by series of video figures, which can be found online at <http://www.etc.cmu.edu/etcpress/NVCVideos/>.

1. THE PRINCIPLES OF ANIMATION

The evolution of animated movement at the Disney studio during the 1930s is pivotal to the formalization of believable and authentic movement parameters. During this era, a core team of animators began to experiment with animated movement. As reported by Frank Thomas and Ollie Johnston in *The Illusion of Life: Disney Animation* (1981), Walt Disney pushed the animators to develop their skills and create a more physically believable animated world. Gradually, a terminology, or language of animated movement evolved, which became known as the Principles of Animation (Johnston & Thomas, 1981). As these precepts are widely known and can be referenced in *The Illusion of Life: Disney Animation*, I will briefly paraphrase them here and apply them in context throughout this chapter:

1. Squash and Stretch refers to the elastic qualities of objects or characters, and how force and gravity can affect their shape. For example, a bouncing ball flattens on impact with the ground, and elongates as it moves through the air. Squash and Stretch is frequently associated with vertical motion and the force of gravity acting on mass. It creates the illusion of weight.
2. Anticipation describes the way a character prepares for an action. When hitting a baseball, the batter will pull the bat back and upwards before swinging to strike the ball.
3. Staging is the visual design of a pose, such that the action of the pose is clearly readable in silhouette. For example, hands are often posed away from the torso.
4. Straight Ahead Action and Pose to Pose are two different animation workflow methods that are used to create loose or controlled results, respectively. Straight Ahead is drawing each movement one frame at a time in chronological sequence. Pose to Pose is creating the start and end of an action, then in-betweening the motion.

5. Follow Through and Overlapping Action are two separate concepts that often appear together in movement. Follow Through motion has a distinctive, wave-like progression, initiating at one end and flowing to the other, such as a waving flag, or the lagging motion of a feather attached to a person's hat. Overlapping action refers to staggering the starting/stopping moments of complex actions, and is often related to cause and effect.
6. Slow In and Slow Out refers to acceleration and deceleration. Particularly for living (i.e. non-robotic) characters, motion always accelerates and decelerates.
7. Arcs describe the curved motion pathways of most organic action, particularly character animation. A golf swing moves along an arc.
8. Secondary Action occurs as the result of primary action. It applies to hair and clothing, or body parts such as the head or hands.
9. Timing is a general term referring to the speed, acceleration, deceleration, and duration of an action. The meaning of an action can vary at different speeds.
10. Exaggeration is when a much broader range of motion is used than would normally occur. For example: a character's jaw drops to the floor in surprise.²¹

The development of these Animation Principles marked a significant evolutionary step in character animation. The principles cultivated a richly detailed, full animation style that promoted the physical properties of objects and characters in motion as the basis for believability. Their impact shifted engagement with the medium of animation away from the novelty of the animated illusion. Audiences could now perceive animated character performances as a seamless representation of believable beings. Through creating celebrity characters such as Mickey Mouse, who appeared in numerous cartoons, a persona evolved that supported the illusion of a “real” character that exists in the world.

While the Animation Principles are valuable historically and still relevant today, the language of animated movement has continued to evolve. Some are critical of the stylistic influence that the principles impose (Bishko, 2007). The following section introduces Laban Movement Analysis, a style-neutral framework of movement concepts.

2. LABAN MOVEMENT ANALYSIS

Rudolf Laban (1879-1958), a Hungarian dance artist and theorist, was at the source of expressionist dance in Europe between 1910 and the onset of World War II. Together with his students and collaborators, Laban was able to distill the ingredients that are part of all movement patterns, formulating a rich and robust movement language that has withstood the rigors of broad applicability. He intuitively understood aspects of the body/mind connection—how intentions affect actions, for instance—that have become contemporary topics among cognitive scientists, somatic practitioners, psychoanalysts, athletes, dancers, and actors. Laban planted the seeds for what has evolved into today's Laban Movement Analysis (LMA).

Laban is most widely known as the originator of Labanotation, a structured notation system for human movement. A lesser-known form of Labanotation, Motif Description, allows for a more thematic and interpretive approach. Motif is used in Laban Movement Analysis as an observation tool, in dance literacy for children, and as a creative tool for choreography. Notation provides a synthesized means of recording what the action is, and how it was performed. In cases where specific actions are analyzed, notation assists greatly in documenting observations. The graphical representation of movement is especially valuable towards observing patterns, constellations, and rhythms of phrasing in movement.

²¹Solid Drawing and Appeal are omitted from this list as they do not refer to movement.

Currently, LMA is a language that applies to all living beings, which, for our purposes, certainly includes animated characters in virtual worlds. LMA provides a conceptual framework through which we can observe, describe, and interpret the intentionality of movement. It possesses one key attribute that the Animation Principles lack: the link between how people move and what their movement communicates to others. It is crucial to understand that this “link” does not constitute a specific mapping of movement to meaning. Rather, it is a contextual framework that supports personal, subjective, and metaphoric inferences perceived by attuning empathically with movement. LMA factors in people’s individual movement habits, cultural traditions, and the immediate circumstances in which movement communication occurs. By always acknowledging the context of movement, LMA is effective for a broad range of applications.

Some examples of limited one-to-one mappings are:

- Smile = happy
- Frown = sad
- Index finger tapping temple, eyes looking upwards = thinking
- Arms folded = angry

Presented as a single image, any one of these examples can be interpreted as suggested. However, when considered as part of a movement phrase, in the context of an individual’s habitual movement patterns, arms folded could indicate things like feeling withdrawn, feeling cold, self-protection, indigestion, waiting patiently, etc. In each of each of these examples LMA has the capacity to separate the movement components from meaning. The separate components of movement are the building blocks of creating authentic expression for a character.



Figure 11-1: Contextual Variations

LMA works within the following fundamental beliefs about movement:

- Movement is an ongoing process of change.
- The change is patterned and orderly.
- Human movement is intentional.
- The basic elements of human motion may be articulated and studied.
- Movement must be approached at multiple levels if it is to be properly understood. (Moore, 1988).

Laban breaks down these multiple levels as three “mentalities” towards movement:

1. That of a mentality plunged into the intangible world of emotions and ideas.
2. That of the objective observer, from outside.
3. That of the person enjoying movement as bodily experience, and observing and explaining it from this angle.

Empathic experience relates to what Laban describes as a compression of these three mentalities: “a synthesized act of perception and function.” (Laban, 1966, p. 7) The process encompasses empathic, bodily *knowing*, and reflection on the knowledge; a registration of the experience.

The following summary of LMA is a detailed overview in the context of animated characters in virtual worlds. My goal is to make this framework, and its methodology of observation, description, and interpretation, accessible and applicable to the creative process of animation. I will also touch on how game design, interaction, and artificial intelligence all impact movement and affect what we experience empathically in virtual worlds.

PRINCIPLES OF LABAN MOVEMENT ANALYSIS

FUNCTION/EXPRESSION

Function and expression are among the core principles of the LMA system. As described earlier, both occur in a reciprocal relationship in the process of movement. Intent drives bodily actions and is expressed through them, while body mechanics enable the realization of intent. Intent organizes the neuromuscular patterning of a movement sequence. The mental preparation of any action always precedes the action, as if the mind is telling the body, “OK, now we’re going to do this.” Athletes take advantage of this interrelationship by practicing visualization techniques to boost their performance.

Whereas function and expression are intertwined in the real world, function always precedes expression for the animator. In creating animation, it takes refined observation skills and practice to animate believable body mechanics. Without the

HOW TO LEARN LABAN MOVEMENT ANALYSIS

The best way to learn LMA is by taking introductory workshops, university courses, or studying in one of several LMA certification programs. This chapter provides a written description of LMA, but text cannot deliver what you will learn from direct movement experience. LMA programs are offered in intensive year-long or modular formats and can be included as a major component of graduate-level studies. Certification programs offer a process-oriented education that allows time for depth, practice, and integration of the material. See <http://www.labanforanimators.com> for a list of courses. For additional resources, see the Appendix at the end of Chapter 5.

latter, the character's expressive movement behavior is unreadable and therefore, cannot create an emotional connection of empathy. However, this is not a suggestion to abandon acting in character animation until one has mastered biomechanics! The clearer an animator is about a character's intent, the more he or she can create functional movement that supports a character's expression. The intent is what we connect with empathically in movement.

INNER/OUTER

Inner/outer refers to how movement reveals the way we exist in relationship with the environment. Inner impulses are expressed outwardly in movement, while circumstances in the environment affect inner experience. When I am inwardly tired or stressed, I become uncoordinated. I start dropping things and even lose my balance. While trying to be in control, my movement reveals that I am not. When outer affects inner, we feel energetic and can have an expanded sense of personal space on a warm, sunny day. Or we can feel tense and agitated in response to noisy neighbors disrupting our quiet living space. Or we can become narrow and shrunken, gathering inwards in response to news of world events. We continuously fluctuate between inner and outer.

STABILE/MOBILE

Movement flows continuously between stability and mobility. Stability offers a resting place. It acts as preparation and initiation for mobility. Mobility arrives at, and concludes in, stability.

Animators are very familiar with the concept of walking as “controlled falling.” The video *Controlled Falling Project*, from acrobatic troupe This Side Up, is full of examples of dynamic transitions between stability and mobility. Most of the stunts involve balancing, moving into and out of moments of balance. [Video Figure 5]

EXERTION/RECUPERATION

Laban built much of his theory around rhythms of exertion and recuperation. He observed these rhythms as parts of movement phrases. We exert our energy to bring ourselves into action, followed by recuperative periods of rest. For example, I exert my energy to mobilize my weight as I stand up from sitting in a chair. An inhalation of breath supports the energizing of my muscles and my intent towards being vertical. Once standing, I pause with a recuperative exhalation. Moments of exertion and recuperation are clearly observable as parts of a larger movement phrase.

My 5-year-old daughter exhibits clear, natural rhythms of exertion and recuperation through her active play. Her preference is for short bursts of exertion followed by longer periods of recuperation. She will initiate chasing games and rough play that engage gross motor patterning. When she has had enough, she will sit down and take up activities that require visual focus, such as playing with a small toy in her hands, or reading a book. This has been her pattern since infancy, and is part of her “movement signature.”

Patterns of exertion and recuperation, portrayed clearly as phrases of movement, help to create credible character movement. This, in turn, makes the character seem more authentic, and impacts our empathic engagement.

THE CATEGORIES OF MOVEMENT

Five categories of movement delineate the full spectrum of LMA's movement parameters: Phrasing, Body, Effort, Shape, and Space.

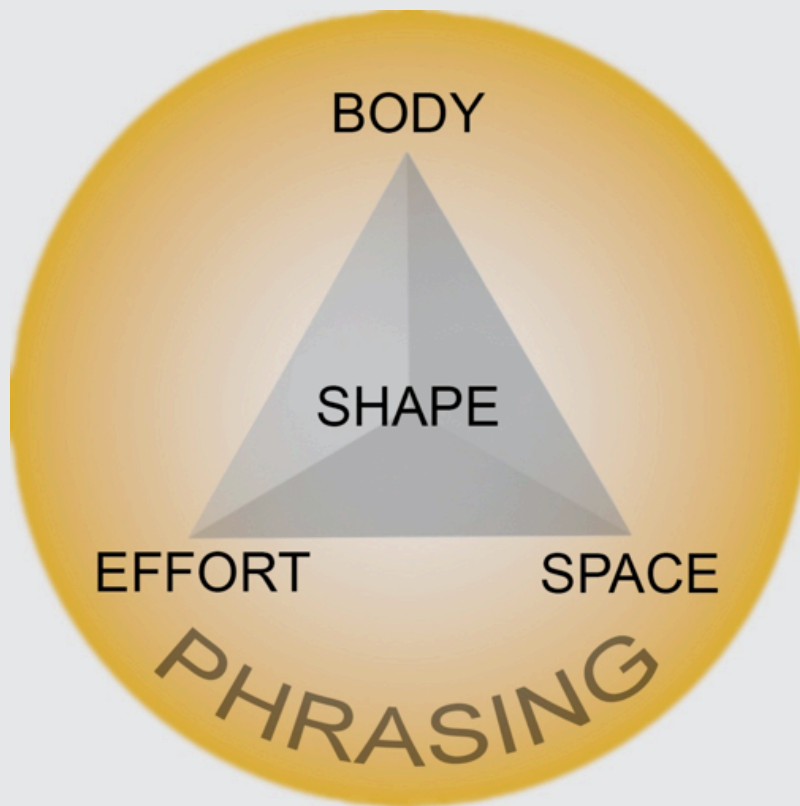


Figure 11-2: Diagram of LMA Categories

PHRASING

Phrasing describes how we sequence, layer and combine the components of movement over time. A movement phrase is like a verbal sentence, or a phrase of music, which represents a complete idea or theme. Just as language communication is organized in sentences, movement communication is organized in movement phrases. Our uniqueness is expressed through our movement phrases: they reveal individualized rhythmic patterns and preferences for how we combine the elements of movement.

Phrasing rhythms arise from breath rhythms. How we modulate our breath within a movement phrase, over the course of an action, is deeply connected with intention. Breath is the most fundamental movement pattern through which we empathically attune with intent.

How one initiates a phrase of movement organizes intent and patterns the neuromuscular coordination of the action. Intention prepares us to move and patterns the movement phrase that follows. (Hackney, 1998).

There are five stages in each movement phrase:

1. Preparation (intent)
2. Initiation (anticipation)
3. Exertion/Main Action
4. Follow Through/Recuperation
5. Transition

Jumping is an action that gives us the chance to observe phrasing very clearly:



Figure 11-3: Diagram of LMA Categories

As discussed above, *preparation* is the mental image that initiates movement. In animation, this may be visible in facial expressions, or subtle body actions such as a slight shift of weight, rising up on the toes, inhaling or exhaling, etc. In Animation Principle terminology, it is the anticipation of the anticipation.

The *initiation* stage is the beginning of the broad motor action that is to follow. Typically, initiation includes a weight shift, specifically where locomotion is involved. Anticipation is often a movement in the opposite direction of the main action; to jump forwards and up, the movement is initiated with a backwards-and-down weight shift in the pelvis.

Jumping is a good example of the *exertion* part of a phrase because our intent is to become airborne, which involves notable exertion. As the legs push powerfully off the ground, the body unfolds towards the spatial goal of forwards and up. This is analogous to the stretch phase of a bouncing ball. The exertion stage of a jump flows into the main action, including the entire arc of the jump.

In the context of phrasing, *follow through* refers not to follow through motion specifically, but follow through of the action. In jumping, the follow through of the phrase comes after the feet have touched the ground: the body continues to arrive afterwards. In this case, the follow through of the phrase does have follow through motion—particularly in the spine, which can flow in a whip-like action, following the abrupt change of direction in the pelvis. This action *recuperates* from the main action. It is a “letting go” of the main action.

In animation terms, the *transition* of a phrase is sometimes referred to as a *settle*, where motion comes to a standstill, or hold. This is a moment where one action has ended and the next is about to begin. The character may have a response to the previous action: “Wow, I really did jump across that creek!” Also, intent towards the next action begins to form. During this pause in the action, we don’t want the character to be completely immobile. Transitions may include overlapping action, such that the character never really comes to a full pause, but a pause or new thought in the intent of the character can be perceived. Our bodies are never completely unmoving; we breathe and move through very subtle weight shifts even when trying to stand still. Animators create moving holds to keep the character alive onscreen. Particularly in 3D computer animation, frozen characters will look either dead, or like robots!

Transitions are among the more challenging moments for animators to create, as they can be subtle. If the character’s intent during the transition is unclear, or has not been attended to, transitions between movement phrases will look awkward onscreen. Transitions may not seem as important as the main action, but if a transition doesn’t read clearly, the believability of the entire movement sequence will suffer. Animator Mike Jungbluth describes the importance of meaningful transitions for interactive games:

Transitions are also when a character is anticipating what they are about to do next, and anticipation is something game design thrives on even if as animators we often have to sacrifice an action’s anticipation in an effort to speed things up. But transitions are pure anticipation of what is about to happen next. (Jungbluth, *Growing Game Animation - Transitions & Player Input*, 2011)

Jungbluth’s article discusses the challenge of creating complete movement phrases in real-time interactive character animation. Preparation, initiation, and transitions are truncated to provide instantaneous character response to the joystick. When movement phrases are incomplete, the essential element of breath, which drives the larger rhythm of phrasing, is missing. Functionality of movement is also compromised. With Ubisoft’s production on *Assassin’s Creed III*, they found a partial solution to the problem: that triggering head movement in response to controller input gives the impression of character intent.²²

Phrasing as narrative

Movement phrases provide a means by which we can perceive the character’s motivation and intent. They are the building blocks of narrative, how movement tells a story. Here is how jumping across stones in a creek could be described as a movement story:

Preparation: Three stones are visible, each one lying progressively farther from the next. The character begins by standing neutrally, and then crouches slightly while visually assessing the distances of the stones and deciding whether crossing the stones is possible.

A second preparation: A decision is made to go for it as the character returns to upright posture.

Initiation: Crouching lower, one leg in front of the other, the character shifts his weight towards the back leg.

Exertion/Main Action: The character pushes off the back leg, jumps to the first stone, landing on one foot.

Follow through/recuperation: Catching his balance, the character places his other foot on the stone in a narrow, upright stance (the stone is small!).

²²See Sidebar: Procedural Pose Modification and Empathy at Ubisoft in Chapter 13

Transition: “I made it! OK, on to the next.” The character adjusts some foot positions.

Preparation: Noting the distance to next stone, the character preplans how to shift his weight to his back leg for the next jump, with so little space for feet, and so on.

From here, a new phrase begins for each jump. The animator can also string together the entire series of jumps as small sub-phrases within the larger phrase: Get across the creek.

Rhythm and uniqueness

People express themselves rhythmically through their movement phrases. Individual phrasing rhythms form uniqueness and personal style. Watching cooking TV shows offer a good opportunity to observe individual rhythms. As they prepare food, the hosts demonstrate with their hands, and typically imbue what could be a set of dry instructions with lots of charm and passion for food. You will see medium-close-up shots and plenty of gesturing that will allow you to read the movement clearly. As you observe, remove yourself from becoming fully engaged in the show, and tune in to the rhythms of the action. It may even help to watch with the sound turned off, although voice carries a lot of rhythm. The interplay between voice and gesture will have an interesting syncopation. Comparing two or more cooking show hosts will allow you to see broad personality differences expressed through phrasing rhythms.

NIGELLA LAWSON VERSUS DAVID ROCCO

Cooking show hosts Nigella Lawson and David Rocco demonstrate a great deal of passion and charisma through their body movement. Lawson’s expression is fluid, flowing, lyrical, and sensuous. Her verbal delivery is intellectual, literary, poetic, and personal. One can imagine that her idea of the perfect day is to entertain in her home, enjoy a hot bath, and curl up with a good book. Rocco is snappy, syncopated, and staccato. His perfect day is out among people, sourcing the freshest ingredients at farmers’ markets, cooking at a friend’s house, drinking wine, and laughing late into the night.

How would you describe the differences in their movement and the rhythm of their phrases? Follow these links for example segments from their TV shows:

Nigella Lawson [Video Figure 6]

David Rocco [Video Figure 7, starting at 4:50]

BODY

The Body category describes structural aspects of the body in motion: which parts are moving or held, how movement flows from one part to the next (which is the essence of Follow Through and Overlapping), how the kinetic chains of the body are being patterned and coordinated, and postural habits from which gestural expression emerges. While this category mostly describes functional (i.e. biomechanical) aspects of movement, its parameters help us observe the degree of ease within the body and the ways in which the body serves authentic expression as the vehicle of its outward manifestation.

Patterns of Total Body Connectivity

The Patterns of Total Body Connectivity (Hackney, 1998) (also known as the Developmental Patterns) are six neuromuscular patterns around which all body movement is organized. As our motor patterns develop from infancy to adulthood, we acquire each pattern in overlapping progression, more or less chronologically as follows:

- The **Breath** pattern is the root of life and movement, and is the foundation of rhythm. Breath connects inner and outer in an ongoing, fluid quality of being; with the self, and within the world.
- The **Core-Distal** pattern connects the core, or navel center to the limbs, head and tail. The limbs find their connection to each other through the core. This pattern underlies our ability to relate self to other.
- **Head-Tail** connectivity is about the internal support of the spine, the relationship of the head and tail in spinal movement, and the full, flexible, three-dimensional mobility of the spine. Our sense of self is reflected in the postural support of the spine.
- The **Upper-Lower** pattern is about the support, yield/push and locomotive ability of the lower body, and the goal-oriented reach/pull of the upper body. This pattern cultivates the ability to cope with gravity; where we claim our stability and mobility.
 - In yoga, the Sun Salutation movement sequence is the practice of Upper-Lower patterning. [Video Figure 8, starting at 2:37]
 - In the poses and transitions between them, observe how the use of Breath, the Core-Distal connection of the naval center to the limbs, and the fluid, sequential motion of the spine support the sequence.
- **Body-Half** connectivity is about the two sides of the body. One half can stabilize, while the other is mobile. Body-Half patterning can be lizard-like. It supports comparing, choosing between one and another, and weighing two sides of an argument.
 - In this Tai-Chi demonstration, the master shows the Golden Rooster, stabilizing the left side against the mobile right. [Video Figure 9]
- The **Cross-Lateral** pattern is about the opposing coordination of right and left. It involves the diagonal connectivity, through the core, of the upper-right to lower-left, and upper-left to lower-right. It facilitates full three-dimensional movement and spiraling action. The development of cross-lateral patterning, needed for crawling and ultimately, walking, develops from the intent to reach for an object and to move towards it.
 - In this sword form of Tai-Chi [Video Figure 10] there are constant transitions between Homolateral and Cross-Lateral patterning.

It's important to understand the Patterns of Total Body Connectivity in the context of building animated movement sequences from scratch. In motion planning, an animator is disassembling the elements of movement that the body and mind have already integrated. Additionally, production for real-time interaction involves creating many short movement clips that are assembled via transition animations. Applying knowledge of the Patterns of Total Body Connectivity assists the process of creating movement that, functionally, flows with the connectivity and coordination that is necessary for believability. It allows sequences of short movement clips to be assembled with a view towards how the body coordinates transitions. Characters imbued with the appearance of connectivity and coordination better reflect the illusion of mental processes that are inherent in body coordination. We empathize more completely with characters that exhibit neurological processes.

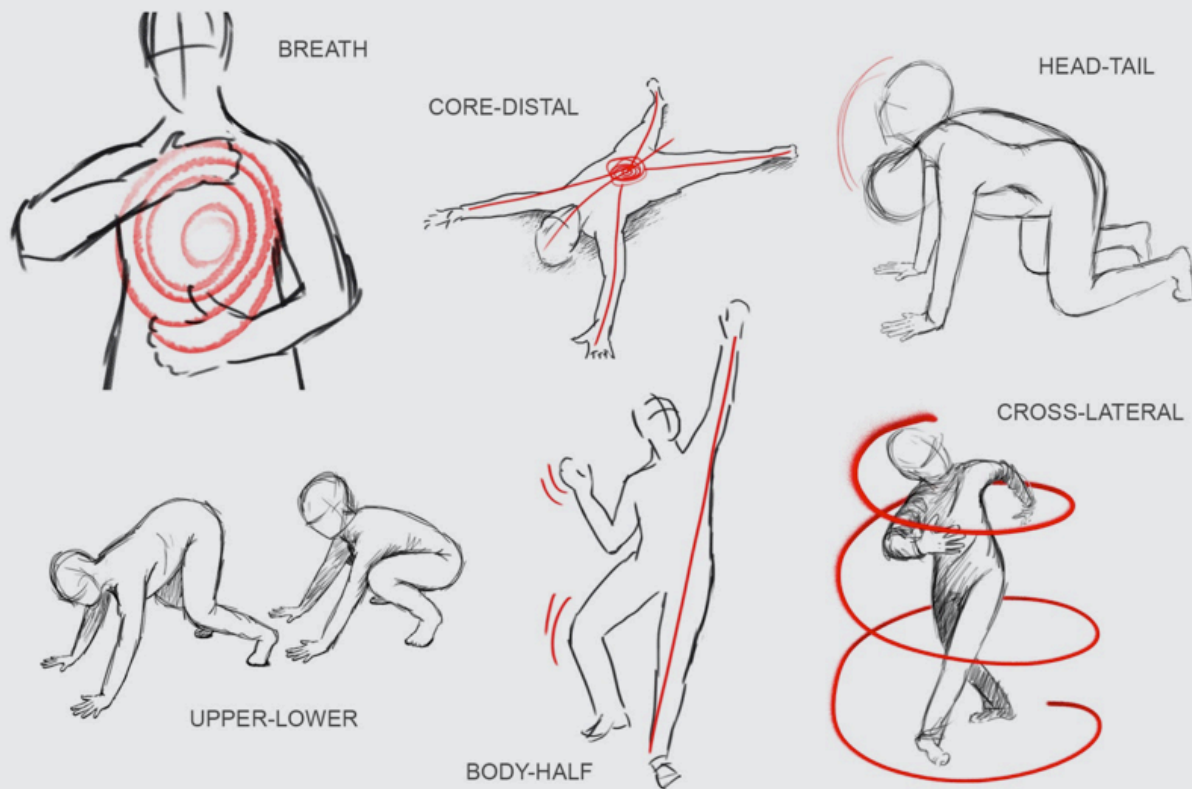


Figure 11-4: Developmental patterns of total body connectivity

Initiation and Sequencing

Initiation and Sequencing are terms that describe where a movement begins in the body, and how it flows through the kinetic chains of the body, from one body part to the next. Imagine that you want to hail a cab. How do you raise your arm? Does your hand lead the way, or do you first lift your elbow and raise the hand last?

There are three ways that movement can sequence through the body:

1. *Simultaneous* sequencing is when active body parts move together. If you are clapping your hands, both hands move simultaneously.
2. *Successive* sequencing occurs when movement flows from one body part to an adjacent body part. Making flying motions with the arms, or moving the spine like a snake, both involve the successive sequencing of joint rotations.
3. *Sequential* sequencing occurs when movement in one body part flows into a non-adjacent body part. Michael Jackson used this type of sequencing when a kick led to pushing up a sleeve, followed by a quick turn of the head. What makes this sequencing Sequential is that these three actions are done in rapid succession and have the feeling of a complete statement, or Phrase. (Hackney, 1998, p. 219) [Video Figures 11 – 13]

To identify which type of sequencing is being used, we have to look at where the movement Initiates. Which body part prepares for the action? Where does the movement flow to next?

These concepts relate closely to Anticipation, Follow Through, and Overlapping, and to the broader concept of movement phrases. Initiation and Sequencing describe bodily phrases, or the components of gesture. In the video examples referenced above, the break dance arm wave is a gestural action that forms a complete phrase, initiating in the shoulder, sequencing down the arm, ending at the hand. In animation, this supports the appearance of coordination and connectivity, which builds the functional/expressive potential of an action or gesture.

Posture and gesture

Character actions are carried out through a series of postures and gestures. Posture is the spinal support of the whole body configuration. It is how a person carries him/herself through life. Gesture is how we express small units of meaning, analogous to words, phrases, and sentences. Gestures occur mostly with our hands; however, we can use any part of the body to gesture, such as using a toe to point at something on the ground, or nodding a head in agreement.

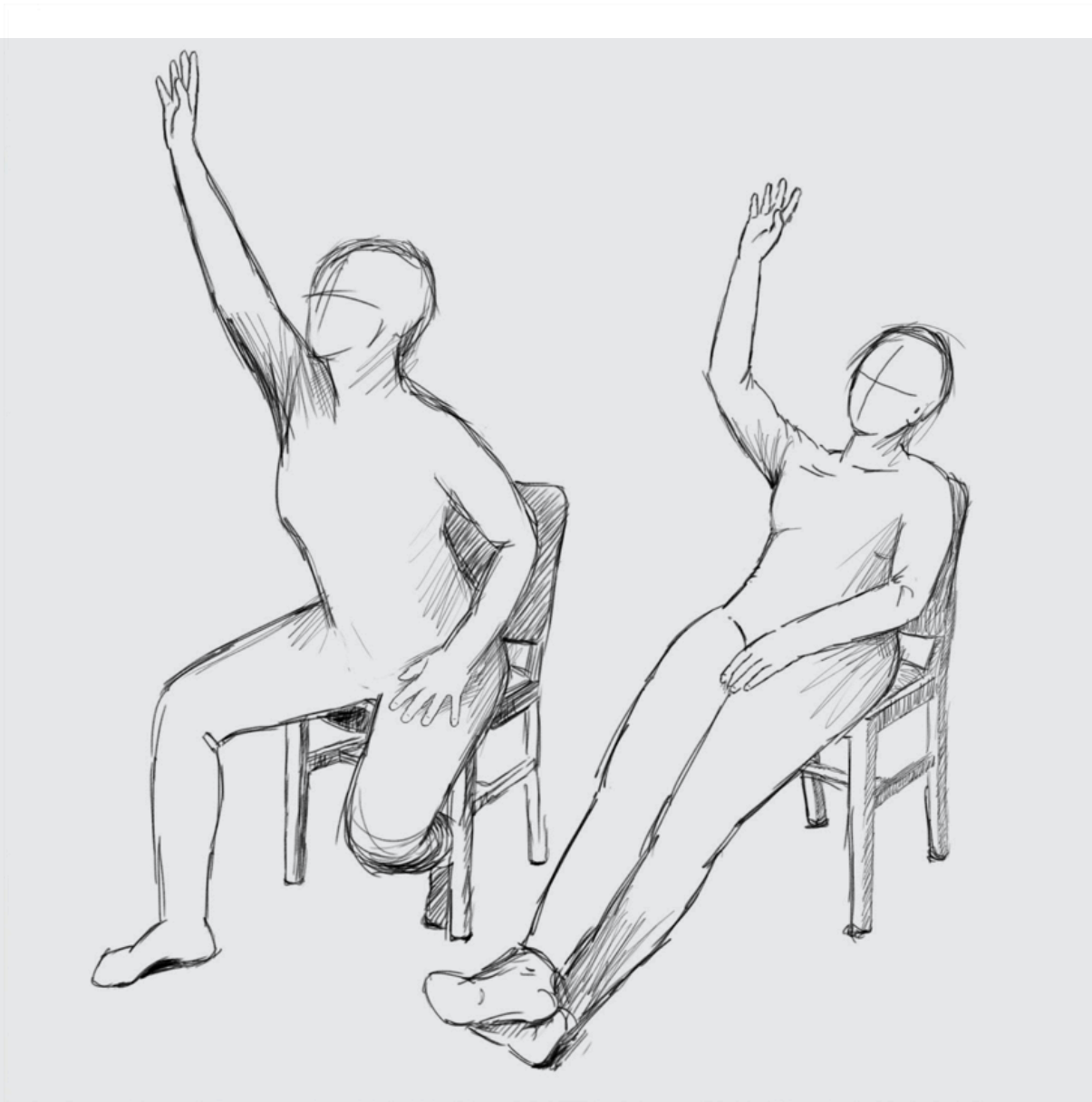


Figure 11-5: Same action, different postures

When posture and gesture are integrated, the whole body says the same thing. Imagine two students in a classroom who each raise their hand. One has a burning question and really wants to be noticed by the teacher. She will sit upright while raising her hand, all in one action. The extension of her arm will seem to arise out of her upright posture an integration of posture and gesture. The other student knows the answer to a question and would like to contribute to the discussion. However, he is concerned that his friends will think it is “uncool” to show an active interest at school. He leans back in his chair and slouches down, only half-raising his arm with a bent elbow. His gesture says one thing, but his posture says another (Davies, 2006, pp. 66-67).

EFFORT

Effort is the inner attitude expressed through movement. It describes the dynamic qualities of how a mover uses their energy. The Effort category delineates four Effort Factors: Weight, Space, Time, and Flow. Each factor represents a continuum of movement qualities. Movement qualities fluctuate along the continuum:

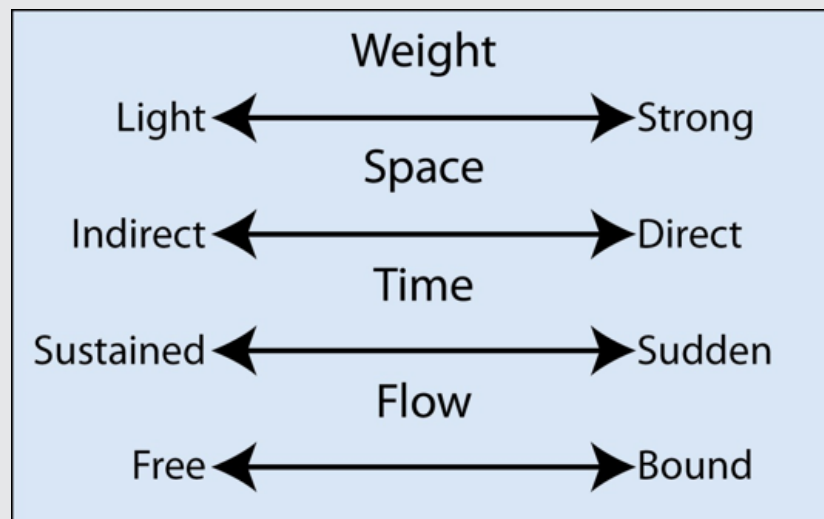


Figure 11-6: Effort Factors

Light, Indirect, Sustained, and Free are considered indulging or accepting qualities. Strong, Direct, Sudden, and Bound are resisting or fighting qualities.

The four Effort Factors of Weight, Space, Time, and Flow are linked with C. G. Jung’s four ego functions: sensing, thinking, intuiting, and feeling, respectively. For example, as we empathically attune with the Strong, Indirect and Quick slashing of a sword, we connect with the sensation of intuitive thought needed to perform the action. These associations lead us to create meaning through metaphor: I sense the strong handling of the sword as I bring my attention to the whole space around me. I intuitively choose the right moment to act.

From these associations, we observe that a mover’s Flow of Weight in Space and Time communicates information about: physical sensations and the agency to mobilize one’s weight with delicacy or force (Weight Effort), the broadness or focus of thought (Space Effort), the intuitive leisureliness or urgency of decisions (Time Effort), and the release or control of feelings (Flow Effort) (Bloom, 2006).

As we move, we express ourselves through continuous change and variation of Effort qualities. Most Effort is observed in subtle combinations of two elements, forming States, and three elements, creating Drives. In the rare case of an extreme and compelling movement, four elements combine in a *full Effort action*. In most movement, States occur as transitions between Drives, which are the dominant configuration.

Summary of Effort

FLOW: Feeling, Progression, “How”.

Feeling for how movement progresses

- Free: external release of energy; going with the flow; unrestrained; unstoppable
- Bound: contained and inward; resisting the flow; precision; control; ready to stop

Examples:

- Throwing a spear: Bound Flow is used to aim precisely and prepare. The spear is released with gradually increasing Free Flow.
- Running a race: Crossing the finish line occurs in full Free Flow, then Binding the Flow is involved to come to a stop.
- Threading a needle: Precision and Bound control of Flow are needed to insert the thread through the eye. Pulling the thread through is relaxed and Free.

WEIGHT: Sensing, Intention, “What”

How you sense and adjust to gravity

Weight Effort is specifically about the intentional use of energy required to move one’s body weight. It does not refer to the mass, or the heaviness of the body.

Active Weight is the intentional use of force in varying degrees:

- Light: delicate, sensitive, buoyant, easy intention
- Strong: bold, forceful, powerful, determined intention

Passive Weight is surrendering to gravity:

- Limp: weak, wilting, flaccid, partial giving in to gravity
- Heavy: collapse, giving up, letting go completely

Weight Sensing:

Activating the sensation of body weight; a buoyant interplay between Active and Passive Weight. “Weight Sensing fluctuates between active and passive weight, finding a yield and release into gravity with a rebound activation.” (Konic, Weight Sensing, 2011)

Examples:

- Light: lifting a delicate teacup, using a feather duster, cradling a soap bubble

SIDEBAR: WEIGHT AND ANIMATION

Weight is constantly addressed when discussing animation, because the illusion of the qualities of weight provides information about the materiality of form in motion. In this video clip from *Winnie the Pooh* (2011), Pooh strokes his jiggling belly, which creates a sense of its materiality. [Video Figure 15, starting at 0:24)

Materiality is intricately bound with intent because the motivation to move and act requires us to mobilize our body mass in constant negotiation with the effects of gravity. This sequence from *Winnie the Pooh and the Blustery Day* (1968) features how a powerful wind affects Pooh, Piglet, and elements of the environment. [Video Figure 16] Each character and object in the scene exhibits particular interactions with both wind and gravity. At one point Piglet is blown right into Pooh’s solid belly, contrasting his lightness with Pooh’s weightedness. The Dance of the Hours scene from

(continued on next page)

- Strong: pushing a piano, stomping a foot in defiance, slamming a door shut
- Limp: a wounded and weakened man lifting his arm. There is no energy in his hand
- Heavy: flopping onto the sofa
- Weight Sensing: In this video of an elderly couple performing a piano duet, the husband keeps the rhythm by weight sensing with his whole body [Video Figure 14].

SPACE: Thinking, Attention, “Where”
Thinking, or attention to spatial orientation

- Indirect: flexibility of the joints; three-dimensionality of space; all-round awareness
- Direct: linear actions, focused and specific; attention to a singular spatial possibility

Examples:

- Indirect: exploring a dark cave with a flashlight; swatting away a cloud of gnats; searching for a lost contact lens on the ground; a lone swordsman fending off a group of attackers who surround him on all sides.
- Direct: hammering a nail; throwing a baseball; pointing; reaching for an object; having a staring contest.

TIME: Intuition, Decision, “When”
Intuitive decisions concerning when

- Sustained: continuous; lingering; indulging in time; leisurely
- Sudden: unexpected; isolated; surprising; urgent

Examples:

- Imagine two people reading a newspaper on the subway: one is enjoying the arts and entertainment section, taking time to read each page, while another thumbs through the headlines, trying to scan as much information as possible before reaching the last stop.
- Sustained: petting a cat; strolling in a garden; gradually decelerating your bicycle as you approach a red traffic signal.
- Sudden: Soccer fans jump to their feet as a goal is scored, *The Flight of the Bumblebee* by Rimsky-Korsakov (1899-1900)²³; rushing to catch a bus

Fantasia (1940) plays with our expectation of heavy, cumbersome elephants, showing them dancing delicately with soap bubbles. At the end of the sequence, the elephants are blown away like leaves in the wind. [Video Figure 17]

We can observe many examples of animated movement, such as this scene from *ParaNorman* (2012) [Video Figure 18] with this specific phrasing (Kestenberg & Sossin, 1977):

Anticipation • Squash & Stretch • Follow Through/Overlapping

This type of phrase tends to develop effectively the intentional activation of weight, and the weightiness of animated characters. Anticipation is the Initiation of the phrase and shows intention. Squash and Stretch creates a sense of the plasticity and elasticity of form, as well as the amount of force taking place during the Exertion and Main Action of the phrase. Follow Through and Overlapping show the impact of the main action and the flowing transference of weight from one body segment to the next. The amount of Exaggeration applied to each part of the movement phrase shows the intentional or expressive activation of weight, as well as the functional activation of force needed to exert the body into action.

²³http://en.wikipedia.org/wiki/File:Rimsky-Korsakov_-_flight_of_the_bumblebee.oga

States and Drives

The range of possible Effort combinations is a rich parameter space. Combinations of two or three Effort qualities are referred to as States and Drives, respectively, with twenty-four State combinations and thirty-two Drive combinations possible.

Flow is considered the baseline Effort factor; it is always present, yet it is not always the dominant movement quality. Rhythmic fluctuations of Bound and Free Flow in infants develop into Weight, Space, and Time as motor patterning evolves (Kestenberg & Sossin, 1977).

Below are summaries of the main characteristics of each State and Drive. The LMA Effort Bank is a website with examples of variations within each. (Konie, 2011)

Table 11-1 – States: combinations of two efforts

AWAKE	REMOTE	STABLE
Space, Time Where, When Thinking, Intuiting	Space, Flow Where, How Thinking, Feeling	Space, Weight Where, What Thinking, Sensing
DREAM	RHYTHM	MOBILE
Weight, Flow What, How Sensing, Feeling	Weight, Time What, When Sensing, Intuiting	Time, Flow When, How Intuiting, Feeling

Table 11-2: Drives: Combinations of three efforts

ACTION	PASSION
Space, Weight, Time (Flow-less) ²⁴ Where, What, When Thinking, Sensing, Intuiting <ul style="list-style-type: none"> • Work, or “action” oriented movement • Performing a task • Without feeling or emotion (Flow-less) • Manual laborers, athletes, heroes 	Weight, Time, Flow (Space-less) What, When, How Sensing, Intuiting, Feeling <ul style="list-style-type: none"> • Emotional • Passionate • No thinking or analyzing • Child in a tantrum, heated discussion, violent raging
VISION	SPELL
Space, Time, Flow (Weight-less) Where, When, How Thinking, Intuiting, Feeling <ul style="list-style-type: none"> • About thinking, ideas • Imagination, envisioning, mental alertness • Without a sense of self, weight-less • Scientists, inventors, architects, designers, philosophers 	Space, Weight, Flow (Time-less) Where, What, How Thinking, Sensing, Feeling <ul style="list-style-type: none"> • Captivating • Engaging, persuasive, seductive • Time disappears. Feeling of timelessness • Magicians, charlatans, Pied Piper, preachers, politicians

Action Drive has become the most widely known aspect of LMA due to its extensive practice within theater. It is also the most accessible area of the LMA system because Laban used descriptive terms to identify the eight permutations of the Action Drive.

Table 11-3: Eight Action Drive Permutations

FLOAT Light—Indirect—Sustained	PUNCH Strong—Direct—Sudden
GLIDE Light—Direct—Sustained	SLASH Strong—Indirect—Sudden
DAB Light—Direct—Sudden	WRING Strong—Indirect—Sustained
FLICK Light—Indirect—Sudden	PRESS Strong—Direct—Sustained

²⁴http://en.wikipedia.org/wiki/File:Rimsky-Korsakov_-_flight_of_the_bumblebee.oga

In addition to these descriptive terms, Laban associated these Effort configurations with specific spatial zones around the body. The Action Drive configurations take the geometric form of a cube. In Figure 2, with the right side of the body active,²⁵ Laban observed a general tendency for Float to be performed in the forward-right-high corner of the cube, Punch to back-left-low, etc.

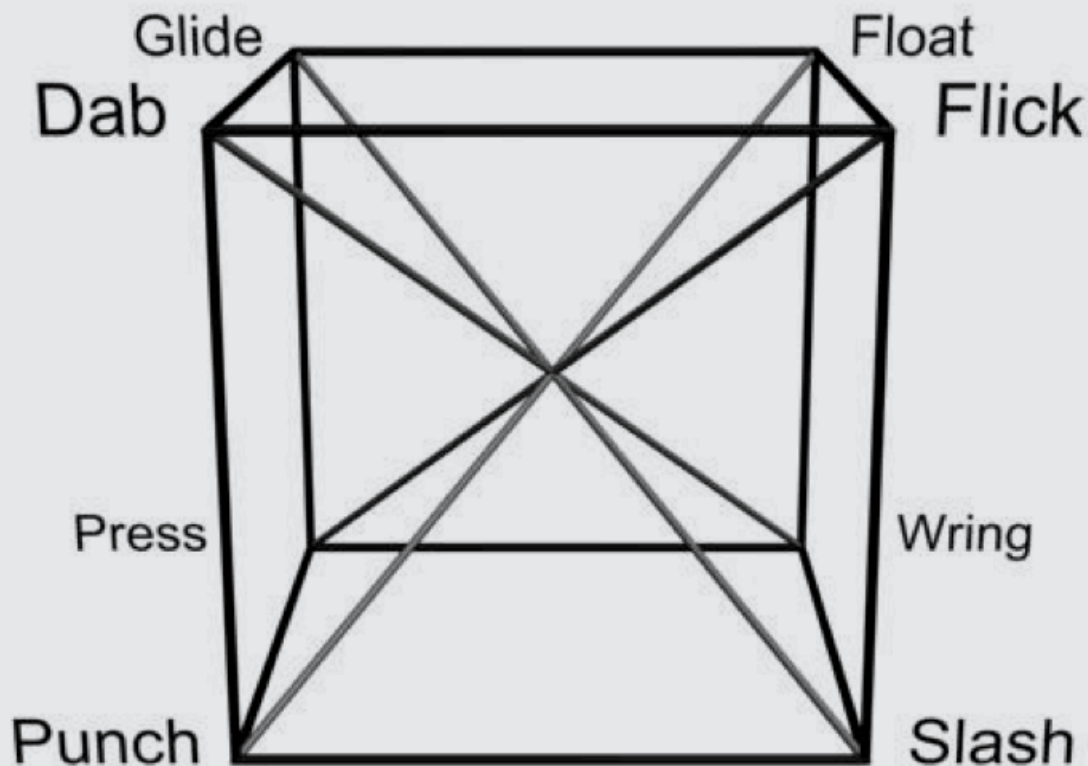


Figure 11-7:- Spatial affinities of the action drive

The affinities of specific qualities of Effort to areas of space is called Space Harmony, and is discussed in further detail in the section on Space. Effort/Space affinities are substantially observable in movement, yet Laban considered that they represent a general tendency, rather than a fixed way of being. For example, gently lowering a kitten to the ground is a counter-affinity to the upward direction of Float. An excited soccer fan can Punch the air upwards in excitement when a goal is scored. Like any aesthetic principle, Effort/Space affinities present one set of options from which many variations are possible.

Action Drive is the only area of the LMA system where movement has been associated with specific actions. It has been an accessible entry point to learning Effort, offering a broad range of movement possibilities for creative exploration. Although the strength of the LMA system is that it avoids mapping movement parameters to specific movements or interpretations, Action Drive seems to be an anomaly within the system. However, theoretically and historically, the principles underlying Action Drive are at the heart of Laban theory. Action Drive bears the unfortunate burden of being the most widely known, yet most misunderstood, aspects of LMA.

²⁵When the left side of the body is active, the spatial affinities switch sides: Float occurs at forward-left-high, Punch at back-right-low, etc.

Applying Effort theory in practice

Effort is a common, if not standard, part of theater training, giving actors a way to embody a variety of characters. Kevin Spacey demonstrated this beautifully when he performed impersonations of other actors on *Inside the Actors Studio*. [Video Figure 19] You can watch him become each character as host James Lipton addresses him with a question, to which he responds in character. As Jimmy Stewart, Spacey uses Direct Space and Sustained Time, with diminished Bound Flow in the moments where his voice hesitates. As Johnny Carson, he becomes Sudden and Indirect, looking away nervously as he evades Lipton's question. At the end of each impersonation, Spacey recuperates into "himself," his natural movement expression. Each impersonation evokes an empathic response in the viewer, parallel to our empathic response to movement qualities in animation.

Animators, as character actors who design performances, can apply Effort to character animation in much the same way that stage and screen actors do. Animators interpret the character performance from the script and storyboard, and receive input from the director until the performance has the right touch. Laban-based actor training can support animators in finding a character performance that is authentic to the story and character.

Developing Effort qualities of movement in animation is all about the design of expressive motion, which experienced animators cultivate intuitively. The Jungian ego functions, Sensing, Thinking, Intuiting, and Feeling, can provide an inroad to thinking about how specific character personalities express themselves in movement. For example, an inexperienced bank robber is going to feel intensely nervous (Feeling: Bound Flow), and maintain a heightened awareness of his surroundings (Thinking: Indirect Space), with quick reactions (Intuiting: Sudden Time) to any perceived threat. Describing the character's intent with language will reveal a treasure trove of words that evoke what Effort qualities a character may embody in movement.

SCRAT: EFFORT OBSERVATION EXERCISE

In the movie *Ice Age* (2002), Scrat is a character with very simple and clear intent: survival. His survival object is an acorn. To practice observing Effort, view Scrat in this opening sequence [Video Figure 20].

Tune in to your empathic feeling of the movement. Generate a list of descriptive words about the qualities of his movement. For example: nervous or panicked. Generate as many words that come to mind.

Next, separate the words according to Effort: Weight, Space, Time or Flow. Format your list as a chart, showing the polar elements of each. Below is a sample chart layout. View your chart and identify which Efforts characterize Scrat's movement. What does this reveal about his intent in terms of Sensing, Thinking, Feeling and Intuiting?

	ACCEPTING	RESISTING
WEIGHT	<u>Light</u>	<u>Strong</u>
SPACE	<u>Indirect</u>	<u>Direct</u>
TIME	<u>Sustained</u>	<u>Sudden</u> <i>panicked</i>
FLOW	<u>Free</u>	<u>Bound</u> <i>nervous</i>

SHAPE

Shape is the one area of LMA that connects directly to creating animation, as it describes the process of shape change over time. In animation, Shape translates to pose design, as well as the fluid interpolation through a series of poses to create movement phrases. Shape reveals how one's inner attitude and relationship with the external environment molds the changing plastic form of the body. Shape change may initiate as an adjustment of the body in relationship to self (just as Squash and Stretch represents the shifting of inner volumes), or as a bridge from self to other through Arc-like (pitching a baseball), Spoke-like (punching a punching bag) or Carving (molding a clay pot) Modes of Shape Change.

How we Shape ourselves through gesture indicates how we feel about our circumstances. Gestures connect the inner world of the self to the outer world of the environment. In this sense, gestures are a vital aspect to the communication of feelings that we can engage with empathically.

Modes of Shape Change

Posture and gesture together form a whole-body shape that is constantly changing its form. Gestures merge in and out of postures through a variety of modes:

- Directional shape change forms a connection from the self to the environment and from the environment to self. Common directional gestures are pointing to objects, or to oneself, or indicating points in space as if an object is there. Directional movement occurs through two types of motion pathways:
 - o spoke-like, moving in a pointed, linear manner, or
 - o arc-like, forming a curved path.
- Carving shape change is flexible and three-dimensional. It evokes an interaction with the environment, as if touching, molding or forming something. A character may be saying, "Let me give you the whole picture," while using both hands to form an imaginary ball in space.
- Shape-flow shape change is about adjusting inner volumes and feelings of self-comfort or discomfort, relating the self to the self. It occurs mostly through the torso and proximal joint relationships. Imagine snuggling your shoulders, neck, and head into your pillow as you go to sleep this is shape-flow movement, adjusting body parts in relationship to each other, as opposed to directional or carving in relation to the external environment. Or, imagine wriggling into a wetsuit. This time, the self-adjusting quality is accompanied by muscular tension instead of relaxation.

Shape-flow breath support describes how the breath supports movement in the rest of the body, visible as shape changes in the torso. The inhalation/exhalation of breath has a quality of presence that brings subtle shades of meaning to an action. For example, the action of raising both arms above the head, supported by an inhalation, has a quality of exaltedness. Releasing them downwards, supported by an exhalation, can seem like "giving up." How would the same action look with the pattern reversed? Breath as support also helps to integrate motor patterning. It provides access to the sensation of core connectivity.

Breath patterning offers a clear way to observe the different stages of movement Phrasing. Recall that the initiation of a phrase launches intention into action and determines how the motor patterning will unfold. Does a movement phrase begin with an inhalation or exhalation? As the breath initiates, how is the mover relating to the environment? As you observe Po in this Kung Fu Panda Holiday special [Video Figure 21], notice clear moments of inhalation (as he starts to battle Tai Lung) and exhalation (when he responds to Master Shifu with disappointment).

Shape Qualities

The body in motion is continuously changing, in a process of Opening into extension or Closing with flexion. Shape Qualities are the ongoing process of shape change, forming an overall configuration of Opening or Closing, in a directional relationship to the environment:

Rising – Sinking
Spreading – Enclosing
Advancing – Retreating

Effort/Shape Affinities

Laban and his collaborators noted that Shape Qualities are frequently clustered with corresponding Efforts:

Light/Rising – Strong/Sinking
Indirect/Spreading – Direct/Enclosing
Sustained/Advancing – Sudden/Retreating
Free/Opening — Bound/Closing

These dual “affinities” represent natural or accessible clusterings of Effort and Shape, yet counter-affinities (such as forcefully punching a fist upwards, or delicately lowering a kitten to the ground) bring texture and richness to the expressive range of movement choices.

The interplay of Flow and Shape forms the basis of elasticity that we can observe in subtle shape changes in the torso. With our breath, we respond to inner feelings of comfort/ safety or discomfort/danger. The resulting fluctuations of tension and release allow the torso to grow and shrink three-dimensionally. The Kestenberg Movement Profile (KMP) is an application of LMA theory that has particular relevance for animation. Based on the observation of qualities of elasticity in infants, the KMP articulates eighteen variations of how the flow of breath supports plastic shape adjustments. It can be viewed as an expanded theory of Squash and Stretch (Kestenberg & Sossin, 1977; Bishko, 2007).

SPACE

LMA considers the ongoing process of movement as it is situated in, and relates to, the spatial environment. The Space category describes the mover’s involvement in the three-dimensional external environment, forming spatial pulls and counter-tensions that stabilize or mobilize the body. The spatial range of movement varies within the Kinesphere, which is the reach space around the body.

Laban observed that people form complex spatial patterns that can be one-dimensional, two-dimensional (planar) or, in the case of three-dimensions, can take on the forms of various polyhedra: the octahedron, cube, icosahedron and dodecahedron. For example, there is a scene in *The Incredibles* (2004) where Bob Incredible’s boss at the insurance company lectures angrily about Bob’s performance on the job. The boss repeatedly punches his fists downwards, on either side of his body, indicating the lower corners of the Vertical plane, thrusts a fist up and forward to a corner of the Sagittal plane, and wipes both hands sideward, suggesting the Horizontal plane. This geometrical component of LMA lies at the core of Laban’s view that inner intention participates in the reciprocal relationship between self and environment through movement.

Off-planar motion is a challenge for animators to explore. Animators must be close observers of natural joint rotations in body poses to create movement that interpolates naturally between poses. There are some limitations inherent to animation software. For example, the default rotary function of character rigs is based on planar rotation expressed in X Y Z coordinate space. Also, software viewports display the figure in a planar format. Animating a walk cycle relies mostly on viewing and posing in the sagittal plane. Animators need to work from multiple viewing angles to consider more subtle rotations.



Figure 11-8: Kinesphere

MOVING SPACE: THE LABAN SCALES

Moving Space: The Laban Scales is an iPhone/iPad application that demonstrates the Space Harmony work of Rudolf Laban. It is an interactive, 3D compendium of the movement sequences of the scales.

<http://www.centermoves.com/moving-space/>

3. APPLYING LMA

This section provides a general overview of analysis principles and practices, demonstrating LMA in the context of character development and acting.

To segue into this topic, here is an example of LMA in action. I was attending a ballet in which three female dancers were choreographed to perform the same sequence. One of the dancers began to catch my attention. As she performed a *grand battement en pointe*, (a high kick with a straight leg), she appeared stiff. The action seemed to require more effort for her than it did for the other dancers. I noticed that she was holding her upper body tightly against the large motion of her leg, which she used to help her stabilize the action. Her upper and lower torso did not seem connected, which required the leg to work hard in a jerky action. Sure enough, this dancer actually fell onstage! I had noticed some functional problems with her movement that affected her level of control and stability. She seemed nervous onstage, perhaps in anticipation of the challenge of performing the *grand battement*, or perhaps causing the upper/lower disconnection that precipitated her fall.

Watching this dancer, I had the sense that something was not working. My observations were led by wanting to understand the problem. The fact that she fell seemed to help answer my question and summarize my observations: that she struggled with upper/lower connectivity in the grand battement. Her lack of connectivity manifested as Bound holding of the upper against the lower, as if trying to hold herself up in space.

ANALYSIS

Analysis practices in LMA can take on many forms, depending on how it is being applied, and what aspects of LMA are particularly relevant for a given field of study. In my example above, my point of inquiry was as a movement problem solver. This is a common application in somatic practices, such as yoga, physical therapy, or in sports training. Other areas of application include public speaking, and management training, where personal presentation, authentic expression and evaluation of others become part of the analysis.

The process of analysis includes the following:

- **Observation:** watching movement, live or from video sources;
- **Description:** using movement notation or writing to record observations; and
- **Synthesis/analysis:** making meaning of what is observed and described, finding patterns.

CONTEXTUAL AND RELATIVE

Movement is elusive to grasp. Its continuous flow and ever-changing nature require us to frame it in a context if we are to study and understand it. With context, we determine the macro and micro frames of reference. For example, stories in films and videogames form a broad context in which character movement takes place. The arc of analysis could span a complete story, a series of scenes, an individual scene, or short action phrases. It may be contextualized around a particular theme, such as analysis of Stability and Mobility, or a look at Phrasing and communication style.

It is often useful to pose a question as a way to contextualize the analysis and hone in on a meaningful outcome of the analysis process. As this section deals with character development, a useful question could be: How does movement serve the character towards reaching a specific goal in this story? The question can relate to specific elements of LMA: What is the character's intention, and how is it revealed through his/her choice of Effort? During the process of analysis, other questions and answers emerge alongside the leading question.

MACRO AND MICRO ANALYSIS

LMA looks at the big picture, but also captures the fine details. You determine the level of analysis when you contextualize: you choose which approach (macro or micro) is relevant for your needs.

Motif Description and Labanotation provide examples of how macro and micro analysis are applied. Motif Description is a form of notation that captures broad themes and allows room for interpretation, yet it can also encompass specific movement details. It is more generative, more poetic. Labanotation is specific: every detail is noted and recorded. It has a documentary purpose, but is also used as a choreographic tool (Fox & Wile, 2002).

Animators are well served by attending to broad, overarching factors that contribute to character movement, and allowing those factors to influence movement choices on a micro level. Animators are notoriously detail-oriented and sometimes have a hard time stepping back from the micro-level of their work to look at the big picture.

OBSERVATION

Observation involves not only our innate movement perceptions, but also the ability to separate one's own bodily experience and physical, expressive way-of-being from what we perceive in observed movement. Movement is multi-faceted, complex, and continuous.

According to Laban scholar Carol Lynn Moore, Observation involves four phases:

1. **relaxation:** setting a relaxed and receptive mood, tuning up one's kinesthetic sensibilities for the concentrated effort of analysis that follows
2. **attunement:** sensing the general configuration of movement and bringing the elements into focus, prior to making decisions about which features are most important

3. **point of concentration:** selecting an aspect of the movement to focus on, while remaining open to subsidiary elements that attract one's attention e.g., choosing to observe the Effort category, or observe for specific occurrences of one Effort quality. Observing fluctuations of Free/Bound Flow, one begins to notice other Efforts used in combination, and gradually changing Effort constellations begin to emerge. Moore says: "The aim . . . is to penetrate the movement event, teasing out the separate elements that comprise it, bringing details into sharp relief, and, in general, elucidating the action through analysis."
4. **recuperation:** taking small, momentary breaks, such as looking away from the scene you are observing, as well as larger breaks to recharge energy and keep the powers of perception fresh. (Moore, 1988, pp. 209-214)

Embodying the action is a useful technique to use during observation. You do what the moving person/character is doing, and repeat the action. This helps you clarify and empathize with the mover more fully. This process engages the empathic aspect of perception, enabling you to feel what it is to be that person.

DESCRIPTION

Observations are recorded on a coding sheet. The coding sheet can be designed to reflect the particular needs of an analysis. It is generally a reference or checklist of LMA elements to watch for. The observer records what he or she sees in LMA terms, as well as the frequency or repetition of habitual movement elements. The observer notes clusters or constellations of which elements of movement seem to occur together. For example, does the mover tend to combine Lengthening in the torso with Light, Upward adjustments of the head and neck? Is there rapid foot-tapping with Bound Flow and holding in the upper body? These descriptions are not so much about the specific actions performed, but about the elements of Body, Effort, Shape, Space, and Phrasing in use.

Animators often act out a sequence that they are planning to animate. They shoot video to reference their acting, providing an accessible analysis tool. This process gives them the most direct access to their kinesthetic expression, allowing them to capture their intuitive sense of how they would like to animate the scene. Poses and timing can be roughly modeled from the reference footage, but the animator will adjust both to highlight specific qualities of the character's expression. An understanding of LMA can help the animator not only access an authentic reference performance, but choose the qualities of the performance that support the animated character's acting. By itself, the video simply provides frame-by-frame documentation of movement, whereas LMA provides tools for synthesis and analysis.

SYNTHESIS AND ANALYSIS

Setting the context and path of inquiry for analysis frames what we are observing. After recording and reviewing observations, aspects of the movement start to "speak" to us. What patterns are noted? What stands out as significant? What is the movement saying?

A fundamental strength of the LMA system is that it does not employ a specific interpretive framework to make meaning of movement. The system supports practitioners in applying LMA to their field, to the context and questions they are asking about movement.

Carol Lynn Moore outlines the entire process as follows:

At the heart of this model is the question, 'Why?,' or the purpose for observing. Around this center are five other structural elements: 1) the role of the observer, 2) the duration of the observation, 3) the selection of movement parameters, 4) the mode of recording impressions, and 5) the process of making sense. (Moore, 1988, p. 224)

Within this chapter, my purpose for observation has been to uncover how function/expression serves authenticity and empathic engagement in character performances within virtual worlds. As an observer, my role involves practical and creative problem solving. My observations range from short movement sequences (less than one minute), to longer sequences that I have summarized in more general terms.

To make meaning, we intuitively allow descriptive, metaphoric language to emerge. Phrases such as “charged with energy,” “sitting on pins and needles,” “taking a stand,” or “down in the dumps” may arise as descriptions of what is observed in movement. For example, once at a conference, I watched a passionate presenter who seemed to be designing geometric shapes in the air around his upper body. He seemed to be visualizing his ideas and constructing them with his gestures, connecting the Quick, Direct and Bound elements of Vision Drive to specific points in space, and then recuperating into Weight Effort in Strong/Quick Rhythm state. Metaphor provides a broad, general and subjective means for making sense of movement observations, and descriptive language places them into context.

MOVEMENT SIGNATURES

Rudolph Laban had an uncanny ability to know people intimately through their movement (Lamb, Warren, in (Davies, 2006, pp. 145-146)). He observed people’s habitual preferences and patterned use of elements from each of the Body, Effort, Shape and Space categories. The unique Phrasing patterns of movement elements constitute one’s Movement Signature.

Describing someone’s Movement Signature involves a process of observation and analysis. For animated virtual characters, the process is generative. This section describes the process of observing for Movement Signature, based on observing a wide sample of motion over time, allowing patterns to emerge. The section on Character Dynamics goes into further detail about generating character development content that leads to the construction of character Movement Signatures.

HOW TO OBSERVE FOR MOVEMENT SIGNATURE

Observing Body

- Is there a preference for certain Developmental Patterns Of Total Body Connectivity? Are any of the Patterns noticeably limited or lacking? For example, Core-Distal connectivity is commonly lacking in contemporary Western culture, where people may feel self-conscious, thus limiting the use of full-bodied expression.
- How does the mover sequence their actions? Do they prefer Simultaneous or Successive sequencing?
- What is the Body Attitude the habitual relationship of the torso and pelvis which forms the default posture that all movement arises from and returns to? Body Attitude can be influenced by body type. For example, someone with a long neck may appear to “hold their head high.”
- Does the character tend to involve their whole body or just individual parts? Does the character integrate gestures with postures? Posture and gesture are integrated when both are saying the same thing simultaneously. Integration can reflect someone who seems genuine and open. Lack of integration can mean that the character is hiding a part of their feelings or keeping their expression more private as opposed to exposing themselves outwardly.
- Is there a part of the character’s body that is stiff or held immobile? This could be the result of a former injury, or simply a repeated holding pattern that has become habitual. Elderly people are often characterized through stiffness and limited mobility, after a lifetime of habitual holding!

Observing Effort

- What does the movement seem to say? Is it about Thinking (Space Effort), Feeling (Flow Effort), Sensing (Weight Effort) or Intuiting (Time Effort)?
- Which particular Efforts seem to stand out? Does the mover use a broad or narrow Effort palette?
- Does the mover show preference for Indulging (Free, Light, Indirect, Sustained) or Yielding (Bound, Strong, Direct, Sudden) Efforts?
- Does the mover repeat any unconscious gestures or qualities? For example, lip biting, spreading or pointing the fingers, or shoulder shrugs. What Efforts stand out in these unconscious actions?
- Do any statements or metaphors emerge about Effort?

OWNING THE DANCE

Michael Jackson's Movement Signature provides an example of applied LMA. The discussion in this paper summarizes Jackson's life history and persona, relating them to his personal movement patterns (Rhodes, 2011).

Observing Shape

- Which Shape Qualities are prevalent? Does the mover prefer Opening or Closing? Rising or Sinking? Advancing or Retreating? Spreading or Enclosing?
- Does the mover involve subtle changes of Shape Flow Breath Support in the torso?
- Does the mover relate mostly to themselves or bridge to the environment?
- Do any statements or metaphors emerge about Shape?

Observing Space

- How much Space does the character use when moving? How big does the character tend to make their actions? The bubble of reach-space around our bodies is called the Kinesphere.
- Where does the mover prefer to go in Space? Up/down, front/back, side/side?
- Does the mover use singular directions? Planar motions? Diagonals?
- Do any statements or metaphors emerge about Space?

Observing Patterns and Phrases

- How does the mover combine their preferred elements of Body, Effort, Shape and Space? For example:
 - Someone uses his hands Simultaneously with Sudden and Direct Effort. Is there repeated use of certain directions in Space?
 - Someone continuously shifts her weight with Free and Indirect gestures. Does she use Homolateral or Cross Lateral body coordination during this action?
 - What are the frequency and rhythm of Phrases? How does the mover transition through Exertion/Action/Recuperation? What Body/Effort/Shape/Space configurations are preferred for Exertion? For Action? For Recuperation?
 - What statements or metaphors emerge about this mover in general? How does this person create meaning through their Phrasing of the elements of LMA?

A Movement Signature is a character profile in movement terms. It describes the essence of how one communicates, and of the individual's way of being. It provides deeply personal information and, when used as a tool for character development, offers a powerful approach to creating empathy with characters.

CHARACTER DYNAMICS

Character Dynamics is the term I use to describe how a range of factors comes together to define a character. Broadly, the idea is that you develop a character within the contexts of:

- their personal history;
- their goals and obstacles;
- the story and present circumstances;
- and in the context of the other characters that they interact with.

All of these factors influence how the character acts, what they express, and ultimately, the way the character moves. This is essential material an animator must know in order to create effective character performances. Character Dynamics builds on what is conventionally known as the character bible:

- o The character **biography** is a list of details about the character's life story. The life of the character is imagined, such as where, when and how they were raised, cultural and ethnic context, beliefs, values, goals in life, etc.
- o What are the character's **goals** and **obstacles**? What motivates the character? What gets in the way? These factors can be broad factors as part of the character biography. They are also, importantly, at the heart of every story, and every scene within that story. How a character chooses their actions within a story context is derived from the influences outlined in their biography, their guiding goals/beliefs/values and personal obstacles, situated within their goals and obstacles of the story and its circumstances. Ed Hooks states that a scene is a negotiation, meaning that a character navigates their obstacles in pursuit of their goals (Hooks, 2003).
- o Social **status** is a reflection of basic survival instincts. Every character interaction is based on the primal need to survive the encounter. Characters continuously modulate their status in relationship to each other to achieve their goals. A character's cultural/social status is one layer of status. Another layer is the status that a character creates for himself in the moment. For example, in the film *Horton Hears a Who* (2008), Kangaroo contracts Vlad the vulture to destroy the clover. Kangaroo maintains status over Vlad by turning down his services. Vlad finally offers to do the job for free, having been manipulated into thinking he has scored the upper hand in the deal. Vlad feels that his status has elevated by landing the deal, and doesn't realize that he has been tricked, which lowers his status even further.
- o What is the **arc** of the character within a story? What happens that may influence changes to the characters goals, obstacles, social status and status relationships?
- o What is the character's **Movement Signature**? How does a particular character move, as the expression of their life history and motivating/inhibiting factors, negotiations with others, and grappling with the arc of their goals/obstacles within this story? What default movement patterns do they bring to their current circumstances? What movement serves them towards achieving their goals and overcoming their obstacles? Does their movement evolve through the arc of the story?²⁶

²⁶See Developing Personality (Bishko, in Furniss, 2008, pp. 60-63) for more details on developing Effort in creating a movement signature.

Game animator Mike Jungbluth feels that a “belief system” needs to be actively embedded into every aspect of a game:

A big part of establishing this belief system and maintaining it is through an exhaustive character bible. Beyond just model sheets and reference for movements, really think about what drives the character forward. What has lead them to the point they are at when the game starts, and where do they draw the line in their world, as to what they believe in? Do their beliefs change or grow as the game progresses?

Do they mind getting their hands dirty or are they reluctant to do so? Both can allow for the same overall gameplay and creation of assets, but being aware of what they believe can make what happens before, during and after all the more meaningful when the animations or dialog matches those beliefs. This goes for not only the character, but the player. In fact, going a step further, this is how we can even begin to color the player’s beliefs, and make them question their own values versus those of the characters in the game. (Jungbluth, *What Does Your Game Believe In?*, 2011)

Character Dynamics is a method that creates a complete belief system for characters, relating their circumstances, goals and obstacles with how they act and move. It is a powerful creative tool for creating empathic character performances.

4. BIBLIOGRAPHY

- Bishko, L. (2007, December 9). *The Uses and Abuses of Cartoon Style in Animation*. Retrieved February 3, 2011, from Animation Studies Online Journal: <http://journal.animationstudies.org/leslie-bishko-the-uses-and-abuses-of-cartoon-style-in-animation/>
- Bloom, K. (2006). *The Embodied Self: Movement and Psychoanalysis*. London: Karnac Books.
- Davies, E. (2006). *Beyond Dance: Laban’s Legacy of Movement Analysis*. New York: Routledge.
- Fox, I., & Wile, C. (2002, March 1). *Names for What We Do*. Retrieved January 11, 2013, from DNB Theory Bulletin Board: <http://dnbtheorybb.blogspot.ca/2010/01/names-for-what-we-do.html>
- Hackney, P. (1998). *Making Connections: total body integration through Bartenieff Fundamentals*. Amsterdam: Gordon and Breach.
- Hooks, E. (2003). *Acting for Animators: A Complete Guide to Performance Animation* (Revised Edition ed.). Heinemann Drama.
- Johnston, O., & Thomas, F. (1981). *Disney Animation; The Illusion of Life* (1st Edition ed.). New York: Abbeville.
- Jungbluth, M. (2011, April 10). *Growing Game Animation - Transitions & Player Input*. Retrieved December 15, 2012, from #AltDevBlogADay: <http://www.altdevblogaday.com/2011/04/10/growing-game-animation-transitions-player-input/>
- Jungbluth, M. (2011, May 25). *What Does Your Game Believe In?* Retrieved December 18, 2012, from #AltDevBlogADay: <http://www.altdevblogaday.com/2011/05/25/what-does-your-game-believe-in/>
- Kestenber, J., & Sossin, M. (1977). *The Role of Movement Patterns in Development*. New York: Dance Notation Bureau Press.
- Konie, R. (2011). Retrieved from The LMA Effort Bank: <http://www.lmaeffortbank.com>
- Konie, R. (2011). *Weight Sensing*. Retrieved January 11, 2013, from LMA Effort Bank: <http://www.lmaeffortbank.com/weight-sensing.html>
- Laban, R. (1966). *The Language of Movement: A Guidebook to Choreutics*. (L. Ullmann, Ed.) Boston: Plays, Inc.
- Moore, C.-L. (1988). *Beyond Words: Movement Observation and Analysis*. London: Gordon and Breach.
- Rhodes, G. G. (2011). *Owning the Dance: Michael Jackson’s Movement Signature*. Columbia University: <http://academiccommons.columbia.edu/item/ac:143893>

12

LOSS OF AGENCY AS EXPRESSION IN AVATAR PERFORMANCE

By Ben Unterman and Jeremy Owen Turner

This chapter examines the role of agency as a tool for non-verbal communication in avatar performances staged in online multi-user virtual environments. These live online events have remediated established traditions of performance art and theatre practices into online spaces such as UpStage and *Second Life*. As with traditional performance forms, particular gestures within virtual worlds are also viewed by the local community as artistic. These gestures are frequently more abstract or codified than everyday gestures, changing their expressive value and significance. The authors focus on two avatar performances *Lines* (2009) and *Spawn of the Surreal* (2007), and explore the ways in which audience agency was manipulated in the creation of aesthetic experience. These artistic events demonstrate the strength of agency as an expressive tool and indicate an intriguing new direction for NVC research even outside of the performance context.

1. INTRODUCTION

Common to both the analog artistic disciplines of theatre and more recently, performance art, there remains a well-documented Western legacy of (en-) coded gestural tropes that spans at least two centuries. Over the course of the 20th century, non-mimetic gestures - gestures which are not based on our real-world gestural experience - became a vital part of modernist experimentation with theatrical form, stretching the limits of human embodiment on stage. Subsequently, these experiments became an important part of the various remediations of the theatrical event, expanding the expressive tools available for the creation of aesthetic experiences. In each case, the technical affordances that enable this remediation allow for different expansions of this idea. The results of these past experiments in the analog world are having an influence on the unique affordances of virtual environments.

In virtual environments, the technologies of digital embodiment and agency can be manipulated to cause a partial ontological disembodiment of the avatar from that of its creator/user. As a result, there are an increasing number of experimental artists from virtual communities that focus on manipulating an “actor’s” level of embodied agency as a new aesthetic affordance that is unique to digitally mediated “virtual” environments.

To illustrate some of these approaches, we will be analyzing two events from the emerging field of avatar performance (also sometimes known as cyberformance, hyperformance, cyberdrama, networked theatre or virtual theatre, among other names), *Lines* (Unterman, 2009) and *Spawn of the Surreal* (Second Front, 2007). As digital remediations of theatrical practice, avatar performances are live artistic events staged using various networked communications interfaces, such as chat rooms, multi-user video games or other virtual environments. Both *Lines* and *Spawn of the Surreal* make use of coded elements within the virtual

environment to control and limit the agency of audience members within the performative event, creating uniquely digital aesthetic experiences.

To properly understand the unique contribution of the handling of non-verbal actor-agency in virtual worlds, we need to first historically contextualize this elusive phenomenon using examples from both the theatre and performance art traditions. With every example, we will address both the issues of the zeitgeist and the critical interpretations that have arisen in retrospect. Specifically, we will initially focus on “real life” performances that either explicitly employs actor automation and/or other non-plausible (inorganic) gestures. Afterwards, we will discuss similar examples in textual and graphical virtual worlds - both from our own personal art-practice as well as other key examples that illuminate this phenomenon. Finally, we will summarize all of these examples and provide avenues for evolving the research significance of this particular theme on “virtual worlds agency” for future possibilities that involve the inclusion of artificially autonomous agents as legitimate collaborative performance-art entities.

2. THE ROLE OF NON-MIMETIC GESTURE AND AGENCY IN RL PERFORMANCE

Within the modernist traditions of theatre and performance art, there is a long tradition of undermining the body’s usual movement patterns. In fact, the emphasis on non-mimetic gesture is one of the phenomena which divided theatre from performance art as the two traditions of embodied performance split from each other in the early to mid-20th century. Our goal here is to introduce some of the key theorists and practitioners in order to illustrate some of the theoretical underpinnings of artistic gesture in virtual environments.

Many of these ideas have their root in “Über das Marionettentheater,” an article glorifying puppetry written in 1810 by Heinrich von Kleist. Not only does he frame puppets as not having man’s moral weaknesses (Malone, 2000, p. 58), but also as not being subject to gravity and thus heirs to more utopian expressions of aesthetic beauty (Malone, 2000). While Kleist’s notions may be over-romanticized, he does nonetheless value non-human gestures above human ones. This call was taken up by scenographer Gordon Craig in 1907 in *The Actor and the Über-marionette*, where he advocated replacing actors with more malleable non-organic figures (Craig, 1983 [1907]). According to Craig, the sole importance of the actor on stage was that they conferred movement on the otherwise static forms of stagecraft, but often did not have sufficient control to be able to create formally consistent work (1983 [1907]).

Many of these same concerns were taken up in some of the formal experiments conducted by Oskar Schlemmer at the Bauhaus Theatre Workshop between 1921 and 1929. In a series of experiments, the workshop attempted to re-create the human body as abstract shapes primarily through the use of costumes. But, as they quickly discovered, because of the ways in which they limited movement, the new costumes forced the actors and dancers to create innovative gestures, which reflected the forms and materials of the costumes themselves. As Schlemmer put it, the costume had the potential to change the nature of the body, changing the mechanical or organic laws to which the body seemed to conform (Schlemmer, 1961), echoing Meyerhold’s insistence on the interplay between the organic and the mechanical (Braun, 1979). The parallels between the Bauhaus experiments and Kleist and Craig’s mechanical actors were further emphasized by Moholy-Nagy’s performance score for a “mechanical eccentric,” an entirely mechanized performance machine (Kirby, 1995), extending the ideas of performance beyond changing the mechanical nature of the body to questioning the centrality of the body itself within performance. While this very physical and mechanical approach was not common in performance art circles, it has been a significant influence on both puppetry and installation art.

A separate line of enquiry in theatrical experimentation saw the roots of effective non-mimetic gesture in the psychic control of the actor. In his “Manifesto on the Theatre of Cruelty,” Antonin Artaud calls for the actor to rid their bodies of their humanity transforming themselves into a true performance instrument (Artaud, 1998: 152). They had to be prepared to act in extreme ways, in order to allow the psyche of their character to emerge “as rays of light” (p. 153), creating a non-logocentric gestural language. Artaud was heavily influenced by a romanticized vision of the transcendental and ritual aspects of Balinese and oriental theatres, which he saw as being more pure than the textually-based western performance tradition (p. 105). He was never able to convert these ideas into reality, however; and most criticisms of his work rely on this fact, although his ideas were very influential in experimental theatre and performance art.

Perhaps the most notable implementation of these ideas from a gestural perspective comes from the work of Jerzy Grotowski’s acting system, which likewise focuses on the actor’s psychic impulses. Echoing Artaud, Grotowski seeks the integration of the actor’s physical and psychic energy to create what he terms “translumination” (Grotowski, 1971, p. 14); he writes that “artificial composition does not limit the spiritual, but actually leads to it [...]. forms of ‘natural’ behaviour obscure the truth” (p. 17). The physical acting in productions, such as *The Constant Prince* (1969), were both striking and compelling, reinforcing the idea of non-mimetic gesture as a potent element of theatrical expression.

As it can be seen here, there are two distinct approaches to creating non-mimetic gesture on stage. The first, as seen in the work of Grotowski and echoed by physical theatre practitioners and dancers worldwide, revolves around the careful training of the human body. By applying rigorous spiritual and physical techniques, the body is able to become more than human, extending its expressivity for artistic effect. Many of these techniques, especially as called for by Artaud, involve the suspension of conscious control through the use of trance states in order to further extend the physical capabilities of the actor’s body. In this way, the actors prepare themselves for the performance not by adopting a character, but by suspending their own internal gestural cues in order to re-embody themselves as abstract expressive forms in space, what Valère Novarina (1988) calls the decomposition of the human body on stage.

But it is perhaps the earlier experiments of Bauhaus, which are more explicit. In that case, gestures were transformed by limiting the agency of the actors by using costumes which only permitted them to move in certain ways. A more abstract, visually oriented theatre was the result – one which replicated the Bauhaus painting and sculptural style on stage. The usual mimesis of gesture undermined in this case by rigid limitations of both form and medium. Regardless of whether the limitations were physical or psychological, the modification of the gestural capacity of the human body was an important tendency in the avant-garde theatre of the early 20th century.

Artists in various domains have furthered this aesthetic. Cybernetic art in particular makes heavy use of the ambiguities of control between people and machines. This is perhaps most striking in the work of Stelarc, especially his performance *MOVATAR* (2000), where a series of electrodes connected to a computer system administered shocks to his muscles, controlling his movements. Using this “inverse motion capture system,” Stelarc questions the relationship between the human body and technology (Scheer, 2002).

Similar explorations of involuntary movement can be seen in modern dance, for example in the work of Merce Cunningham. His choreography for “Untitled Solo” (originally produced in 1953) was based largely on chance operations, recreating the body as “an inorganic ‘assemblage’ of parts” (Copeland, 2004, p. 184). According to Copeland, this piece served as a key inspiration for later Cunningham choreographies, which made use of computer technology to recombine elements of the organic machine, which is the human body (2004). “Untitled Solo” effectively serves to connect the physicality of embodied performance with the ideologies of production found in media production, including uses of editing from film and remix from analog and digital audio production. It is this link between the body and the technologies of production, which is salient when it comes to our discussion of avatar performance.

These examples of non-mimetic gestures brought on by a loss of agency mark a very important and challenging step in the aesthetic presentation of the human body. These modalities are also present in avatar performances and other forms of online expression based on the appearance of liveness and representations of the human body. As has been seen in several of these examples, there has been a link between these experiments and technologies from the beginning, in particular in the case of the Bauhaus Theatre Workshop and the choreographies of Merce Cunningham. As we will see in the following examples, the interplay between embodiment and technology can take similar forms in virtual environments.

3. A SHORT HISTORY OF TEXTUAL AND 2D AVATAR PERFORMANCE

Since we have just gone through a brief overview of the use of agency-challenging tropes within analog or “real life” performance art traditions, we will now focus in more detail on “avatar” performance art in virtual worlds. Although there have been documents of numerous avatar performances over at least the past decade, we will limit our examples to those that allow audiences some measure of gestural agency. This important aspect of interaction allows individuals to influence the event using mimetic representations of real-life nonverbal communication or more abstract forms of nonverbal communication using a gestural device such as a mouse. We begin with avatar gestures within the text-only virtual worlds such as MOOs, following that with some examples from 2-dimensional events staged in The Palace and UpStage. These strategies seen in these examples serve as context for our more detailed discussion of the avatar performance Lines.

Text-only virtual environments presented special challenges for the expression of emotion and non-verbal communication online; challenges which needed to be met in order to create compelling and expressive performative events. Many strategies were used to accomplish this, including some which undermined user agency within the virtual environments.

Before addressing these techniques, we should take a minute to stress the importance of user agency within these environments. Media theorists have long held that agency is particularly important within digital environments. Lev Manovich, for instance, states that users’ abilities to influence the presentation of information is “the most basic fact about computers” (2002, p. 71), while authors such as Janet Murray emphasize the importance of agency and interaction in the creation of new digital storytelling techniques (1999). In addition, there have been well-documented examples of the importance of agency within virtual communities. Perhaps the best known of these is the virtual rape of avatars in LambdaMOO through the use of code, which attributed unwanted actions to other users’ avatars (Dibbell, 1994). Because of the importance of agency to both narrative and social interactions online, the unwanted loss of agency may well be the strongest single emotional tool available to artists online.

Most examples of this transition are not nearly as traumatic as those experienced in LambdaMOO. In some cases, such as the Schweller Theatre in ATHEMOO, which is modeled on traditional physical theatre spaces, the loss of agency is an important part of the presentation of avatar theatre. The primary audience space, the Main Floor, limits audience members’ communications to whispers and the only actions they are allowed are: clap, laugh, boo and shout (Schweller, 1999). A second audience space, the Balcony, allows greater range of actions, but is effectively segregated from the Main Floor so as not to disturb the viewing of that portion of the audience (Schweller, 1999). In this particular case, limiting the agency of audience members is seen as a benefit as it allows them to enjoy the performance more without the distractions inherent in usual online interactions.

But the manipulation of agency in text-based environments can involve more than the simple removal of agency from participants. One such environment was MetaMOOphosis, which was intended as the site of improvisations based on Kafka's short story *The Metamorphosis*, written in 1915. In order to participate with the story, participants need to choose and put on "costumes" which influence the ways in which they are able to interact with the environment. In fact, participants are not able to "say" or "emote" (the two basic MOO actions) without putting a costume on (Sacks, 1999), relegating them to the role of passive observers of the scene. Once a user selects a costume, their visible identity changes and they are given access to various tools, which help make their experience of the event unique (Sacks, 1999). In a world where everything is based on textual description, the costume serves as a way to enhance participants' agency within the narrative environment in order to emphasize elements of the story as it emerges.

These two are very different approaches to limiting user agency both have their roots in more traditional approaches to theatrical audience reception. In fact, the limits placed on audiences through both social constraints and the spatial design of theatrical space contribute to the mediation that is at the core of the theatrical experience (Blau 1987). That the implicit behavioural limits placed on audience members become explicit in at least some of the remediations of the theatrical event is certainly to be expected as new art forms emerge. Likewise, the adoption of costume and role identification as limits and extensions of participant behaviour are common in interactive theatre elements and live-action role playing games. These limits to user agency avoid the discomfort felt in other areas precisely because they adhere to established paradigms of behavioural control. The familiarity of these models allows the limits they place on verbal and non-verbal communication alike to become part of the artistic experience associated with these events.

In fact, as avatar performances moved into more graphical media, the lack of clear audience and participant conventions proved to be difficult. Early experiments in 2D graphical avatar performances using The Palace platform certainly experienced a lot of issues associated with this. During a performance of Avatar Body Collision's *Dress the Nation* (2003), for instance, performers felt obligated to remove a participant whose drawings and comments were interrupting the show itself. Consequently, this occurred in spite of the artists' commitment to creating work, which encouraged audience participation. The lack of other options in dealing with "inappropriate" audience participation became a significant preoccupation of this group and were part of the reason for the creation of UpStage, the first open source platform designed specifically for avatar performances.

UpStage is a 2D graphical chat with special options designed to aid in avatar performances. Like the Schweller Theatre in ATHEMOO, it has slightly different interfaces for audience members and for actors, although these do not appear as a literal translation of theatrical spaces. The agency of audience members is significantly limited in this environment, as they are not able to affect the visual space at all, and are never identified as individuals within the chat log. When actors speak, the words are displayed in speech bubbles next to their avatar and are spoken using a voice synthesis program. The words also appear next to the avatar name in the chat panel in black text. In contrast, audience members are never named and all of their text appears in grey - slightly faded. This establishes a very traditional performance hierarchy, where the actors on stage are prioritized over the largely anonymous audience members. Like previous examples, limits to audience agency are used in order to highlight the work of art and the artist, encouraging a very presentational, non-interactive style of avatar performance.

In UpStage, the patterns of non-verbal communication usually follow in a similar presentational style. Each avatar can contain several different images or videos, which are usually used to illustrate different moods or poses of a single figure. Avatars can also be positioned and moved across the visual space and these movements are coordinated to create the choreography of a specific performance. The non-verbal aspects of communication in this environment thus borrow heavily from traditional visual arts, in particular collage and painting with occasional references to very basic animation. Significant attention is paid to compositional elements as well, as this is the primary form of non-verbal expression in this environment.

Music and other audio are also used extensively as both a mimetic and non-mimetic support for the stage action. Both mood and plot elements are frequently indicated in this way, adding expressive depth to performances which are mostly lacking in gestural cues of emotion and meaning.

4. CASE STUDY 1: *LINES*

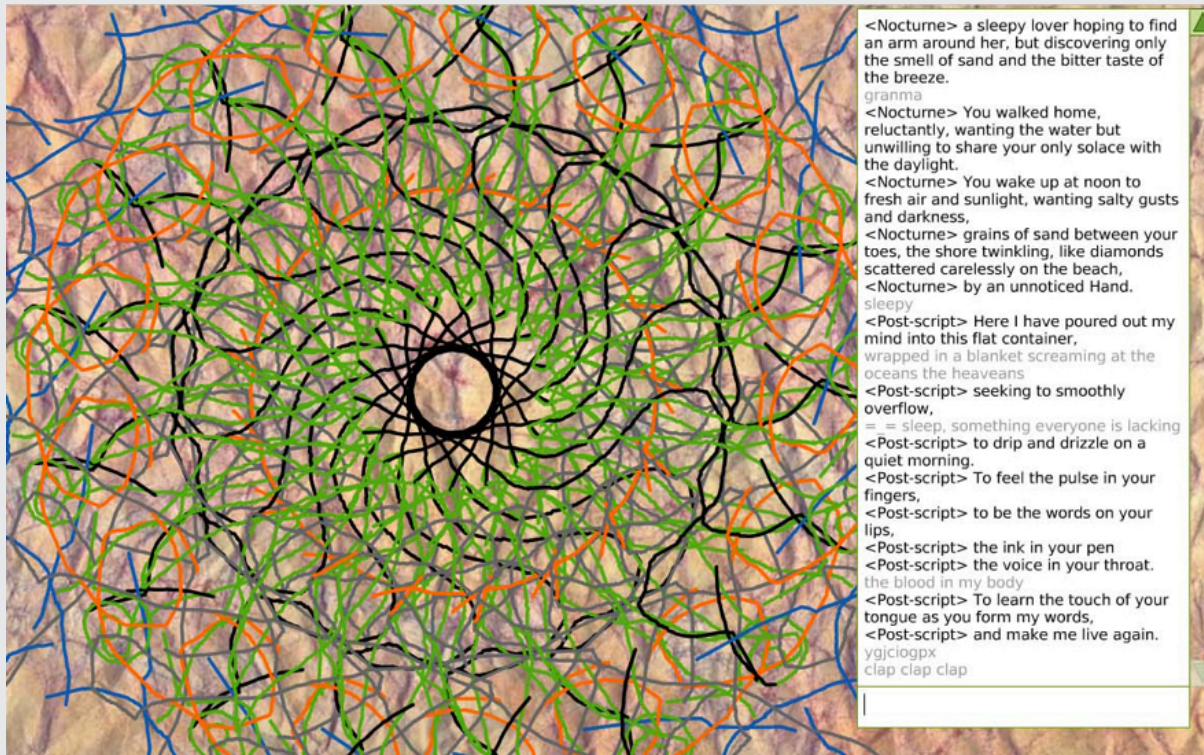


Figure 12-1: A Screenshot from *Lines*

Created by Ben Unterman, Daniel Silverman and Don Masakayan in 2009, *Lines* [Figure 12-1] is an experiment in interactivity in avatar performance using the UpStage environment. Whereas previous events liminally engaged audience activity and opened dialogue between the audience and performers, *Lines* actively encouraged explicit audience creativity. To accomplish this, the artists paired a series of poems with one of the most basic forms of computer interaction: onscreen line drawing.

The script for *Lines* contained a series of poetic meditations on city spaces written by Silverman and adapted for online presentation. These descriptive poems were written in the second person and represented what the reader or audience member had experienced during the performance section. Intentionally, this text was written as something to be appropriated and interpreted by the audience. Certainly, the explicit recognition of the interpretive act is historically typical of both interactive performance and digital storytelling. Regardless, the artists were very much conscious of the inherent challenges that arise when disparate disciplines and worlds converge in real-time.

In order for the artists to create an interactive experience, which mirrored the self-reflexive text, the artists utilized mandalas as way of projecting the meditative aspects of the text into the visual realm. Rather than providing the audience with pre-composed images specifically selected to reflect the text, the artists

gave the audience members a generous amount of creative leeway to contribute their own. To facilitate the user's agency in this creative process, Masakayan created a very unconventional drawing tool for making mandalas.

Before the hyperformance commenced, audience members were encouraged to experiment with the drawing tool, thereby familiarizing themselves with the potential to influence the visual appearance of the stage. Once the show officially started, however, the audience's perceived participatory agency was immediately subverted by the temporarily occluded affordances of the performance interface. During the first act, the lines on the screen no longer followed the cursor location, but were rotated 180 degrees around the screen's centre. This shift away from the expected manual functionality towards more autonomic gestural responses indicated to the audience that they were moving away from a literal world into a more metaphorical one. In this way, audiences were forced to re-evaluate the creative position of their own agency within the performance space.

Building on this defamiliarization of the audience's use of the drawing tool, the artists gradually resumed "control" of their creation by copying the spectators' initial drawing three times around the screen's centre. This introduction of rotational symmetry into their intended visual output mirrored the construction of the text itself. The artists provided the audience words to possess as personalized modular content, but simultaneously, these artists also hi-jacked the audience's creative agency by taking control of their own creation. In other words, the artists subverted the audience's will, their textual creation for the performative rights to the interactor's visual contribution. With each successive scene, the rules of the drawing tool evolved and the number of rotational copies of the audience's drawing increased until every line they put on their screen was displayed thirteen times around the screen's centre. Consequently, the audience's gestalt interactive output was a mandala made up of their personal visual interactions with the piece (and the performance residue of their gestural negotiations with the interface).

With the introspective nature of the text in mind, each mandala was only ever seen by its creator. Essentially, neither the artists nor each audience interactor ever got to see the others' creations. Ultimately, *Lines* epitomized an interactive experience, which could be simultaneously characterized as: individual and communal, intentional and accidental, connected and defamiliarised.

Throughout this hyperformance, the audience provided the artists with specific gestural feedback that directly addressed their reactions to the interactive structure. Through the use of these interactive strategies, the artists created a unique gestural field to be explored both communally and individually by participants in the work. Each audience member was performing gestures shaped by the performance text which were captured, in a very literal way by the mouse movements and translated into marks on the screen. Additional expressive content was added to these gestural traces by the modifications triggered by the artists, creating an expressive feedback loop between the participants in this event.

The evolution of this gestural object brings to mind descriptions of many analyses of the reception and interpretation of artistic artefacts. Eco states: "... every reception of a work of art is both an interpretation and a performance of it, because in every reception the work takes on a fresh perspective of itself" (1989, p. 95). Each additional gesture and the subsequent modification of those gestures adds complexity to the audience's understanding of the expressive nature of these gestures which are at once reflections of self and other. The end result is a non-verbal artefact of the audience's performing of their performance of the creative act.

In taking a very formalist and symbolic approach to the visual artefact (and in forcing the audience to do the same), the artists create a locus of meaning which transcends the usual reliance of non-verbal communication on the re-creation and expansion of easily understandable gesture. Historical reflections of this approach can be seen both in the visual formalism of Bauhaus performances as well as in the more

formalized symbolic gestural forms of classical dance or Kabuki theatre. The symbolism of both the verbal and non-verbal expressions work together in this piece in order to shape not the literal interpretation of the work but rather the range of interpretive meaning that the audience was able to attribute to their performance of the work.

In many ways, the creative appropriation of agency became even more important than the visual and textual artefacts in *Lines*. The constant changes in the rules of interaction effectively prevented the conscious mastery of the visual system, and the production of mimetic meaning within the visual field. Comments from the audience during these performances were tightly focused on this element of the narrative experience, alternating, in many cases, between the joy at discovering their creative agency and their frustration at having that agency removed. This was the clearest indication that both the textual and visual elements of this work became subordinate to the flow of expressive control between audience and artist.

Approaches to theatrical reception help us understand this process somewhat. Of particular interest is Blau's re-interpretation of Brecht's call for alienation within theatrical performance. As he points out, alienation itself is not sufficient to create meaningful works for an audience. Instead, the artist needs to strive to create a state of illusive immersion, even as the audience is alienated from it by the knowledge of its artificiality (Blau 1990). Along these lines, theatrical meaning emerges from the inherent tension between these two seemingly contradictory states. Wilson makes similar remarks about digital spaces as well, emphasizing the ontological overlap "between 'acting' and being 'acted upon'" in virtual environments (Wilson, 2003, p. 3). In *Lines*, the alternation between feelings of agency and control and the re-shaping of that agency beyond our ability to affect it becomes the source of much of the dramatic tension of the piece.

The manipulation of agency within the specific artistic context of this work also highlights the essentially ludic nature of this transaction. In addition, it shifts the nature of the play from being mimetic to being rooted in a struggle for control which is at once collaborative and competitive. This creates tension between what Caillois has identified as the two fundamental modes of play: *paidia* (fantastical invention) and *ludus* (rule-based, structured play) (Ryan 2009). Using the categories Marie-Laure Ryan extrapolated from this division (2009), *Lines* forces the audience to alternate between the playable story of the drawing experience and the narrative game of the struggle for control of the narrative meaning. The tension between these two modes of play and action, which replaces the more traditional forms of dramatic tension created through the arrangement of narrative elements in a story. In this observation, there are echoes of Murray's analysis of the ways that game structures are interpreted as representing symbolically meaningful dramatic structures (Murray 1997).

In the case of *Lines*, the alternation of different states of agency and control highlights the fundamental contradictions inherent in virtual environments where our desire for freedom and play are arbitrarily stopped by the restrictions of the environment. Participants are thus faced with a decision of how to approach understanding the work, yet another layer in the increasingly recursive systems of control. Ultimately, the expressivity of *Lines* emerges not from textual or visual sources, but from the manipulation of each participant's sense of agency.

5. AVATAR PERFORMANCE IN 3D VIRTUAL ENVIRONMENTS

A wide range of 3D virtual environments have also been used as sites for avatar performances. Theatrical and musical performances were accomplished even in early multi-user VRML environments, such as Digital Space Traveler. Subsequently, a large number of events have been staged in both game-oriented

virtual spaces as well as those which have a more social orientation. Many of these events actively challenge the uses we make of these spaces, commenting on the nature of digital embodiment and interaction. As these environments became more complex, gesture and other forms of non-verbal communication became much more important sources of artistic inspiration and expression.

While most 3D virtual environments focus on creating realistic gesture, there are a few key examples, especially from *Second Life* where non-mimetic gestures became key building blocks of performance events. These performances break away from traditions of avatar and environment design, which stress the reproduction of the physical world, using this shift to emphasize abstract and remediated nature of performance within these spaces. In this way, they draw on many of the theatrical traditions of gestural formalism and ritualistic approaches to performative expression.

These practices effectively break down the avatar-body schema, forcing people to address the avatars as code rather than interpreting them in conventional ways. Effectively, these works play with the inherent ambiguity in the perception of avatars remarked on by Taylor who posits that avatars are alternately perceived as code, objects, agents or even reflections of self (Schultze et al 2009), emphasizing the ephemerality of our connection with our digital representations (DiPaola, Turner & Browne, 2011). By foregrounding this complex relationship, these performances question the ontological role of the avatar and, as a consequence, the contextual limits imposed by physical reality on non-verbal expression.

Most of the key examples in this field were produced in *Second Life*, largely due to the ready availability of a programming language (LSL - Linden Scripting Language), which enables the modification and creation of original gestures. These tools are widely used by performers, for instance to create gestures for ballet or commedia dell'arte reproductions. But there are several examples of avatar performance, which take the use of code and scripting in entirely different directions.

Within some of their live *Second Life* musical performances, Avatar Orchestra Metaverse has created interesting examples of non-mimetic gestural performances. *Fragula* (2007) was a performance of a piece by Björn Eriksson, which made use of coded “instruments”²⁷ which produced not only sound but gestural effects. Each of the sounds produced by these instruments initiated a change in how the avatar was displayed, causing the performers’ avatar bodies to involuntarily roll, hover, jump, flip, prance and contort across the performance space. As mentioned earlier, “musical” online performances like *Fragula* pay explicit attention to how the avatar can also be treated as a partially embodied gestural “instrument” - that is subject to the whims of the composer. Rather than containing meaning themselves, these gestures become subordinated to the musical arrangement, which is the *raison d’être* of the performance.

Using non-mimetic gestures as figurative elements within visual and sculptural works, Alan Sondheim makes extensive use of avatars, which seem to be out of control. For example, in 2009, he created an avatar called “Julu Twine,” a formal visual experiment, which he described as a “punctum for camera and movement” (Sondheim, 2009). Julu Twine was made up of an elaborate system of inter-related avatars and objects, and was stripped of as many human associations as possible to see if it could effectively blend into its environment, acting as both figure and ground within the composition of the work. Not only does this transformation of a (digital) body into abstract figure echo the Bauhaus Theatre Workshop experiments, it also illustrates the tension between alienation and immersion proper to both the theatrical event and (as seen above) its digital equivalents by breaking down not only the meanings associated with gestural expression but also our idea of what it means to be embodied in a virtual environment.

²⁷Designed by Andreas Mueller (“Bingo Onomatopoeia” in *Second Life*)

Also, important within the framework of our discussion of agency in virtual worlds are the works of “code performance” created by Gazira Babeli. As the name suggests, these works use code rather than avatar action as the expressive form of the work. Most of these, such as *Grey Goo* (2006) involve the creation and manipulation of objects within the space. On occasion, however, the works challenge the agency of other avatars within the virtual space; such is the case with “Don’t Say” (2006), a scripted tornado, which shakes a user’s avatar until they apologize (<http://turbulence.org/blog/archives/003987.html>). That a loss of agency and punishment are equated is certainly not limited to virtual worlds (as Foucault (1995) has more than amply illustrated), but become expressed in very different ways in the coded reality of *Second Life*. As we will see in the following case study, these virtual environments allow us to enact quite literally Artaud’s call to attack the theatrical audience.

6. CASE STUDY 2: SPAWN OF THE SURREAL



Figure 12-2: A screenshot from *Spawn of the Surreal*

While it is not the only one of their performances to do so, *Spawn of the Surreal* (2007) by Second Front [Figure 12-2], an avatar performance troupe founded in 2006, provides us with a clear example of the ways in which the audience’s gestural agency can be manipulated as part of an avatar performance event. First performed as part of the Chaos Festival, the show began as Second Front invited participants to sit in the audience space of a makeshift theatre.

Initially, Second Front pretended to be ushers and announced that the performance would start momentarily after everyone had been seated. However, the seats injected animation-code into their avatars’ default poses and appearance. After enough ushered audience members had filled the pre-requisite seats, a trigger was activated by Second Front’s programmer that forced the audience members to mutate into flailing deformed avatars. Even after standing up from their seats, the audience members were unable to control their awkward Cubist gestures.

Consequently, audience members engaged in a “sort of improvised dance” (Quaranta, 2007) and clumsily ventured out of the theatre into a non-consensual co-performance with Second Front - who voluntarily shape-shifted into similar mutated avatars. Other than basic navigational control, the audience members were unable to express their usual gestural palette. In fact, the audience members could only control the full range of their gestural agency by manually logging out of *Second Life* and re-booting back in. Generally speaking, the scripted prank-seats almost acted as “interpassive objects” except that “transferral of activity [...] onto another being or object” that “consequently ‘acts’ in one’s place” (Wilson, 2003, p. 2) was non-consensual. Ultimately, Second Front’s *Spawn of the Surreal* employed deception as a performative trope in order to subvert the consensual nature of an interpassive ritual. Once the activated chairs limited the audience’s animation and posing abilities, the audience members unwittingly surrendered their avatar bodies as surrogate selves. These interpassive surrogate selves merged with an interpassive device (the scripted chair) and surrendered to an artist-controlled agency that acted in their place on their behalf.

Whether they wanted to or not, the audience members participated in this performance “via this mediating virtual [...] object” (Wilson, 2003, p. 2). Specifically with regards to the interactions in this performance, the audience members’ avatars became an alterior surrogate for their habitual virtual selves. In a non-performative situation in *Second Life*, “a feedback loop is created between user and avatar whereby part of one’s self is extended or projected onto the screen, enacting a dynamic of agency by proxy” (Wilson, 2003, p. 2). However, with this performance, the avatar/user’s typical sense of static virtual identity and “dynamic” agency was reversed. During *Spawn of the Surreal*, the audience avatars’ identities became chaotically fluid while their agency regressed into a pre-scripted stasis. The loss of their agency caused some avatar users to feel temporarily ontologically separated from a direct identification with their avatar since their avatars no longer looked like the intended proxy-versions of their projected selves.

After experiencing this frustrating agency-loss for the duration of Second Front’s performance, the audience members would have to eventually come to the realization that “...the avatar self in *Second Life* enters into a new kind of agency, one that is no longer self-referential because there is no unified, singular identity in which to refer to; rather, the avatar identity is relational to whatever object context it finds itself in[.]” (DiPaola, Turner & Browne, 2011, p. 6). Much in the same way that the Brecht theorized that alternation between alienation and immersion would expose the true nature of the theatrical event, the subversion of their agency within *Second Life* highlights Manovich’s observation about virtual worlds where “the subject is forced to oscillate between the roles of viewer and user, shifting between perceiving and acting, between following the story and actively participating in it” (Manovich, 1995, p. 7).

While the audience’s experience of agency is similar in some ways to what was created in *Lines*, there are significant differences as well based on the intangible nature of the connection between users of virtual environments and their avatars. An increasing number of intermediated theatre productions also use the deconstruction and fragmentation of embodiment as aesthetic statement, reflecting what they see as the increasing infiltration of the body by technology (Lehman, 2006). What makes avatar performances such as *Spawn of the Surreal* unique within a performative context is their ability to extend that deconstruction of the human form to the staged embodiment of the “audience.”

In many ways, this echoes some of the more controversial goals expressed by Artaud and his desire to bring the audience into the middle of the event. In addition to the trance-like gestural approach seen above, Artaud also called for the audience to become surrounded and attacked by the action, effectively becoming part of the theatrical event. This tendency was most actively embodied by the Fluxus movement and the Happenings scripted by Alan Kaprow, which emphasised the audience’s participation in the performative action itself. The addition of the audience to the gestural dance of *Spawn of the Surreal*, while involuntary, also follows in this artistic tradition. By attacking the audience through their agency - the one near universal feature of embodied virtual worlds, this work is able to create an effective emotional reaction through the manipulation of the core attributes of the mediated body. As in staged performance art, the saliency of the artwork comes directly from the subversion of the most closely-held expectations of the audience.

7. LOSS OF AGENCY AS NON-VERBAL COMMUNICATION

Whether expressed within two- and/or three-dimensional “virtual” worlds, non-verbal avatar gestures – when contextualized as artistic performance – appear to reflect the limitations of the available gestural palette within each world’s user-interface. As shown from the selected examples above, artists can either use this palette to restrict their own avatars’ agency and/or have the freedom to compose the “audience’s” gestural potential for collaborative agency. Sometimes, a prop (scripted object) is used as a proxy catalyst for gestural interaction but in most cases, it is the affordances of each proprietary user-interface that determines the range of non-verbal communication possibilities.

Unlike analog theatrical situations, the fact that each user/artist has to negotiate a navigational and gestural interface less intuitive than what is available in their own biological body schema, leads to a particular avatar agency that is much more malleable and instrumental by oneself and others. It is no surprise then, that artists within such virtual worlds can spend as much time on abstract and representational gestural design/composition as on the manipulation of visual aesthetics. The manipulation of agency itself becomes one of the most significant artistic tropes for conveying a uniquely “artistic” expressivity when relying on non-verbal gestures for aesthetic communication between avatars in a virtual world.

The emergence of techniques for direct manipulation of agency has implications beyond the specific context of avatar performance. There is already a significant body of work on the roles of interactive strategy on the emergence of narrative (Ryan, 1999; Murray, 1997) and flow (Csikszentmihalyi, 1996). But most of these studies focus on the ways that interaction and agency can increase the salience of other, more logocentric forms of expression. What these examples suggest is that, at least within the context of virtual environments, the role of agency can also be examined as an expressive strategy in its own right.

Both *Lines* and *Spawn of the Surreal* use the loss of agency as a digital *verfremdungseffekt* (alienation effect), reminding participants of the mediated nature of their experience. While this may be desirable in certain artistic contexts, the expressive range of this tool needs to be expanded and adapted for other uses. Certainly, versions of coded agency-modification to express other emotional states is easily imaginable, using the direct modulation of our (and others’) experience of virtual environments to convey the kinds of information usually transmitted by gestural communication. There is no shortage of work to be done to understand the relationships between emotional experience and agency, but there is a lot of promise inherent in the successful use of agency manipulation in shaping audience experience within these two artistic examples.

Eventually, this discussion may also be expanded to discussions of the expressivity of artificially autonomous agents in virtual environments, and the ways in which they act as “surrogate selves” with their own proxy/agency issues and interpretations. At the time of writing, there are already numerous examples of agent-driven performances by artists such as Gazira Babeli, Selavy Oh, Aland Sondheim and Moxmax Hax (Max Moschwitzer). Further, these artists are exploring the potential for mixing ontologies and gestural attributions through the implementation of hybrid performances where the boundary between a user’s avatar and agent dissolves. Therefore, by transcending the discussion of manually controlled avatars as gestural subjects, the qualitative relationship between a user and his/her virtualized representation would simultaneously become more clarified (i.e. the differences between analog and virtualized subject become gradually alienated from each other in an explicit manner) and yet ontologically ambiguous.

If avatars are increasingly seen as disembodied “puppets” by many NVC scholars (Ventrella, 2010), it would certainly make sense to delve deeper into issues of artificial agency and its implications for Non-verbal gestures as “art performances” in Virtual Worlds. Perhaps one day, researchers will also discover a method of quantitatively measuring the precise degrees to which gestural agency oscillates towards and away from the user’s virtualized intentionality. At this point, however, the discussion of agency within the context of avatar performance affords us an intriguing window into this possible expansion of our understanding of non-verbal communication beyond simple gestural mimesis and into a realm which exploits all of the expressive potential of the virtual environments into which we project ourselves.

8. BIBLIOGRAPHY

- Artaud, Antonin. (1998 [1924-1937]). *Le Théâtre et son double*. France : Folio essais.
- Blau, H. (1987). *The Eye of Prey: Subversions of the Postmodern*. Bloomington: Indiana U. P.
- Blau, H. (1990). *The Audience*. Baltimore: John Hopkins U. P.
- Braun, Edward (1979). *The Theatre of Meyerhold : Revolution on the Modern Stage*.
New York : Drama Book Specialists.
- Copeland, R. (2004). *Merce Cunningham: the modernizing of modern dance*.
New York & London: Routledge.
- Craig, E. G. (1983[1907]) “The Actor and the Übermarionette.” In J. M. Walton (ed.) *Craig on Theatre*.
London: Methuen.
- Csikszentmihalyi, M. (1990). *Flow: The Psychology of Optimal Experience*. New York: Harper and Row.
- Dibbell, J. (1993). “A Rape in Cyberspace: How an Evil Clown, a Haitian Trickster Spirit, Two Wizards, and a Cast of Dozens Turned a Database into a Society.” *The Village Voice*, December 23, 1993.
- DiPaola S, Turner J, Browne P,. (2011). “Binding the pair: making a historical case for avicentric self-expression within 3D virtual communities”, *International Journal of Web-Based Communities*, 7(2), 157-173.
- Eco, U. (1989). *The Open Work*. Cancogni, A. (trans.). Cambridge, MA: Harvard University Press.
- Grotowski, Jerzy (1971). *Vers un théâtre pauvre*. France : Éditions L’Age d’Homme.
- Kirby, M. (1995). *Happenings: An Introduction*. In Sandford, M. R. (ed.) *Happenings and Other Acts*.
London & New York: Routledge. 1-25.
- Lehman, H.-T. (2006). *Postdramatic Theatre*. New York & London: Routledge.
- Malone, P. M. (2000). “Cyber-Kleist: The Virtual Actor as Über-Marionette.” in Voigts-Virchow, Eckart (ed.), *Mediatized Drama/Dramatized Media*. Trier, Germany : Wissenschaftlicher. (57-65)
- Manovich, L. (2002). *The Language of New Media*. Cambridge,Mass: MIT Press.
- Murray, Janet. 1997. *Hamlet on the Holodeck: The Future of Narrative in Cyberspace*.
New York: The Free Press.
- Novarina, V.. (1989). *Le théâtre des paroles*. Paris : Editions P. O. I.
- Quaranta, Domenico. (2007). *Remediations: Art in Second Life*. Re-published online in HZ
Journal #11. <http://www.fylkingen.se/hz/n11/quaranta.html> - Accessed online on May 22, 2011.
Originally published in Italian. “Flash Art Italy”, issue 266, October - November 2007,
pp. 114 - 116: “Rimediazioni. L’arte in *Second Life*”.
- Ryan, M.-L. (1999). *Immersion vs. Interactivity: Virtual Reality and Literary Theory*.
SubStance, 28(2), 110-137.
- Ryan, M.-L. (2009). *From Narrative Games to Playable Stories: Toward a Poetics of Interactive Narrative*. *StoryWorlds: A Journal of Narrative Studies*, 1, pp. 43-59.
- Sacks, R. (1999). *The MetaMOOphosis: A Visit to the Kafka House – A report on the permanent installation of an interactive theatre work based on Franz Kafka’s Metamorphosis*.
In S. A. Schrum (ed.) *Theatre in Cyberspace: Issues of Teaching, Acting and Directing*.
New York: Peter Lang Publishing. 159-174.

- Scheer, E. (2002). What Does an Avatar Want? Stelarc's E-Motions. In Zylinska, J. (Ed.), *The Cyborg Experiments: The Extensions of the Body in the Media Age*. New York: Continuum International Publishing Group. 81-100.
- Schlemmer, O., Moholy-Nagy, L., Molnár, F., & Wensinger, A. S. (1967[1961]). *The Theater of the Bauhaus*. Wesleyan University Press, Middletown, Conn.
- Schultze, U. & Leahy, M. M. (2009). The Avatar-Self Relationship: Enacting Presence in *Second Life*. Presented at the Thirteenth International Conference on Information Systems, Phoenix, Arizona. ICIS 2009 Proceedings.
- Schweller, K. G.. (1999). Staging a Play in the MOO Theater. In S. A. Schrum (ed.) *Theatre in Cyberspace: Issues of Teaching, Acting and Directing*. New York: Peter Lang Publishing. 147-157.
- Second Front (2007). *Spawn of the Surreal* [Avatar Performance]. Performed as part of the Chaos Festival.
- Sondheim, A. (2009). Personal correspondence with Jeremy Turner.
- Unterman, B. (producer and director). (2009). *Lines* [Avatar Performance]. Performed as part of the 090909 UpStage Festival.
- Ventrella, J. (2011). *Virtual Body Language*. Petaluma: Eyebrian Books.
- Wilson, L. (2003). *Interactivity and Interpassivity: A Question of Agency in Digital Play*. Melbourne: RMIT Press.

SECTION IV

ANIMATING AND PUPPETEERING

13

13. EMPATHY IN VIRTUAL WORLDS: MAKING CHARACTERS BELIEVABLE WITH LABAN MOVEMENT ANALYSIS

By Leslie Bishko

Editor's Note: This chapter is the third part of a larger piece on *Empathy and Laban Movement Analysis for animated characters*. We have broken it into three smaller chapters throughout the book, but it may be read as part of a single work. The other two parts can be found in Chapters 5 and 11. These chapters are supplemented by series of video figures, which can be found online at <http://www.etc.cmu.edu/etcpres/NVCVideos/>.

1. CHARACTER MOVEMENT

Now we are prepared to discuss character movement in virtual worlds, using the movement vocabularies of the Animation Principles and Laban Movement Analysis to look for empathic effectiveness in terms of functional/expressive believability and authenticity. This section touches on movement in keyframe animation, motion capture, interactive console games and real-time puppeteering.

KEYFRAME ANIMATION

Keyframe animation, the process of designing motion pose by pose, creates empathy through the animator's creative process. The process draws on an animator's knowledge of the Animation Principles, video reference, embodiment of a character by acting out a scene, and personal sense of movement styles. When used well, keyframe animation methods enables stylistic variety, going beyond typical cartoon movement styles (Bishko, 2007).

How to Train your Dragon (2010) is a keyframed animated feature film that includes some meaningful character performances. The following Laban Movement Analysis of the scene where Toothless, a dragon, permits the lead character Hiccup to touch him outlines the elements of believability and authenticity in movement terms. This scene contains a tender, empathic exchange between the characters. It can be viewed in Chapter 6 on the DVD, beginning at 31:06.

Contrary to the character's introduction as the most fearsome and elusive dragon around, Toothless emerges as pet-like, variously incorporating movement characteristics of a cat, dog, horse, and even a bat. He combines power and speed with cuteness, playfulness and intelligence. (Asay, 2010)

As the scene opens, Hiccup is seated on a rock drawing in the sand with a stick. Hiccup's body weight is slightly Passive, with face resting in one hand and elbows braced on his knees. Mentally, Hiccup appears to be in Remote State (thinking and feeling): his gaze is Directed to his drawing with an ongoing Free flow. With Light and Direct touch (Stable state, thinking and sensing), he creates a drawing of Toothless in the sand.

With alert, upright posture, Toothless observes Hiccup drawing, following his line with a Direct/Bound gaze (Remote State, thinking and feeling). Next, he copies Hiccup. In a swirling dance, using a tree held in his mouth, he expresses himself through creating a winding, looping piece of abstract art. Overall, Toothless moves Freely and Indirectly in the space surrounding Hiccup, varying Sustained time with occasional pauses (Vision Drive: Flow, Space and Time, combining feeling with thinking and intuiting). At one moment, he stops, makes a Quick/Direct glance at Hiccup, and then punctuates his drawing with a Direct/Light poke. Toothless playfully romps around creating his art with Indirect flexibility in his body, Free Flow and moderate Strong Weight (Spell Drive, thinking, feeling and sensing). His body locomotion has a weighted quality, while the drawing action is less weighted, communicating more strongly through the elements of Vision Drive.

Hiccup is in the center of the drawing. He stands and starts to move outwards, demonstrating to Toothless that he will avoid placing his foot on the artwork. With Bound Flow arms held out at his sides, Hiccup draws on an all-around Indirect awareness of the artwork that surrounds him, looking, turning and stepping Lightly (mirroring Toothless' Spell Drive). The Bound Flow used in the upper body serves to hold him up from the ground, creating Lightness and a sense of buoyancy. In this sequence, Hiccup is mirroring the dragon's Spell drive. He stops without knowing that Toothless is standing close behind him.

Hiccup chooses this moment of close proximity to extend his hand towards Toothless. The dragon recoils, slightly Retreating with a sideways turn of the head. With an exhalation, Hiccup closes his eyes, drops his head and shoulders, and turns his head away from the dragon, sensing into himself with slight Passive Weight and Shape Flow mode of shape change. By letting go of Space and eye contact, Hiccup shows the dragon his trust. He makes himself vulnerable. Then he extends his hand towards the dragon again. This gesture is Free and Light. It is Directed towards the dragon spatially, but is more about feeling and sensing. His hand maintains Free/Light, Advancing towards the dragon, while his body is held in a retreated stance, anticipating with Bound flow.

After a pause, the dragon advances his face towards Hiccup's hand, and they connect. This is a powerful moment. There is a weighted quality to their touch: the moment of impact is Sudden and Light, with continued, gentle firmness (diminished Strong) as they maintain contact. The impact sends a shudder through Hiccup's body. After a few moments, the dragon retreats. They open their eyes and make eye contact. Toothless shakes his head and snorts, in a horse-like manner, and then darts away.

This scene is an adrenaline moment (Hooks, 2003) in the film, and an emotional turning point in the plot as trust emerges between Hiccup and Toothless. These two characters share the language of movement as their only method of communication with each other. A series of reciprocal actions build the development of their trusting relationship. In this scene, the mirroring of action is a continuation of the previous scene in which the dragon shares a fish with Hiccup, then returns Hiccup's smile. The dragon communicates his empathic connection to Hiccup by imitating him.

The animation takes the connection between Hiccup and Toothless a step further by repeating the movement qualities within the scene. Hiccup initiates the scene in Remote and Stable states, which evolve into a full Spell drive in the dragon's motion. Then Hiccup responds to the dragon with Spell drive. Their Efforts acknowledge each other, which serves to build trust between Hiccup and Toothless. The trust is communicated through touch.

This scene is a testament to the aesthetic integration of movement design with storytelling. It stands out. Audiences empathize with the heightened emotions of the scene as the result of sensitive and effective animation of movement qualities. The use of LMA to articulate the strengths of this scene demonstrates that LMA can be used creatively to develop authentic character performances that we engage with empathically.

MOTION CAPTURE

Motion capture has become highly integrated and blended within animation production in feature films and games. As a result, production methods are becoming more fluent, and the resulting movement is less distinguishable from other methods. Feature films like the *Lord of the Rings* trilogy (2001-2003) sparked debates about whether the character performance of Gollum could be attributed to Andy Serkis, the actor who performed the role of Gollum in a motion-capture suit, or to the animators who manipulated over eight hundred animation controls to accentuate and refine Serkis' performance. Like rotoscoping, its predecessor, motion-captured movement delivers a high level of realism, yet it is not always successful at engaging viewers empathically. Mention of the "uncanny valley" usually comes up in discussions of whether motion capture is successful at empathic believability (Dambrot, 2011).

Ever since rotoscoping was used for Disney features such as *Snow White and the Seven Dwarfs* (1937), animators have been exaggerating rotoscoped motion to make it more believable. It is widely known that rotoscoped characters do not appear to have believable qualities of weight. Animators who work with motion capture are tasked with applying animation principles to exaggerate the recorded motion, to create more believable weight. In this sense, both rotoscoped and motion-captured movement are analogous to the use of sound recording. Recorded sound in film and music is aesthetically integrated with the sonic landscape through processing, mixing and editing (Furniss, 2000).

Whether it's for feature films or interactive games, the choice to use motion capture is often based on the desire for realistic action, and nuanced, actor-driven performances. It is a stylistic choice for movement qualities distinctly different from the styles of motion created through keyframe animation.

Recent developments in motion capture have raised the bar in capture of the face. The film *Avatar* (2009) made use of a fully immersive production environment in which actors' performances were captured in context with virtual props and sets. The motion capture methods included facial capture through small cameras mounted inches in front of the actors' faces, enabling capture of facial details in context with full body capture. *LA Noire* (2011) employed a technology for facial capture called MotionScan¹, in which the capture data creates animated geometry and textures, bringing a new level of realism to facial animation in games. However, body and facial motion were captured separately, which shows. There is a gap between the realism of the facial animation and the level of detail in the body. The heads appear to be lively and expressive puppets attached to stiff bodies at the collar line.

With both of these examples, as I watch, I find myself very aware that actors' performances are driving what I see. This carries a strong element of intentionality in the movement. I also find that the heightened level of detail gives the characters a certain quality of presence, which is engaging on the level of authenticity. Like watching *Madagascar*, it takes time for me to empathically attune to, and eventually believe in the movement. The fluidity and integration present in *Avatar* cultivates my empathic experience, whereas the disconnectedness in *L.A. Noire* feels fragmented; I engage with the characters from the neck up. *Avatar's* immersive, virtual production methods are also successful at believably integrating the characters in the environment.

¹<http://depthanalysis.com/motionscan/>

LMA contributes concepts that can influence choreography for motion capture and how motion-capture talent is directed during a shoot. In this sense, LMA can be applied to the motion capture process much the same way as it is used in dance or theater. In the Empathic Toolset section, I provide a discussion of how LMA concepts can be layered and blended with motion capture and other animation methods.

INTERACTIVE CONSOLE GAMES

Compared to rendered animation and visual effects, real-time virtual environments have significantly fewer ways to cheat believability. The challenges of creating authentic, functionally/expressively believable movement include managing how the player drives movement through gameplay input. Jungbluth believes the player is another layer of the process:

The buttons and commands the player inputs are the moments when they get to directly speak to the game. That is the moment they are communicating with characters in the game, both those they are controlling and those they are interacting with.

Often times, it is the designers or programmers thinking actively about what those input commands will be, but when you start to think with an animator's knack of performance and personality, there is a world of possibilities that we can use to help push the emotional feel of that conversation. (Jungbluth, 2011)

The discussion below focuses on character intention and movement issues in games. Many of these game developers' issues have already received significant attention in production and character animation research. Therefore, my goal in this section is to describe these issues using LMA terminology, applying LMA as a potential solution. This section begins to outline the effectiveness of LMA in creating greater character authenticity, engagement, and empathic user experience.

PROCEDURAL POSE MODIFICATION AND EMPATHY AT UBISOFT

By Jay White

The following is a summary and discussion of procedural techniques employed by Ubisoft on Assassin's Creed III, presented for Vancouver ACM SIGGRAPH by Simon Clavet on October 17, 2012.

Simon Clavet, Animation and Physics programmer at Ubisoft, believes that character animation in virtual worlds is about to undergo a paradigm shift. Just as we saw a jump between 2D and 3D sprite animation (Super Mario Brothers to Doom), and another shift to animated 3D meshes, he believes that we are on the cusp of a new type of virtual world where characters are not hand-animated, but move according to programmed subroutines. Characters will locomote, emote, and respond to their environment predominantly through procedural animation, with some manual animation or motion capture layered on top for style.

In *Assassin's Creed III*, Ubisoft has integrated procedural movement into most of the characters' actions, including pelvis adjustment if the character is carrying additional weight; acceleration banking where the body leans inwards, lowering temporarily to push off for a change of direction; foot IK, including real-time raytracing from the feet to detect uneven ground planes and decide when to plant a foot, and heuristic methods to decide which foot should be planted and locked, and convincing procedural blends between forward movement and strafing. Clavet also demonstrated games where bipedal and quadrupedal characters were moving and acting solely through procedural movement.

(continued on next page)

Game production practices can hamper efforts towards creating believable movement. The challenges in production are rigorous and sophisticated, requiring production tasks to be divided among teams of creative and technical personnel. The outcome is that through combinations of techniques and many iterations of development, multiple people create the continuous motion of any given character. It is a challenge to achieve fluid movement when the flow of action and the spontaneity of intent are fragmented in the process of production.

The most important way that LMA can contribute to integrating a production is through its use as a common language of movement. This enables a team to define clear character movement signatures and ensure that all areas of production contribute to making believable characters.

The following Fable III gameplay example shows problems that are typical in game animation [Video Figure 22]: jerky transitions, heavy breathing during idle sequences, and sudden changes of orientation and foot sliding. The large character that enters the scene has a lack of torso movement, and does not move as we would expect. His gestures have small range but his voice is big, creating a mismatch in characterization.

In LMA terms, some commonly known functional/expressive movement problems in game animation are:

Lack of intent/phrasing. Phrases signal a character's intent and form the linguistic units of movement. Due to the expectation of immediate responsiveness to game controls, the Preparation of Phrases is either minimal or non-existent in many games. As a result, characters do not signal intent well. Ubisoft has begun to address this in *Assassin's Creed III* by having characters turn their head in response to player input. (See Jay White's Sidebar on Procedural Pose Modification and Empathy at Ubisoft)

Ubisoft understands the need for an anticipatory action to lead a movement phrase, but they also feel that the first priority is to create a responsive, tactile, physical connection between the player and the virtual avatar. They are solving this problem by having the avatar's head immediately move in response to game controller manipulation, which anticipates the motion of the body. Players are okay with a delay in the body movement if they see this signal that the avatar is thinking about the movement that will occur.

Clavet posits two ways that procedural animation can help make humans feel empathy for virtual characters during real-time interactions in virtual worlds:

The first is that it is possible to catalogue the physical movements that are associated with an emotion, and to integrate these movements with keyframed animation to help promote empathy. For example, Clavet works with the assumption that if a person is disgusted by something, they will move away from it. So a procedure that rotates a character's neck backwards, distancing it from an object, will make viewers feel that the object disgusts the character.

One can imagine how this neck movement could be one point on a continuum, where the spine would rotate backwards at increasing levels of disgust, and eventually the character could take a step back, then even turn and run, all procedurally. One might also imagine that a whole catalogue of these procedural emotion-continua could be created, called upon, and layered in combination with each other. This could result in a near-infinite range of externalizations that would simulate a being that perceives and responds to its environment.

(continued on next page)

Another intent problem is motion that seems to happen without reason, or seems out of context. Examples in this scene from Grand Theft Auto 4 include running over furniture, and breaking into a run in a confined space. [Video Figure 23]

It is physically impossible to instantly run at full speed.

Lack of acceleration/deceleration is obvious when characters start to run and change orientation instantaneously. This is a glaring problem because in many games, running is the player character's main action. It is partly a phrasing problem, where the preparation to run is missing. However, characters need to vary their running speed, specifically when starting and stopping, in order to appear natural (Johnston & Thomas, 1981). Changes in orientation must also be gradual. Jungbluth explains that these issues are also a production pipeline problem:

Pathing is something most animators never pay much attention to, as its implementation happens by any number of designers or programmers throughout various levels and can change at the drop of a hat.
(Jungbluth, 2012)

Acceleration/deceleration is also missing in first person camera animation, noticeable in this video of Halo 4 gameplay. [Video Figure 24] First person camera represents head turning combined with eye movement, and includes the brain's visual perception. Even the slightest amount of acceleration and deceleration would improve our empathic engagement in this first person gameplay.

Lack of coordination. Patterns of Total Body Coordination, transitions, and connectivity are aspects of coordination.

- **Patterns of Total Body Coordination:** The animation process requires disassembly of the elements of movement that the body and mind have already integrated. For example, in daily life we walk easily, yet it is a complex action that evolves through the participation of all six Patterns of Total Body Coordination. Because our brains have been wired to integrate the patterns, to take them apart for the animation process puts you into the pre-motor cortex of the brain (Hackney, 1998), as if you are forgetting how to move properly. Because animators aren't informed about the Patterns of Total Body Coordination, they sometimes create uncoordinated movement, even though they themselves are coordinated in their own bodies.

Secondly, Clavet claims that a purely procedural struggle between multiple objectives, if it follows the laws of physics, gives a character the impression of undergoing an internal struggle that creates believable intention. For example, the character Esther might be programmed to try to stand upright and to look at a ball at the same time. As the ball moves behind her back, Esther reaches a point where she wants to watch the ball but will lose her balance if she continues to do so. Esther appears to make a decision - she stops watching the ball temporarily, twists around to the other side, then continues following the ball.

Clavet presented (unpublished) video clips illustrating these two methods were fairly convincing for short scenes and controlled circumstances. But I wonder how convincing the character would be when a new thought and/or a new external stimulus appeared. How does the procedure decide which external factors to respond to? How would it decide between competing internal emotions?

It seems to me that empathy would break down after watching a purely procedural character for any extended period of time, because we would start to notice patterns that do not reflect the ways that individual subjective experience affects the immense diversity of human actions.

- **Transitions** between motion clips are not always matched to the movement. A broad palette of transition animations is needed. The transition clips not only have to match but they must blend well. Transitions also need some degree of preparation or follow through—they form the phrases of movement.
 - Algorithms would need to detect whether a motion clip is at the start or end of an action to blend with the right transition. For example, in the opening of this sequence from *Fable III* [Video Figure 25], a character takes a bow. He starts with a wide stance, and then takes two steps to place his feet into position for the bow. After the bow, the same two steps are reversed, returning him to the original stance. The motion has excessive weight shifting, and the reversed steps are obvious. It can be done properly by using different animations for the weight shift preceding and following the bow.
 - Motion clip blending needs to occur on a per-channel basis in order to customize the motion of each body part. Interpolation of whole-body states limits connectivity.
- **Connectivity** relates to Patterns of Total Body Connectivity, the connectedness of the body and the coordination of connectedness. There is a lot of animation (in games or film) that does not involve the core, yet connectivity is all about the core. Typically, we see lots of arm waving and head-turning without attention to spinal movement and the graded rotations of the individual vertebrae. Because the Patterns of Total Body Connectivity evolve towards developing full movement potential in relationship to gravity, connected motion supports the illusion of gravity.

Weight errors arise from a combination of factors, such as weight shifting, sliding feet, connectivity in transitions, phrasing issues, and timing for interaction as opposed to character intent.

- Game animation is full of locomotion, which is all about **weight shifts**. Effective weight shifts hinge on the subtle interplay of pelvis and foot positions, timing of feet in relation to pelvis, and rate of acceleration/deceleration. *The Fable III* example shows weight shift transitions in over-simplified directions: forwards/backwards, side/side and along diagonal pathways, which gives the movement a mechanical appearance. Because weight shifts occur frequently in transitions, motion trees of transitional animations need to accommodate many weight shift possibilities, supported by the game engine's motion clip sequencing and blending.
- Sliding feet in game animation is a well-known error. Lack of groundedness through the feet seems to break the most important visual cue of physicality: that the character exists in a gravitational relationship to the environment. Ubisoft is currently working with procedural methods to solve this problem (see Procedural Pose Modification and Empathy at Ubisoft, below).
 - Character contact presents challenges for film animation, and the same challenges are amplified for games. LMA concepts to support believable character contact are:
 - Weight Effort, varying the degree of pressure in the quality of touch from light to firm, as well as the full spectrum of Effort possibilities;
 - Shape Qualities, describing the way characters mold and accommodate their body shape to another, supporting the relation of self to other;
 - Mode of Shape Change, describing how characters change their body shape over time.

Idle sequences are made of cycles, which are always obvious. Cycles are an efficient solution for idle sequences because idling is meant to be a background action. However, like the ticking of a clock, cycles

of motion in idles tend to attract our attention. Common problems include heavy breathing, overactive weight shifting, and fidgety actions such as adjusting shoulders, muscle flexing, neck rolls, foot tapping and looking around. We experience a character idling as behavioral motion: nervous, fidgety and overacted. Ways to improve idles include creating subtler motion, less frequent actions, and assigning small gestures to environmental triggers. For example, characters can be triggered to turn towards or away from each other, or turn their heads to follow other action in the scene.

Improvements to **facial animation** are emerging; through motion capture techniques such as the MotionScan technology used in *LA Noir*. However, procedural layering of contextual variation may not be possible with a technique such as this, which bakes motion data into the geometry. Ubisoft's practice of initiating action with a head turn can be extended to the eyes as well, because the eyes always look first, before the head turns. The gaze direction, head, and neck turn, if they are broad enough, can be supported by spinal rotations, sequencing in a chain of action along the spine.

Exaggeration issues include over exaggeration, or oversimplified action such as limited use of body parts, lack of weight shift or gesture during dialogue. Physics based animation for secondary motion is sometimes over exaggerated and stands out too much. Stylized costumes that don't move properly with character body motion also limit believability. Secondary animation should not attract attention away from the main action. Our eye is drawn to these errors because the fact that they are errors makes them stand out. Existing solutions for many of these problems take a layered approach, blending keyframed or captured motion with procedural animation. It is possible to create authentic, believable movement that we experience empathically by modeling LMA parameter spaces within game engines.

REAL TIME PUPPETEERING

This section describes a personal account of my own visit to a real time virtual world: *Second Life*. Movement issues in *Second Life* are similar to those of real-time games, with the exception that all elements in-world are active in real time. This means that everything that moves is being driven by a person on the other end of a controller, so movement problems are applied to everything.

I ventured into *Second Life* to meet some experienced participants who could demonstrate custom animations. I interviewed them about movement, animation, identity and self-representation. My avatar used the default movement palette and stood out in sharp contrast to the other avatars I interacted with and observed.

In our chat, my colleagues discussed how they became good friends despite the limited non-verbal communication in *Second Life*. What they are lacking, empathically, in physical motion, they gain through language and dialogue. One person described how, upon meeting people in real life whom she knows well in *Second Life*, she finds something in their movement is recognizable. I attribute this to choosing custom animations for an avatar because you have an empathic attunement with animations that reflect you stylistically. We also discussed the need for better movement interfaces. It becomes clear to me that intuitive interfaces and self-designed animations could lead *Second Life* participants to a deeply immersive experience. I am curious about what identity, relationships and "reality" could be like when the process of movement communication is both intuited empathically and consciously puppeteered.

2. AN EMPATHIC TOOLSET

LMA, as a theory of movement, has much to contribute towards creating authentic, intentional, believable, expressive and characterized movement in virtual worlds that is perceived as meaningful through a process of empathic attunement. This section summarizes how LMA informs creative development, solves

movement problems, and supports the creation of character movement we can connect with empathically. I believe that LMA can be applied successfully in two ways. As a conceptual toolset, LMA helps production teams communicate about movement on multiple levels, and work towards creating empathic characters. This process can be extended to embedding LMA concepts into software tools.

INTERACTIVE CONSOLE GAMES

Since the late 1970's, groundbreaking researchers such as Dr. Tom Calvert², Dr. Norman Badler³, and Dr. Don Herbison-Evans⁴ have been seeking computational solutions for Laban's theories. (Calvert, 2007; Chi, Costa, Zhao, & Badler, 2000) We can learn valuable lessons from their research by looking at the methodologies, successes and gaps in their work.

Laban was able to grasp the elusiveness of movement and create a theoretical framework about it. The framework itself has an accessible structure, which is an appealing feature to those who seek computational solutions to movement problems. Yet, this is precisely where LMA runs the risk of being misinterpreted. People naturally gravitate towards the areas of the system where mappings seem apparent, specifically Effort affinities with Shape and Space. Misinterpretation of Laban's material occurs in several ways:

- interpreting the system as quantifiable;
- reading about LMA through indirect sources;
- partial knowledge of the system;
- lack of consultation with certified Laban Movement Analysts;
- limited access to LMA training.

My personal experience of the system is that when it is applied contextually, subjectively, metaphorically and poetically, it offers relevant and meaningful insight. I have demonstrated this approach through the descriptive examples included in this chapter. I believe that computational applications can be successful through methods that begin with, and stay rooted in, the qualitative aspects of the system.

Within Laban studies, there is a firm belief in experiential, body-oriented pedagogy. Ultimately, you must experience and embody this method to grasp its nuances and applications fully. That is an important part of the methodology itself.

There is a large gap in the literature since Laban's original publications. LMA is practiced broadly across somatic fields, and in recent years, more practitioners are publishing scholarly research. However, there are few definitive reference texts on the system as it is currently taught in the certification programs.

Another obstacle to the awareness of LMA is the simplified way it has been taught in theater. Laban gave descriptive names to the eight Effort configurations that constitute the Action Drive, which has made it one of the most accessible parts of the system. Action Drive is standard training in theater, but typically all that theater students are exposed to. People are generally unaware that there is more to the system.

²<http://www.cs.sfu.ca/~tom/>

³<http://www.cis.upenn.edu/~badler/>

⁴<http://donhe.topcities.com/>

TOOLS FOR CREATIVE DEVELOPMENT

LMA excels as a creative tool in dance and theater. Choreographers find creative inspiration through LMA, and performers embody elements of the system to become articulate and versatile movers. For characters in virtual worlds, we can take a similar approach through the following practices:

- LMA provides a robust, descriptive language of movement that links function with expression. In production, this language can serve as an immensely practical tool to build a common understanding of movement among production teams. It can help clarify and unite separate areas of production towards a common goal.
- LMA can be applied as a creative tool for game design that ripples out to all creative areas of production. For example: imagine a game about two warring cultures. Based on their history and values, each culture emerges with their own generalized movement signature. Game strategy and story can be developed based on how each culture approaches combat through their particular movement patterns.
- Gameplay for motion-based interfaces can take this concept further, embedding procedural analysis of player motion into game content. LMA can influence mappings between player motion and in-game motion elements.
- LMA can function within the narrative structures of plotline and circumstances to support the development of complex and intriguing characters. The elements of Character Dynamics can help to build believable characters that move in a way that is congruent with their personal profile and story context.
- A movement signature for characters develops empathy. Motion libraries can be associated with specific characters, according to their movement signature, which can be modified according to gameplay. As characters overcome obstacles and move through the phases of a game, their movement signatures can evolve to reflect how they manage the next set of goals. For example, if a player character's status begins as low and progresses to high, their Body Attitude may adjust accordingly. Their use of space may range from near to full reach space.

TOOLS FOR CORRECTING FUNCTIONAL/EXPRESSIVE MOVEMENT PROBLEMS

The section on interactive console games surveys a range of functional/expressive movement problems in virtual worlds. Tools for addressing these problems are both conceptual and software-based:

- Phrasing concepts can address functional/expressive movement problems that arise from user input response (how movement initiates), and transitions between motion clips. Use of proper phrasing and transitions will clean up functional movement problems that limit the expression of intent.
- Modeling the Patterns of Total Body Connectivity within motion trees is essential towards building coordinated movement phrases, and the illusion of neurological processes in characters. The principles of connectivity can help determine the proper sequencing of motion clips. These patterns are the root from which other aspects of the LMA system can be articulated and refined in character animation.
- LMA contributes refined articulation of Weight qualities, and of movement elements that contribute to creating the illusion of weight in animation. These concepts can be applied to weight shifts, locomotion, and character contact through keyframe, motion capture, physics-based and procedural methods for addressing weight.

- As a theory of movement, LMA is a robust and detailed delineation of function and expression that animators can explore creatively as designers of character motion. It can clarify characterization and stylistic choices for the animator working individually, and within a team.

TOOLS FOR PROCEDURAL ANIMATION

Computational models of LMA parameters run the risk of creating fixed mappings on several levels. The computational solution must successfully represent an LMA concept under broad contexts. For example, a model of Indirect Space Effort must be transferable from one character's motion to another, and include controls for contextual modification. Indirectness could be swatting away mosquitos or sword-fighting 3 opponents.

If modeling a parameter is successful, the rules for applying it must be variable and contextual. For example, a character navigating a cave may use Indirectness to maintain a full, all-encompassing awareness of the dark surroundings. He may continuously wave a flashlight about to survey the space. If every character that enters the cave moves in the same way, the mapping becomes apparent, and empathy is broken. However, another character with a different movement signature may move through the space with Directness. Movement signatures contribute to the rule space for how each character interacts with the environment.

Computational models of LMA parameters will not generate all character motion. Rather, they can layer and blend with keyframed, motion captured, physics-based and procedurally generated motion to cultivate nuanced qualities that create empathic engagement. The integration of user input with LMA rule spaces may offer rich potential for personalizing the empathic experience even further, especially in the case of fully immersive virtual worlds.

3. CONCLUSION

Empathy is inherent to the gestalt experience of all virtual world elements: story, environment, circumstances, and character movement. Given the context in which we experience characters, we become accustomed to limitations in movement because perceptually we fill in the blanks (Sheldon, 2004, pp. 114-115, 118-119). Our empathic experience of character movement that lacks intent speaks to us loudly, because we know it in our bodies. How we make characters move in virtual worlds renders our empathic experience of intent and authenticity. LMA provides a theoretical framework that is rooted in the body, and articulates the nuances of our empathic experience in virtual worlds. Through LMA, we understand movement both generatively and experientially, and can tease apart the elements that fuse together in the empathic experience. It holds up as a framework for creating character movement that reflects intent, feels authentic, and is therefore believable.

Some of the barriers to creating intentional movement in real time virtual environments are already eroding. Many functional problems are being solved through procedural solutions. LMA concepts can help us push procedural solutions further, by using it to examine the blurring line between intentional and behavioral movement. Future research in this area can explore things such as procedural phrasing, and the construction of parameter spaces that hold up under variable contexts.

LMA is effective because it bridges our embodied knowledge with the ability to create intentional, authentic, believable movement. It is a system, yet it evades being formulaic. It holds up under contextual variation, and subjective interpretations. It works in a highly relational format, not with absolutes.

These are strengths of the system that are rooted in human experience. LMA articulates universal truths about movement. It helps us pin down and comprehend the elusiveness of movement. LMA offers a groundbreaking solution to the complex problems of how movement communicates in virtual worlds.

4. BIBLIOGRAPHY

- Asay, P. (2010, March 25). *Chris Sanders' Very Own Personal Dragon*. Retrieved December 13, 2012, from plugged'n: <http://www.pluggedin.ca/upfront/2010/chrissandersveryownpersonaldragon.aspx>
- Bishko, L. (2007, December 9). *The Uses and Abuses of Cartoon Style in Animation*. Retrieved February 3, 2011, from Animation Studies Online Journal: <http://journal.animationstudies.org/leslie-bishko-the-uses-and-abuses-of-cartoon-style-in-animation/>
- Calvert, T. (2007). Animating Dance. *GI '07 Proceedings of Graphics Interface 2007* (pp. 1-2). New York: ACM.
- Chi, D., Costa, M., Zhao, L., & Badler, N. (2000). The EMOTE Model for Effort and Shape. *Proceedings of the 27th annual conference on Computer graphics and interactive techniques* (pp. 173-182). New York: ACM Press/Addison-Wesley Publishing Co.
- Dambrot, S. M. (2011, December 29). *Social robotics: Beyond the uncanny valley*. Retrieved January 12, 2013, from Phys Org: <http://phys.org/news/2011-12-social-robotics-uncanny-valley.html>
- Furniss, M. (2000, April 1). *Motion Capture: An Overview*. Retrieved February 22, 2011, from Animation Journal: <http://www.animationjournal.com/abstracts/essays/mocap.html>
- Hackney, P. (1998). *Making Connections: total body integration through Bartenieff Fundamentals*. Amsterdam: Gordon and Breach.
- Hooks, E. (2003). *Acting for Animators: A Complete Guide to Performance Animation* (Revised Edition ed.). Heinemann Drama.
- Johnston, O., & Thomas, F. (1981). *Disney Animation; The Illusion of Life* (1st Edition ed.). New York: Abbeville.
- Jungbluth, M. (2011, April 10). *Growing Game Animation - Transitions & Player Input*. Retrieved December 15, 2012, from #AltDevBlogADay: <http://www.altdevblogaday.com/2011/04/10/growing-game-animation-transitions-player-input/>
- Jungbluth, M. (2012, September 10). *Growing Game Animations: State Changes and Pathing*. Retrieved December 9, 2012, from #AltDevBlogADay: <http://www.altdevblogaday.com/2012/01/25/growing-game-animation-state-changes-and-pathing/>
- Sheldon, L. (2004). *Character Development and Storytelling for Games*. Boston: Thompson Course Technology.

14

AVATAR PUPPETEERING: DIRECT MANIPULATION OF AVATAR JOINTS FOR SPONTANEOUS BODY LANGUAGE

By Jeffrey Ventrella

Avatars in virtual worlds provide affordances for self-expression and some degree of animated body language. But in most commercial virtual worlds, avatars are ill-suited for spontaneous movements intended to communicate subtle reactions and attitudes that are coverbal with text or voice chat. Compared to the variety and degree to which we can manipulate items on a standard desktop, or build expressive 3D structures in *Second Life*, the avatar trails behind. It should be the most manipulable and expressive object in a virtual world. This chapter argues in favor of interfaces that use physical simulation (a variant of ragdoll physics) as a substrate for a puppeteering system that permits the avatar to become highly reactive and manipulable, not only with a user's touch, but with other factors in the environment as well...including other avatars. A system developed by the author for the *Second Life* avatar, which *almost* became a permanent commercial feature, is described, with some technical details. It is contrasted with newer interfaces using the Kinect for full-body gestural inputs. Analyses of the advantages and disadvantages of these different modes of input are explored.

1. INTRODUCTION

While talking on the phone or texting with a friend, it is impossible to give your friend visual signals that indicate understanding, affirmation, confusion, or levels of attention. These indicators are typically provided by head motions, facial expressions, hand movements, and posturing. In natural face-to-face interaction, these signals happen in real time, and they are coverbal; that is, they are often tightly-synchronized with the words being exchanged.

You may have had the following experience: you are exchanging texts in an online chat with a friend. There is a long period of no response after you send a text. Did you annoy your friend? Maybe your friend has gone to the bathroom? Is your friend still thinking about what you said? One problem that ensues is cross-dialog: during the silent period, you may change the subject by issuing a new text, but unknowingly, your friend had been writing some text as a response to your last text on the previous topic. You get that text, and – relieved that you *didn't* annoy your friend – you quickly switch to the previous topic. Meanwhile, your friend has just begun to respond to your text on the *new* topic. The conversation *bifurcates* – simply due to a lack of nonverbal signaling. Some visual language would help.

New communication media are emerging which have the potential to provide some of the body language that is missing from our remote communications. In this chapter, I will look at one communication medium in particular in which virtual embodied communication is starting to get a good workout: avatars in 3D

virtual worlds. I am not implying that text users are going to start using avatars as soon as the technology is ready. Text abstraction and asynchrony can be a good thing when one wants to stay loosely-tethered to a conversation. But, considering that so many avatar-based virtual worlds use text chat, we have a good example for a media enhancement to text chat that offers virtual embodiment.

While virtual worlds show great promise as environments for self-expression, nonverbal signaling in avatars is currently quite poor. Most popular virtual worlds do provide avatar animations that can be played instantly, but they are usually pre-animated by an unknown artist, and they are generally contrived, with predictable expressions. They certainly serve a purpose...but if virtual worlds are to hold promise for giving people expressive virtual bodies to reconstitute natural language – to offer the tools for language generation that allows users to engage the semiotic process – some more sophisticated puppeteering interfaces need to be provided. I believe that the technology to achieve this is within reach, but the motivation has been missing.

To put this in context, consider the following two interaction scenarios:

1. Consider a standard desktop interface, which includes visual icons, files, folders, and all the familiar windows and dialog boxes that most of us are accustomed to manipulating – so accustomed in fact that we are usually not conscious of what we are doing. We move, resize, drag-and-drop, and layer windows; we click icons, we twiddle sliders. It may be a stretch to call this “nonverbal language” – it is more like manipulating a workspace. The point here is the level and the variety of manipulation that the user can do in real-time.
2. Now consider an avatar in *Second Life*, which plays a walk cycle when the user moves through the virtual world; or a fly animation when the user is flying, and so on. The avatar can play one of several predetermined animations. With some effort, and less-than predictable results, the avatar can be made to set its gaze to something in the world. Scripts can be written to cause the avatar to play animations in response to certain triggers, but these are limiting and not well-suited for spontaneous language genesis.

The point: avatars are not very manipulable or expressive in real-time – especially compared to our real bodies, and maybe even when compared to standard desktop interfaces. Our brains control several hundred cognitive puppet strings when we communicate with each other – mostly on an unconscious level, and sometimes consciously. How can we reconstitute some degree of natural brain-to-body puppeteering using keyboards, mice, Wii controllers and Kinect sensors?

In this chapter I will only consider inputs available through mice, keyboards, and touchscreens. Why not consider Kinect? Two reasons: (1) The recent excitement about full-body gesturing inputs, as enabled by the Kinect, is fueled by a high degree of early-adopter energy and commercial hype. When the hype settles, we will be able to take a good look at the potential uses – of which there will be many. I prefer to wait until the hype is out of the way. (2) Many dedicated *Second Life* and *World of Warcraft* users spend long hours in-world. If they had to use a Kinect to make all of their avatar gestures over the span of several hours a day, they would likely run out of energy. Alternatively, they could stick with it and become totally buff. While there's nothing wrong with getting a lot of exercise, it seems unlikely that a user will want to use his or her own arm and head motions to control an avatar for very long.

For puppeteering full-body gestures in avatars over extended time, some kind of abbreviated input might be more appropriate (such as bending your index finger to make the avatar nod, or “walking: your index and middle fingers on the table to make your avatar do a dance move or to adjust proxemics). If that is not convincing enough, consider the fact that many people are ill or physically disabled, and so they would

have no choice: they would need some alternative to literal, real-time, fully-body gesturing. The notion of using “virtual body language” as a way to reconstitute natural body language, using some kind of *encoding*, is covered extensively in (Ventrella, 2011).

Several recent systems have been developed that use combinations of desktop-range motion capture with standard desktop input devices, including explorations in digital puppetry by Mapes et. al (2001). The need for pragmatic technical approaches to enhancing presence and communication in virtual settings is noted by Allmendinger (2010), for instructional purposes.

NOD YES IF YOU AGREE, NOD SIDEWAYS IF YOU’RE NOT SURE

Imagine being in a virtual world and seeing your avatar approximately from behind, in the standard third-person view, hanging out with other avatars. Imagine that you had the ability to mouse-click on your avatar’s head while you were chatting with your friend’s avatar. Then, using the mouse, you could make a single “yes” nod (rotating the head up just a bit and then right back down). Imagine now that you might want to nod “yes”, rotating up...only this time much more slowly than before, to indicate that you are *starting* to agree, but that you are not quite there yet. Then, as soon as your friend utters a concluding word, you pop a quick affirming nod back down. Imagine being able to cock your head to the side to show confusion or to make a coy gesture – seconds after you friend made a compliment. If he made another compliment, only this time with lewd overtones, you might choose to look to the side just a bit, giving off a subtle but important signal, so as not to encourage him anymore. These are small motions indeed, but they can have strong impact when considering that they are a part of our natural, multimodal social interactions – where *timing matters*. Of the many classes of nonverbal signaling, I am referring here mostly to spontaneous reactive gestures, not necessarily *emblems* (which act as visual words, and which may be just as easily “played” using a canned animation and a semantics-oriented interface).

Just this one example I have given (being able to click on your avatar’s head and affect any kind of head rotation you want – with any rhythm you want) has enough expressive potential to justify a system I call “avatar puppeteering” (Ventrella, 2008). Such a system was prototyped as a feature for the *Second Life* avatar (involving all avatar joints, not just the head). This chapter covers the subject of expressive avatar puppeteering, with insights from my experience of having implemented a user-interaction scheme to such a degree that it nearly became a part of the commercial product. While this feature showed promise, there were also many aspects that were problematic – both technical and conceptual. And so it makes a good case study for designers of nonverbal communication interfaces. Mistakes and dashed dreams make great background material for design research (especially when it also shows lots of potential and there is lots of motivation to get it right).

FACING THE CHALLENGES

The proposed system only covers the nonverbal channel of direct-manipulation for gross-body motion. I am not discussing ways to puppeteer facial expression, for instance. This is obviously an important aspect of body language, but the reason I am not addressing it is subtle: a group of avatars in a contemporary virtual world in many cases will be viewed from a third-person point of view, where faces occupy a small fraction of the field of view – *even though they may be the center of attention*. With a group of two or more expressing avatars, faces generally attract saccades – the foveal focus of the eye, whereas body posturing and nonsymbolic motions are most easily picked up in the periphery of the human visual system – often unconsciously, yet with strong effect. This is the nonverbal domain that I am most interested in addressing – and it is also one of the most neglected. It is like the musical soundtrack in a film that drives the mood and narrative flow, and it is often forgotten when we recall the way an actor delivered a line.

Another point: it is questionable as to whether direct-manipulation interaction maps very well to the subtlety and complexity of faces, whereas it is easier to imagine such a mapping to raising your hand, shrugging your shoulders, or tapping your foot. If we were talking about virtual worlds in which avatar faces were always shown up-close and full-screen, then it would be more appropriate to talk about *facial puppeteering*.

BACKGROUND

As stated above, research in full-body gestural interfaces using motion-capture is currently quite active. Solutions that are intended to alleviate the missing body problem I alluded to earlier include a system by Schreer, et al (2005), for broadcasting real-world body language in telephone call centers. While I could cite the many examples of research in this area, it is somewhat irrelevant: this chapter is mostly concerned with puppeteering within the limited gestural space of mice, keyboards, and touch screens, using direct-manipulation. In this space, there has been research to explore adding nonverbal social signaling to an otherwise verbal communication medium. One example is a system that uses very simple avatars (almost icons) which is interfaced with the Skype text chat system to provide a layer of visual signaling while multiple users communicate with text (Seif El-Nasr, et al, 2011). Using an experimental interface, the users can adjust directional gaze and posture, and to trigger specific signals along the lines of, “my turn to talk”, or “this is getting boring”, or yea, I agree”. This and other nonverbal puppetry systems use *procedural animation* – a form of animation that is generated on-the-fly with running software, guided by user-interaction. Many such procedural techniques have been explored by Ken Perlin, including Improv (1996), a system for generating real-time behavior in virtual actors.

Direct manipulation of virtual objects dates back to the very dawn of computer graphics, with Ivan Sutherland’s SketchPad (1964). The basic idea is that the user touches the screen with a special stylus, which serves as a manipulator for computer graphical objects. Using a mouse makes this activity a bit less direct (and if you can remember the first time you used a mouse: it was probably awkward...because of this indirection). However, there can still be a mapping of human motion input to the responsiveness of the graphics – and that accounts for discoverability and motor learning. Recent advances in touch screens have enabled commercial products that allow a whole new generation of direct manipulation interfaces – closer to Sutherland’s original input scheme.

The use of forward dynamics (Newtonian mechanics) for simulating the physics of human bodies is used often in games, and has acquired the name “ragdoll physics”. Figure 14-1a A ragdoll in itself is innocent and cute, but virtual ragdolls have taken on a sinister connotation, often associated with pain. This is due to the use of this technique for showing game characters tumbling to horrific deaths, as shown in the game Stair Fall [Figure 14-1] in which the violent death is accompanied by spattering blood. But the basic technique of ragdoll physics can also be used for more subtle effects. NaturalMotion, a company in Oxford, England, uses techniques whereby physics-based motions are enhanced to make the characters look alive (Figure 14-1b). This is achieved by using several AI-like techniques that cause the characters to self-apply corrective forces so they don’t fall down, or to hold their hands out to brace themselves when falling.

The use of direct manipulation on ragdoll-animated virtual humans has several challenges, when considering its use for controlling a character in real-time. This includes the unpredictability associated with having several joints moving in response to the user moving one control joint. Constraints can be applied to alleviate this problem, but the corrective constraints can often be more complicated than the problem itself. Many animators would prefer to stick with the familiar 3D interactions used commonly in 3D modeling and animation packages. But ragdoll physics, with some added constraints, does show promise as a basis for intuitive puppeteering, due to the directness of joint manipulation (control point movement corresponds directly to joint motion, and so it has a short learning curve for beginners). That is part of the motivation for the avatar puppeteering system described here.

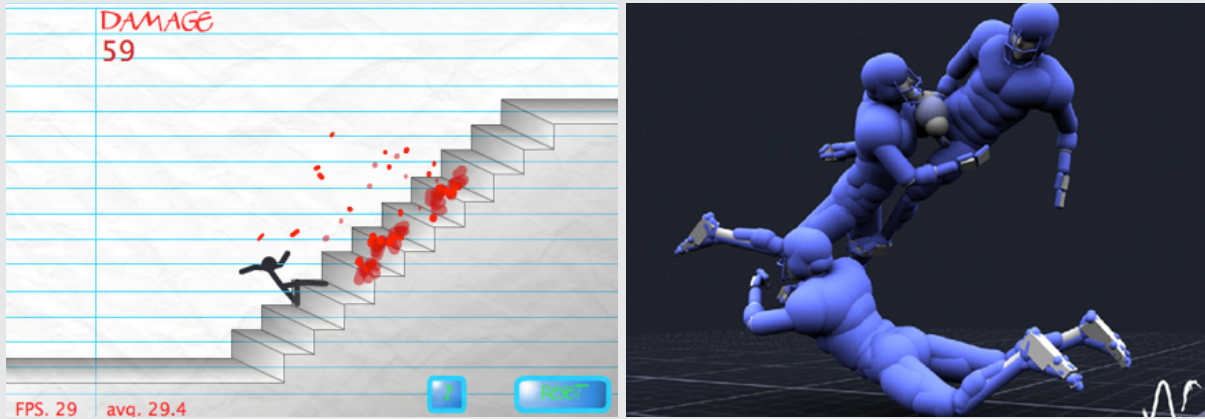


Figure 14-1: Ragdoll Physics (a) a scene from the computer game *Stair Fall*, (b) a promotional image from NaturalMotion

2. CASE STUDY: PUPPETEERING THE SECOND LIFE AVATAR

The avatar for *Second Life* uses a hierarchical skeleton – like what is used in most game and film characters. This is shown in Figure 14 2a. In hierarchical modeling, the elbow is a child of the shoulder, and the wrist is a child of the elbow, and so on down the chain of joints. This makes for efficient representations, since it is not necessary to store the positions of all the joints at any given time: only the rotations of joints is needed.

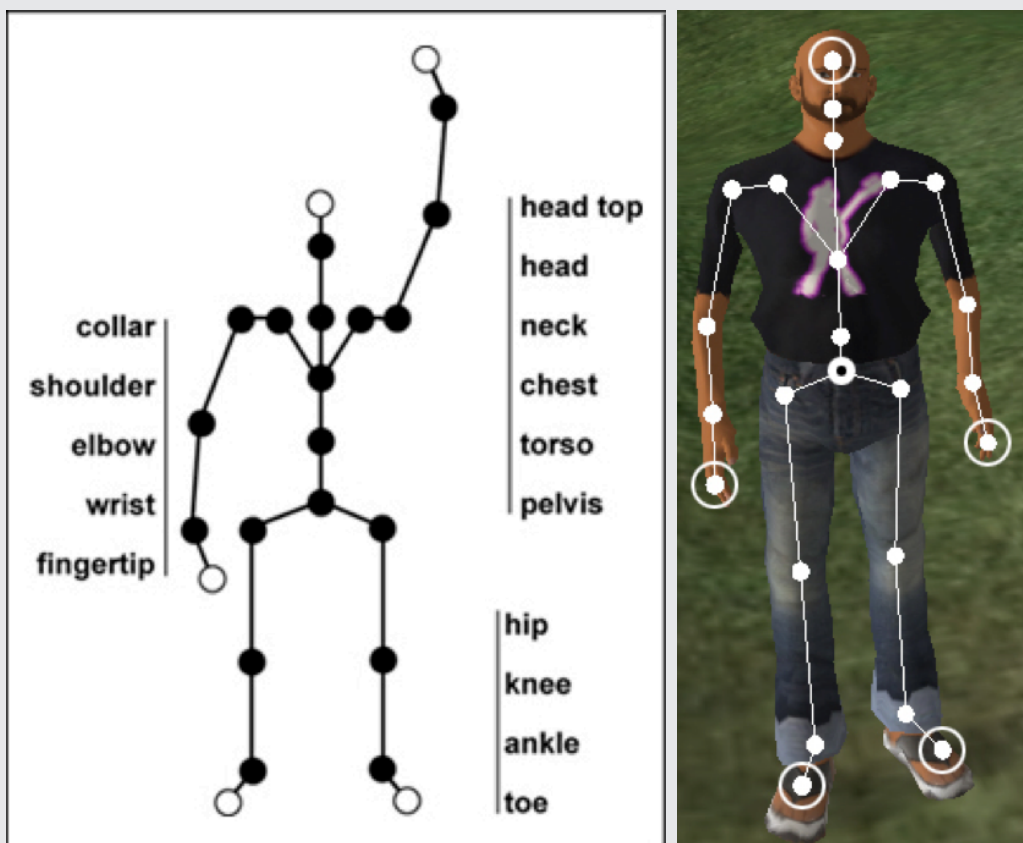


Figure 14-2: (a) Skeleton schematic with end-effectors; (b) superimposed onto a *Second Life* avatar

But there is a problem with *trivial hierarchy* (based on a tree topology where every joint is constrained by exactly one parent), especially if the goal is to make social virtual worlds in which avatars need to interact intimately with things in the world, including other avatars. Simple hierarchical modeling in itself cannot accommodate two avatars holding hands, two avatars aiming their heads to gaze at each other, or for the user to pick up an avatar's hand and wave it in the air in rhythm to the music. For this, the avatar's hand must be constrained by *both* the user's cursor manipulator and the arm. This can be solved using *inverse kinematics* (IK). With IK, the parent-to-child directionality of the skeleton is overridden. For instance, you would use IK to place a toe on the virtual ground in an exact location (and not poke through). The ankle, knee, and hip joint rotations are adjusted. (IK is commonly used for adjusting a regular walk cycle animation to accommodate an irregular terrain). Ragdoll physics offers its own variation on inverse kinematics: instead of calculating the mathematics of proper rotations to cause the desired end-joint to be in a specific position, the joint is “forced” into place over a span of time, and all the physical forces trickle through the skeleton - as if you held a ragdoll above the floor, and grabbed its toe and touched it to the floor.

PHYSICS FOR EXPRESSION

To achieve the goal of more fluid, environmentally-connected responsiveness in the avatar, a physics layer was built on top of the existing avatar code for *Second Life*, which I call the *Physical Avatar*. This layer enabled a user to tug at individual joints of an avatar—essentially to puppeteer, using a direct manipulation user interface. Postponing higher-level hierarchical controls, the Physical Avatar system attempts to solve the direct-manipulation issue of spontaneous body language. By replacing the avatar's hierarchically-arranged joints with balls connected by spring forces, the whole body stays together, due to the simultaneous forces of all the springs continually acting on all the balls. Figure 14-3a shows how the left hip and left knee balls are held together using a thigh bone spring force. This spring force is highly dampened so that it behaves like a semi-rigid bone. When pulling on the elbow ball, as shown in Figure 14-3b, the neighbor balls (shoulder and wrist) get pulled along. Everything is connected with equal priority.

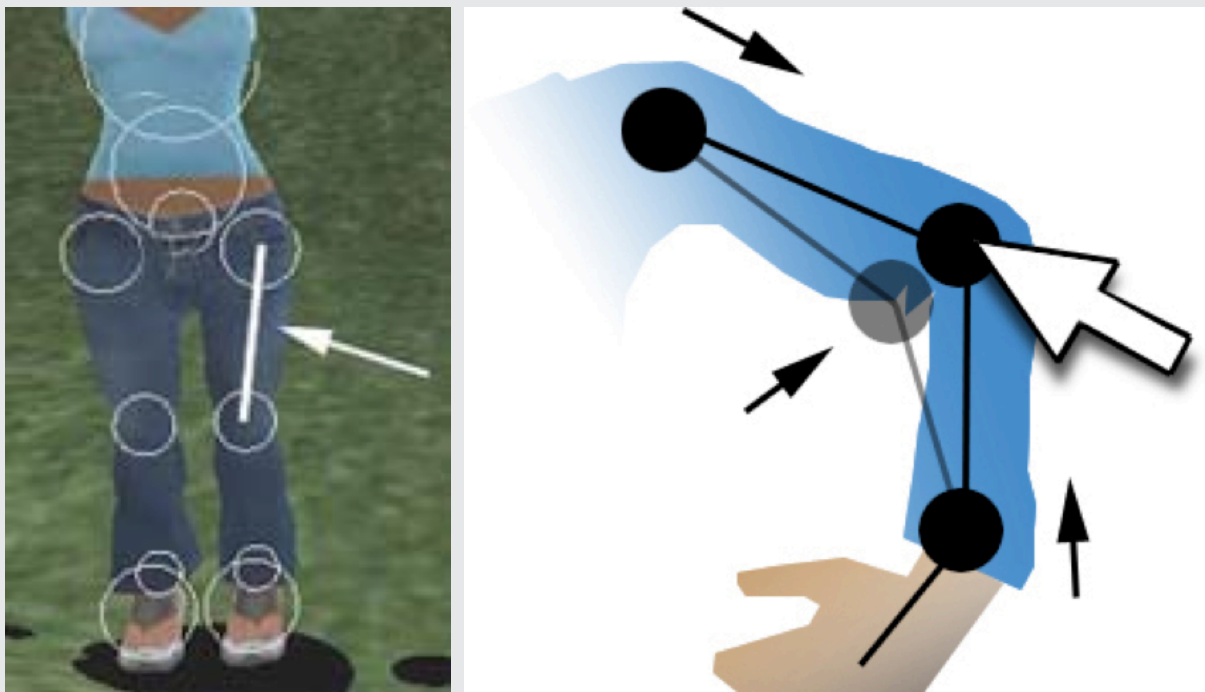


Figure 14-3: (a) left thigh bone represented as a rigid spring, (b) pulling an elbow ball acts on several other balls

This removal of *hierarchy* by using spring forces is considered a necessary step in order to re-introduce hierarchy—only this time with more expressive potential.

INTERACTION

The user interaction scenario is as follows: if the user holds down the CONTROL key and then hovers the mouse cursor over his/her avatar, translucent dots appear over the joints as the mouse cursor passes over them, as shown in Figure 14-4a. (two connecting spring forces are superimposed for illustration) This provides affordance and discoverability. If the user then clicks on a dot, the system instantly constructs a balls-and-springs rig that replicates the exact world-coordinate system positions associated with the hierarchical skeleton. Then, as the user drags the mouse cursor around, still holding the mouse button and CONTROL key down, the avatar joint (now represented as a ball in 3D space) gets moved around, and because it is connected by springs to neighboring joints, they get moved as well. What has just been described is a simple ragdoll layer of representation (without gravity effects so that the body stays suspended in space).

At the point in which the user clicks the mouse cursor on a joint, the distance D between the 3D camera viewpoint and the joint is calculated and stored. As the mouse cursor moves, a vector of length D is projected out from the camera viewpoint to the joint location. The vector serves as a projection into the 3D scene, resulting in a position in the scene: as the mouse moves, the joint moves along the surface of a sphere of radius D with the camera viewpoint at the center. Figure 14-4b illustrates a user moving the left hip joint using the mouse cursor. The root position is also shown in the illustration. This is the non-skeletal position that represents the avatar's location in the world coordinate system.

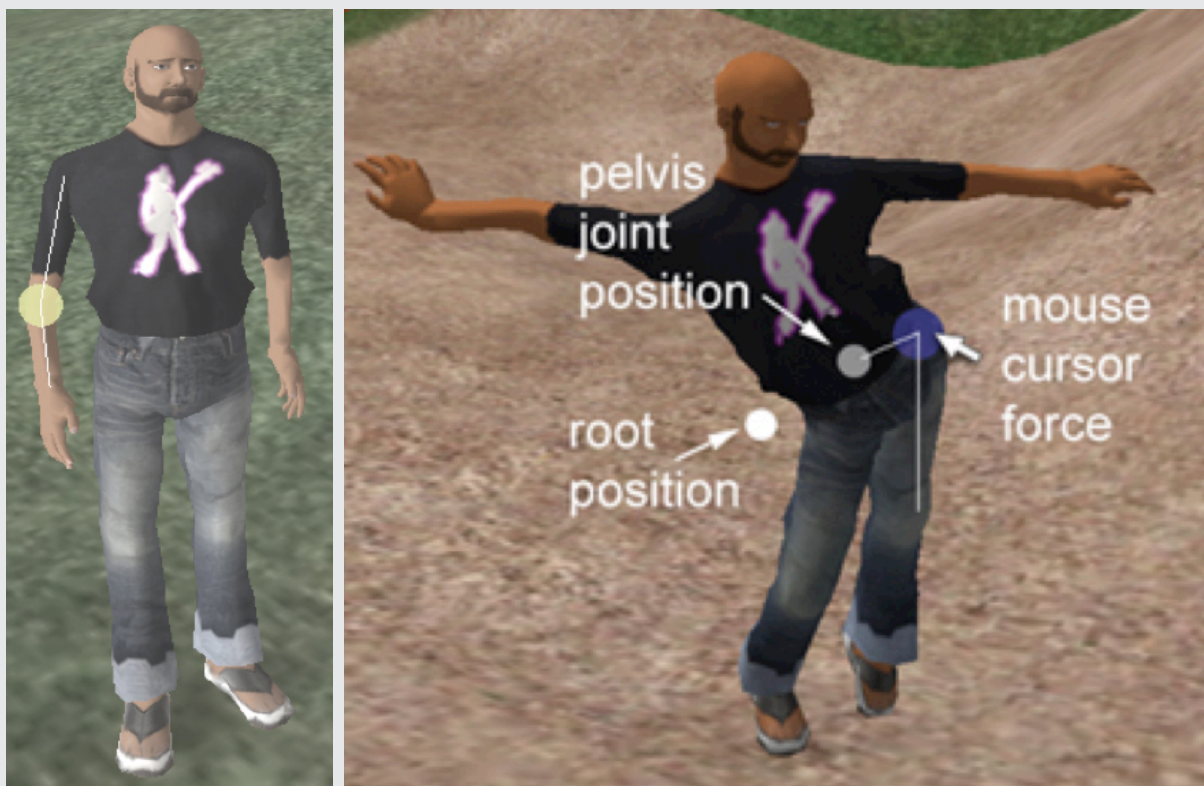


Figure 14-3: (a) left thigh bone represented as a rigid spring, (b) pulling an elbow ball acts on several other balls

SUBVERTING HIERARCHY TO TAKE-IN THE BIG BEAUTIFUL WORLD

In standard hierarchical animation, a rotation in the ankle joint causes the foot to pivot and this causes the toes to move to a new position. There is essentially no need to specify a toe joint position because all the geometry can be determined completely from knowing the ankle rotation and position. But as a consequence of not having a toe joint, it is not possible to “grab” the toe and move it in order to rotate the ankle. That requires inverse-kinematics, as contrasted with forward-kinematics, in which the positions of joints are determined in the direction from parent-to-child (i.e., shoulder, elbow, wrist, fingertip), as in typical hierarchical modeling.

Figure 14-5 illustrates the idea that forward-kinematics, using standard hierarchical modeling, requires parent joint information in order to determine child position and angle, whereas inverse-kinematics starts from a child joint, and adjusts parent joints accordingly. Inverse kinematics is required, for instance, to constrain the hands of two avatars to touch each other, as illustrated in Figure 14-5b. Physical Avatar is basically a variant of inverse-kinematics, achieved through physical simulation (forward dynamics). In order to grab a toe, the tip of a hand (a “fingertip” joint), or the top of the head, five extra joints (end-effectors) were added to the skeleton. This brings the number of avatar joints from the default of 19 up to 24, as shown in Figure 14-2.

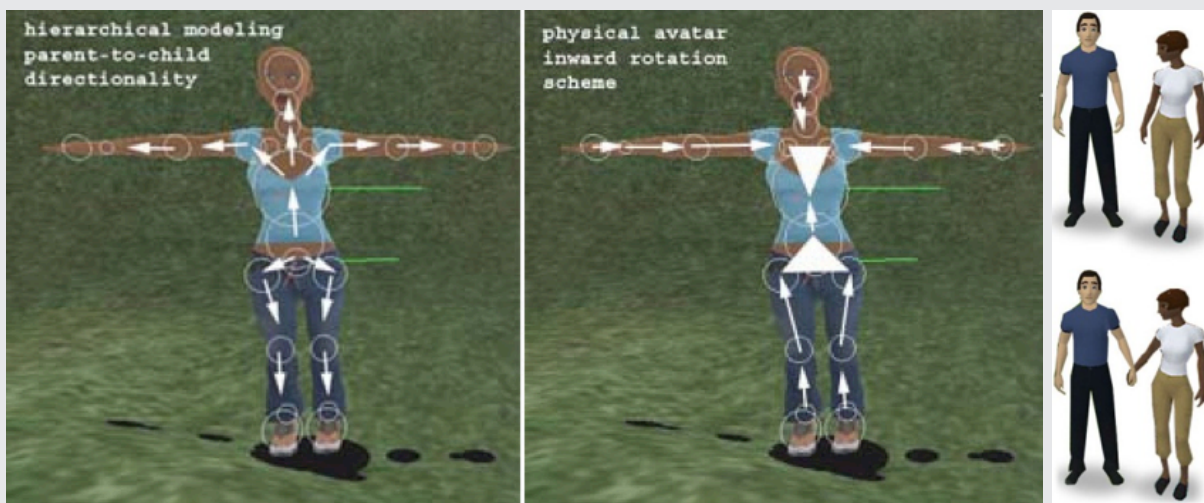


Figure 14-5: (a) Directionality differences between hierarchical forward-kinematics and inverse-kinematics; (b) avatars holding hands

In order for Physical Avatar to be fused with the existing hierarchical animation system so that it can be turned on and off at any time, the system must perform three main tasks in real-time:

1. convert the hierarchical skeleton into the physical representation of balls and springs
2. calculate forces (from user manipulation, or from other sources such as avatar-to-avatar contact, gravity, wind, collisions, emotional state, and many other possible effects)
3. convert back into the hierarchical representation, so it can be rendered with these modifications.

This third task is nontrivial. It requires re-generating all the joint rotations (represented as quaternions, in the case of the *Second Life* avatar) by traversing the hierarchy, using the “default pose” (shown at the top-right of Figure 14-7) as a reference. Critical for this final stage is calculating coherent rotations in the

pelvis and chest regions. The pelvis and chest joints each have three child joints. Consider the pelvis: the ancestor of all the other joints. In an avatar animation, the torso and hip joints are rotationally “fused” to the pelvis joint (the combination of these four joints constituting a non-elastic tetrahedron, which roughly corresponds to a real pelvis, and a real pelvis connects the spine to the thigh bones). This is illustrated in Figure 14-6.

In order to calculate a coherent rotation in the pelvis, we must secure its three children in relation to each other, so that they act as one rotational body. This is done by adding three special spring forces which constrain the two hip joints to each other, and the torso to the hips. Like the bone-springs associated with the standard avatar bones, these springs are highly dampened and so they are quite stable. The same technique is applied to the chest region. Even if they shift slightly, the resulting rotation is always legitimate, by way of cross-products and normalizers that generate a legitimate rotational representation for every time step in the simulation.

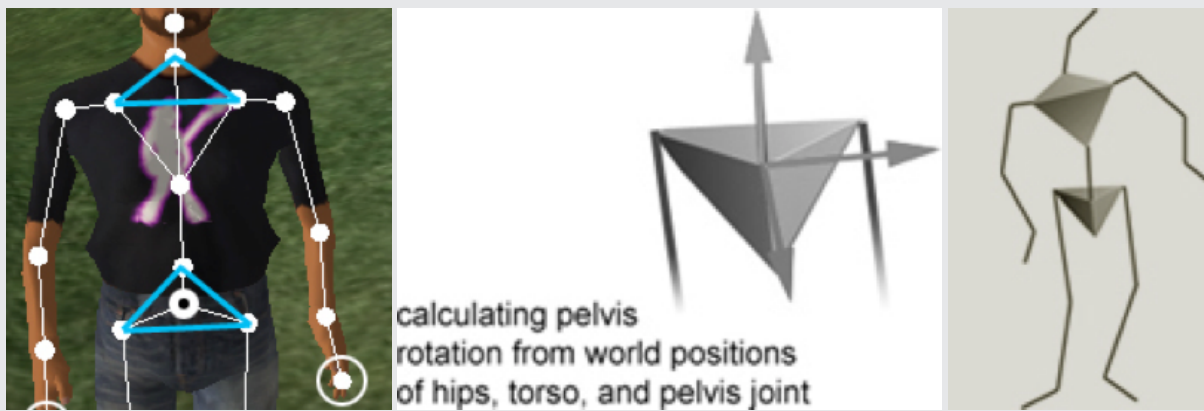


Figure 14-6: Extra spring forces determine tetrahedra for calculating rotations

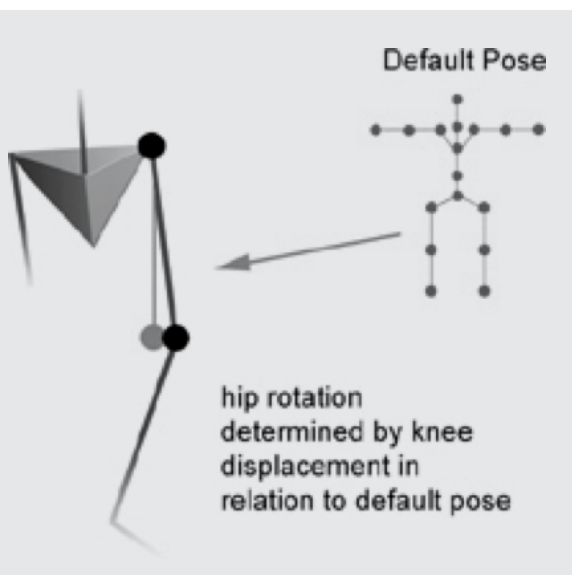


Figure 14-7: The default pose local positions are used to reconstruct rotations

DEFAULT POSE AS REFERENCE

In the default pose (sometimes called “T-pose”), the avatar’s arms are extended out and the joints are mostly straight. In this pose, all the joint rotations are set to zero (the identity matrix). Since all rotations are parent-relative, the default pose represents normalized joint rotations – used to compare the rotational offsets caused by manipulation. In quaternion terms, the difference between the local position of a manipulated joint and the local position of where the joint would be if it were in its default pose determines a local arc sweep (a spherical linear interpolation, or *SLERP*). This value is used as the new rotation of the joint. The set of all such rotations constitute the “puppeteered” avatar that is rendered.

If you are skilled at 3D math, you may be able to extrapolate any missing pieces from my limited explanation. If you are *not* a math geek, then you may have glazed over in some parts. It would suffice to say that step 3 above is complicated, and not-perfectly implemented...but mathematically tractable. You can think of it as a black box: it converts the all-important ragdoll model into an all-important standard hierarchical model.

A NOD TO 3D MODELING INTERFACES

I have not yet addressed my original desire to click on my avatar's head and make it nod yes or no. That is because these gestures cannot be easily mapped to the view surface – in the typical case when your avatar is standing or sitting: oriented upright in the scene. Clicking on an avatar joint (like the end-effector at the top of the head) and moving it with a mouse can only drag it around on the view plane (affecting the classic Indian *head waggle*, or a *yes-nod*, depending on the angle of the camera view/cursor vector relative to the head). On the other hand, nodding no requires a rotation around an axis that is generally not parallel to the line of sight. To address this, an interface was prototyped allowing the user to press a set of special control keys to specify different axes of rotation.

Clearly, these extra controllers add complexity to the interaction. These kinds of interaction problems have been solved to a great extent in advanced 3D modeling and animation packages (including the prim building tools in *Second Life*). It is valid to ask whether these advanced interaction techniques will make their way to the common user, who is not trained in advanced 3D modeling interfaces. This complexity would be one reason for arguing in favor of intuitive full-body gestural interfaces, to avoid the problem of arbitrary mapping from 2D to 3D. But I am not yet willing to give up on the possibility that we can still design an intuitive puppeteering system that allows rotational manipulation. One good way to test this hypothesis would be to implement an avatar version of the current prim building interfaces of *Second Life*. Since there are already many residents who are well-versed in building complicated and beautiful objects in *Second Life*: rotating, scaling, translating, etc., it would seem that these interfaces could be mapped to the avatar, for the purpose of real-time expression.

3. APPLICATIONS

One application that was explored with the Physical Avatar system was a scheme by which a user could create a series of full-body poses, save each one in memory, and then use them as keyframes for generating whole animation sequences. The benefit of this way of making avatar animations is that the user doesn't have to leave the virtual world, open up a separate animation software package, build an animation using an alien avatar, and then import that alien's animation back into the world. This scheme allows a user to construct an animation *in-situ*: in the context of all the surrounding ingredients that give that animation meaning.

Another potential application explored was symmetric puppeteering, whereby moving one joint causes the joint on the opposite side of the body to move in the same way, as shown in Figure 14-8.

Take this idea to another level: imagine clicking and dragging the mouse to generate more complex actions that involve several joints, coordinated to generate more complex whole-body language expressions.

These examples of applications that were explored – to varying degrees – show that the puppeteering scheme developed could be expanded in such a way to enable higher-level controls. For instance, a user can build a gestural phrase out of microgestures and then record that as an avatar animation. This would give users the ability to construct unique body language expressions of increasing complexity and utility.

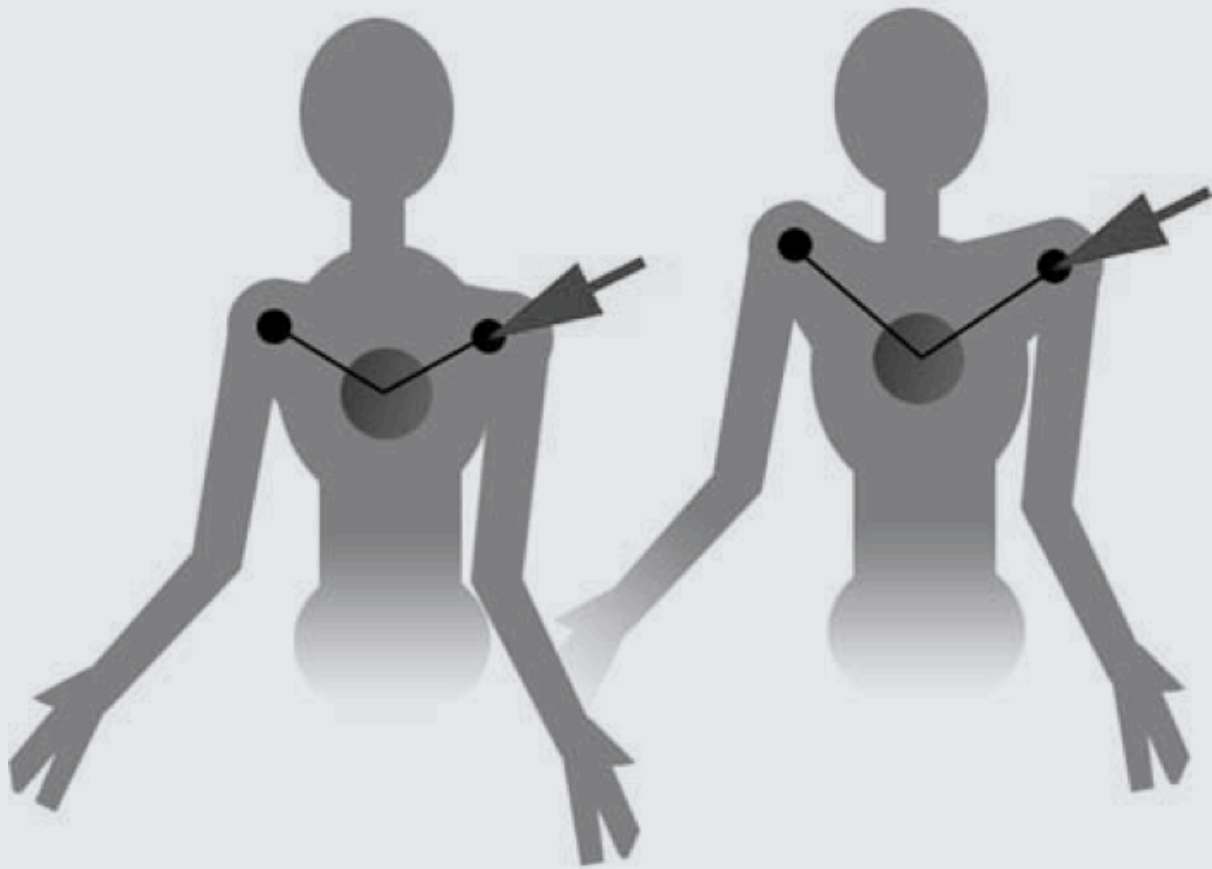


Figure 14-8: Symmetrical puppeteering: creating a shrug by interactively moving only one joint

4. CHALLENGES

There are many more considerations not covered here regarding required constraints and motor control mechanisms so that these higher level motions appear human. To this end, the Physical Avatar employed many constraints on joint movement (for instance, not letting the knees bend backwards). Another challenge: without sentence queues, as exemplified by the virtual humans of NaturalMotion, all of these motions would appear as the equivalent of pushing and prodding a corpse. This was one of the most challenging and complex aspects of designing the system. These problems would become even more challenging as we try to design responsive behaviors on higher levels. The experiments done with the Physical Avatar were rudimentary in this regard.

NETWORKING

There is another important aspect of the Physical not described yet: a substantial software component that distributes representations of puppeteering actions across the internet. This is what enables multiple remote users to witness each other's puppeteering activities, across the vast expanses of the internet (otherwise, there would be no communication :) The problem is to send the appropriate avatar joint data up to the server and then to broadcast it down to all the clients (viewers) that have instantiations of the puppeteered avatar visible. Several techniques were explored, including sending only the position of the manipulated joint and relying on identical physical simulations on all viewers to generate the same behavior. This technique,

while the most efficient in terms of internet communication, was the most unstable, being subject to drift among the viewers, and the problem of physical simulations not being totally in-sync. A collection of viewers cannot easily maintain identical, synchronized physics simulations throughout the puppeteering activities. A less computationally-efficient technique involved sending a continual stream of the entire set of avatar joint rotations resulting from puppeteering actions up to the server, and then down to all the viewers. This solution, while conceptually less elegant and computationally less efficient, proved to be more stable.

The networking aspect of *Second Life* avatar puppeteering is complex in itself, and worthy of an entire paper, or even a Ph.D. dissertation (which I would not be the one to write) This is perhaps the most important reason why the Physical Avatar was not able to make it as a feature in *Second Life* – there were many problems related to the particular quirks of the *Second Life* code for managing internet communications. The Physical Avatar code had to contort itself, as it were, to accommodate these quirks.

5. A THOUGHT ABOUT FUTURE PUPPETEERING INTERFACES

I want to offer a thought experiment as a way to explore possible future interfaces for embodied communication: compare two activities that we all engage in: (1) talking face-to-face and interacting and expressing in the company of others; (2) writing words and drawing pictures on pieces of paper. The first activity is old – older than technology itself, by most definitions. The second activity uses traditional technologies: the cognitive technology of writing and reading, and the physical technology of paper and pen. The ergonomics of this second activity are very different than the first, but it is a mode that we find ourselves in while communicating with others more and more all the time. Computing and the internet have enhanced this mode greatly, not only in terms of writing, reading, watching, and making pictures, but also in terms of speaking, listening, and exchanging information in many forms. With advances in touch screens, ambient computing and virtual embodiment, this second activity will become even more enhanced, and it will start to incorporate more of the levels of expressivity that we associate with the ancient form-factor of face-to-face communication. However, (and this is the crux of my thought experiment) will humans eventually tire of this limited form factor and return entirely to the ancient modality of full-body interaction? After all, full-body gestural interfaces are advancing at such a high rate that the media between us will become effectively invisible. Why bother with paper, pens, books, and touch screens?

It is likely that we will always continue using portable media in which we give and receive information through small, sometimes-mobile, flat visual surfaces – using our hands, our eyes, and sometimes our voices. As long as this is the case, body language (which is certainly never going away) will evolve new manifestations, new abbreviated forms of puppeteering – building out from the ergonomics of reading, writing, watching and drawing. And despite this limitation, full-body expression will still be experienced *vicariously* and *virtually*, via avatars. In this chapter, I make the assumption that manipulating visual information on a flat visual surface is a permanent future human activity. We have a rich history and tradition of ways to manipulate visual information in this way, and so puppeteering our virtual selves is likely to evolve as an extension of these actions.

6. CONCLUSION

Ragdoll physics, with the enhancements of various anatomical constraints and sentience-queues associated with living beings, creates a sense of immersion within a virtual environment. Multiple avatars that have communication affordances creates the experience of copresence – the locus of virtual body language.

Creating a physical connection between these avatars (and the user's touch) – enhances this sensation further. This chapter proposes a system for puppeteering one's body language in real-time, using a physical model as the substrate – as the foundation for connecting the virtual body to the environment, and potentially to other avatars. The goal is to scale up to allow very expressive – and possibly very intimate – interactions.

7. BIBLIOGRAPHY

- Allmendinger, K. Social Presence in Synchronous Virtual Learning Situations: The Role of Nonverbal Signals Displayed by Avatars. *Educational Psychology Review*, v22 n1. 2010.
- Mapes, D. P., Tonner, P., Highes, C. E. Geppetto: An Environment for Efficient Control and Transmission of Digital Puppetry. *Virtual an Mixed Reality – Systems and Applications. Lecture Notes in Computer Science*, vol. 6774. 2011.
- Perlin, K. Goldberg, A. Improv: A System for Scripting Interactive Actors in Virtual Worlds. *SIGGRAPH Proceedings, Addison Wesley*. Pages 205-216. 1996.
- Sutherland, I.E. "SketchPad: A Man-Machine Graphical Communication System," in *AFIPS Spring Joint Computer Conference*. 1963. 23. pp. 329-346.
- Seif El-Nasr, M., Isbister, K., Ventrella, J., Aghabeigi, B., Hash, C., Erfani, M., Moric, J., Bishko, L. Body Buddies: Social Signaling Through Puppeteering". *HCI International, 2001 conference proceedings*
- Schreer, O. et al. Real-time Avatar Animation Steered by Live Body Motion. *Image Analysis and Processing, 13th International Conference Proceedings*. pp.147-154. 2005
- Ventrella, Avatar Puppeteering (web page: www.avatarpuppeteering.com), 2008
- Ventrella, J. *Virtual Body Language*, ETC Press. 2011

15

AUTOMATION OF AVATAR BEHAVIOR

By Hannes Högni Vilhjálmsson

The graphical virtual worlds that sprang up in the mid-90s were truly inspiring. Finally the cyberspace that was promised to us by science-fiction was taking shape in front of our eyes. We could finally shed our real-life carcasses and enter a world of endless possibilities in shiny new digital bodies. The avatars we would embody would allow us to share this experience with others, just as if they were right there with us face-to-face.

1. CHAT VS. AVATARS

But something was not quite right. While we would now see all these avatars in the environment, they were suspiciously quiet and still. They hardly stirred, except for the occasional “slide” in or out of the room. It would have been easy to mistake the place for a wax museum, and yet, the place was literally bursting with excited conversations about this new frontier. The problem was that in order to actually notice the conversations, one had to open a chat window. The owners of the avatars essentially parked their bodies somewhere in the 3D environment and then started communicating with their fellow cybernauts using the traditional text interface (see Figure 15-1).

There were two modes of operation here. One where the user navigated their avatar around the virtual world to explore it, play in it or even construct it, and another where the user let go of the avatar and started communicating with other users through a separate modality. The avatar was hardly more than a graphical token that was associated with your location, rather than an embodiment that would deepen the social experience.

This split between the avatars and the conversations between their users even persisted when users were given buttons to play short animation sequences such as “waving” and “laughing”. Of course users would play with these means of nonverbal communication, but once the typing of chat messages commenced, these buttons would hardly get used. One reason may be that using them would require letting go of the keyboard, therefore disrupting the flow of chatting. It required a certain level of dedication to explicitly control the avatar while at the same time composing meaningful contributions to an ongoing conversation. Therefore, it was common to see still avatars where people were in fact interacting. This sort of behavior has since been studied more closely, contributing further evidence that puppeteering of expressions and gestures simply takes too much effort to be frequently used (Seif El-Nasr et al. 2011; Shami, N.S. et al. 2010).



Figure 15-1: Worlds Chat was the first 3D chat environment on the Internet, launched in 1995 by Worlds Inc. It was a real breakthrough, but the avatars felt like chess pieces standing silently around the rooms, while conversations took place in the box below (picture from Bruce Damer)

2. IDLE ANIMATION

Recognizing that motionless avatars were lifeless and kind of creepy, it wasn't long until idle animation loops were added, which required no input from the users. These would essentially provide the illusion that the avatars were alive while users were doing things like chatting. We may take these idle movements for granted now, but it is a big deal to realize that users could not be trusted to keep their avatars alive by constantly animating them. It was not enough to give the users what we could call “button puppets” and expect a continuous performance that would bring them to life.

Idle loops really breathed life into the first avatar worlds, and the wax museum effect was gone. Instead, we had what seemed like crowds of people going about their business. Some of them would look around expectantly while others had a ponderous face, as if lost in thought. This was a first step, albeit a simple one, towards avatar automation. Still, things were not quite right. These animations were typically played cyclically or randomly, regardless of what was in fact going on. So, if one were to start a conversation with another user, through the chat window, the avatars would go on displaying the same nonverbal behavior as before the conversation started.

This could lead to mixed signals at best and downright inappropriate signals at worst. For example Alpha World, one of the more popular worlds (see Figure 15-2), used the action of looking at one's watch as an idle animation loop for some of its avatars. Few things are more socially demoralizing than watching your listeners repeatedly checking their watches while you are telling them something important. The knowledge that these were idle animations was not even enough to shrug off the strong nonverbal signal.

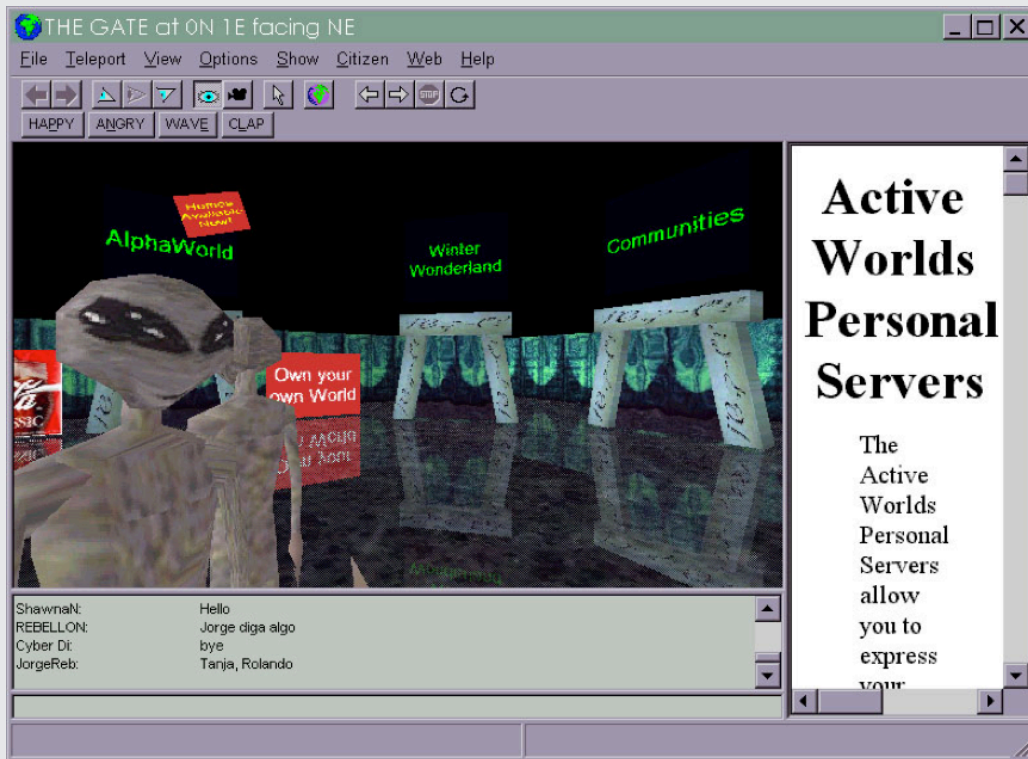


Figure 15-2: AlphaWorld (later ActiveWorlds) was one of the first 3D online environments, released in late 1995, to feature articulated button controlled avatars, and a set of idle animations. The “looking at watch” idle animation was a particularly memorable animation (picture from Bruce Damer)

3. LOCOMOTION ANIMATION

Bringing avatars to life through automating animation needs to be done carefully. Any nonverbal behavior exhibited by the avatar will be interpreted in the current context, both the social and environmental. Therefore, it makes sense to have the animation reflect the activity the avatar is currently engaged in. Perhaps the most readily accepted avatar animations that tie into the context are locomotion animations. Here the looping animations are chosen based on the user’s intended mode of locomotion, for example walking, running or flying.

It is interesting to note that it is not the user that is choosing the animation to play during locomotion, but rather, the user may simply be pointing in the desired travel direction, while the avatar takes care of producing continuous body motion. This feels relatively natural to the user of the avatar, even though the automation may be adding various details to the behavior. For instance, when a user presses a button to go forward, the avatar may at first take short steps that gradually build up to a fast pace, slowing down again while crossing a river, and then transitioning smoothly over to a standing posture when the user lets go of the forward button.

Locomotion control for avatars, and corresponding continuous animation, has seen great advances in the last decade, mostly driven by the need to let game world avatars seamlessly traverse ever more complex environments made possible by more and more powerful graphics hardware. By using only one or two buttons players may have their avatars produce a wild dash through a crowded street, a daring jump from a pair of barrels up towards a window, where they barely manage to grab a hold of the ledge and pull themselves up into the safety of the room within.

4. ENGAGEMENT WITHOUT MICROMANAGEMENT

The fluid sequence of life-like motion and interaction with the environment that users see in their game world avatars arguably creates a stronger sense of being in the environment than if they were required to directly specify the animation to be played from moment to moment. Avatar micromanagement is something that games have tried to avoid when it distracts from core game play. Instead of micromanaging the avatar body, the level of control is placed at the level of player intent. For example, one might merely signal the desire to “jump” to produce all the corresponding preparation, execution and termination animation sequences.

What about social situations and communication? We were stuck with avatars that somehow were left out of conversations or displayed behavior that was in little relation to what was going on when people were attempting to communicate. Would it be possible to make the avatars provide social and communicative cues without requiring the users to execute specific animations? Similar to the locomotion scenario, might it be the case that the more interactivity with the social environment that the avatar can portray, the greater the sense of actually being there in the presence of others? It may seem counterintuitive that letting an avatar provide communicative signals on the behalf of the user would lead to a more integrated social experience. This became the subject of a research project called BodyChat at the MIT Media Lab in 1996 (Cassell and Vilhjalmsson, 1999) and later also became part of the so-called “Avatar Centric Communication” approach in *There* (Ventrella 2011).

5. SOCIAL ANIMATION WITH BODYCHAT

BodyChat was a graphical chat system that automated some of the communicative behavior one would expect to see in a normal face-to-face setting. The automation was not random, but instead it was triggered by a combination of two things: (1) The social goal of the user; and (2) what was going on in the environment. For example, the user could click on another user’s avatar to indicate that their goal was to start a conversation with them. This would result in a prolonged glance and a smile towards the other user, and if the other user was available for a chat (as indicated by an availability toggle), the other user’s avatar would reciprocate the glance and smile (see Figure 15-3). This would then trigger a salutation sequence that would conclude with another smile and a head nod when the users would bring their avatars together.

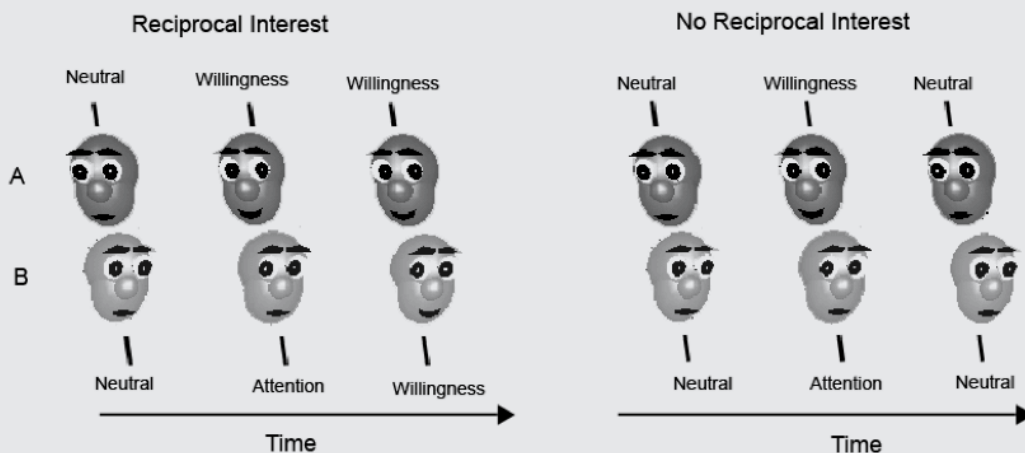


Figure 15-3: Avatars of two users, A and B, demonstrate a sequence of expressions that result from A wanting to start a conversation with B, first when B is available (left) and then when B is not available (right).

This whole process was modeled after documented human behavior, in particular studies on human greetings (Kendon, 1990). During chat, the avatars in BodyChat would animate eyes, eyebrows and head based on keywords and punctuation in the chat message, based for example on research on the relationship between facial displays and syntax during speech (Chovil, 1992). In Figure 15-4 we see from a first person perspective how another user's avatar is indicating that user's goal of talking to us. However, we have chosen to be unavailable (the toggle on the lower left), so our avatar will respond only with a quick glance, dismissing further interaction.



Figure 15-4: Another user's avatar displays willingness to chat with us in BodyChat. Notice that since we are not available for conversation (toggle on lower left), our avatar will dismiss the request with the appropriate expression.

6. MORE CONTROL WITH AUTOMATION?

To understand the effect of automated behavior on the social experience, a controlled user study was conducted. Four different versions of BodyChat were compared between subjects: *automatic* - Autonomy was responsible for all animated behavior (other than just moving the avatar around the environment); *hybrid* - In addition to the automation, all animations were also available to the user through a menu; *manual* - Automated animation turned off, but the animation menu was available; and *none* - Where avatars could not animate. For the sake of our discussion here, we will concentrate on the comparison between the automatic, hybrid and manual conditions. A complete discussion can be found in the original paper (Cassell and Vilhjalmsson 1999).

Subjects were tasked with entering a virtual environment with their avatar, meet other people and learn as much about them as they could. Unbeknownst to the subjects, all other avatars were under the control of a confederate that followed a strict behavior and conversation protocol. For example, the protocol dictated that the confederate produce similar nonverbal behavior as the subjects where manual control was possible. The confederate also provided scripted personal facts in response to questions from the subjects.

The results showed that the automation contributed to a better user experience. First some behavioral measures: conversations in the autonomous condition were significantly longer (a mean of 1111 seconds) than those in the manual (mean of 671 seconds) or hybrid (mean of 879 seconds) conditions. This can be taken as an index of the interest that subjects had in pursuing conversational interaction with people, when they were using the autonomous system. Moreover, subjects in the autonomous condition remembered more facts about the people they interacted with (a mean of 5.2) than did subjects in the manual condition (mean of 3.8) or the hybrid condition (mean of 4.5 facts). This can be taken as an index of how engaged subjects were in the conversation, perhaps because their attention was not divided between participating in the conversation and controlling the avatar.

The results of a subjective questionnaire indicated that the avatars in the purely automated condition were judged significantly more natural than both the manual and the hybrid versions. Furthermore, the system in the purely automated condition was judged to support significantly greater expressivity than the manual version, and somewhat greater than the hybrid version although not to a significant degree (see Figure 15-5). The measures of naturalness and expressivity are aggregates of several questions that probe these factors as detailed in (Cassell and Vilhjalmsson, 1999). It was clear that the automation was adding something to the experience.

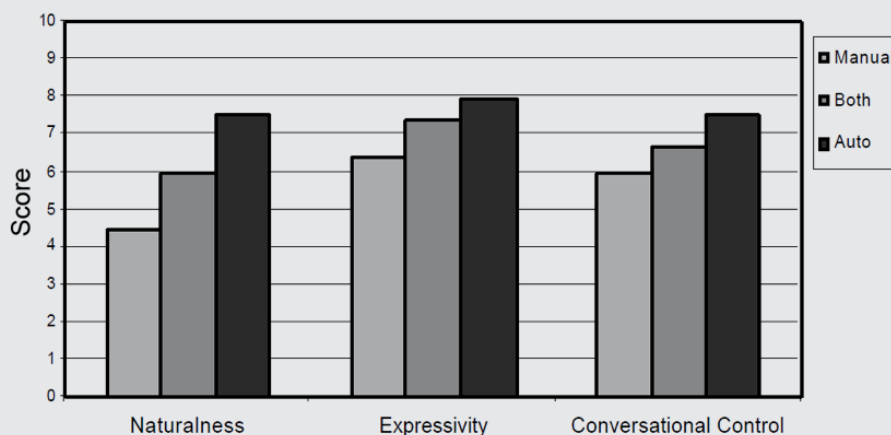


Figure 15-5: Subjective between-subject scores of avatar naturalness, expressivity and conversation control across manual, hybrid and fully automated animation conditions

The most controversial result however, regarded the issue of control of the situation. Subjects were asked (a) how much control did you have over the conversation? and (b) how much control do you think the other users had over the conversation? An ANOVA, and subsequent post-hoc t-tests, analyzing these two questions together revealed that users of the autonomous system considered the conversation significantly more under users' control than did the users of either the manual or the hybrid systems. This may sound counterintuitive because the nonverbal conversational behaviors were not under their direct control at all. However, one could argue that since the users were freed from the overhead of managing nonverbal behavior, they could concentrate on steering the course of the conversation itself. The fact that the users felt more in control of the conversation in the purely automatic version versus the hybrid version, where manual control is added as an additional feature, may indicate that the mere knowledge of a menu introduced some kind of a distraction.

Thinking too much about what you are doing with your face and hands while having a conversation with someone will make it harder for you to conduct yourself naturally. Try it the next time you have a face-to-face conversation. Decide ahead of time that you will not show any gesture or facial expression except a thumbs-up gesture and a smile, but you have to pick the right moment to use them. Pay attention to what happens to the rest of the conversation, especially with regards to timing.

We sometimes see odd synchrony between the spoken channel and the gestural channel in amateur actors or new politicians that have just been coached in the use of gesture. They are conscious of what they consider proper gesturing, but they may not always get the timing right, for example when making a point, they might throw a clenched fist into the air just a moment too late, essentially losing the momentum and breaking the flow. Good actors and good public speakers no longer need to micro manage their behavior, as they have already internalized their chosen speaking style. Their gestures and facial expressions arise from the same place as their spoken words – a place that represents their communicative intent. It is almost as if the realization of that intent simply grabs a hold of the body, and animates it in the manner that is most effective, without conscious effort.

7. FACE-TO-FACE COMMUNICATION

It doesn't take professional speakers to coordinate multiple communication modalities. When we engage in casual conversation with friends or when we interact with a cashier at the grocery store, we spontaneously produce an elaborate multimodal performance. Behaviors that range from posture and facial expressions to our tone of voice and words spoken are woven together into coherent and seamless communication. Are these behaviors serving any particular purpose? Do they somehow contribute to the success of our everyday conversations and transactions? To answer that question, we must first understand the process of face-to-face communication and what kind of activity underlies its success.

The study of face-to-face communication has been undertaken by many different fields, and typically fields that straddle boundaries between more traditional disciplines. These include discourse analysis, context analysis, sociolinguistics and social psychology. Scientific enquiry has revealed patterns of behavior, coordinated between all participants of social encounters. While the patterns are many and varied, the processes they represent generally fall into two main categories: *interactional* and *propositional*. In essence the former deals with establishing and maintaining a channel of communication, while the latter deals with providing effective communication across that channel (Cassell et al. 1999; Cassell et al. 2001). There is more going on in face-to-face interaction, but these processes provide fundamental building blocks.

On the interactional side, two important functions are *turn management* and *feedback*. Properly managing turns is necessary to ensure that everyone is not speaking at the same time, and can therefore be clearly

heard. Turns are requested, taken, held and given using various signals, often exchanged in parallel with speech over nonverbal channels such as gaze, intonation, and gesture (Duncan 1974; Goodwin 1981). Taking or requesting turns most often coincides with breaking eye contact (Argyle and Cook 1976) while raising hands into gesture space (Kendon 1990). A speaker gives the turn by looking at the listener, or whoever is meant to speak next, and resting the hands (Duncan 1974, Goffman 1983; Rosenfeld 1987).

Speakers often request feedback while speaking and expect at the very least some sign of attention from their listeners. Feedback requests typically involve looking at the listeners and raising eyebrows (Chovil 1991). To request a more involved feedback, this behavior can be supplemented with pointing the head towards the listener or conducting a series of low amplitude head nods ending with a head raise (Rosenfeld 1987). Listener response to feedback requests can take on a variety of forms. Brief assertion of attention can be given by the dropping of the eyelids and/ or a slight head nod towards the speaker. A stronger cue of attention may involve a slight leaning and a look towards the speaker along with a short verbal response or a laugh. A speaker's ability to formulate messages is critically dependent on these attentive cues, and therefore, even if only one person is speaking, everyone present is engaged in some kind of communicative behavior.

The propositional side deals with what we say and how we make sure those who are listening pick it up correctly. In addition to the content itself, there are three types of communicative functions that play an important role: *emphasis*, *reference* and *illustration*. Emphasis signals to listeners what the speaker considers to be the most important contribution of each utterance. It commonly involves raising or lowering of the eyebrows and sometimes vertical head movement as well (Argyle et al. 1973; Chovil 1991). A short formless beat with either hand, striking on the stressed syllable, is also common (McNeill 1992). Reference is most commonly carried out by a pointing hand. The reference can be made to the physical surroundings such as towards an object in the room, or to imaginary spots in space that for example represent something previously discussed (McNeill 1992). References through pointing are also made towards the other participants of a conversation when the speaker wishes to acknowledge their previous contributions or remind them of a prior conversation (Bavelas et al. 1995). Illustration is the spontaneous painting with the hands of some semantic feature of the current proposition. The particular features may lend themselves well to be portrayed by a visual modality, such as the configuration or size of objects, or the manner of motion (Kendon 1987; McNeill 1992).

From this overview we can see that the processes of interactional and propositional management are each carried out by a series of communicative functions. Each function in turn is realized through one or more nonverbal behaviors. The behaviors are not arbitrary, their function is often clear. It is good to keep in mind that human communication evolved in a face-to-face setting, and therefore the body has had an integral communicative role for most of human history.

8. ARE COMMUNICATIVE FUNCTIONS SUPPORTED IN MEDIATED ENVIRONMENTS?

It is a relatively recent development that humans can have conversations without their own bodies being present. So, what happens when the body is removed? The nonverbal behaviors are no longer available to support the crucial functions we just described. This sometimes leads to difficulties in communication. For example, think about a meeting where a couple of participants are only present through a voice conference system. What invariably happens is that the physically present participants tend to dominate the discussion while those on the phone have a hard time synchronizing their turns with the rest of the team. The body proves to be an advantage.

In the 1960s, AT&T introduced the Picturephone, a combination of a telephone and television that was meant to make remote conversations feel like they were truly face-to-face. The technology did not catch on, and even though available bandwidth and cheap hardware has made video mediated communication (VMC) very accessible, this mode of communication has still not become as commonplace as expected. A number of studies attempting to explain the slow adoption have shown that while VMC provides many important benefits over audio-only, it is also hampered by important limitations and in some cases may introduce negative artifacts that compromise the interaction.

Some of the benefits provided by VMC include the availability of nonverbal feedback and attitude cues, and access to a gestural modality for emphasis and elaboration (Isaacs and Tang 1994; Doherty-Sneddon, Anderson et al. 1997; Isaacs and Tang 1997). Seeing evidence of attention and attitude may be the reason why VMC has been shown to particularly benefit social tasks, involving negotiation or conflict resolution. In fact, groups that communicate with video tend to like each other better than those using audio only (Whittaker and O’Conaill, 1997). However, benefits for problem-solving tasks have been more evasive (Doherty-Sneddon, Anderson et al. 1997), and there one of the greatest limitations of classic VMC may play a role: The participants are not sharing the same space, they are each trapped in their own window.

Many important communicative functions break down when participants don’t have a common frame of spatial reference, especially for group conversations (Isaacs and Tang 1994; Whittaker and O’Conaill 1997; Neale and McGee 1998; Nardi and Whittaker 2002). For example, turn-taking and judging the focus of attention becomes difficult when gaze direction is arbitrary. Pointing and manipulation of shared objects becomes troublesome and side conversations cannot be supported.

Variations on the classic video conferencing system have been developed to address some of the limitations. For example, “video surrogates” can be created by physically embedding two-way video units in a room for each remote participant. In some cases, such surrogates are even attached to robotic platforms. These systems can provide some level of gaze awareness and increased social presence (Inoue, Okada et al. 1997; Yankelovich et al. 2007; Adalgeirsson and Breazeal 2010), but often rely on static configurations or manual control, which misses the dynamic and spontaneous nature of fully embodied communication. Besides, specialized hardware solutions can become preventively expensive.

How about avatars in virtual worlds then? They certainly have one advantage over video: They all occupy a shared virtual place. But can they support the crucial communicative functions? As we have already mentioned, we cannot expect the users of the avatars to manually produce all the right nonverbal behaviors that normally are produced unconsciously face-to-face.

If we intend to capture the spontaneous movements of the users and transfer that movement over to the avatar (e.g. using computer vision or other real-time tracking technology), we run into a similar problem to that of the transmitted video: the user behind the screen doesn’t occupy the same space as the user’s avatar. Therefore, any spatially related behaviors may not read correctly on the avatar’s body.

This could be addressed by fully immersing the user in the virtual environment through a head-mounted display, but the user would still be confined to the immediate physical space and facial expressions may be difficult to read when the eyes are completely covered. Furthermore, the user’s performance may not be “large” enough for the virtual world, both literally in terms of movement ranges, and in terms of style befitting a world that might be larger than life.

9. AUTOMATING COMMUNICATIVE BEHAVIOR

That brings us back to avatar automation. Can communicative nonverbal behavior be automated in avatars to the extent that crucial communicative functions are being served? If the avatar could exhibit the needed spontaneous behavior and demonstrate full immersion in the social environment, we could finally talk about a virtual encounter that simulates meeting face-to-face.

The key to making this possible is to let go of the notion that the avatar is a mere puppet. Instead, we can think of it as an autonomous agent that serves as our primary interface to the virtual world and to its other inhabitants.

A couple of very early experiments with avatar automation include Comic Chat (Kurlander, Skelly et al. 1996) and Illustrated Conversation (Donath 1995). In Comic Chat the avatars were not animated models, but characters in a comic strip that got automatically generated, frame by frame, from an ongoing text chat (mainly based on keywords in the text) and a special emotion selection wheel. The real genius was that the generated comic strip gave a strong visual impression of a face-to-face group interaction. In Illustrated Conversation users were represented by their portraits on the screen, but each portrait was picked from a set of photos taken from different viewing angles. The system automatically picked photos that would result in gaze alignments that properly reflected who was attending to whom. While the avatars in these systems had very little articulation, they hinted at the power of machine augmented expression.

The BodyChat study showed that automation is useful, but how closely do we need to model actual human communicative behavior to get the benefits of automation? Two early studies demonstrated the importance of using principled approaches to automated animation, rather than resorting to the much cheaper randomized behavior.

The first study compared how two subjects interacted with each other in an audio only condition, random avatar gaze condition, algorithmic avatar gaze condition and through a video tunnel (Garau, Slater et al. 2001). The timings for the gaze algorithm were taken from research on face-to-face dyadic conversations and based on who was speaking and who was listening. A questionnaire assessing perceived naturalness of interaction, level of involvement, co-presence and attitude toward the other partner showed that the algorithmic gaze outperformed the random case consistently and significantly. This suggested that for avatars to meaningfully contribute to communication it is not sufficient for them to simply appear lively. In fact, the algorithmic gaze scored no differently than the video tunnel with regards to natural interaction and involvement, demonstrating at least subjectively that even crude and sparse (only gaze) but appropriate behavior in avatars brings the interaction closer to a face-to-face experience.

Another study showed the impact of principled automation on actual task outcomes in a collaborative scenario. The study compared a random gaze avatar with an algorithmic gaze avatar where a subject had to collaborate with double-blind actors on constructing syntactically correct permutations of sentence fragments (Vertegaal and Ding 2002). The subjects in the algorithmic gaze condition gave significantly more correct answers than in the random gaze condition, showing that appropriate communicative behavior helps us get things done.

If we now think about our avatar as an autonomous agent, what capability does it have to possess to mediate communication usefully? Minimally it requires two things to do its job: a *model* and a *context*. The model describes the important communication processes and the rules that underlie appropriate behavior, for example a face-to-face model would need an algorithmic representation of turn-management and a rule for mapping the function of taking the turn into a coordinated behavior of averted gaze and increased gestural activity. The context represents the information that is needed to correctly interpret what is going on in the social interaction, and therefore choose what parts of the model are relevant at any given time. For example,

a part of the context needs to keep track of whether your avatar is already engaged in conversation with someone, and what kind of a social situation you are in.

10. THE SPARK SYSTEM

BodyChat took the first steps towards turning avatars into agents with communication skills, but the focus there was on automating only a few interactional functions, so crucial propositional functions were left out. The full range of interactional and propositional function support was introduced in a system called Spark (see Figure 15-6), developed in 2003 (Vilhjalmsson 2004; Vilhjalmsson 2005). Spark consisted of virtual world clients and a central server that not only synchronized the clients, but also analyzed all chat messages using real-time natural language processing. The idea was to give the avatar agents plenty to build their communicative behaviors on.

The bulk of the communication *model* in Spark was represented by a series of language and event processing modules inside the server that each kept track of a critical communication process for every ongoing conversation. For example, there were separate modules for turn-taking, grounding (e.g. feedback), visual and textual reference (e.g. for pointing), emphasis, illustration (e.g. for elaborate hand gestures) and topic shifts.

The *context* got represented by three important structures: discourse model, domain knowledge and participation framework. The *discourse model* was a dynamic structure that reflected the state of the ongoing conversation. A key component was the discourse history, essentially a list of objects referenced so far in a conversation. The discourse model also included a list of objects visible in the immediate environment, since those are considered part of what is shared information. The *domain knowledge* was a static structure that described the ontology of a particular domain that related to a conversation. While helpful for resolving ambiguities and for suggesting richer semantics for gestures, conversations about topics not covered by the domain knowledge base would still generate behavior. Finally the *participation framework* kept track of the status of every person in a particular gathering, such as whether they are speaking, being addressed, general listeners or not attending to the conversation at all.



Figure 15-6: Three users discuss how to solve a route planning puzzle, while their automated avatars express relevant communicative behavior, both interactional and propositional. The old tree gets a pointing gesture (the arm has already retracted a bit in the screen shot) when it is first mentioned, since it is clearly visible to everyone.

The users of Spark communicated only via text messages, but everything that got written, was processed by the system. The processing progressed through a couple of major steps: functional annotation and behavioral annotation. The first step identified the communicative functions, interactional and propositional, that would need to accompany the delivery of the message. For example, if this was the first time the user “spoke”, a turn taking function would need to be added, and if the message contained new information, that information would get associated with an emphasis function. In the second step, the functionally annotated message would be given to the user’s avatar for delivery. The avatar would then transform the functional annotation into behavioral annotation, according to the model mapping rules (see Table 15-1). This would ensure that every communicative function would produce a corresponding supporting nonverbal behavior.

The avatar would finally deliver the message after a typical processing lag of about 2-5 seconds (either as scrolling text or synthesized audio) along with a fully synchronized embodied performance. The textual chat and the animated body would become completely integrated.

Table 15-1: An example of functional annotations automatically added to chat messages by the Spark server and the corresponding behavior generated by the avatar agents upon delivery

FUNCTION ANNOTATION	BEHAVIOR ANNOTATION
EMPHASIS_WORD	HEADNOD
EMPHASIS_PHRASE	EYEBROWS_RAISE
GROUND_REQUEST	GLANCE (ADDRESSEE)
TURN_GIVE	LOOK (ADDRESSEE)
TURN_TAKE	GLANCE_AWAY
TOPICSHIFT	POSTURESHIFT
REFERENCE_TEXTUAL	GLANCE (LAST REF. SPEAKER)
REFERENCE_VISUAL	GLANCE (OBJECT)
CONTRAST	GESTURE_CONTRAST
ILLUSTRATE	GESTURE_FEATURE

11. IMPACT ON COMMUNICATION

The goal with Spark was to turn the relatively narrow channel of text into a fully embodied face-to-face experience, where the benefits of having a body during communication would become clear. The only way to find out whether this goal was achieved was to carefully study the impact of the Spark avatars on a communicative situation. A study was devised where groups of three users would use Spark to solve a visual puzzle. The puzzle was in the form of a map and the group's goal was to decide on the shortest path between two locations, given information about various hazards on the way.

In half of the sessions, users would be represented by Spark avatars; in the other half, no avatars were visible. In both types of sessions, an interactive puzzle map and chat window were available. Since it had already been established that randomly animated avatars perform very poorly compared to principled animation (Colburn et al. 2001; Garau 2001; Vertegaal and Ding 2001), the avatars were compared to having no avatars, instead of “dumber” random avatars. Comparing this new breed of avatars to what people can achieve with text chat, would instead address a widely known and used communication medium without any distraction.

Two other conditions, crossed with the avatar versus no-avatar conditions, were the use of synthesized speech versus scrolling text. Apart from noting that people typically didn't like the synthesized voices, this part of the study won't be discussed further here. However, the fact that each subject could only be assigned to 2 instead of all 4 conditions in this 2x2 design (although balanced for order effects and only assigned to

adjacent cells) due to practical constraints, , made the analysis of the data more difficult and contributed to lower power than with standard within-subject experiments. Nevertheless, some clear results emerged.

The 14 subjects that tried both an avatar system and a system without avatars were first asked to compare the systems on a 9 point Likert scale from a high preference for no avatars to a high preference for avatars along 6 dimensions including which system was “more useful”, “more fun”, “more personal”, “easier to use”, “more efficient” and “allowed easier communication”. One tailed t-tests showed that the preference for avatars was significant ($p < 0.05$) for all but the “easier to use” question where no significant preference either way was found. These subjective results clearly indicated that people found the avatars compelling and helpful. Of particular interest is that the avatars delivered this greater subjective impact at no extra cost to usability.

To test whether the Spark avatars improved the overall process of conversation, compared to text-only messaging, 11 different measures of quality of conversation process were taken. Seven were objective behavioral measures from the chat logs, including the portion of utterances without explicit grounding (i.e. verbal verification of reception), portion of questions that got replies, portion of non-overlapping utterances and portion of on-task utterances. Four were subjective Likert scale questionnaire measures, including sense of ability to communicate and sense of control over conversation. All but one measure was found higher in the avatar condition and a t-test of the grand mean (across all 11 normalized measures) showed that indeed it was significantly higher ($p < 0.02$) in the avatar condition than in the non-avatar condition, supporting the hypothesis that the Spark avatars were improving communication.

To test whether the Spark avatars would improve the outcome of the collaboration, compared to text-only messaging, 8 different measures of the quality of task outcome were taken. Two were objective measures, one being the quality of the map route that the subjects chose together and the other being the completion time (which ranged from 5 to 40 minutes). Six were subjective Likert scale questionnaire measures including, “How well did you think the group performed on the task?”, “How strong do you think the group’s consensus is about the final solution?” and “How much do you think you contributed to the final solution?” Again, all but one measure was higher in the avatar condition, and again, a t-test of the grand mean (across all 8 normalized measures) showed that it was significantly higher ($p < 0.02$) in the avatar condition than in the non-avatar condition, supporting the hypothesis that the Spark avatars were improving the collaboration in various ways, although interestingly the resulting solution to the puzzle was not significantly different.

12. BALANCE OF CONTROL

How do you know how much of the avatar behavior in general should be left up to automation? The short answer is that it depends entirely on the context of use. But for each context there are several factors that need to be considered. Perhaps the most important thing to have in mind is that ultimately the users should feel in absolute control of the situation they are dealing with, which possibly may be achieved through greater automation at the behavioral level. For example, being able to tell your avatar that you wish to avoid certain people may free you from having to worry about accidentally inviting them to chat by making an unexpected eye contact.

There are other factors to consider as well. First of all, the avatar may have access to more resources than the user to base its behavior on. These resources include the virtual world in which the avatar resides. Beyond what is immediately visible, the avatar may even be able to use senses not available to the human user. In the example above, the avatar would be able to know whether the person you are trying to avoid is standing

behind you and therefore would not make the mistake of turning around to face them. Time is also a resource, and sometimes it is crucial that an avatar reacts quickly to a situation. A time delay from the user to the avatar could force control over the situation out of the user's hands.

Related to the resource of time, the avatar can maintain consistent continuous presence in the virtual world even if the link from the user is a discrete one. The discreteness may be the result of a physical link that can only support control commands in short bursts, or it could be that high cognitive load requires the user to multi-task. In either case, delegating control to the avatar may ensure that the remote operation is not interspersed with abrupt standstills.

Although an avatar is meant to be a representation of a user, it does not necessarily mean that the avatar can only mimic what the user would be able to do. In fact, the avatar is an opportunity to extend the capabilities of the user, even beyond the capability of being in a remote place. For example tele-operated robots, which in a sense are physical avatars, may be able to perform operations such as changing a valve at super-human speeds. The user, or operator, may therefore want to leave the execution up to the robot after making sure it has been maneuvered into the right spot. Similarly, in a social setting, an avatar could have certain nonverbal behavior coordination skills programmed that are beyond what the user would be able to orchestrate. A user could for example choose an avatar that knew how to produce the gestural language of a riveting speaker, leaving the exact control of that skill up to the avatar itself.

13. THE FUTURE OF AUTOMATED AVATARS

At the beginning of this chapter, the point was made that introducing automation to avatar animation was a way to bring them alive, but then we saw that making the automation purposeful and driven by a model of face-to-face communication could significantly augment the social experience. However, it is unlikely that the creators of virtual worlds would be able to spend considerable resources on developing the social skills for these new kinds of agent-based avatars. Therefore, what has to happen for this new generation to take hold is that such skills need to be packaged and made available as plug-ins or engines that can be dropped into any virtual world without too much effort.

Consider how the availability of 3rd party physics engines has made the automated animation of physical behavior ubiquitous in virtual worlds today. The next revolution will likely happen in the area of life-like social animation, based on the rules of face-to-face and social behavior. Linking an avatar up to a set of virtual sensors that perceive the social environment and a social behavior generator might become part of the virtual world development process. This packaging is the subject of an ongoing research project at CADIA (Pedica and Vilhjalmsson 2010).

Another shift we may experience, along with the greater emphasis on skillful and complex animation of avatar abilities, is a greater demand for customized behavior that sets the avatar apart from other avatars. Such customization would still have to comply with the general behavior model, but there is room for endless variety. We can expect this based on how much time and effort users spend on customizing their avatar appearance. If given the choice, users are likely to tweak how their avatar walks, greets and makes a point. There is an obvious issue with complexity, but future user interfaces can help manage that by making behavior customization intuitive with the right level of abstraction.

When you enter a virtual world in the future, you will be struck by the richness of movement exhibited by the avatars around you, but perhaps more importantly, you will notice that the movement reveals patterns of coordination, giving you an instant read on the social situation evolving around you. Just like in real life.

14. BIBLIOGRAPHY

- Adalgeirsson, S. and Breazeal, C. (2010) "MeBot: a robotic platform for socially embodied presence". In *Proceeding of the 5th ACM/IEEE international conference on Human-robot interaction (HRI '10)*. ACM, New York, NY, USA, 15-22.
- Argyle, M. and Cook, M. *Gaze and Mutual Gaze*. Cambridge University Press, Cambridge, 1976.
- Argyle, M., Ingham, R., Alkema, F. and McCallin, M. *The Different Functions of Gaze*. *Semiotica.*, 1973
- Bavelas, J.B., Chovil, N., Coates, L. and Roe, L. *Gestures Specialized for Dialogue*. *Personality and Social Psychology*, 21 (4). 394-405. 1995
- Cassell, J., Bickmore, T., Billinghurst, M., Campbell, L., Chang, K., Vilhjalmsson, H. and Yan, H., *Embodiment in Conversational Interfaces: Rea*. in *CHI*, (Pittsburgh, PI, 1999), ACM, 520-527.
- Cassell, J, Bickmore, T., Campbell, L., Vilhjalmsson, H., and Yan, H. "More Than Just a Pretty Face: Conversational Protocols and the Affordances of Embodiment", in *Knowledge Based Systems* 14: 55-64., 2001
- Cassell, J. and H. Vilhjalmsson , "Fully Embodied Conversational Avatars: Making Communicative Behaviors Autonomous." *Autonomous Agents and Multi-Agent Systems* 2(1): 45-64. 1999
- Chovil, N. *Discourse-Oriented Facial Displays in Conversation*. *Research on Language and Social Interaction*, 25 (1991/1992). 163-194.
- Colburn, A. R., Cohen, M. F., Drucker, S. M. , LeeTiernan, S. , and Gupta, A. "Graphical Enhancements for Voice Only Conference Calls," Microsoft Corporation, Redmond, WA, Technical Report MSR-TR-2001-95, October 1, 2001 2001.
- Doherty-Sneddon, G., A. H. Anderson, et al. (1997). "Face-to-Face and Video-Mediated Communication: A Comparison of Dialogue Structure and Task Performance." *Journal of Experimental Psychology: Applied* 3(2): 105-125.
- Donath, J. (1995). "The Illustrated Conversation." *Multimedia Tools and Applications* 1(March): 79-88.
- Duncan, S. *On the structure of speaker-auditor interaction during speaking turns*. *Language in Society*, 3. 161-180., 1974
- Garau, M., Slater, M. , Bee, S. , and Sasse, A, "The Impact of Eye Gaze on Communication using Humanoid Avatars," presented at *CHI 2001*, Seattle, WA, 2001.
- Goffman, E. *Forms of Talk*. University of Pennsylvania Publications, Philadelphia, PA, 1983.
- Goodwin, C. *Conversational Organization: Interaction between speakers and hearers*. Academic Press, New York, 1981.
- Inoue, T., k.-i. Okada, et al. (1997). "Integration of Face-to-Face and Video-Mediated Meetings: HERMES." *Proceedings of ACM SIGGROUP'97*: 385-394.
- Isaacs, E. A. and J. C. Tang (1994). "What video can and cannot do for collaboration: a case study." *Multimedia Systems* 2: 63-73.
- Isaacs, E. A. and J. C. Tang (1997). "Studying video-based collaboration in context: From small workgroups to large organization." *Video-Mediated Communication*. K. Finn, A. Sellen and S. Wilbur, Lawrence Erlbaum Associates, Inc.: 173-197.
- Kendon, A. *Conducting Interaction: Patterns of behavior in focused encounters*. Cambridge University Press, New York, 1990.
- Kendon, A. *On Gesture: Its Complementary Relationship With Speech*. in Siegman, A.W. and Feldstein, S. eds. *Nonverbal Behavior and Communication*, Lawrence Erlbaum Associates, Inc., Hillsdale, 1987, 65-97.
- Kurlander, D., T. Skelly, et al. (1996). "Comic Chat." *Proceedings of SIGGRAPH'96*: 225-236.
- McNeill, D. *Hand and Mind*. The University of Chicago Press, Chicago and London, 1992.

- Nardi, B. A. and S. Whittaker (2002). "The Place of Face-to-Face Communication in Distributed Work." *Distributed work: New ways of working across distance using technology*. P. Hinds and S. Kiesler. Cambridge, MA, MIT Press.
- Neale, D. C. and M. K. McGee (1998). "Making Media Spaces Useful: Video Support and Telepresence." HCIL-98-02. Hypermedia Technical Report. Human-Computer Interaction Laboratory, Virginia Tech
- Pedica, C. and Vilhjalmsón, H. (2010). "Spontaneous Avatar Behavior for Human Territoriality", in *Applied Artificial Intelligence*, Volume 24 Issue 6, July 2010, 575-593 Taylor and Francis, Inc. Bristol, PA, USA
- Rosenfeld, H.M. *Conversational Control Functions of Nonverbal Behavior*. in Siegman, A.W. and Feldstein, S. eds. *Nonverbal Behavior and Communication*, Lawrence Erlbaum Associates, Inc., Hillsdale, 1987, 563-601.
- Seif El-Nasr, M., Isbister, K., Ventrella, J., Aghabeigi, B., Hash, C., Erfani, M., Morie, J.F. and Bishko L. (2011), "Body Buddies: Social Signaling through Puppeteering" in *HCI 14*, Vol. 6774, Springer, p. 279-288.
- Shami, N.S., Cheng, L., Rohall, S., Sempere, A. and Patterson, J. (2010) "Avatars Meet Meetings: Design Issues in Integrating Avatars in Distributed Corporate Meetings", in *Proceedings of the 16th ACM international conference on Supporting group work (GROUP '10)*. ACM, New York, NY, USA, 35-44.
- Ventrella, J. (2011) *Virtual Body Language*. EyebrianBooks.
- Vertegaal, R. and Ding, Y., "Explaining Effects of Eye Gaze on Mediated Group Conversations: Amount or Synchronization," presented at *CSCW 2002*, New Orleans, LA, 2002.
- Vilhjalmsón, H. "Animating Conversation in Online Games", in M. Rauterberg (ed.), *Entertainment Computing ICEC 2004, Lecture Notes in Computer Science 3166*, pp. 139-150, Springer, 2004
- Vilhjalmsón, H. "Augmenting Online Conversation through Automated Discourse Tagging", 6th annual minitrack on Persistent Conversation at the 38th Hawaii International Conference on System Sciences, January 3-6, 2005, Hilton Waikoloa Village, Big Island, Hawaii, IEEE, 2005
- Whittaker, S. and B. O'Connell (1997). "The Role of Vision in Face-to-Face and Mediated Communication." *Video-Mediated Communication*. K. Finn, A. Sellen and S. Wilbur, Lawrence Erlbaum Associates, Inc.: 23-49.
- Yankelovich, N., Simpson, N., Kaplan, J. and Provino, J. (2007) "Porta-person: telepresence for the connected conference room" *CHI Extended Abstracts 2007*: 2789-2794

16

SYNTHESIZING VIRTUAL CHARACTER BEHAVIORS FROM INTERACTIVE DIGITAL PUPPETRY

By Elena Erbiceanu, Daniel Mapes, and Charles E. Hughes

1.

INTRODUCTION

A major challenge in creating interactive virtual environments is getting the user engaged and interested in the experience, providing the vital connection to the story that is needed to achieve a level of immersion that triggers affect responses. Although verbal communication is straightforward in conveying information, it is the nonverbal behaviors that create the symphony of cues that humans are inherently trained to recognize in person-to-person interaction. For instance, observing a character that is verbally communicating the same story while exhibiting different gaze behaviors (e.g., fast gaze shifts with no sustained eye contact versus prolonged, lingering eye contact) can trigger completely different affective responses, varying from untrustworthy / hiding something, to casual communication, and all the way to triggering an empathetic response and identifying with the character's emotions. For example, a character telling a secret is far more convincing when it leans across the table to whisper to you.

ELENA ERBICEANU ON HER METHODS

Our first stage of designing a role playing experience in Geppetto involves observation of the real characters to be modeled. This is typically an on-site study, e.g., with middle school children, and involves collaboration with members of the subject population. When this collaboration is not possible due to practical, ethical or epistemological concerns, the process is done via subject-matter experts.

Our actors and our subject population collaborators are then asked to act out various emotions for each virtual character, according to his/her persona, while being recorded. The videos are then analyzed to extract the most expressive key frames, which are then used to mimic the behaviors previously recorded to model the virtual character. We do this iteratively until everyone is happy with the articulation of the various emotions.

(continued on next page)

In addition to inherently conveying a wide range of believable emotions, nonverbal cues lend authenticity of the characters, not just in displaying human characteristics, but as members of specific populations, cultures or age-groups by employing cultural, social or generation-specific gestures, memes and attitudes. Moreover, gestural cues can enhance or dampen the strength of the message being verbally conveyed, which is often apparent with public speakers and theatrical performances, and could even negate what is being said entirely. Simply said, nonverbal cues add a whole new dimension to communication.

The design of virtual humans that exhibit believable, natural behaviors is a key aspect of virtual worlds that rely on triggering an emotional/empathic response. These kinds of responses are important for a large class of applications that involve social training or situations, including medical, educational, entertainment and military training scenarios. Although expressive behaviors can be scripted for most of the characters appearing in a virtual world in a similar fashion to the animation industry, when direct interactivity with the user comes into play this becomes a daunting task, especially for open-ended conversation.

This chapter describes an initiative to ‘borrow’ the ability of natural interaction from a human that plays the role of the virtual character. The main elements of nonverbal communication (NVC) are identified, captured and transferred in real-time onto the character in the virtual world, thus obtaining human-like interactive virtual characters. We describe several incremental versions of the Geppetto digital puppetry system, which aims to provide an intuitive way to control a character’s nonverbal communication cues by a theatrically-trained human while conversing with the user. Furthermore, we explore the avenue of encoding these behaviors into intelligent agents that control the virtual characters when the human puppeteer gives up his control. We propose a mixed initiative paradigm to capture, learn and create behaviors for groups of interactive virtual characters.

Once the pose sets are decided upon, we do rehearsals of the experiences. These rehearsals are observed by subject matter experts, with adjustments that can lead back to further videotaping, modeling, rigging and more rehearsals.

Often two or more puppeteers collaborate during rehearsals. This allows each to understudy the other, as well as define the persona of each character even further, achieving consistency in performances. That is, the characters have essentially the same personalities, independent of who is puppeteering.

Once rehearsal is completed, we start play testing in order to do formative evaluation. This is where we start to adjust pacing and often add or tone down aspects of characters based on observed reactions – empathy, engagement, etc.

The final stage is delivery. How much adjustments, if any, can be made in this phase depends on the purpose of the experiences. If it is a controlled study, then the system is locked down and the performances are constrained. If it is a training activity, then we adjust as often as necessary and quite commonly based on each customer’s needs.

This active research is still evolving, embedding new and exciting control paradigms as soon as they become available and informally investigating their usability through dialogue with the trainers/puppeteers.

The data gathering for behavior synthesis of the background characters is designed to be transparent to the puppeteers, which are being observed in this process but do not have to change their actions. The accuracy of the synthesized behaviors will be discussed with the puppeteers

(continued on next page)

2.

VIRTUAL CHARACTER BEHAVIORS

Also known as embodied conversational agents (ECA), virtual characters enabled with the power to fully interact with the user on a more elaborate social level present an open problem. There is a

notable difference between a Spoken Dialogue System, which limits the domain of the conversation significantly by focusing on a specific task and relies mostly on computing the linguistic dimensions of communication, and an ECA who is endowed with an identity and higher-level social skills that are more similar to human-human communication. Advances in speech recognition, natural language processing and expressive speech synthesis can bridge the gap between a Spoken Dialogue System that can carry a basic conversation and an ECA. However, the latter encompasses a whole new dimension of communication that needs to be addressed: the expression of emotion and intent through nonverbal cues.

We define the *NVC behavior* of a virtual character as the sequence of actions over time that manifest on both an auditory and a visual level creating a stream of non-verbal communication. Note here that this chapter views language content, i.e. the actual sentences words and ideas being conveyed, as verbal communication, while speech and voice traits such as pitch, speech rate, pauses and inter-breath stretches are considered as both verbal and nonverbal cues. Another way to look at it is as the script of a theatre play: verbal communication represents the set of ideas to be expressively spoken by the character, e.g., the lines, while the congruence of all factors that are implicitly detected by a human observer, e.g., the director's comments of how the piece should be acted, adds the elements of nonverbal communication. From a visual perspective behaviors include the set and qualifying traits of upper-body or full-body poses and motions, micro-gestures such as facial expressions, gaze activity, etc.

We identify two aspects in achieving believable nonverbal communication for an interactive ECA: (1) constructing an explicit set of NVC tools that the character can employ in conversation; (2) deciding when to use each of these cues, accounting for the human end-user's actions and emotional state and for the state of the experience. A successful combination of these two factors results in the character conveying a direct or indirect message to the user, as well as a believable emotional state. Although the first part of the problem (constructing the explicit set of NVC cues) can be designed on a per-scenario basis with experts in nonverbal communication, it is hard to find a seamless solution to the second (choosing and seamlessly blending the animations considering time and other factors to make characters behave human-like). A mismatch in timing, even though the gesture is adequate, can make the behavior look very unnatural. An inexperienced public speaker trying to emphasize a word with a wide arm motion, but doing that just barely too late, will miss the entire power of that gesture.

A viable solution to achieving a human-like interactive ECA is to 'borrow' the difficult communication aspects of open natural one-on-one interaction, both linguistic and nonverbal, from a human counterpart that is playing the role of the character. This means that a trained puppeteer controls the voice, body pose and motions, posture and basic facial features of the virtual character that is currently talking to the user. This control can be achieved through capture modalities that range from explicit one-to-one body and facial tracking to abstract ones that employ various human-computer interface devices where a button can trigger a smile and raising a hand can make the character stand up from his seat. Provided

in a similar iterative process until it can reach a prototype state. Then, user studies will be performed from the perspective of the trainee, to assess the believability of the background characters, whether they move unnaturally, too much or not enough, whether they seem to maintain their perceived emotional state and act according to their personality.

that the mapping from the puppeteer's actions to the character's movement happens in real time, the timing issues and adequacy of the gestures for the current state of the conversation are practically solved. If the system that captures and reproduces these gestures onto the ECA is flexible enough, the interaction can be powerful and feel completely natural, achieving a much more engaging experience.

Last but not least, we emphasize one of the main characteristics that drive the ECA's behavioral patterns: emotion. If we can capture the correlation between the behavior of a character that is being played out by a human puppeteer and the emotion they are trying to convey, then we could use this correlation as a behavioral model for that character to partially recreate the nonverbal communication given a desired emotion. How would the introvert student character behave if they were anxious? Can we trigger a believable anxious state of the ECA without the input of a human puppeteer? We propose the idea of an intelligent agent that can take over the control of the virtual character as soon as the user switches focus to another character or object in the world, maintaining the emotional state that was induced during interaction and acting appropriately according to the generally established identity of the character.

3. TRANSFERRING NVC FROM A HUMAN TO A VIRTUAL CHARACTER: DIGITAL PUPPETRY FOR INTERACTIVE PERFORMANCE

Using nonverbal communication to convey emotion is widely studied in theatre and dance, and researched in scientific disciplines such as neuroscience, psychology and cognitive sciences, but has yet to thoroughly permeate research and practice in virtual worlds. We introduce the notion of puppeteering as a solution for profound, believable interaction with a virtual character, where a human puppeteer controls the virtual character's voice and body language in the experience. In the prototype phase of designing an experience we employ interactors: people who have theatrical training and experience in dramatic improvisation, which helped hone their skills in nonverbal communication. These professionals create fully developed characters that get the users engaged and emotionally connected. In creating the characters, they develop the background story and persona of each character, relationship within the group as well as detailed nonverbal communication aspects of that character, including preferred body poses and gestures during interaction, defining facial expressions and so on. Based on this prototype of the experience, an interactor can easily enact the character in live interaction with the end-user. A digital puppetry system captures the actions of the human puppeteer and transfers them onto a virtual character in real time. The ultimate goal is to create a system that is easy enough for subject matter experts to learn and manipulate, so they become the trainers/puppeteers who can effectively manipulate the virtual characters that are interacting with the user.

The *Gepetto* virtual reality digital puppeteering system evolved over time transitioning from one way of capturing non-verbal behaviour to another. The first method we used was motion capture, which is straightforward to use requiring the person controlling the character to simply act the part; this results in full freedom but is quite physically demanding for the puppeteer. The second is hand puppeteering of the virtual character's body, which restricts the set of gestures and motions that the character can perform and requires a training and adjustment phase. However, we developed a well-designed set of poses that puppeteers can use. Thus, once they learn the interaction system and the set of poses it is just as powerful and less demanding than the motion capture system.

The result is an efficient one-to-many digital puppetry system, where a single human puppeteer controls multiple avatars, one at a time, throughout the interactive experience. The puppeteer is able to "step in" and "step out" of a character and perform the corresponding part, while the out-of-focus characters remain in

an expectative state. The disconnect between the puppeteer who is controlling the virtual character and the actual audio/visual experience that is presented to the user, who potentially does not even notice that there is a human behind it, allows us to run the interactive experiences remotely, the communication data being transmitted live over the Internet. The fact that the puppeteer and users going through the experience do not have to be co-located, or even on the same continent, gives this approach significant capabilities for training and education scenarios.

4. THE GEPPETTO DIGITAL PUPPETRY SYSTEM

The Geppetto digital puppeteering system employs simplified models of nonverbal communication that are learned from experts in interactive performance (Wirth, 2011); the resulting experiences have believable, human-like virtual characters that can interact with the user in natural conversation. The interactive performance experts are involved on two levels: (1) designing a meaningful pose set for each virtual character (the set of body poses and facial expressions that can be mixed to create believable nonverbal communication) and (2) controlling the mixture of poses for a character in real-time while interacting with the user. The first happens in the design phase of a new experience and is vital to providing authenticity to the characters, building their persona. The puppeteers freely play out the roles and decide the key poses that are most meaningful and expressive for a particular experience. The second level is the active control of the character in real-time while the experience is running, when the puppeteer decides which nonverbal behaviors to adopt for the character currently being played out.

Geppetto was successfully used for a range of very different virtual reality experiences. First, a teacher training project that is currently still running and receiving a lot of interest with its virtual classroom populated by children exhibiting a variety of personality types commonly encountered in middle schools. Second, a cultural training scenario where, with the drawback of having to use the same spoken language in both settings (English), authenticity of the non-verbal communication was vital to create a meaningful experience. The third project, targeting peer resistance skills development was even more challenging since it was vital for the adolescents to be highly immersed in the experience, and they are a lot more sensitive and attentive to ‘fake’ behaviors. Whether the virtual scene is at school or in a friend’s home, the young participant must make decisions that could lead to risk, e.g., agreeing to go to an unchaperoned party or accepting the advances of an older boy. Peer pressures can engender responses that result in perceived isolation, risk or acceptance. Verbal and non-verbal communication from the puppeteered avatars reinforces personality types and provides in-character reactions to the participant’s decisions and her ways of expressing these decisions.

These studies have different control paradigms that evolved over time. However, the common thread to all three is the use of a human subject matter expert, trained to act as a puppeteer, to decide in real time the exact actions of the virtual characters that are interacting with the user. The presence of the human controller is essential to making the interaction natural in a non-constrained context – every single run of an experience is unique, leading to quite different conversations on the same topic. In the following, we provide vignettes of these three projects that showcase the capabilities, as well as the limitations of digital puppeteering: teacher training, cultural awareness in negotiation, and helping middle-school girls to deal with peer pressure. We discuss the main idea of each project, a few details of the control paradigms for the virtual characters, as well as the range of nonverbal communication they can exhibit.

TEACHER TRAINING: TEACHME

Motivations. The TeachME (Teach Mixed Environment) experience provides pre-service and in-service teachers the opportunity to learn and practice teaching skills in a virtual middle school classroom under self- or external supervision, but without affecting “real” students during the learning process (Dieker, M., Hughes, & Smith, 2008). The scenarios must be very flexible to include specific student behaviors that the trainee must encounter and subject knowledge that they need to focus on. The teacher-trainee, typically in the pre-service phase of preparation, must quickly assess and address the needs and motivations of each student, observing both verbal and nonverbal communication aspects in the student’s behaviors. Digital puppeteering allows practically unique, open-ended conversations with each character, making the experience fun and challenging to the teacher.



Figure 16-1: Puppeteer controlling avatars in teacher training.

System design. The classroom is composed of five mixed-gender virtual students that span various races and ethnicities, each with its own prototypical behavior. Each time he or she approaches an individual student, the puppeteer, located remotely, takes control of that student, providing nuanced interaction (verbal and non-verbal) specific to that character.

Overall, the concepts developed and tested with the TeachME environment are based on the hypothesis that performance assessment and improvement are most effective in contextually meaningful settings.

Capturing the puppeteer’s actions in this experience relies on two different types of direct tracking for both the head and upper body movement of the puppeteer [Figure 16-1]. Due to the need for higher accuracy of position and orientation for the *head movement*, we employ three retro-reflective markers on the puppeteer’s headset and track them as 3D points using a two-camera infrared (IR) motion capture system.

Body movement is coarsely tracked with a Microsoft Kinect (“Kinect,” 2011), mapping the puppeteer’s movements one-to-one to the virtual character’s body, allowing for coarse control of the position and orientation of the body and limbs of the character.

Facial features are controlled through foot buttons. The set consists of four basic facial poses: smile, frown, wink and mouth open that can be simultaneously blended to convey a minimal set of facial visual cues.

Switching control between characters is either done automatically, based on the teacher’s movements (e.g., if the teacher moves backward after talking to one of the students, the camera zooms out to allow the full perspective of the classroom; if she moves explicitly toward a student, the camera zooms in to the close perspective), or with the use of an Ergodex game pad (*DX1 Input System*, 2008). The Ergodex is a configurable keypad customized to allow the puppeteer to both rapidly switch control to any of the five middle school age virtual characters and provide a custom key interface to control the scenario: advancing to different story levels, very basic camera control and even group actions like general laughter in the classroom.

Additionally, bi-directional audio and uni-directional video capture (of the trainee) is transferred through a Skype connection. There is no additional processing of the sound between the puppeteer (trainer) and trainee, since our professionally trained puppeteers are very adept at switching between particular character voices, but one can envision using voice modulation to alter the trainer's voice when using subject matter experts to control the characters. The puppeteer uses the video feed of the trainee to conduct the interaction appropriately.

Observations. Due to the one-to-one mapping between the puppeteer's movement and the body of the ECA, the virtual students of the TeachME classroom were able to exhibit a large range of nonverbal communication. The students had the ability to change body posture to lean forward when trying to pay attention; look to their right if they are talking about their classmate or point their arm in the direction of an object that is subject to the conversation; wink, smile or frown as a result of the teacher's words etc. Having a human being performing this control in real time while the interaction is happening was essential to achieving this level of natural communication.

The human-like behaviors caused a high level of engagement of the teachers being trained, who would forget soon after starting the experience that they are talking to virtual avatars. Throughout the project each virtual student had developed its own persona and gained the sympathy of the users. They were not regarded as robotic characters; on the contrary the teachers going through the experience got fully engaged and even somewhat emotionally attached to the characters' story. This engagement was so strong that it has led quite a few teachers to feel disappointed when they failed to get through to a specific child. That disappointment motivates them to want to quickly reenter the TeachME environment to try again to be successful with all the personality types represented by these virtual students.

From the puppeteer's perspective, this control paradigm was the most natural to use. There is almost no cognitive load when controlling the ECA by simply acting out the role. However, the range of motion of the puppeteer must be equal to the character's, which implies a level of physical exertion that can become tiring. The button- and pedal-based interfaces that generated the basic facial expressions and character switching were simple to use and did not require extensive training.

CROSS CULTURAL SCENARIO: AVATAR

Motivations. The Avatar project needed to allow adult users to practice their cross cultural communication skills (Barber, Schatz, & Nicholson, 2010), however with the drawback of having to use the same spoken language overall (English). Thus, a key goal was the creation of virtual characters with authentic and natural nonverbal communication. The study was based on a comparative experience between two different scenarios: one in the Pashtun region of Afghanistan [Figure 16-2] and one in an urban setting of New York. Both scenarios contained virtual characters with the same three personalities: one elder leader, a young idealist, and an antagonist. In each scenario the participant would practice negotiation with the group, which was controlled, one character at a time, by the puppeteer/trainer.

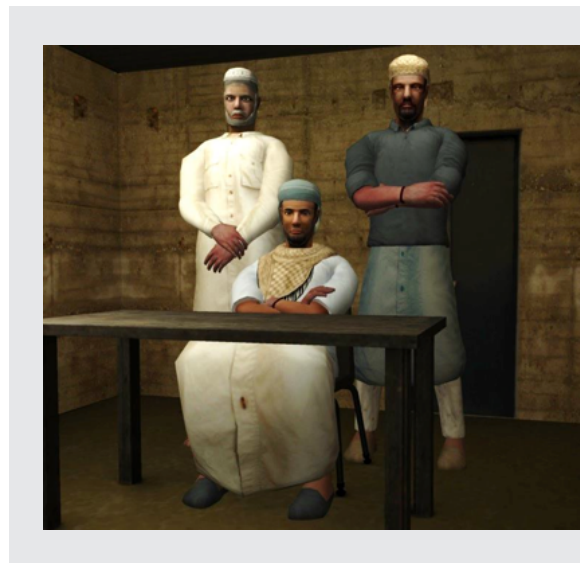


Figure 16-2: Pashtun avatars in cross-cultural training

System design. Due to the existence of cultural differences in the non-verbal communication that conveys an affect or emotion of the virtual characters (A. Kleinsmith, De Silva, & Bianchi-Berthouze, 2005; A Kleinsmith, De Silva, & Bianchi-Berthouze, 2006), using a direct motion capture of the puppeteer's body to control the characters would require extensive cultural training of the puppeteers as well as excessive cognitive load. Moreover, basic body tracking works well on a coarse level (head, body and limbs); however, it is missing the fine-grained details like the wrist and finger positions. These fine-grained details of nonverbal communication are necessary to seamlessly convey the appropriate cultural differences and bring authenticity to the scene.

Thus, a second control paradigm had to be designed for this experience that constrained the space of possible poses that a virtual character can assume, thus lowering the risk of gestures that are not culturally appropriate, as well as allowing more refined and complete poses. In the design phase of the scenario, the interactive performance experts and cultural experts defined a set of full body end-poses for each virtual character [Figure 16-3]. These poses are characteristic of each character's non-verbal communication patterns, defining his behavioral persona from a social and cultural point of view. The poses can be modeled as finely-grained as needed, from hand and body motions up to facial and hand and finger details.

Intuitively, the reverse mapping from each character pose to a control action by the puppeteer, uses an empirical approach that leads the character using its wrists as control points. To determine the most natural location of the puppeteer's wrists if they wanted to embody that character, each characteristic pose was presented to the puppeteer, who was then asked to mimic the pose. The pairs of real world wrist locations and character's wrist locations were then used to determine the motion space in which the puppeteer operated.

During the actual experience, the puppeteer moved freely in the real space. At each point in time, the location of his/her wrists was tracked; the closest two end-poses to the corresponding wrist position in character space were retrieved and, using a weighted blend, the current character pose was obtained. This resulting pose could be anywhere between the two end-poses, e.g. if the two end-poses were: stand straight



Figure 16-3: Key poses for a single avatar

with arms down along body, and hold arms straight to each side perpendicular to the side of the body, the blended pose could be: arms hanging down sideways at a 45 degree angle to the body. In addition to the two current end-poses, short-term historical data was also used in the blend with weights decaying to zero over time.

The same type of dual infrared camera motion capture system was used to track the puppeteer's head orientation and the location of the wrists. Facial poses are controlled using the same five-finger glove metaphor from TeachME. However, the Ergodex game controller was replaced with a WiiMote held in the non-gloved hand.

Observations. The new control paradigm allows the creation of virtual characters that are constrained to exhibit only culturally appropriate gestures, although there is a possibility of the blending process to generate undesirable gestures, these can easily be avoided by modifying the end pose set and/or 3D mapping, yet are still freely controlled by the puppeteer. The gestures are selected seamlessly using retro-reflective markers held either in the puppeteer's hands or placed on their wrists, while the puppeteer moves the puppet's head directly by their own head movement. Since the puppet poses and the controlling hand locations are mapped nearly one-to-one, this system is almost as intuitive to use as raw motion capture but with fewer motion artifacts and more pose details.

The gestural communication in this paradigm is significantly more natural compared to the generic motion capture of the body movement, since it allows extremely detailed stances, up to the exact positioning of each finger. The puppeteer cannot only lean across to whisper a secret, she can place her hand next to her mouth to hide her whispering from nearby people. The gestures become so natural that they are seamless, to the point that the end users do not have to actively exercise their cognitive skills to understand the implicit tension and progress in the current negotiation, as they would in a classic training scenario where they are merely told or asked to read that same information.

There were still a few downsides to this approach. Although the mapping of the body motion of the puppeteer is almost one-to-one at a coarse level with the resulting pose, some poses were ambiguous enough to be hard to select, such as the difference between a hand being held palm up or palm down. This stems from the limitations of using only two wrist markers for the tracking, as well as from the difficulty in finding a natural mapping that discriminates well between poses. Another hindrance is the need to maintain individual calibration settings for each puppeteer as they changed shifts.

DRAMA-RAMA

Motivations. The DRAMA-RAMA project employs an interactive virtual environment to help middle school girls develop and hone skills to resist/avoid peer pressure [Figure 16-4]. The need for human-like interaction with the virtual characters is accentuated by the fact that the target end-users, adolescents, are sensitive and attentive to 'fake' behaviors. Nuances in the body language of the character are important in conveying veiled messages throughout the scenario, e.g. the 'cool' virtual boy Xavier (Figure 16-4, right) fakes shyness when trying to talk the end-user into giving him a kiss for his birthday present.

System design. The previous control paradigm used in Geppetto had allowed complex gestures and poses for the virtual character. However, the high-level body motion was still mapped almost one-to-one, requiring the puppeteers to fully act out the parts. This motivated a transition to a new control interface through which the trainer/puppeteer can simply express intentions for the nonverbal communication to be exhibited by the virtual character instead of exhibiting the actual gestures. This new paradigm balances the need for freedom in the control of the ECAs with the need to reduce the physical strain on the puppeteer.



Figure 16-4: The characters from DRAMA-RAMA are playful cartoon renderings rather than realistic ones, based on strong feedback from our SMEs (9th and 10th graders who had recently graduated from middle school). Each has its own strong personality ranging from the shy, to the wise, to the troublemaker.

As in the previous experiment, detailed characteristic poses were defined for each virtual adolescent to match his/her style of communication, according to the background story and personality assigned to the character by the subject matter experts and interactive performers in the design phase. The nonverbal communication experts identified subsets of these poses that were intended to be used together. Thus, the poses for each character were grouped into bins based on the behavior mode within which the puppeteer would be operating (e.g. seated poses vs. standing poses). To select between characters and behavior modes for each character, the puppeteer moves a retro-reflective marker to the corresponding bin on a reference sheet while seated at a desk. The retro-reflective marker is tracked as a single 3D point using the same IR dual-camera system from the previous iteration of Geppetto. Once a character behavior bin is chosen, the selection of the *body pose* is a hybrid between motion capture and a game controller. The marker is moved within a virtual 3D volume above the desk surface, relative to the character reference sheet.

- For ease of use, the characters are placed on the reference sheet in the same order as they appear on screen relative to each other. Each character has one to three associated pose bins; these are represented as white disks. The arrangement of the pose bins on the control map sets the standing poses at the top of the map (farthest from the puppeteer, when the reference sheet is placed horizontally on the desk in front of him), followed by seated variations.
- A character with the desired pose bin is selected by lifting the marker from the control map at the corresponding disk; once passed above a fixed height (the activation threshold) the character and pose bin are selected.
- The end pose is triggered by moving the marker freely in the 3D volume above the desk. The representative end-poses for the character are mapped to points in this volume; at each moment in time the two closest poses to the marker are blended to

determine the current active pose; the closer the marker gets to a characteristic pose point, the higher the weight of that pose is in the blending.

- Placing the marker below the activation threshold returns the system to selector mode and releases the current puppet to an automated non-puppeteered state (typically called non-player controlled characters or NPC's).

The *head orientation* is directly tracked by the same IR camera system measuring the control marker, with a group of retro-reflective markers placed on a cap that the puppeteer is wearing (the system tracks both the single-point pose selector and the multi-point cap at the same time; it is easy to differentiate between the two by location history and expected distances/relation between the points). This data is used to modulate the head position in the current active pose.

Facial pose blending is triggered with a Logitech G13 advanced gameboard that replaced the Ergodex as a simple button interface, allowing a more complex facial morphing than the basic poses triggered previously with the foot buttons, as there are more keys to employ. Specifically, we now have smile, frown, sneer, wink and open mouth that can all be blended to create combinations. These poses are weighted by the length of the key press and have a soft decay to allow rich blending. Blinking is automatic. An alternative lip tracker using 4 IR markers attached around the puppeteers mouth was explored but the puppeteers preferred the simplicity and control afforded by the game board.

The bidirectional audio for the puppeteer and user is similarly transmitted through a Skype connection, as well as the video stream of the user going through the experience.



Figure 16-6: Puppeteering avatars in a demonstration session.

Observations. This new control paradigm is closest to classic puppeteering than all the previous approaches Geppetto employed, in that it allows control actions instead of the puppeteer having to exhibit the actual behavior of the virtual character. Using blended end-poses affords a great deal of control over the details in the body language of the ECA, while triggering these poses by mere hand movement reduces the physical strain from the puppeteers at the cost of some preliminary training to learn the system.

One important aspect of how the control scheme is designed is the support of natural blending between poses and gestures. For the facial features, it creates weighted blends involving all facial poses simultaneously (e.g. the character winks with a half-open mouth). Having multiple full body poses mixing over time allows Geppetto to transcend the plastic look often associated with digital puppets, thereby bringing the characters to life. For the body poses, the use of the continuous 3D volume as pose selector allows the system to practically yield any number of intermediate poses that the puppeteers can freely decide to use to convey a deeper emotional meaning to the character's actions. By using a well-designed set of end-poses, the blended poses generated by the interpolated points in the 3D volume are close to the natural transitions a human would make, although at times a little theatrical.

From the perspective of the end-user, this translates to human-like virtual characters that allow a high level of immersion in the experience. When talking to Xavier alone, some of the middle-school girls that participated in the study even exhibited flirtatious gestures without realizing, such as playing with their hair or touching their neck and smiling.

From the puppeteer's point of view, the use of the marker as pose-selector gave a lot of flexibility with minimal physical strain. The cognitive effort was increased for the training period; however after a few days the puppeteers were able to use this control method efficiently. Moreover, having the character switching and pose selection controlled through the same marker interface allowed the puppeteers even smoother transitions from one character to another, creating more fluid group behaviors.

5. CAN WE AUTOMATE BEHAVIORS FOR VIRTUAL CHARACTERS?

In each of the scenarios we described in this chapter Geppetto employs one human puppeteer to control a group of virtual characters. Although multiple ECAs are on-screen during the experience, only one character is controlled at any point in time by the puppeteer – the one that is interacting with the user – while the remaining out-of-focus characters are in an automated expectative state. The solution of having multiple puppeteers in an experience is possible, but in fact unnecessary, because the scenarios discussed above often require complex personal interaction, which is usually done by a single SME. Desirably, the characters that are currently not in focus would exhibit toned down, mostly non-verbal versions of their behavior when they were under the puppeteer's control.

Thus, several questions arise. Can we infer the parameters of the behavior of the ECAs from their past behavior under the control of the puppeteer? What are the factors that define the state of the character at any point in time and how can we measure these factors? How do we create an intelligent agent that learns in real time what the human puppeteers are doing when embodying a virtual character and tries to simulate plausible off-focus non-verbal behaviors that stay within the realm of expected actions and reactions by that character?

CAPTURING EMOTION

There has been extensive research in the field of affective computing. However, the problems of modeling, analyzing and understanding natural human behavior still remain challenging for a computer to solve (Gunes, Schuller, Pantic, & Cowie, 2011). Emotions are inherently complex, hard to define and often hard to recognize in their subtle forms. Yet even the simplest models of emotion-driven behavior can play a critical role in the believability of virtual humans, and thus the successful immersion into a virtual experience (Lester & Stone, 1997; Van Mulken, André, & Müller, 1998).

Emotion space. Gunes et al.(2011) identify in a recent survey the fact that there is no agreement of how to model dimensional affect space (e.g. continuous vs. quantized) due in part to the multitude of possible communicative cues and modalities. Thus, the problem of affect recognition has been reduced to spaces that range widely in complexity. The description of discrete categories, e.g. happiness, sadness, fear, anger, disgust and surprise, is one of the more natural ones (Ekman, 1971, 1982). The definition of affect as a two-class problem (positive vs. negative or active vs. passive) appears both in the domain of audio (Schuller, Vlasenko, Eyben, Rigoll, & Wendemuth, 2009) and visual signals that focus mostly on facial feature recognition (Nicolaou, Gunes, & Pantic, 2010).

Rather than relying on a discretized categorical space that attempts to distinguish between a fixed number of independent emotional states, one school of researchers advocates the use of a dimensional description of the human affect, where the affective dimensions are highly inter-related. The most widely used model in this respect is a circular configuration that spans the two-dimensional space of *arousal – valence*: the *Circumplex of Affect* (Russell, 1980). An extension of this model is the 3D *PAD emotional space*, which includes the new dimension of dominance-submissiveness next to the pleasure-displeasure, and arousal-nonarousal dimensions (Mehrabian, 1996). There are many more variations of these models, taking into account expectation (Fontaine, Scherer, Roesch, & Ellsworth, 2007), intensity (McKeown, Valstar, Cowie, & Pantic, 2010), etc.

Modalities. Zeng et al. discuss the need for multi-fusion approaches (combining the input signals from multiple modalities of affect expression) for emotion recognition, as well as current methods and details on using feature extraction for each of the communicative modalities in spontaneous (as opposed to posed) displays of affective states (Zeng, Pantic, Roisman, & Huang, 2009). Several theoretical and empirical studies also demonstrate that incorporating both vocal and visual expression for the perception of affect leads to better results compared to the use of single modalities (Russell, Bachorowski, & Fernandez-Dols, 2003; Russell & Mehrabian, 1977).

One interesting aspect of multi-modal natural communication is that a great deal of information can be conveyed even in fleeting glimpses of expressive behavior; thin slices of behavior of under five minutes have been shown to be almost as powerful as wider windows of observation (Ambady & Rosenthal, 1992). This encourages the belief that, using only a small window of behavioral factors, a simplistic behavior model could be constructed and updated in real time, as the puppeteer is embodying the virtual character.

An older study analyzing motion picture film records suggests that the nature of an emotion is not conveyed only by a standalone, static body part, but by the synchrony of motion occurring in one or multiple body parts (Ekman & Friesen, 1967). This suggests that the idea of analyzing the transition between full-body poses, rather than tracking each individual limb, is feasible for affect assessment, which fits very well with our puppeteering paradigm.

THE VALENCE-AROUSAL MODEL

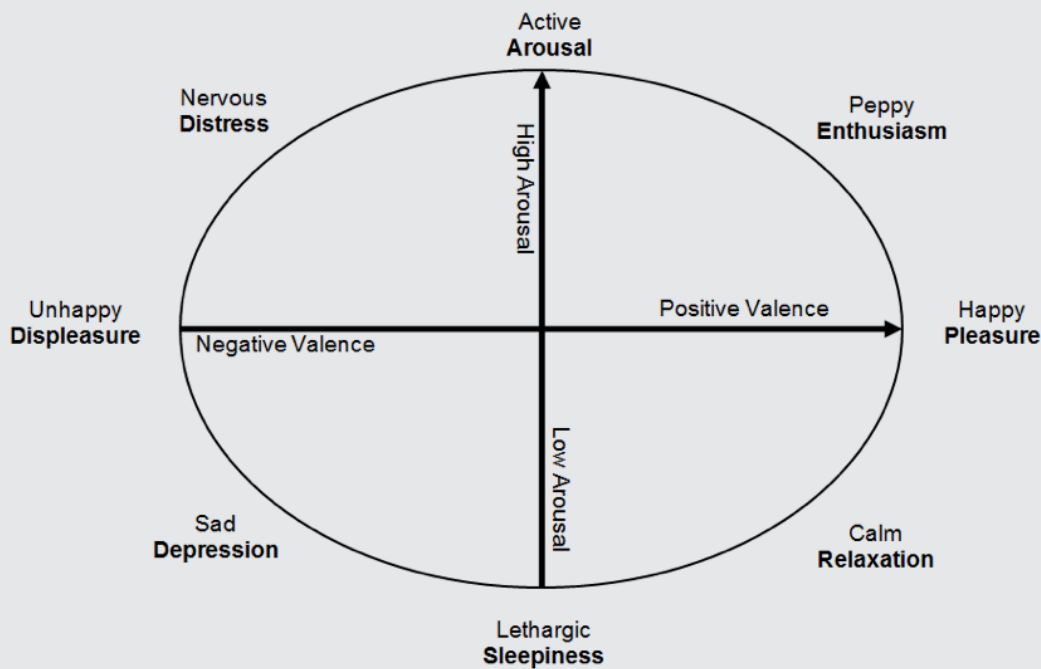


Figure 16-6: The Arousal-Valence Model.
Emotions are intuitively mapped in the circular continuum defined by the two axes.

The Circumplex of Affect, or the valence-arousal model (Russell, 1980), was suggested more than thirty years ago and is still one of the most commonly used spaces for emotion representation. It relies on the notion that emotions are not discretely bounded, but flow into each other in a natural circular fashion. It defines two bipolar entities that create the affect continuum: arousal (relaxed vs. aroused) and valence (positive/pleasant vs. negative/unpleasant).

For this study we choose the valence-arousal model because of its power to define and generate a seamlessly innumerable set of subtle emotions that flow into one another, and because of its similarity to our pose space, which is essentially a continuum of blends determined by a limited set of basic poses. Another reason for choosing the Circumplex model is that subsequent cross-cultural studies reveal that the same circular structure still holds for the emotion space determined by Estonian, Greek, Polish and Chinese subjects in their respective native languages (Russell, Lewicka, & Niit, 1989). The power to span across language and culture is very attractive to our problem domain of defining behaviors independent of localization.

The valence-arousal space straightforwardly embeds eight of the basic affect concepts in a circular fashion: pleasure (0°), excitement (45°), arousal (90°), distress (135°), displeasure (180°), depression (225°), sleepiness (270°) and relaxation (315°) as seen in Figure 16-6.

BEHAVIOR ANALYSIS AND GENERATION

The two problems of behavior analysis and synthesis are intertwined. Any observed data factors could potentially be used for a rough purely syntactic approximation of the non-verbal communication of the virtual character, independently from the emotion model. The same data can also be a meaningful input to the affect analyzer that assigns a semantic meaning to the behavior. The hope is that in the process of

pursuing this research a correlation can be found between the affect state and the actions of the character, such that we can trigger behaviors specific to the persona of a character by simply choosing a desired emotional state.

First, there are a few important differences between our domain and many of the current related research in the affective computing field:

- Our goal is to approximate the affect of the virtual character at a moment in time, in the hope of identifying and reproducing a similar behavior, and not to explicitly label a specific emotion. The affect analyzer shows a point in the circumplex model that is roughly estimating the character's emotional state. For visual inputs we do not need to compute facial features from images or video – our system already provides this data, because the puppeteer must explicitly input these in the process of controlling the character (press a button to smile, another to wink etc.)
- We do not want to rely on language-specific methods – linguistic sentiment analysis approaches do not apply here, nor is there a need for natural language processing. The only use of audio is for its acoustic and prosodic messages that reflect the way the words and sentences are spoken.

The proposed *affect analysis* process happens in real time, while the puppeteer is embodying the virtual character (Figure 16-7). As mentioned before, previous research suggests that small windows of time can be enough to observe emotions, thus the affect analyzer updates the predicted emotional state of the character only based on recent observed data; the actual window size that is sufficient to characterize the affect state of the character will be determined empirically. Both audio and visual/motion modalities are used in affect analysis.

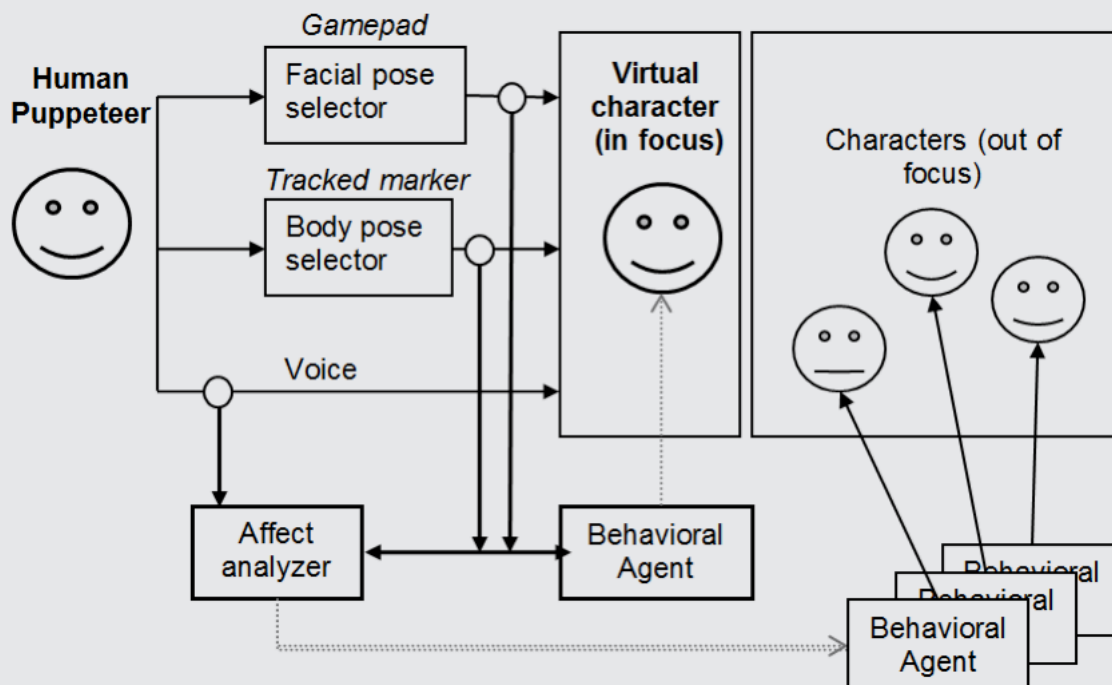


Figure 16-7: Affect analyzer plug-in and control scheme for the behavioral agents. Both the affect state and the agent for the current character are updated in real time based on the puppeteer's actions. Once out-of-focus, the agent absorbs the current affect state and takes over in controlling the character, disconnected from the puppeteer.

Audio cues. Several research groups are focusing on mapping audio input to a continuous emotion model. The most prominent correlation is the one between *pitch* and arousal (Calvo & D’Mello, 2010). Additionally, speech rate, pitch range and high frequency energy (Scherer & Oshinsky, 1977) are indicators of higher arousal, as well as shorter pauses and a higher breath frequency. A combination of faster speaking rate, low pitch and large pitch range, and longer vowel durations may be positively correlated with valence (Schröder, Heylen, & Poggi, 2006) and could potentially vary depending on specific languages.

Visual cues. The digital puppeteering system provides most of the visual communication aspects explicitly due to the actions of the puppeteer: *facial features* are generated with button presses; *body poses* are generated with the pose selector. Thus, for data acquisition we only need to intercept these explicit signals and analyze how they happen over time.

We envision the virtual character starting in a neutral affect state (this can further be switched to a predefined affect state according to his/her persona, once we obtain a mapping for plausible behaviors given affect). Once the puppeteer “steps in” a character we start tracking the aforementioned audio/visual factors and update the state toward positive/negative valence and arousal accordingly. The effect of the puppeteer’s actions is weighted over time, with the most recent having the higher impact and the older actions slowly phasing out until they do not influence the current emotional state. Once the puppeteer “steps out” of the character, the emotional state gets locked in place.

At the same time, for the purpose of *behavior synthesis* the system should capture and analyze the following factors:

- the frequency-weighted set of poses used during the current behavioral slice;
- whether there are any meaningful sequences of poses being used;
- the minimum, maximum and mean amplitude of each pose in the set, i.e. the blending weight calculated in the pose selector algorithm;
- the speed of transition between poses.

In conclusion, after observing in real time how the puppeteer controls the character, the data gathered over the most recent window of time should be analyzed by the system in order to synthesize a plausible behavior for that character’s persona given his/her current emotional state. For example, let us assume that when the puppeteer was embodying the character the general characteristics of his actions were: very limited or even lack of motion, body pose mostly leaning back, a generally low tone of voice and a low speech rate. In this case, the affect analyzer should have determined a low arousal state and positive valence corresponding to a relaxed state. On the other hand, the behavioral agent will generate a model of the corresponding nonverbal communication of the character that can be used to control their actions while out-of-focus; this model must maintain similar body postures and motions at a comparable speed and amplitude with what has been observed. Alternatively, if the puppeteer generates swift motions and wide amplitude body poses while raising the pitch of their voice and talking fast, the affect analyzer should signal an excited emotional state and the agent should use a similar speed in switching body poses.

The end goal is to obtain a correlation between affect and behavior slices that could potentially lead to further advances in both affect recognition from behavior and behavior synthesis given an emotional state and previous data defining the persona of the ECA. The system could then automate basic nonverbal behaviors that are tailored to both the character and their affect state, setting the grounds for the next steps in ECA behavior automation – adapting the affect (and thus the behavior) of the automated characters to match in real-time the state of the experience, the other automated characters, and the actions of the end-user.

6. CONCLUSIONS

Based on our observations of the described working experiences in digital puppeteering, the Geppetto system is a well-suited environment for capturing data on how a human-like virtual character would behave. We envision this research direction to set the base for extensive studies of the implicit nonverbal communication that happens in human interaction. Recorded sessions of digital puppeteering with trained puppeteers can generate data that can be further used to analyze human behaviors. Applying such knowledge to live sessions with the puppeteers can lead to the design of real-time interactive autonomous virtual characters, and thus, to more believable, more immersive and more effective virtual reality experiences. In the following, we discuss some of the avenues for improvement in this fresh and promising research domain.

The arousal-valence model was chosen as the initial study for this problem; however, another promising approach is to employ appraisal-based componential models of emotion (Grandjean, Sander, & Scherer, 2008). These alternative representations are based on the hypothesis that emotions are generated through continuous, recursive personal evaluation of both the subject's own internal state and the relevant state of the outside world with respect to that subject. Such an approach would feasibly support the need for short-term memory of the character's actions that determine its current affect and which will influence its subsequent actions.

Given the similarities between the continuous pose-space of the latest digital puppetry control paradigm, where a sphere-like 3D volume is used as the pose selector, and the circular emotion-space continuum of the arousal-valence model, it would be interesting to see whether there can be a possible mapping between the two. Can the poses be rearranged in the selector-space such that finding the right pose corresponds even more naturally to the slice of the emotion-space we are currently in?

We currently address the direct use of the puppeteer's voice and motion/pose analysis to approximate the current affect state of the character. Due to the ease of use of current body tracking technologies like the Kinect, following the user's motion as well is only one step away. Consequently, the avenues of active listening by the intelligent agents controlling out-of-focus characters are also open, leading to even more realism, believability and immersiveness of the interactive experience.

Although not covered in this step of our research, additional modalities can be considered to complete the available data for affect recognition: biosignals varying in levels of invasiveness from micro EEG brain monitoring systems to simple galvanic skin response readers, thermal imaging etc. These could be used for the puppeteer as well as for the user, where appropriate.

Once the affect recognition is set in place it could be used as real time feedback for puppeteers while they are running the experience. Thus, they would be able to see how their current behavior can be interpreted and possibly correct their actions if that was not the desired effect. Moreover, the training scenarios might require a character to get to a specific emotional state, in which case the underlying framework could use the affect recognition screen to assess and display the current goals for the puppeteer to attain.

The ultimate goal is to create intelligent agents that can not only learn the emotional state and the associated nonverbal communication of a virtual character from the human puppeteer that controls it, but synthesize that character's behavior after the puppeteer has released control. Furthermore, these agents could be used to monitor, advise and even train inexperienced puppeteers to control the corresponding characters. Given a mapping between the characters' affect states and behaviors, simple and intuitive scenario authoring can be done even without the need for a puppeteer, e.g. setting up a small classroom with a student who is anxious to get the teacher's attention, two shier ones who are following him interestedly, another who is easily distracted by every noise and movement around and yet another one who looks afraid of being asked

about his homework. Even though human-like open-ended conversation with a virtual character is still far from solved, we are undoubtedly getting closer and closer to creating virtual humans that can make us forget we are in a virtual world.

7. BIBLIOGRAPHY

- Ambady, N., & Rosenthal, R. (1992). Thin slices of expressive behavior as predictors of interpersonal consequences: A meta-analysis. *Psychological Bulletin*, Vol 111(2), 256-274.
- Barber, D., Schatz, S., & Nicholson, D. (2010). *AVATAR: Developing a military cultural role-play trainer*. Paper presented at the 1st International Conference on Cross Cultural Decision Making(CCDM).
- Calvo, R. A., & D’Mello, S. (2010). Affect Detection: An Interdisciplinary Review of Models, Methods, and Their Applications. *IEEE Transactions on Affective Computing*, 18-37.
- Dieker, L., M., H., Hughes, C. E., & Smith, E. (2008). Implications of Mixed Reality and Simulation Technologies on Special Education and Teacher Preparation. *Focus on Exceptional Children*, 1-20.
- DX1 Input System*. (2008). ERGODEX. Retrieved from <http://www.ergodex.com/>
- Ekman, P. (1971). Universals and cultural differences in facial expressions of emotion. *Nebr. Symp.Motiv.*, 207-283.
- Ekman, P. (1982). *Emotion in the human face*: Cambridge University Press, New York, 2nd edition.
- Ekman, P., & Friesen, W. V. (1967). Head and body cues in the judgment of emotion: a reformulation. *Perceptual and motor skills* 24, no. 3 711-724.
- Fontaine, J. R., Scherer, K. R., Roesch, E. B., & Ellsworth, P. (2007). The world of emotion is not two-dimensional. *Psychological Science*, Volume 18, Issue 2, 1050-1057.
- Grandjean, D., Sander, D., & Scherer, K. R. (2008). Conscious emotional experience emerges as a function of multilevel, appraisal-driven response synchronization. *Consciousness and Cognition*, Volume 17, Issue 2, 484-495.
- Gunes, H., Schuller, B., Pantic, M., & Cowie, R. (2011). *Emotion Representation, Analysis and Synthesis in Continuous Space: A Survey*. Paper presented at the Intern. Workshop on Emotion Synthesis, rePresentation, and Analysis in Continuous space (EmoSPACE 2011), Santa Barbara, CA, USA.
- Kinect. (2011). *Microsoft Xbox*. Retrieved from <http://www.xbox.com/en-US/kinect>
- Kleinsmith, A., De Silva, P. R., & Bianchi-Berthouze, N. (2005). Recognizing Emotion from Postures : Cross-Cultural Differences in User Modeling. *Emotion*, 50-59.
- Kleinsmith, A., De Silva, P. R., & Bianchi-Berthouze, N. (2006). Cross-cultural differences in recognizing affect from body posture. *Interacting with Computers*, 18(6), 1371-1389.
- Lester, J. C., & Stone, B. A. (1997). *Increasing believability in animated pedagogical agents*. Paper presented at the Proceedings of the First International Conference on Autonomous Agents, Marina del Rey, CA.
- McKeown, G., Valstar, M. F., Cowie, R., & Pantic, M. (2010). *The SEMAINE corpus of emotionally coloured character interactions*. Paper presented at the IEEE International Conference on Multimedia and Expo, Singapore.
- Mehrabian, A. (1996). Pleasure-arousal-dominance: A general framework for describing and measuring individual differences in temperament. *Current Psychology: Developmental, Learning, Personality, Social*, vol 14, 261-292.
- Nicolaou, M. A., Gunes, H., & Pantic, M. (2010). *Audio-Visual Classification and Fusion of Spontaneous Affective Data in Likelihood Space*. Paper presented at the 20th International Conference on Pattern Recognition (ICPR).
- Russell, J. A. (1980). A Circumplex Model of Affect. *Journal of Personality and Social Psychology*, Vol. 39, No. 6, 1161-1178.
- Russell, J. A., Bachorowski, J., & Fernandez-Dols, J. (2003). Facial and vocal expressions of emotion. *Ann. Rev. Psychol.* 54, 329-349.

- Russell, J. A., Lewicka, M., & Niit, T. (1989). A cross-cultural study of a circumplex model of affect. *Journal of Personality and Social Psychology*, Vol 57(5), 848-856.
- Russell, J. A., & Mehrabian, A. (1977). Evidence for a three-factor theory of emotions. *J. Res. Personality*, 11, 273-294.
- Scherer, K. R., & Oshinsky, J. (1977). Cue utilization in emotion attribution from auditory stimuli. *Motivation and Emotion, Volume 1*, 331-346.
- Schröder, M., Heylen, D. K. J., & Poggi, I. (2006). Perception of non-verbal emotional listener feedback. *Speech Prosody*, 1-4.
- Schuller, B., Vlasenko, B., Eyben, F., Rigoll, G., & Wendemuth, A. (2009). *Acoustic emotion recognition: A benchmark comparison of performances*. Paper presented at the IEEE Workshop on Automatic Speech Recognition & Understanding, Merano.
- Van Mulken, S., André, E., & Müller, J. (1998). *The Persona Effect: How Substantial Is It?*
- Wirth, J. a. N., Anne E. and Mapes, Dan and Ingraham, Kenneth E. and Moshell, J. Michael. (2011). *Interactive performance: dramatic improvisation in a mixed reality environment for learning*. Paper presented at the 2011 International Conference on Virtual and Mixed Reality: Systems and Applications, Orlando, FL.
- Zeng, Z., Pantic, M., Roisman, G. I., & Huang, T. S. (2009). *A Survey of Affect Recognition Methods: Audio, Visual, and Spontaneous Expressions*. . Paper presented at the IEEE Trans. Pattern Anal. Mach. Intell.

SECTION V

STUDYING NONVERBAL
COMMUNICATION IN
VIRTUAL WORLDS

17

A FEW CHOICE ANIMATIONS: NONVERBAL COMMUNICATION THROUGH PRODUCTION AND CONSUMPTION IN SECOND LIFE

By Jennifer Martin

In an open-air club, a woman lounges in a casual but welcoming pose, ankles crossed and leaning languidly on an elbow. Across the room, a man taps his foot, seemingly impatient. Nearby, a woman tosses her hair in a flirty way, while a man responds to her apparent advances by stretching and puffing out his chest. Despite their resemblance to offline individuals and situations, these people are actually avatars that exist within the world of *Second Life*. Furthermore, despite the resemblance of their nonverbal communication to actions and gestures that would commonly be seen offline, these actions vary significantly from their offline counterparts in terms of how they are enacted and used within the virtual world.

Nonverbal communication has long been studied as an important element of human interaction and communication. Research suggests that it can be expressed in a number of ways including voice quality, motion, touch, and use of space (Duncan, 1969), as well as less embodied forms, such as appearance (Richmond & James, 2008). Research indicates that this form of communication is important in regards to conveying a variety of messages including intimacy, power, identity, and deception (Knapp & Hall, 2009). Nonverbal communication research in virtual worlds is relatively recent addition to this body of literature. However, the field is undergoing rapid development, with emerging work on body language and nonverbal behaviour (Antonijevic, 2008; Ventrella, 2011), nonverbal social norms (Yee, Bailenson, Urbanek, Chang, & Merget, 2007), and spatial social behaviour (Friedman, Steed, & Slater, 2007).

Virtual worlds are notoriously limited because everything in the world must be included by developers or, in the case of worlds like *Second Life*, by its users (Burns, 2008). User-generated means of communicating provide a supplement to the limited pre-made options included by the world's developers. The majority of nonverbal communication that exist in *Second Life* is created by users and is often given or sold to other residents. Consequently, the availability of user-created augmentations to virtual life and communication mean that production and consumption become important and highly adaptable tools for engaging in and studying nonverbal communication in virtual worlds like *Second Life*.

1.

NONVERBAL COMMUNICATION IN SECOND LIFE

Nonverbal communication is based on the idea that even in the absence of verbal communication, “While we are in the presence of another person, we are constantly giving signals about our attitudes, feelings, and personality” (Knapp & Hall, 2009, p. 4). The rise of digital technologies, however, has called into question whether nonverbal communication is possible through computer mediated communication (Lo, 2008). While there are differences between online and offline interactions, these concerns are somewhat reduced within environments like *Second Life*, where nonverbal communication between embodied avatars is made possible through in-world production and consumption of animation.

Offline, nonverbal communication can take several forms including communication directly through the body itself, often known as body language (Schefflen, 1972) and communication through its various trappings or accouterments, such as clothing, accessories, vehicles, and other items with which it can be associated (Shang, 2008). Online, nonverbal communication through accouterments is very similar to its offline counterpart. The accessories around a body offer information such as stylistic preferences, group affiliation, politics, and even resistance to mainstream society (Hebdige, 1991). Given the focus on objects that are associated with but external to the body, this process is largely the same whether it is engaged virtually or offline.

Body-based nonverbal communication is a different matter. Offline, nonverbal communication is highly embodied. The body physically responds to other people and situations by moving in particular, often unconscious ways (Buck & VanLear, 2002). Conversely, when using an avatar online bodily responses are not automatic.

JENNIFER MARTIN ON HER METHODS

While virtual worlds often mimic offline life, especially in terms of more social environments, this mimicry does not necessarily mean that the underlying features of the virtual are the same as its offline counterparts. Such is the case with virtual nonverbal communication. Although this form of communication within a world like *Second Life* may look very similar to its offline counterpart, it is fundamentally different in terms of why it is possible, how it is used, and even the meanings that it has for virtual world residents. Furthermore, assuming that the surface similarities between online and offline go deeper does not offer a complete picture of what is going on within the virtual world or why.

From a methodological perspective, conventional approaches to studying nonverbal communication are not as effective in a world like *Second Life* because the communication itself is different. Rather than unconscious body-based nonverbal communication, *Second Life* relies on communication that is intentional and product-based. Given that in-world nonverbal communication is heavily based on creation and acquisition, methodological approaches should take into account work on offline nonverbal communication but should also rely more heavily on methodologies more commonly found in studies of consumption, production, and presumption. By relying on methodologies more common to this field, it is possible to come to a greater understanding of how nonverbal communication is enacted virtually, especially in terms of user needs and preferences.

(continued on next page)

Despite its importance and usefulness, forms of nonverbal communication available to the avatar are limited in *Second Life*. There are 134 internal animations built into the world. Of these, only 16 gestures are built into the viewer's communication interface, which allows residents to easily access movements like blowing a kiss or waving. While some of the remaining animations are defaults that are automatically used, such as walking, others are triggered through more complicated processes and thus are not widely used. Moreover, the forms of nonverbal communication available as internal animations are largely undesirable to residents, with most animations are described as clunky and awkward (Rymaszewski et al., 2008). For residents, this means of conveying meaning is limited in scope and problematic in practice. However, since *Second Life's* content is almost exclusively user-created, residents have the option of creating their own nonverbal communication. For those who do not have the technical skills, creators are able to give away or sell what they have developed, and there is a wide variety of nonverbal communication options available.

Examining what elements of virtual life are important for people and how they engage with them is one of the fundamental ways of understanding how residents make meaning in their virtual lives. Looking at nonverbal communication as an intentional rather than unconscious makes clear what elements of nonverbal communication are most valued; the elements of communication created, purchased, and used by residents are significant, as is how they are used. This importance is also marked by the effort put into communicating nonverbally, its cost, frequency, and what forms of are most commonly sought and enacted. Virtual nonverbal communication is important not because it reveals elements of unconscious feeling and perception, as with offline communication, but because its intentionality reveals what elements of communication are most important to residents within the virtual world.

2. THEORETICAL AND METHODOLOGICAL CONSIDERATIONS

While related research on nonverbal communication remains important, studying communication practices in *Second Life* also requires a broader understanding of the roles that production and consumption take in social life. Different fields offer analyses of the practices and effects of production and consumption, both in general and as they relate to virtual spaces. In economics, production is viewed as the process of creating goods or services that meet needs (Kotler, Armstrong, Brown, & Adam, 2006). Consumption also serves needs, but does so through the purchase and use of goods and services (Gough, 1994; Princen, 2001). These considerations also extend to virtual spaces, where production and consumption are studied in terms of user-generated content (Bruns, 2009) and the uses and values attributed to virtual goods by those who purchase them (Martin, 2008).

Both sociology and anthropology offer considerations of production and consumption that engage not only practices, but also meanings and effects. Qualitative and quantitative research both point to the lived experience and social effects of production and consumption. Economic sociology, for instance, considers the causes and effects of economic phenomena while taking into account how economic relations function within already existing social relations (Granovetter, 1985). Production is examined in terms of how its processes are affected by social forces (Zafirovski, 2002), while consumption is seen most recently through a lens that is intended to show a balanced view of consumption practices that includes the ways in which cultural products can be adapted to the needs of consumers (Campbell, 1995). Both practices can be read in a similar way through the lens of economic anthropology. This approach explores human behaviour

as it relates to economic practices, and considers how humans meet needs through both consumption for personal use and consumption for exchange (Polyani, 1944). Production and consumption are considered in terms of their role in social life, with goods helping to fulfill social obligations (Douglas & Isherwood, 1996 (1979)). In this social role, consumption is also seen as a means of including and excluding people from a group (Bauman, 2007; Veblen, 1979 (1899)).

These facets of everyday life are also engaged through cultural studies research. Theories of the consumer society, for instance, explore the idea that consumption has supplanted production in the formation of identity and social status (Baudrillard, 1998; Bauman, 2007). In related work, scholars also explore the degree to which production and consumption have become linked in society. The idea of the prosumer, for instance, acknowledges individuals who both produce and consume (Ritzer & Jurgenson, 2010). Given the current focus on the role of prosumers in digital culture and new media, this body of work has also focused on virtual worlds (Bruns, 2009; Herman, Coombe, & Kaye, 2006; Jones, 2006; Kücklich, 2005) with an eye to the ways that users are increasingly being used as generators and consumers of digital content in a variety of ways, which includes nonverbal communication.

The relation of production and consumption to the field is recognized within the body of literature on nonverbal communication, although they are associated with augmenting the body rather than body-based forms of communication. By consuming, individuals are able to signal qualities as varied as wealth, politics, social status, musical preferences, and ideological ideas (Goffman, 1959; Hebdige, 1991). However, this perspective excludes body-based forms of communication that play an important role in communication. Unlike consumption that augments the body, the translation of body-based nonverbal communication into *Second Life* is not as direct or successful as product-based forms. This failure is due to the fact that virtual bodies in *Second Life* are not directly responsive to the feelings of their users, nor are they equipped with the ability to easily engage in body-based nonverbal communication.

Production and consumption are implicated in multiple means of nonverbal communication within *Second Life* and are more involved in in-world communication practices than they are offline. As such, the study of some nonverbal communication should at least consider if not actively employ methodological considerations more common to work on consumption and production. Because these characteristics are not often, if ever, implicated in offline body-based nonverbal communication, there are few methodological standards from which to consider these practices and existing research on offline nonverbal communication can only minimally inform similar research into virtual worlds. There is, however, research from within consumption studies, both offline and with respect to virtual worlds, which asserts that consumption practices have readable meanings (Fiske, 1989a, 1989b; McCracken, 1986). Given their relationship, considering consumption practices in terms of their readability offers a starting point from which to examine the practices and meanings associated with nonverbal communication within the virtual world.

Given that production and consumption are widely and directly implicated in its practices, the study of virtual nonverbal communication should consider how these elements of virtual life support communication, as well as what they can tell us about the communication practices and preferences of individuals. While studying virtual nonverbal communication does rely on a different methodological framework than previous communication research, it is also one that offers researchers a different set of advantages. The fact that nonverbal communication is intentional in the online world offers the opportunity to examine what values it is ascribed, and which facets of this communication are most important to residents based on what elements they choose to engage, how they are deployed within the world, and which communicative tools are produced and consumed for their use. Furthermore, the degree to which these tools are used is suggestive of the importance of nonverbal communication. Just because an animation or pose is developed or even acquired does not necessarily mean that it is used. Therefore, what items are created or consumed the most is also suggestive of which is most valued by residents.

Beyond the world itself, *Second Life* has a wide range of official and unofficial venues that make it possible to examine what elements of nonverbal communication are most important to residents, why they are significant, and how they are used within the world. Formal systems include the review function of the official marketplace where residents can evaluate their purchases, as well as its organizational system, which allows users to see what items for sale are bestsellers and most highly ranked (LindenLab, 2011b). Other systems include unofficial forums, webpages, and blogs³², all of which provide additional information on how residents acquire, use, and view their preferred forms of nonverbal communication.

3. CREATING AND USING COMMUNICATION

Nonverbal communication in *Second Life* exists in a few forms and is largely made possible through the use of code created by Linden Labs, residents, and even other computer programs such as Poser or Blender, in the case of animations. When applied to an object or an avatar, this code governs how it behaves within the world (Moore, Thome, & Haigh, 2008). Nonverbal communication that relies on code includes in-world animations, poses, and gestures. Animations cause avatars to move in a series of motions, such as walking or dancing. A pose functions similarly to an animation, but tends to have a static component where the avatar is moved into a position and then stays there. Gestures are like macros and can combine multiple elements together, such as poses, animations, sounds, and text chat.

Code that is used to create nonverbal communication can function in two different ways. First, it can be applied directly to the avatar, so that the script, animation, or pose is always available. For instance, if a walk animation is applied in order to override the default walk, the avatar will always use the new walk animation so long as it is active. Second, code can be applied to objects so that the object can affect avatars that interact with it, such as when that object is in use, or when the resident authorizes the object to animate their avatar. For instance, a couch might offer three different poses which only affect an avatar when it is sitting on that particular piece of furniture. Poseballs – objects that were originally in the shape of a ball, and that can animate groups of people – also make use of object-embedded communication, with avatars animated with poses or animations when they click on the poseball itself, or when they are in proximity and allow themselves to be animated. While residents can directly purchase animations and poses for their avatars, they can also purchase objects that will animate their avatars, or use them in-world.

Despite the universal ability to create, not all residents consider themselves to be creators, or have the skills necessary to engage in the somewhat complex process of creating poses or animations. As a result, residents who do not consider themselves to be competent enough to make some or all of what they might want or need and who wish to communicate nonverbally must acquire or consume to make this possible. While they can be difficult to initially create, many of these forms of communication do not require a large amount of technical knowledge for the average resident to use. Once they are acquired, avatar-based animations and poses can simply be attached to the avatar in a way similar to virtual clothing or accessories. Using external devices that act on the avatar usually requires little more than clicking on the object, agreeing to allow the animation, and in some cases selecting the preferred animation from a list of possible choices. Some animation sets, however, are more complicated to use than others, and may have a learning curve. Similarly, different poses and animations can interact with each other and interrupt the intended communication. Using poses or animations effectively – that is, without awkward movements or conflicting animations – can require some skill and knowledge to do well.

³²Although comments were published on publicly available sites, residents posting to forum and marketplace threads have been anonymized in this work.

4. PRODUCTION, CONSUMPTION, AND COMMUNICATION

Production and consumption are expected elements of product-based communication. Robertson states that, “Products vary in the degree to which social-symbolic meaning is important. Cars and clothing are both products which are high in visual display and recognized in our society as ‘saying something’ about a person” (Richmond & James, 2008, p. 3). What is perhaps more unusual is the fact that even nonverbal communication that is associated more directly with the body – its expressions, movements, postures, and gestures – are also tied into the system of production and consumption within *Second Life*. Nonverbal communication is largely associated with the body, in particular its facial expressions, postures, gestures, and movements (Davis, 1973). Because these elements of communication are limited, they enter into *Second Life* and are engaged by residents within a system of production and consumption.

Production and consumption are both necessitated and facilitated by *Second Life*. First, integrated nonverbal communication is integrated with virtual bodies in problematic ways, which necessitates working around the available options. Nonverbal communication that is automatically available through avatars provides limited options. These forms are also considered to be unattractive, awkward, and even unusable by residents, with an official *Second Life* guide pointing out that, “One of the first things that people notice when they start out in *Second Life* is how awkward and artificial the default avatar animations look” (Moore, Thome, & Haigh, 2008). Consequently, there is a need for code that helps to change these conditions. Second, since the world facilitates and even relied on user creations, there is a great deal of freedom for residents to actively change the world. Unlike many other games and virtual worlds that are restrictive in terms of what can be added to the world or the avatars within, *Second Life* residents are able to create pieces of code in the form of animations or scripts that change how their avatars are able to communicate.

Consumption is also made possible through the fact that *Second Life* is a world with a virtual economy. Residents are able to make, sell, buy, and trade virtual goods within the world in exchange for a virtual currency called the Linden dollar, or Lindens for short. This economy is linked to offline economies by virtue of the fact that the in-world currency can be exchanged for offline currencies, with one United States dollar exchanged at a variable rate that is usually around 250 Linden dollars (LindenLab, 2011a). Within this economy, production and consumption take two forms: free and paid. The linking of online and offline economies makes it possible for virtual goods to have a recognizable and relatively concrete value, which allows for and even encourages their sale as a way to make money (Freedman, 2007). To this end, some *Second Life* residents have turned their in-world businesses into very profitable careers, with some even deriving their offline incomes from this work (Hof, 2006).

While the virtual economy functions in ways very similar to offline economic systems and market economies, *Second Life* also has what is sometimes referred to as the “freebie economy” (Tomsen, 2007), a selection of thousands of virtual goods that, as the name suggests, are made available to other residents for free. The reasons underlying the availability of free goods are multiple: giving away virtual goods is linked to the production of lower-quality goods that are unlikely to sell; older or out of style items have been devalued; enticing residents into buying items by offering samples or other free items; or making available and showing off a resident’s creative labour (Llewelyn, 2008). These goods may not be consumed in a way that is dependent on spending money, but even though these items are free, residents still need to identify a need, find the appropriate goods, and obtain them in a manner that is related to consumption.

The creation of animations and scripts that allow for nonverbal communication – whether they are free or for profit – offers a form of communication that is flexible, individual, and highly varied. The official *Second Life* Marketplace alone offers residents over 60 000 animations (LindenLab, 2011b). When objects

that include animations are included in this count, it rises to over 150 000 . This wealth of communication options available through the marketplace reveals the popularity and importance of animations, scripts, and poses for residents, as well as the wide variety of possibilities for engaging this type of nonverbal communication. Through examining what nonverbal communication practices are consumed and how they are used, it can be seen that nonverbal communication is acquired and used by *Second Life* residents for three main reasons. First, it provides residents a way to customize their avatar, to make it their own, and to ensure that it is the most ideal representation of them within the virtual world. Second, it makes it possible for certain forms of commerce that benefit residents to be practiced within the world. Finally, it facilitates communication and connections between residents. Although these uses are all markedly different, by using nonverbal communication, *Second Life* residents are able to meet their myriad needs in the virtual world.

COMMUNICATION AND CUSTOMIZATION

Examining the multiple discourses around the production and consumption of nonverbal communication reveals that, in some ways, nonverbal communication is not necessarily used for communication, at least in the sense of conveying a message to others. *Second Life* residents often spend a time and effort making their avatars exactly what they want and ensuring that they are highly individual if not completely unique. One resident suggests that, “There are so many people in SL who look absolutely identical. When you look unique it really stands out” (Resident 1). While this is often done through changing physical appearance and clothing, animations and poses can also be an important element of customization as a way to ensure that the avatar fully matches the resident’s preferences and expectations.

Use of poses and animations is one way to customize the avatar. Although other residents can see these changes – reaffirming their use in nonverbal communication – they can also be for the benefit of the resident. Residents regularly discuss avatar customization and creating or consuming in order to customize is frequently cited as being for the benefit for the resident. Benefits range from building a brand (Resident 2) to improving the appearance of their avatar (Resident 3). Residents frequently discuss the pleasure they get from creating and changing their avatar to suit their needs within the world. Some residents change their appearance multiple times a day, with one noting that, “Since I have been in SL I have constantly changed my avatar: shape, skin, clothes and accessories day by day. I love the creativity inherent in this world and try and use it to my advantage as much as I can” (Resident 4). Making the avatar exactly what is desired by the user down to the subtle nuances of a coy head tilt, demure foot shuffle, sexy hip thrust, or determined stride serves as a point of pride for many *Second Life* residents, which exists apart from the communicative power of animations within the world.

To this end, about half of the animations available in the marketplace are designed for avatars to use individually, with many available options (LindenLab, 2011b). Many animations are designed to appeal to the general population, and the rate at which they are consumed suggests that they are widely used. Packages of avatar animations, for instance, are frequent marketplace bestsellers. The most popular individual animations are focused on changing basic avatar animations, and offer different options for walking, running, dancing, sitting, and posing. Many animations or sets have upwards of 3000 reviews, which suggests that not only are these animations used by their purchasers, but also that residents value animations enough to review and comment on them through the marketplace.

Animations are available to meet almost any resident preference. While general sets are available with a range of basic animations, there are also sets that are more targeted to match particular identity preferences. For instance, the marketplace has animations specific to residents who use wolf or mechanical avatars, as well as animations designed to mimic particular styles of movement from salsa dancing to Harajuku girl poses. Animations and poses intended to fit even the smallest of niche markets are also not only available, but actively acquired and used by residents. Although child avatars are a relatively small niche market in

Second Life, two sets of marketplace animations have 69 and 20 reviews respectively (Fride, 2011a, 2011b). While these options are less frequently reviewed, the fact that they are available is indicative of a virtual world where production and consumption can meet almost any need, even for a relatively small subset of the population.

Within the marketplace, product reviews give an indication of what residents value when acquiring nonverbal communication. Many reviews are focused on personal preferences, with those looking for a specific style – such as cute, sexy, or macho – referencing how the animation accomplishes what they want. There are, however, general trends that indicate preferred elements of nonverbal communication that transcend the particulars of style. Animations, gestures, and poses that appear to be natural and subtle receive the best reviews and the most sales. After buying ORACUL's Pure Lady, one resident writes that, "The standing animations are subtle and lady-like. It's the best one I've tried and the price is great. Thank you" (Resident 5). This sentiment is shared across the marketplace's most well-reviewed products along with additional comments highlighting the importance of responsiveness to different needs and situations. Speaking about an animation override (AO) set, a resident writes that, "I've owned it for some time now and just love it. You can go from 'relaxed' to 'xxx' with a click of the mouse" (Resident 6), with other residents agreeing that their nonverbal communication should be easily engaged and adaptable to different situations.

There is also an element of distinction that is mentioned in reviews, with residents noting forms of nonverbal communication that let them stand out. One appreciative resident writes that, "This is one of the premier AO's in SL i think, with its excellent lifelike animations it will sure make your Avie stand out from the rest, plus it has an excellent hover animation which doesnt make you look like your being crucified with your arms out to the side" (Resident 7). This focus also takes the form of reviews that mention wanting to be distinct from new users, and many criticisms of nonverbal communication come from residents who feel that their purchases have failed in this regard. In a review from a resident who is unhappy with a product, they write that, "I tried this AO now for 3 days and i am not really happy with it. By changing from one pose to another i often get for a few seconds a newbie stand. It is really not cool if you try to look sexy to stand then like a newbie" (Resident 8).

COMMUNICATION AND COMMERCE

Second Life allows residents to create or purchase the forms of nonverbal communication that they want or need to best represent themselves. However, in some instances production and consumption is necessary for residents engaged in in-world professions that require high quality, up-to-date animations and poses. This reliance is most clearly seen with residents for whom body language is an important element of their job, such as escorts, dancers, and models. These jobs rely heavily on the communicative power of the virtual body and require dedication to the establishment and upkeep of nonverbal communication. Modeling, for instance, is a profession where top in-world models are expected to maintain the avatars as well as their walks, poses, gestures, and other body-based communication. Speaking on the pressure and expense of maintaining a working avatar, one former *Second Life* model writes that, "First you have to attend expensive modeling schools, participating in all kinds of contests. Spend a LOT OF MONEY, buying skins, hairs, poses, outfits, etcetc and then you work as a model for awhile until you realize you can't afford this" (Voff, 2011). As a result, consumption becomes necessary in order to work.

Although there are many high-quality, highly rated free animations and poses available to residents for free, workers who depend on their in-world employment may need to acquire better quality or even specially customized animations if they are unable to create them themselves. In both cases, these practices engage both production and consumption in the nonverbal communication of residents. Even animations, poses, and gestures that are for sale need not be expensive and can be easily bought for less than a few hundred Linden dollars (or a few USD). Some options, however, cost more. Vista Animations is a *Second Life* creator that was nominated by models for the Best of *Second Life* Awards in the category of best animations (Summerkat, 2008). Their "The Perfect Lady" animations sell for 2100 Lindens (\$9.24 USD) on the

Second Life Marketplace (Barnes, 2011). When compared with ORACUL's "Pure Lady" or the "Automatic Facial Animation", both of which are popular, well-rated, and sell for 300 Lindens (\$1.58 USD) and 500 Lindens (\$2.43 USD) respectively (Banjo, 2011; Papp, 2011),³³ these animations are more costly. They can, however, offer additional benefits that are valued by those who depend on animations for specific purposes.

Finally, for those in professions that require current and highly individual nonverbal communication, exclusive custom animations and poses are possible, but can also be significantly more expensive.³⁴ If individuals want options that will not be available to any other residents, making or commissioning their own nonverbal communication is an option. If they wish to purchase these items, the price can be expensive, with some sellers on the official *Second Life* forums recommending that, "you charge them about what you might expect to make on it in say a six or eight month period" (Resident 9). While more expensive than the options available through marketplace sellers, commissioning a custom animation increases the possibility that a professional or any other resident will be able to acquire the nonverbal communication that they want or need.

5. COMMUNICATION AND COMMUNITY

Given the visual nature of the world and its social and interactive components, using nonverbal communication can also be read as a communicative practice that engages others.

While the use of nonverbal communication indicates preferences about how the avatar moves and behaves within the world (Linden, 2010), it also serves as a way to facilitate communication and connection with other residents. As a social environment, *Second Life* has a large focus on forming relationships with other avatars. These relationships are frequently friendly and sociable, although there is also a great deal of room in the virtual world for more intimate connections from online relationships to cybersex.

Creating and using animations and poses allows residents to more effectively engage and interact with others. Using an enthusiastic dance animation at a bar, lounging casually while chatting, or energetically beckoning someone over can all signify engagement with other residents. While verbal text-based communication is extensively used within the world, nonverbal communication enables additional communication without necessitating further typing, or breaking up a conversation with elements that are intended to be more descriptive of what the avatar is feeling or doing. Interfaces, such as those using a heads-up display (HUD), can be designed to facilitate using different animations effectively and fluidly, undermining customary online communication where, "Good typing skills and creativity are fundamental for the scene to come off well" (Waskul, Douglass, & Edgley, 2000). The use of animations and poses allows residents to visually represent the ways that they are feeling – either to one other resident or many – in addition to eliminating some of the difficulties and extensive typing found in text based description of any complicated activity.

Animations are also used to enhance in-world relationships, both in terms of facilitating their development and maintaining them once they are established. Romantic and sexual partnerships are one example, with many residents relying on widely available forms of nonverbal communication to enable and affirm their relationships. Residents have developed thousands of relationship-based animations. As with individual options, couples animations are widely available, with over 4000 couples animations and over 5000 erotic animations accessible through the marketplace (LindenLab, 2011b). There is also a wide range of options available, from a playful animation for "the airplane game" where one partner lies down and the other balances on their raised knees as if they were flying (Parx, 2011) to a 36 minute seduction sequence (Tinkel, 2011).

³³These prices are current at the time of writing, but are subject to change at any time.

³⁴For further considerations of performance art and theater in virtual worlds, see Chapters 9, 10, 11, and 12 in this book.

These animations can be used to both develop and maintain relationships. Flirting animations can be used to convey interest and availability early in a relationship. Once interest is established, nonverbal communication can also be used to maintain relationships through physical gestures and activities that are not immediately available. Speaking about an avatar, one article on virtual infidelity notes that, “He fell pretty hard for his avatar sweetie. They bonded intellectually, emotionally, and yes, thanks to *Second Life* animations, even physically” (Kalning, 2007). In this instance, the use of nonverbal communication adds a depth of communication and connection that would not necessarily be available through text or the existing in-world options

Nonverbal communication is also used to facilitate group interactions. In some instances, animations embedded in objects are placed in public locations for use by multiple avatars. One of the more common examples of publically available animations is poseballs. While residents can buy poseballs for their personal use, they are commonly seen on *Second Life* land that is open to the public, as well as in private businesses such as shops or clubs. Basic poseballs can animate one avatar, or multiple avatars with the same animation. More advanced poseballs may offer uses choices about which animation they would like to use, making it possible for multiple residents to use the poseball in different ways simultaneously. Given the fact that any avatar can use them, embedded animations and poses are useful for fostering group activities. By offering attractive and entertaining animations to customers or visitors, business and landowners can make their spaces more attractive. Furthermore, for residents who have not sought out poses or animations for their avatars, poseballs and other items with embedded animations offer an opportunity to engage in nonverbal communication without having to buy or otherwise acquire an animation or pose.

Finally, nonverbal communication is also used to establish and signal group membership, especially for groups that interact in prescribed ways and that depend on nonverbal communication as part of their interactions and identities. For instance, members of fashion-related groups may require up-to-date poses and animations to best show their clothes to others. This communication is also important to role-playing groups. Gorean role-play, for example, tends to involve submission and relies on animations that convey the group’s ideals and practices. One resident recounts how, “My training is progressing as Kajira (my mistress said though, she trots ahead) and I slowly realize that I am missing animations. Although I have the standard animations such as Tower and Nadu etc., but not much else to express myself as can Kajira” (Resident 10). In this instance, not having relevant animations is a detriment to involvement with a particular community, as well as with another individual. By acquiring animations and other forms of nonverbal communication, residents are better able to interact with and establish their affiliations with particular groups.

6. ADDITIONAL CONSIDERATIONS

Residents are able to use the array of animations and poses to construct a repertoire of communication. In many instances, the meanings associated with these forms of communication are fairly straightforward and easily communicated to other residents of the world. There are, however, miscommunications that happen through misuse and misinterpretation. While residents may think that they are communicating in particular ways, the intentional nature of nonverbal communication means that it can be misused by residents and misread by others. For the former, nonverbal communication can be misused relatively easily and selecting the wrong animation could mean responding to a joke with a bored yawn rather than a laugh or giggle. Moreover, nonverbal communication has been recognized as being subject to misunderstanding, (Andersen, Hecht, Hoobler, & Smallwood, 2003), a quality that is not necessarily lost despite the intentionality of virtual communication. Furthermore, this type of communication can be misleading or even dishonest since it can be used at any time and for any reason. While nonverbal

communication offline is seen as a form of communication in which it is difficult to hide true feelings, the same communication online is not limited to underlying emotions. As such, the resident can use it to communicate what they want, rather than what they actually feel.

Beyond direct communication, these practices communicate more generally. Locating and using an animation can convey information about a resident that has little to do with the animation itself. From a status perspective, the differences between those who purchase expensive animations and those who rely on more inexpensive or even free options are not necessarily apparent. Even expensive animations have problems, and there are very good quality examples available for free. Therefore, animations are not necessarily indicative of status in the sense of residents having in-world purchasing power. There are, however, differences between those who acquire good quality animations and those who rely on the awkward defaults or other inferior examples. As one guide to *Second Life* suggests, “you need to stop walking like a duck if you want to look like a seasoned *Second Lifer*” (Jewell, 2007). Well-made animations are valued and are relatively easy to recognize, especially in comparison to less desirable options. While these meanings are linked to nonverbal communication, they are not a direct result of the animations themselves, but rather a result of their use within the world more generally.

This status also extends to the use of animations more generally. While animations are not overly difficult to use, there is a learning curve associated with many *Second Life* activities. The use – and especially the practiced use – of animations suggests that a resident is serious enough about their avatar and skilled enough in their *Second Life* to find and use an animation. Individual animations are not the only factor in social status, but they do play an important role in establishing residents within the community and showing their commitment to it through their willingness to acquire and use animations.

7. CONCLUSIONS

With the body taking an increasing role in virtual nonverbal communications, its associated objects, accouterments, and accessories assist in conveying meaning to others, including personal preferences and group affiliations. In terms of more body-based nonverbal communication, because *Second Life* content is almost exclusively user-created, residents with the necessary skills are able to develop scripts, animations, and items that allow for a wide range of nonverbal communication. Furthermore, since the world also supports an active economy, these user creations can be sold to other residents who are not able to develop their own ways of conveying meaning beyond the limited possibilities for nonverbal communication within the world. Consequently, an avatar’s flirty walk, bored expression, angry stance, or dejected pose are likely to be a script or animation made, found, or bought by the resident.

Examining how these elements of nonverbal communication function in *Second Life* allows for a greater understanding of the social and communicative life of the world. With fundamental elements of offline communication largely missing from offline interactions, the introduction of effective nonverbal communication allows residents to compensate for these limitations. Because these practices are intentional and linked to a system of production and consumption, they also indicate more than simply how an avatar is feeling in the moment. They reveal what elements of virtual life residents value through the kinds of communication they make, prefer, and are willing to seek out or even pay for. They suggest preferences around identity and group membership through the chosen types of nonverbal communication and how they are used. They indicate a resident’s proficiency with the world and their willingness to adapt to the norms and expectations of the environment by locating and learning how to use nonverbal communication that is not built into the world itself. In these instances, nonverbal communication happens not only through the application itself, but also in the meanings and significance that can be attributed to what is created and chosen and how it is used within the world.

Using this knowledge, it is also possible to alter, design, and construct virtual worlds to better meet the needs of their users. Production and consumption meet many of the communication needs in *Second Life* that are not addressed by what developers have included in the world. However, the advantages and capabilities of production and consumption are only available in worlds that allow for user-generated content. In most virtual worlds associated with games and many of the more tightly controlled social environments, nonverbal communication is initially limited and, in order to ensure a tightly controlled environment, cannot be created by users. With millions of people active and interacting with each other in virtual worlds from games to social environments, effective communication is an important element of this experience which an understanding of what people want and why will help to establish.

Given the positive effects associated with online interaction, the facilitation of effective communication becomes even more important. The benefits associated with online identity creation are well established, including creating an ideal self (Bessi re, Seay, & Kiesler, 2007) and working through personal issues (Turkle, 1995). Increasing the ability to identify with an avatar through nonverbal communication also stands to increase the possibility of obtaining these benefits, and may also help to increase the effects. Furthermore, with people relying on these worlds not only for entertainment, but also as sites that support valued interpersonal connection such as friendship and romance (Yee, 2006), nonverbal communication becomes an important element of developing and sustaining these relationships. In these instances, developing more effective nonverbal communication in digital environments could increase access to beneficial features of online life.

Enabling residents to produce and consume allows them to move beyond many of the limits of more conventional computer mediated communication, especially in terms of the lack of nonverbal communication (Lo, 2008). Reliance on production and consumption factors into many elements of virtual life, and is especially important in regards to nonverbal communication. Much of the in-world nonverbal communication that happens within *Second Life* is linked with production and consumption by virtue of the fact that it is otherwise unavailable, or available but only in an unsatisfactory or otherwise limited form. Despite the limitations in terms of pre-made nonverbal communication available to *Second Life* avatars, there are nearly limitless possibilities for nonverbal communication within the world. In turn, residents can make use of these offerings in varied and highly personal ways in order to ensure that they are able to communicate effectively both verbally and nonverbally within the world.

8. BIBLIOGRAPHY

- Andersen, P. A., Hecht, M. L., Hoobler, G. D., & Smallwood, M. (2003). *Nonverbal Communication Across Cultures*. In W. B. Gudykunst (Ed.), *Cross-cultural and intercultural communication*. Thousand Oaks, CA: Sage Publications.
- Antonijevic, S. (2008). From Text to Gesture Online: A microethnographic analysis of nonverbal communication in the *Second Life* virtual environment. *Information, Communication, and Society*, 11(2), 211-238.
- Banjo, O. (2011). Auto Emote - Automatic Facial Animation HUD Retrieved May 6, 2011, from <https://marketplace.secondlife.com/p/Auto-Emote-Automatic-Facial-Animation-HUD/122757>
- Barnes, V. (2011). XMAS PROMO! Vista Animations- The Perfect Lady AO u Retrieved May 6, 2011, from <https://marketplace.secondlife.com/p/XMAS-PROMO-Vista-Animations-The-Perfect-Lady-AO-u/298448>
- Baudrillard, J. (1998). *The Consumer Society: Myths and Structures*. London: Sage.
- Bauman, Z. (2007). *Consuming Life*. Malden, MA: Polity Press.

- Bessière, K., Seay, A. F., & Kiesler, S. (2007). The Ideal Elf: Identity Exploration in World of Warcraft. *CyberPsychology & Behavior*, 10(4), 530-535.
- Bruns, A. (2009). *Blogs, Wikipedia, Second Life, and Beyond: From Production to Producership*. New York, NY: Peter Lang.
- Buck, R., & VanLear, C. A. (2002). Verbal and Nonverbal Communication: Distinguishing Symbolic, Spontaneous, and Pseudo-Spontaneous Nonverbal Behavior. *Journal of Communication*, 52(3), 522-541.
- Burns, A. (2008). *Blogs, Wikipedia, Second Life, and beyond: From production to producership*. New York: Peter Lang.
- Campbell, C. (1995). The Sociology of Consumption. In D. Miller (Ed.), *Acknowledging consumption: a review of new studies*. New York, NY: Routledge.
- Davis, F. (1973). *Inside Intuition: What We Know About Nonverbal Communication*. New York: McGraw Hill.
- Douglas, M., & Isherwood, B. (1996 (1979)). *The Word of Goods: Towards an Anthropology of Consumption*. New York: Routledge (Basic Books).
- Duncan, S. J. (1969). Nonverbal communication. *Psychological Bulletin*, 72(2), 118-137.
- Fiske, J. (1989a). *Reading the Popular*. Boston, MA: Unwin Hyman.
- Fiske, J. (1989b). *Understanding Popular Culture*. Boston: Unwin Hyman.
- Freedman, R. (2007). *How to Make Real Money in Second Life: Boost Your Business, Market Your Services, and Sell Your Products in the World's Hottest Virtual Community*. New York: McGraw-Hill.
- Fride, M. (2011a). MrS - CHILD BOY AO Retrieved May 5, 2011, from <https://marketplace.secondlife.com/p/MrS-CHILD-BOY-AO/645996>
- Fride, M. (2011b). MrS - CHILD GIRL AO Retrieved May 5, 2011, from <https://marketplace.secondlife.com/p/MrS-CHILD-GIRL-AO/645997>
- Friedman, D., Steed, A., & Slater, M. (2007). *Spatial Social Behavior in Second Life* Paper presented at the IVA 2007: 7th International Conference On Intelligent Virtual Agents, Paris, France.
- Goffman, E. (1959). *The Presentation of Self in Everyday Life*. New York, NY: Anchor Books.
- Gough, I. (1994). Economic Institutions and the Satisfaction of Human Needs. *Journal of Economic Issues*, 28(1), 25-66.
- Granovetter, M. (1985). Economic Action and Social Structure: The Problem of Embeddedness. *American Journal of Sociology*, 91(3), 481-510.
- Hebdige, D. (1991). *Subculture: The Meaning of Style*. New York, NY: Routledge.
- Herman, A., Coombe, R. J., & Kaye, L. (2006). Your *Second Life*? Goodwill and the performativity of intellectual property in online digital gaming. *Cultural Studies*, 20(2-3), 184-210.
- Hof, R. D. (2006, May 1). Virtual World, Real Money. *Business Week*.
- Jewell, S. (2007, September 2). Get a new walk. Retrieved from <http://virtuallyliving.wordpress.com/2007/08/26/stop-looking-like-a-newbie-step-3-get-a-new-walk/>
- Jones, D. E. (2006). I, Avatar: Constructions of Self and Place in *Second Life* and the Technological Imagination. *Gnovis*, 6.
- Kalning, K. (2007). Is a virtual affair real-world infidelity? Or are in-game chats and animated sex just harmless experiments? . *Technology and Science* Retrieved September 1, 2011, from http://www.msnbc.msn.com/id/18139090/ns/technology_and_science-games/t/virtual-affair-real-world-infidelity/#.TmVUx5jna6Q
- Knapp, M. L., & Hall, J. A. (2009). *Nonverbal Communication in Human Interaction*. New York: Centage Learning.
- Kotler, P., Armstrong, G., Brown, L., & Adam, S. (2006). *Marketing* (7th ed.). Upper Saddle River, NJ: Prentice Hall.
- Kücklich, J. (2005). Precarious Playbour: Modders and the Digital Games Industry *FibreCulture*(5).

- Linden, T. (Writer). (2010). Don't know what an Animation Override is You need to see this - *Second Life* Viewer 2.3 Beta Tutorial YouTube.
- LindenLab. (2011a). LindeX Exchange: Market Data. *Second Life* Retrieved September 2, 2011, from <https://secondlife.com/my/lindex/market.php?>
- LindenLab. (2011b). *Second Life* Marketplace Retrieved May 2, 2011, 2011
- Llewelyn, G. (2008). The Hard Facts About the *Second Life* Economy. *Gwyn's Home* October 13. Retrieved June 7, 2009, from <http://gwynethllewelyn.net/2008/10/13/the-hard-facts-about-the-second-life-economy/>
- Lo, S.-K. (2008). The Nonverbal Communication Functions of Emoticons in Computer-Mediated Communication. *CyberPsychology & Behavior*, 11(5), 595-597.
- Martin, J. (2008). Consuming Code: Use-Value, Exchange-Value, and the Role of Virtual Goods in *Second Life*. *Journal of Virtual Worlds Research*, 1(2).
- McCracken, G. (1986). Culture and Consumption: A Theoretical Account of the Structure and Movement of the Cultural Meaning of Consumer Goods. *The Journal of Consumer Research*, 13(1), 71-84.
- Moore, D., Thome, M., & Haigh, K. (2008). *Scripting Your World: The Official Guide to Second Life Scripting*. Indianapolis, IN: Sybex.
- Papp, d. (2011). ORACUL-Pure Lady AO Retrieved May 6, 2011, from <https://marketplace.secondlife.com/p/ORACUL-Pure-Lady-AO/572760>
- Parx, P. (2011). Zambooki Gift Couples Air Animation - Airplane Game Retrieved May 5, 2011, from <https://marketplace.secondlife.com/p/Zambooki-Gift-Couples-Air-Animation-Airplane-Game/1733497>
- Polyani, K. (1944). *The Great Transformation*. Ypsilanti, MI: Beacon Press.
- Princen, T. (2001). Consumption and its Externalities: Where Economy Meets Ecology. *Ecological Economics*, 31(3), 347-363.
- Richmond, V., & James, M. (2008). *Nonverbal behavior in interpersonal relations*. New York: Allyn & Bacon.
- Ritzer, G., & Jurgenson, N. (2010). Production, Consumption, Prosumption: The nature of capitalism in the age of the digital 'prosumer'. *Journal of Consumer Culture*, 10(1), 13-36.
- Rymaszewski, M., Au, W. J., Ondrejka, C., Platel, R., Van Gordon, S., Cezanne, J., . . . Rossignol, J. (2008). *Second Life: The Official Guide*. Indianapolis, IN: Wiley Publishing, Inc.
- Schefflen, A. E. (1972). *Body Language and the Social Order; Communication as Behavioral Control*. Englewood Cliffs, NJ: Prentice-Hall.
- Shang, D. (2008, January 21). The current economy within SL. Retrieved from <http://forums-archive.secondlife.com/327/04/249819/5.html>
- Summerkat. (2008, May 2). Official List of Nominees for the Best of SL Award 2008. Retrieved from <http://runwaykidz.wordpress.com/2008/05/10/official-list-of-nominees-for-the-best-of-sl-award-2008/>
- Tinkel, I. (2011). ***Seduction FULL Scene. Victory - 36 minutes sequenced complete seamless scene - HUD and rug Retrieved May 5, 2011, from <https://marketplace.secondlife.com/p/Seduction-FULL-Scene-Victory-36-minutes-sequenced-complete-seamless-scene-HUD-and-rug/1502827>
- Tomsen, K. (2007, May 1, 2011). Latest trends in SL economy (shopping). Retrieved from <http://forums-archive.secondlife.com/142/15/230021/3.html>
- Turkle, S. (1995). *Life on the screen: Identity in the age of the internet*. New York, New York: Simon and Schuster.
- Veblen, T. (1979 (1899)). *The Theory of the Leisure Class*. Kingsport, TN: Kingsport Press/Penguin Books.
- Ventrella, J. (2011). *Virtual Body Language*. Halifax, NS: ETC Press.
- Voff, U. (2011, May 2). My Experience as a SL Model. Retrieved from <http://voffuggla.blogspot.com/2011/02/my-expererience-as-sl-model.html>

- Waskul, D., Douglass, M., & Edgley, C. (2000). Cybersex: Outercourse and the Enselfment of the Body. *Symbolic Interaction*, 23(4), 375-397.
- Yee, N. (2006). The Psychology of MMORPGs: Emotional Investment, Motivations, Relationship Formation, and Problematic Usage. In R. Schroeder & A. Axelsson (Eds.), *Avatars at Work and Play: Collaboration and Interaction in Shared Virtual Environments* (pp. 187-207). London: Springer-Verlag.
- Yee, N., Bailenson, J. N., Urbanek, M., Chang, F., & Merget, D. (2007). The Unbearable Likeness of Being Digital: The Persistence of Nonverbal Social Norms in Online Virtual Environments. . *The Journal of CyberPsychology and Behavior*, 10, 151-121.
- Zafirovski, M. (2002). The Social Construction of Production: An Application of Economic Sociology. *M@n@gement*, 5(2), 147-174.

A MICROSOCIOLOGICAL PERSPECTIVE ON NON-VERBAL COMMUNICATIVE STRATEGIES IN MMORPGS

By David Kirschner and J. Patrick Williams

In tandem with technological advancements in virtual world design, video games have evolved from turn-taking affairs where individuals or groups hoped to clear the screen in record time or better a previous performance, to immersive, interactive environments where hundreds or even thousands of players acting in concert develop expertise in a variety of semiotic domains as they strive to achieve individual- and group-oriented goals. The newest generation of video games, Massively-Multiplayer Online Role-Playing Games (MMORPGs), has exploded in popularity in the past decade along with advances in graphics technology, computer processing power, and the continual spread of high-speed internet access. MMORPGs such as *World of Warcraft* (WoW) are popular virtual environments of epic scope where people engage in distinctly *social* play. MMORPGs are culturally-rich virtual environments that persist 24 hours a day, 365 days a year, and are populated by unique 3-D characters that players control using a computer keyboard and mouse. Players are immersed via linkages between themselves, the fantasy world, and other players in much the same way as in table-top role-playing games (Fine 1983; Waskul and Lust 2004), except that in MMORPGs players are typically displaced physically from those with whom they play. Like table-top role-playing games, one of the main design structures in MMORPGs leads players to level up or increase their character's status, abilities and power through ever more upgraded armor and weapons, fighting skills, and recognition of accomplishments (Bainbridge 2010a; Barnett and Coulson 2010; Klasttrup and Tosca 2009).

DAVID KIRSCHNER AND J. PATRICK WILLIAMS ON THEIR METHODS

How do players align their actions in order to achieve collective goals? What role does the user interface play in facilitating this? What kinds of interpretations of objects and events do players make? How do players subsequently communicate meanings to one another? These were the research questions that underlay our study of World of Warcraft (WoW), a Tolkienesque, fantasy-based virtual gameworld that has boasted more than 12 million current subscribers worldwide (Blizzard Entertainment 2004; 2011). Using ethnographic methods including participant observation, in-depth interviewing, writing personal diaries and field notes, and collecting audio and video recordings, we participated in the social world of raiding. We had been playing WoW recreationally since 2006 and were familiar with many aspects of the gameworld, having created multiple characters and played them to

(continued on next page)

In both genres, players' abilities to improve and progress are channeled through a social filter: "as a player gains in levels, quests become increasingly difficult to accomplish alone, reaching a point where a coordinated group of players is required to move further" (Ducheneaut 2010:135). Unlike single- and multi-player games, MMORPGs remove simple pattern recognition and the amount of time one plays as key determinants of success, promoting instead the ability to engage in successful coordinated action with other players. In MMORPGs, collaborative forms of play emerge through participation in fantasy-cultural milieus where players socialize each other to play in ways that are structured by game designers (Bartle 2003; Salen and Zimmerman 2004). MMORPGs are thus sociologically interesting because of their socially-centered design as well as because of the ways people interact with and through them (see Bainbridge 2010b; Ducheneaut and Moore 2004; 2005; Nardi 2010). And yet, while scholars have attended to concepts such as community, identity, and cultures of play, the microsociological means through which players' coordinated action constructs the social fabric of MMORPGs has been largely overlooked or downplayed. Game and virtual world designers must implement user interface (UI) features supporting modes of communication that facilitate the social activities players engage in. As part of a response to the perceived gap in the literature, this chapter looks specifically at the user interface, how it supports nonverbal communication between the game and the player, and how it facilitates coordinated action among players. Using video-recorded gameplay, screenshots, and interview data, we build upon Mead's (1934) conception of social action as group-oriented behavior comprised of smaller individual acts. We frame the achievement of coordinated action in MMORPGs through analysis of the visual and auditory dimensions of World of Warcraft's UI. Designers and researchers alike may find the analysis useful to implement practically or build upon theoretically, as we discuss how these communicative interface dimensions, which support the collective activities constituting MMORPG gameplay, enhance one another, and thus the potential of computer-mediated social activity itself.

the maximum available level. However, neither of us had more than fleeting experiences raiding. Therefore in October 2009 we went in search of a stable group of players to raid with. Social groups are a generic feature of MMORPGs. In WoW, like-minded players band together to form "guilds," which provide social and technical support to players (Williams, Ducheneaut, Xiong, Zhang, Yee and Nickell 2006). After a couple of weeks of searching and a stint in a short-lived guild, the second author was referred to the leader of an established raiding guild, The Cleaning Crew, and in November, 2009, we were interviewed and subsequently invited to join.

This guild has become a "third place" (Steinkuehler and Williams 2006), a research site, and the locus of our engagement with the game to the present. We joined as novice raid members and eventually took on dual identities in the guild as players/researchers. To research social interactions in an MMORPG, one must actively participate in its culture, for "you cannot observe a virtual world without being inside it, and in order to be inside it, you have to be 'embodied'" (Pearce 2009:196). During approximately 50 hours of play with other guild members between November 2009 and January 2010, we developed reputations as relatively good and trustworthy raiders, which, along with our researcher identities, facilitated our becoming active in guild life. We chatted regularly with other guild members via text and voice chat and, between January and April 2010, participated regularly in scheduled weekly guild raids on WoW's "end game" areas. After three years of casual play, we found the familiar gameworld to be unfamiliar again. Through regular and sustained interaction with guild members we learned to (re)define gameplay from

(continued on next page)

1.

RAIDING IN THE WORLD OF WARCRAFT

Although there are many ways to play MMORPGs, successful “raiding” is considered by many to represent the pinnacle. In game lingo, raiding refers to a process whereby groups of players enter the most difficult and challenging areas of the gameworld and, through careful planning and coordinated action among their characters, learn to overcome powerful computer-controlled enemies called “bosses.” Raiding is the ideal version of collaborative MMORPG gameplay and represents the most complex form of simultaneous interactions among players and between players and the game. Unlike other parts of MMORPG worlds, raid areas are intentionally structured so that groups must progress through series of bosses, each with a unique set of abilities and conditions designed to frustrate players’ efforts. No matter how skilled players are, a mistake by a single individual will often result in the group being wiped out and having to try again. Raiding thus requires that players constantly (re)define situations by considering their own knowledge, goals and actions, and by learning to anticipate, interpret and efficiently respond to actions initiated by the game itself, while also taking into account the imagined knowledge, goals and actions of other players. More than any other form of MMORPG play (with perhaps the exception of advanced player-versus-player content), raiding requires that players commit to maximizing their knowledge of character classes and to learning the most efficient methods of play.³⁵

³⁵Only top-level characters are able to participate in raids, and the “top level” changes over time as the developer releases new content to maintain players’ interest in the game. In WoW, the average time it took a player’s character to achieve maximum in the original version of the game was nearly 400 hours of gameplay (Ducheneaut, Yee, Nickell, and Moore 2006; for a more recent discussion see Lewis and Wardrip-Fruin 2010). Beyond the commitment needed to reach top level, raiding itself is time-consuming, usually requiring 3-5 hours of dedicated playtime per gaming session as well as additional time spent preparing, e.g., reading strategies, watching videos, acquiring food, potions, enchantments, scrolls, and other items to increase one’s offensive and defensive capabilities.

a raider’s perspective, thus providing us with a “doubly privileged form of contact” with the MMORPG and its players (Prus 1996:20).

We collected data through three digital media: textual, vocal and visual. First, we recorded our text-based chat in multiple channels between November 2009 and April 2010 and sporadically afterward, resulting in more than 600,000 words of in-game interaction between ourselves, other players, and the game itself. The chat logs represent naturalistic gameplay interspersed with research oriented discussions and brief informal interviews.

Second, we recorded vocal communications sent through a voice over IP (VoIP) program popular among gamers because it allows talking to replace typing as the primary method of verbal communication. We recorded vocal data from January 2010 through April 2010 for every raid, in addition to other non-raid conversations that appeared relevant to our research questions. For example, if guild members began discussing boss strategies, events from a recent raid, or guild-related issues, we began recording. In total, we recorded 36 VoIP sessions averaging 153 minutes each. Third, we recorded video of raid boss encounters, often from two perspectives, using a video capture program that also records and preserves in-game audio and external voice communication, producing richly layered data. In the end we analyzed 52 videos averaging 4 minutes and 49 seconds. Watching videos of ourselves later (i.e., observing the participant observers) was not only a reflexive exercise, but allowed us a very detailed examination of audio and visual aspects of player interaction. It also facilitated our thinking about player interactions in raids as we could literally watch raid members’ role performances. Finally, we interviewed

(continued on next page)

We have uploaded a video at <http://www.youtube.com/watch?v=BEcs4KZdWGk> to help illustrate some of our descriptions of the UI as they pertain to raiding in WoW and suggest that readers pause to watch the video now.

As with most advanced groups in MMORPGs, raids need a flexible combination of players to succeed. That flexibility is shaped by a core feature of MMORPGs—character class. In fantasy MMORPGs, a character class is an archetype such as warrior, priest, or hunter, each of which may specialize in one of several areas of expertise that define the character’s primary role. There are three primary class roles in MMORPGs: tank, healer, and DPS (an acronym for “damage per second”). Two seconds into the video [0:02], the user moves toward the flurry of activity on the top of the screen, where there is a paladin (class) tank (role) surrounded by enemies. Tanks, like the military vehicle from which the role takes its name, are heavily armored characters that hold an enemy’s attention, or aggression (“aggro”), so that other players can concentrate on performing their roles with minimal interference. Since tanks take constant damage, they need constant healing. Looking on the right-hand side of the screen, notice two anthropomorphic trees with green light emanating from their hands. These are druid (class) healers (role). Healers are tasked with keeping tanks and other raid members alive. Some healers use powerful single-target healing spells focused on tanks, while others use area-of-effect (“AoE”) spells to heal many allies at once. DPS refers to characters that are responsible for damaging the boss. Most player characters on-screen are DPS, including the one recording the video. DPS are typically subdivided into some combination of melee versus ranged and physical versus magical damage. As this description suggests, WoW is designed with diverse sets of reciprocal role opportunities during play.

three senior guild members using the VoIP software, which averaged 85 minutes. The guild leader, two assistant guild leaders, and two high-ranking guild members were typically responsible on a week-by-week basis for organizing raids. We designed a semi-structured interview to draw heavily on their own understanding of the nature of game design, raiding and player motivations to give us a “top-down” perspective on building raid groups and promoting/supporting effective coordinated actions. We analyzed data from all of these sources using a combination of inductive and deductive strategies, relying on open coding and existing literature to develop an understanding of how players utilize various aspects of the user interface in their interactions.

The most difficult aspect of raiding is not eliminating bosses per se, but rather coordinating player action during encounters. Only when players are able to synchronize their characters’ respective role performances can they defeat bosses and obtain rewards. Many players will encounter the same boss dozens of times before successfully defeating it. Further, raid bosses “respawn” every week, offering repeated opportunities for groups to hone strategies and build teamwork in order to smoothly progress through the raid area. Weekly repetition results in routinization of interaction where “respective identities and roles [become] essentially given and unproblematic, so that negotiation is mainly a matter of all recognizing the governing occasion or situation” (McCall 2003:331). Nevertheless, due to the random make-up of many raids, and the general complexity of boss encounters, raids continue to showcase the emergent nature of social action, warranting an analysis of the interface through which players interpret and act upon symbolic communications involving the game itself and other players.

In the remainder of this chapter, we look specifically at the technological interface through which WoW players interact with other players and with the game. The interactions can be broadly classified as player-to-game, game-to-player, and player-to-player, where the game mediates, i.e., player-to-game-to-player. Through the UI, players receive, transmit, and interpret information in dynamic situations. Our goal is to identify some of the particular bits of technoculture used in these situations, while highlighting the role that they play in mediating social action.

2. THE USER INTERFACE

Output devices (e.g., screen, speakers) project images and sounds that compose raid boss encounters, while input devices (e.g., keyboard, mouse, microphone) are the means through which players communicate with the game and other players. WoW's UI (see Figure 18-1) mediates and organizes the symbols that players use in coordinating their actions. Consider the short video extracted from one of our play sessions, where we are engaged with a raid boss named Lady Deathwhisper. The goal of the encounter is to first destroy a magical shield that surrounds and protects her while fending off a series of minions that she summons to aid her, and then to eliminate her. She regularly casts a green, circular area-of-effect spell called Death and Decay, which damages any character standing in it, and randomly takes control of a single player's character, making it hostile to the raid for a short time. That character must be subdued, but not killed, by other players until the effect wears off. While destroying her shield, the raid must also deal with her minions putting a curse on magic-using characters, preventing them from casting the same spell more than once every fifteen seconds. In the video, you can see a group of characters standing and moving around in various directions. You can see projectiles of various colors and shapes moving back and forth between the raid group and its enemies. You can see numbers appear, which represent the damage inflicted by our character on some of those enemies. You can see some colored effects surrounding certain characters and perhaps even notice that a few characters have red writing above their heads, marking them as enemies. Readers familiar with the game might understand that, for example, the healing druid becomes mind-controlled at 0:07 (represented by purple discoloration, swirling chains and character growth) and subsequently gets frozen in a block of ice by an ally. It is hard to say much more because the UI is turned off initially. However, 23 seconds into the video [0:23] the visual UI is turned on and things become quite different. In fact, those unfamiliar with MMORPGs are likely overwhelmed by the appearance of so much information at once. Note also that the audio is muted, so there is still another (hidden) layer of information not being dealt with yet. Still, non-MMORPG players are likely to understand little of what is occurring. Players, however, must learn to interpret this dizzying array of output data from the UI as they play. The UI is more than just graphical; it serves a pragmatic function, allowing players to connect to their characters and thus to “construct meaning and interpret [cues] as a [series of] orderly event[s]” (Hung 2009:7). Without the UI, players would be unable to orient to the situation, losing their ability to interact meaningfully with the game and with one another.



Figure 18-1: World of Warcraft's User Interface with customized content

VISUAL

Every aspect of *WoW*'s gameplay has a visual component, whether it is the potency of a particular weapon, chatting with friends, or being injured during a boss encounter. For example, the mouse cursor is both a symbol and a tool, similar to that on a typical computer screen, except that it changes shape when the player interacts with different game objects to represent situationally relevant information. The default cursor appears as a gauntleted hand. When a player sees another character, she can quickly ascertain whether it is a friend or an enemy by moving the cursor over the unidentified other. The gauntlet may change to a sack of coins, signifying a friendly merchant, or change to a sword, signifying an enemy (first seconds of video), or to one of a number of other meaningful symbols. In addition to the mouse cursor conveying specific information, it allows the player to target objects in the game world, orient the camera, and so on. Thus the mouse cursor offers one of the most basic parts of the UI that players use to construct a definition of the situation.

WoW characters rely on dozens of skills and abilities. For example, a healer has different spells for healing one or multiple targets instantly or over time, or for cleansing diseases or curses. To use a skill or ability, the player pushes a specific keyboard button or clicks the appropriate icon located along the bottom or right side of the screen on the “action bar” (see Figure 18-2). Notice the variety of feedback communicated from game to player that his action was, is, and soon again can be, carried out in-game, beginning at 0:31 in the video clip. The border of the icon representing the ability currently being used illuminates. When this character casts the “Haunt” spell at 0:34, notice its icon (the third from the left on the bottom row) becomes grayed out, meaning the spell cannot be cast again. This gray gradually recedes back into color, signaling its availability for use. Tactile interaction with the mouse and keyboard translates to character actions in the game. Interacting with one object (e.g., a targeted enemy) through the use of another object (e.g., an icon in the action bar) through the use of yet a third object (e.g., mouse or keyboard) is not always an easy task, and novice players especially have trouble making connections between the input commands and their character's actions (Hung 2009). The UI aids in the learning process with visual cues, helping players organize visual data in an orderly way.



Figure 18-2: Screen blurred to highlight action bars locations across the bottom and right side of the screen.

The character portrait is situated in the upper left-hand corner of the screen. It shows a picture of the character's face, health and special energy bars, level, name, and any friendly or harmful effects currently affecting the character. Clicking on any friendly or enemy character brings up a portrait of that target to the right. When a character receives damage or healing, the color-coded amount is displayed on the portrait, and when a character has aggro the outline of the portrait turns red (damage received and aggro are both visible in Figure 18-3). The UI allows players to arrange windows that are condensed versions of the character portraits showing much the same information, enabling players to see the statuses of other players. Healers, for example, need to see such information in order to perform their roles effectively, while other players, such as DPS, only need to see their portrait and that of the targeted enemy.



Figure 18-3: A healer's UI, showing this character portrait in the top left, the character portrait of his target just to the right, and miniature character panes of all other group members below.

Visual representations of actions, statuses and notifications of events provide both obvious and subtle cues that players interpret and act toward. When Lady Deathwhisper casts Death and Decay, a large, circular, green, bubbling animation appears on the ground. A character standing inside the green circle takes damage, visible numerically above its head and on the character portrait. Players must perceive and interpret either the green circle or the flashing numbers (or both) as indicative of damage being taken by their character to act accordingly. One of the first lessons players learn is that “fire is bad.” Here “fire” is symbolic of any harmful visual anomaly on the ground. Watch the players on the left side of the screen at 0:42 when Lady Deathwhisper casts Death and Decay. Players learn that when they stand in these effects, they take damage, die, and are often reprimanded by others for failing to move. Players should be discerning, however, and interpret a red circle with the same animation as a friendly spell, which is harmless to them.

Text, far from relaying only typed messages, has a variety of visual significances. Players first perceive and interpret the color, size and location of text, and allocate their attention accordingly to read and act on it. Players type to each other in chat windows, which work like an instant message service with various “rooms” players can join that are categorized by group membership or location and represented by different colors. Typing “/raid” allows one to type to all members of the raid group in orange text, typing “/whisper” followed by a player's name sends a private message to that player in purple, and so on. Raiders use the

orange raid channel to give instructions, review strategy, and engage in communications that most raiders assign high priority. The raid leader has the ability to strategically place text in a more prominent, central location via a “raid warning,” which bursts onto the center of the screen of each raid member, accompanied by a loud sound. Similar warnings conveyed by a custom game modification appear prominently in Figure 18-3. “If you need to emphasize something, you put it in raid chat. I don’t know why, but people seem to follow instructions from raid chat a lot more than they do instructions from [voice chat]. So, if I have an important point to make, I’ll either put it in raid chat or raid warning...” (Xeky, raid leader interview). The game itself also conveys information to players in the chat window. During boss encounters, various relevant and extraneous conversations co-occur in the raid channel, the guild channel, and in various active private channels. The multi-layered and colorful flow of textual interactions draws players into various roles that may support or weaken situational role performances.

Visual animations are the most persistent means through which the game communicates to the player and the primary, continuous means through which players communicate actions to other players, mediated by their characters. Each animation is symbolic—it must be interpreted and acted toward. When Lady Deathwhisper curses someone, a player must remove it. When she dominates a character’s mind, another player must counter it. These are the most effective responses players may have in these situations, and not responding quickly and precisely leads to characters dying. Players act on the basis of what has happened before and what will happen next, where their character is, where it is going or what it is moving away from, and so on. Game designers intend players to respond to cues given off by boss maneuvers in specific ways, and raid success hinges on players’ responses within boundaries defined by the game rules. Collective responses must be individually learned, communicated to other players visually or sonically, and then practiced together.

AUDITORY

WoW’s designers have implemented game mechanics with equal representational emphasis on visual and auditory dimensions, but because of the genre’s visual primacy (i.e., everyone must be watching the game to play), audio-use is less universal than video. Specifically, some players choose to disable or minimize in-game sound, or to replace it with music. Nevertheless, sound in *WoW* has rich symbolic value, providing information and orientation toward events in which characters’ visibility may be impaired or which occur off screen, as well as supporting what is visible. Two types of audio are worth distinguishing. Game-to-player audio refers to sound files that the software plays in connection with in-game events. For example, when a tank’s shield blocks an enemy attack, the game produces a sound that mimics a sword hitting a shield. Jørgensen (2008) argues that audio works simultaneously “as support for gameplay by providing the player different kinds of information that needs to be comprehended... [and] also by providing an understanding for how the game *should be played*, and *how to behave* in a specific in-game context” (emphasis added). Player-to-player audio, which is vocal interaction among members of a guild, raid, or other group, serves a similar purpose. Players interact with the game’s audio content and other players’ utterances, learning to correctly identify, interpret, and orient themselves to specific audio cues in order to effectively coordinate their behaviors. In other words, neither type of audio is produced or interpreted in isolation; both produce specific, situational meanings. Players without access to audio miss these communicative acts, which can negatively impact their ability to coordinate their actions.

Game-to-player audio tends to complement the role of visual data by providing an additional symbolic resource on which players may draw when defining actions and events during gameplay, though it sometimes fills in informational gaps from lack of visibility. For example, if a player hears the crackling sound of Lady Deathwhisper’s Death and Decay from behind her character outside her field of vision, she can imagine the event occurring and take precaution not to step backward into the spell. If the player does not know what Death and Decay sounds like, or does not know the source of the ominous crackling sound, the next recognizable sound may be that of a dying character. Bosses are vocal in communicating their

actions to players. All sounds Lady Deathwhisper generates are what Jørgensen (2008) calls “proactive... demanding evaluation or action on the part of the player.” When Deathwhisper casts Dominate Mind, she yells, “You are weak, powerless to resist my will!” In addition to hearing this through the computer’s speakers, the message appears in text chat in red letters, ensuring that players who disable the sound are still able to interact with the boss. For each taunt, the game conveys meaning through verbal content, vocal presentation (yelling), and nonverbal color and placement of text. When players hear or read this taunt, they may or may not interpret it as signaling that spell. Players who understand its intended meaning know what is happening without needing visual confirmation, whereas players who cannot differentiate this taunt from Deathwhisper’s other taunts, or do not draw a connection between this taunt and Dominate Mind, will remain unaware of the dominated character if it is out of their visual field. The importance of this interpretational moment is recognizing that utterances (indeed, all sound effects) directly correspond to symbolic phenomena that orient players toward in-game phenomena.

Once learned, players interpret “proactive sounds” Jørgensen (2008) along a hierarchy of urgency, prioritizing them based on their character’s role, the type of action or event the sound represents, and the imagined consequences of the resulting effects on characters. When Lady Deathwhisper announces her Dominate Mind attack with its corresponding visual cues, DPS players should become aware that one of the raid’s members is now targetable for attack. Players must ensure that, as they are switching among multiple targets, they do not accidentally attack and kill a fellow player. Meanwhile, for those few characters with the ability to subdue a dominated character, they must locate and prevent her or him from attacking other raid members. Determining the urgency of, and possible responses to, audio and visual cues requires a shared understanding of their potential meanings and taking into account how other players ought to respond to them. A good example of this situation is in the video at 0:07 when Lady Deathwhisper casts Dominate Mind on the healing druid. As the druid enlarges, a hunter (class) directly to its left freezes it in ice. Notice the blue line shooting from the hunter to the druid at 0:09. Player-to-player audio enables characters to quickly negotiate role performances such as this, but we will set aside analysis of verbal interaction in this chapter.

Demonstrating knowledge of the intricacies of boss encounters is necessary for those who want to continue raiding. Players whose characters do not perform their roles efficiently, often defined through their actions and their communicative and interpretive competencies via text and voice channels, may be put on a “no invite” list for future weeks if not kicked from the raid immediately. The more difficult the fight, the more important voice chat is, and the less extraneous conversations occur. The more routine a fight, the less players rely on verbal communication. Raiders are expected to do their homework each week before a raid begins by reading online guides and watching videos put up on websites such as Tankspot (www.tankspot.com). This is especially important for newer raiders, who have relatively less experience in large groups that require complex coordination. Online guides, however, function only as a means of anticipatory socialization (Thornton and Nardi 1975) and do not prepare one for the adrenaline-fueled, emergent nature of gameplay through which one learns and eventually personalizes her role(s). The communications that raid members engage in during boss encounters serve as important resources through which players practice and learn how to better interpret the various visual and auditory channels of the encounter.

3. FROM USER INTERFACE TO COORDINATED ACTION

In MMORPGs such as *WoW*, arriving at a shared definition of the situation is necessary for successful gameplay. The emergent nature of coordinated action is highly salient in raids, where learning the mechanics of the encounter encompasses this process as knowledge is refined over time to fit together

with the knowledge of other players. Knowledge finds its form in action, and boss encounters are resolved through the alignment of individual players' lines of action. *WoW* is interesting because of the mixture of emergent social action with the technological determinism that is coded into its design structure. Defeating bosses is predicated on the coordinated actions of players who are connected to one another through keyboards, cables, and data streams. Players must direct their characters to act according to a shared definition of the situation, performing actions in a highly rationalized sequence while constantly imagining the expectations and behaviors of their fellow players. In this way, social action becomes almost algorithmic. Game design uses algorithms (i.e., sets of instructions for carrying out procedures via a finite series of steps) as a method of providing obstacles to player's goals. Players must learn how to pragmatically interpret the algorithms that underlie any particular boss encounter (indeed, any encounter at all between the player's character and the virtual environment) and then act on that knowledge in concert with other players. Further, players must subsequently each develop an algorithmic play style that maximizes the potency of their class-based role(s).

In a game characterized by social interactivity, we noticed that raiders were constantly striving to maximize efficiency, with the result that players' interaction with the game sometimes eclipsed interaction with other players. If players' actions become mechanized out of an algorithmic imperative, then to what extent are players actually interacting with each other through the use of the UI versus solely with the UI itself? The social activity of taking the role of other players is replaced with the single-player mentality of top scores or fastest reaction times. For example, a healer describes a successful role performance: "I mean, just watch people's health bar. When the health bars go down, you fill it back up" (Vaid, interview). The "people's" importance is far surpassed by that of the health bar. During boss encounters, player-to-player and player-to-game interactions are always role-to-role interactions. The "who" matters only insofar as players are engaged in coordinating role performances. "That was my job, just to keep the tanks alive. If the tank dies, you know, probably my fault. Yeah, if the tank dies, the raid dies. So, my fault" (Vaid, interview). It is important for healers to be able to make the distinction between the tanks' and other characters' green health bars. The distinction is more or less necessary depending on how routine the encounter is and how well the team as a whole coordinates their actions. Difficult or less practiced situations require greater attention to detail because of their emergent characteristics.

"Gameplay is not a feature designed into the game alone, but an emergent aspect of interaction between the game system and the player's strategies and problem solving processes" (Jørgensen 2008). Because coordinated behavior is a requirement for successful end-game play, one's actions are equally oriented toward other group members and the game itself. The act of filling up green bars cannot be conceived of as being asocial because here one is filling green bars for the purpose of keeping teammates alive. The green health bars of allies, like the red health bars of enemies, are symbolic representations of the shared fantasy within which players interact with each other as well as with social objects in the game itself. Being self-conscious of the act and imagining its effects on, and the potential responses to it by, other people makes it a social act (Mead 1934). A healer can imagine that if she stops healing, player characters will die and blame her, and likewise can imagine that if she continues healing, then player characters will be able to continue performing their roles and may praise her. Continuing to perform one's role under the assumption that everyone else is doing the same is the fundamental process underlying smooth social interaction. Consider this example: A cursed healer and a dying mage (a class that can remove curses) are running toward one another. The mage knows he is almost dead and may assume the healer is running toward him to heal because she notices the mage's depleted health bar. The healer knows she cannot heal the mage since her healing spells are temporarily unavailable due to the curse, and may assume that the mage is running toward her because he sees that she is cursed and intends to remove it. Each is making an assumption of the other's intentions through role-taking, imagining the other player is acting toward the visual symbols of the low health bar and icon representing a curse. The positive resolution to this scenario may eschew visual interaction and instead rely upon textual or vocal interaction. The healer may type or say, "decure me!" Everyone is engaging in the social act of playing the game, but as in the example above,

players interact consciously with others in mind, and sometimes they do not. Conscious recognition of other players is often muddled because all player-to-player interaction in *WoW* is computer-mediated, which masks aspects of other players. An exception would be of players sitting in the same room talking together, a potentially more complicated situation in which players must negotiate multimodal spaces, interacting visually, verbally and nonverbally, face-to-face and hardware-mediated, with physically and digitally co-present others (Keating and Sunakawa 2010). Yet even without physical co-presence, players still negotiate among multiple communicative layers, in some ways made more difficult by the absence of proximal bodily cues and therefore making it more important to be able to interpret non-verbal communicative acts in shifting contexts.

A haphazard or lackadaisical approach to collaborative gameplay will likely negatively impact everyone involved, which can promote indignation, bickering, or a disbanding of the group. To avoid such problems, players must be able to imagine encounters from multiple perspectives, taking themselves as objects, and trusting that other players possess a similar level of reflexivity in their role performances. Since each boss encounter is uniquely algorithmic, players are required to perform their roles within a set of continually negotiated situational demands. Interpreting visual and auditory symbols from the game and other players allows for player creativity and problem solving. When players are able to interpret symbols and define game objects in the same way, or to understand that each player is defining an object in a way appropriate to his or her role, and the interpretations of these meanings lead to joint action for a collective goal, then a shared definition of the situation exists and joint action will likely be successful, as collectively and situationally defined by raid members. Voice communication and user-created game modifications are two technical innovations that standardize audio and visual cues, and to some extent, the meanings and intended interpretations of those cues. As shared meanings become more common, raids become more successful.

4. COMPUTER-MEDIATED COMMUNICATION, VIRTUAL WORLDS, AND COORDINATED ACTION

In MMORPGs such as *World of Warcraft*, players are bombarded with information that they must make sense of and act back on if they wish to master the intricacies of collaborative play. The amount and frequency of data streaming onto the screen and through the speakers sometimes make it difficult to decide what is important. Some players choose to rely on certain media within the UI over others for most of their information and communication needs. Perhaps the most oft-cited theory for explaining why people prefer one communicative medium over another is the psychologically-oriented media richness theory, which holds that the richest media, the one that most reduces uncertainty and equivocality among interactants in a situation and best captures the essence of face-to-face interaction, is best suited for a particular moment of communication (Daft and Lengel 1986). Media naturalness theory introduces an evolutionary augmentation to richness, claiming that the human brain handles face-to-face interaction best and that communication with lower degrees of “naturalness” (e.g., communication that is non-face-to-face, asynchronous, or that has expressive cues filtered out) poses cognitive problems. Media like email are both less rich and less natural, while “super-rich virtual reality media” like online video games are less natural because they are too rich, i.e., they possess excessive stimuli (Kock 2004:340). Media richness and naturalness theories operate under assumptions that face-to-face interaction is ideal and are thus biased in how they frame non-face-to-face communicative environments and channels, seeing new media environments particularly as inherently problematic. Yet, it is new media environments such as virtual worlds with and within which a growing number of people regularly interact.

In this chapter we have taken a more open approach to the significance of informational communicative media. Our approach is that of symbolic interactionism, a sociological perspective that emphasizes the social aspects of meaning-making, interpretation and context in human activity. Symbolic interactionism emerged from American pragmatism and from the work of scholars such as George Herbert Mead, Charles Horton Cooley, and John Dewey. Herbert Blumer (1969), a student of Mead, rigorously developed a set of premises upon which symbolic interactionism has come to rest. First, people act toward things on the basis of the meanings those things hold for them. Second, the meanings of things arise out of social interaction with others. And third, people handle and modify meaning through an interpretative process in dealing with the things they encounter. The symbolic interactionist approach has more in common with communication theories such as “media synchronicity,” which focuses on “the extent to which individuals work together on the same activity at the same time; i.e., have a shared focus” (Dennis and Valacich 1999:5). Unlike media richness and media naturalness, which focus on users’ optimized media choices, media synchronicity highlights communication performance, or how media enable users to achieve synchronicity and successful communication. Here communication consists of two basic processes: conveyance (the transmission of new information and its processing and interpretation through individual cognition) and convergence (the discussion of subjective interpretations to reach intersubjectively shared meaning) (Dennis, Fuller and Valacich 2008).

Conveyance and convergence processes often blend or co-occur in hyper-action-oriented settings within online games. Media synchronicity is not rigidly defined for a given medium in a given situation (Kahai, Carroll, and Jestice 2007), but is instead situationally emergent, simultaneously shaping and being shaped by the interplay of player, technology, and context. MMORPGs and other computer-mediated environments support a host of overlapping communication channels. The interactionist approach espoused here focuses not on *why* gamers choose among communication media, but on *how* they effectively deal with specific communicative media during gameplay. The benefit of a focus in communication performance highlights the intersubjectivity of communication, where “meaning derives from interactive interpretation by multiple persons, not simply from the cognition of a single individual” (Miranda & Saunders 2003:88).

Computer-mediated communication (CMC) research has already focused extensively on business and organizational concerns in interaction and interpretation in areas such as trust-building and virtual teamwork (e.g., Jarvenpaa, Knoll, and Leidner 1998; DeLuca and Valacich 2006). One goal has been to understand which media are best suited for facilitating such activities to result in better performance, higher productivity and more effective problem-solving. CMC scholarship has also been useful in analyzing communication and coordinated action among groups in MMORPGs. In the 21st century especially, there have been increasing calls across disciplines for researchers to explore communicative potentials and performances in virtual worlds (e.g., Castronova 2006; Davis et al. 2008; Bainbridge 2010b; Montoya, Massey, and Lockwood 2011). Research has attended to analyzing the functions and uses of these modalities, beginning with text-based communication in multi-user domains and moving on to the increasing support of voice chat in MMORPGs (Mortensen 2006), and even into the futuristic technology of real-time video representation in virtual worlds (Van Broek, Lou, and Van den Broek 2011). A number of studies have looked at player preference for, and in-game performance with, text versus voice chat in virtual worlds. A study of PC first-person shooter games found that “where gamers used text, their gameplay as a [team] was less cohesive. [The players] had more trouble coordinating strategy and their scores were lower... they explained that text communications...[were] something they now dislike, much preferring the social experience of being able to talk” (Halloran, Rogers, and Fitzpatrick 2003:138). Another study of text and voice chat among *WoW* guilds concluded that “the social impact of playing *WoW* with only text was clearly negative” (Williams, Caplan, and Xiong 2007:443) because text chat did not simulate presence or promote the stronger social bonds that voice chat did. Being able to hear teammates increased feelings of group membership, perceived likeability, and willingness to trust, all of which are important for cooperative gameplay (Chen 2009).

A notable trend is that new developments in MMORPG communication supplement rather than replace previous developments, resulting in multi-modal communicative environments. Emoticons, for example, are widely used in graphically rich virtual worlds for reasons similar to their continued use in SMS, instant messaging, and email: to efficiently express emotions (Walther 2006). In most cases, the expressive potential of characters is limited. Smiles, winks, and nods are communicated through text rather than through the visual manipulation of facial features. Characters' bodies are, however, able to communicate information. Manninen (2001), for example, observed that players of the first-person shooter game, *Counter-Strike*, communicated to one another through character movements, such as pointing and moving rapidly back and forth to tell a teammate to move in a specific direction. In that study, players had access to text chat and were co-present, yet they often persisted in using computer-mediated gestures rather than talking or typing to one another.

In another example gleaned from our data that shows how different modes of communication mutually enhance one another, players use their characters' bodies to practice a boss fight. The raid leader, through voice and text chat, instructs a player-character to stand in the center of the room, pretending to be the boss. He then positions everyone else in their respective places to initiate the exercise. Some fights allow players to position their characters appropriately relative to one another before initiating the encounter, but this practice exercise of an unfolding fight highlights the performative aspect of characters and the rich visual information their movements and positioning communicates. The raid leader explains the fight step by step, and instructs player-characters to move in relation to one another, based on their roles, and the "boss," in accordance with how it moves and acts in a real fight. In this way, characters' bodies become visual learning tools so that the raid team can practice the intricate coordinated movements necessary for succeeding in a complex encounter without the interference of real enemies. This creative communication is sometimes an orthogonal, unintended outcome of design decisions which we can identify through observation and analysis. These communicative tools, and the ways players utilize them, can then be meaningfully integrated back into virtual worlds.

5. CONCLUSION

The confusing chaos of visually and sonically dazzling effects, bosses, minions, and other players' characters running this way and that, all emoting and creating noise must in pragmatic terms be made meaningful, interpreted, and acted toward. Players do not simply choose the richest or most natural medium through which to communicate. Rather, in raiding (a situation demanding high synchronicity), players engage with specific parts of the UI as the situation unfolds. For those who put in the long hours, the day-to-day coordination of actions among raid members results in evolving sets of interpersonal and digital-media competencies, i.e., expertise, that include navigating across the computer screen, the ability to collaborate by recognizing the visual representations of other players' actions, and overall awareness of digitally-mediated situations (Reeves, Brown, and Laurier 2009). Each aspect of gameplay (moving, using abilities, reading a map, and so on) relies on skill sets integral to social roles, which players learn, refine and modify over time. A holistic sense of expertise develops as players practice "chaining together these small actions into temporally and componentially longer sequences of action" (Reeves et al. 2009: 214). By the time a player is raiding, he or she needs to have developed a thorough understanding of game mechanics, media literacy vis-à-vis the user interface, and a knowledge base that includes the various roles expected of her character's class and specialization.

Certainly virtual worlds, including MMORPGs, should not be viewed as a single medium, but rather as the interplay of physical and digital interfaces (mouse, keyboard, UI) that comprise multiple ways of communicating. None of the communication channels surrounding virtual worlds exist in isolation; each layer of the whole is symbolic and interpretable. And while players need to learn to navigate screens

full of visual and auditory stimuli, they are fundamentally being socialized into doing so through their interactions with other players, the game, and with the UI itself. This is not dissimilar to people in everyday life using other technological interfaces, from learning to navigate the sights and sounds of city streets via map applications on a smartphone to becoming socialized as an environmentally-friendly driver by striving to maximize the vehicle's energy efficiency according to the digital display of miles per gallon on the dashboard. The role of the UI as an agent of socialization in virtual worlds ought not be glossed over. Its expressive potential across platforms allows human creativity to flourish in everyday work and play. Designers of MMORPGs, virtual worlds, and other digitally-mediated environments largely take this into account already, as the interfaces they implement provide a range of communication channels to users, ideally facilitating interaction and community-building. Some researchers are already implementing novel methods of nonverbal communication in virtual worlds (see Innocent and Haines 2007). An understudied realm of especially high productivity in interface design is custom game modifications that users create and share in order to improve the gaming experience (e.g., Kow & Nardi 2010; Sotamaa 2010). Players themselves are flows of design creativity that developers tap into. Blizzard, developers of *WoW*, designed and implemented new UI features over the years to help players complete quests and learn to defeat raid bosses, drawing direct influence from prior user-developed game modifications. There is a symbiotic relationship between researchers, developers, and players. Each of them relies on the other to contribute to the continual development of novel UI designs and methods of CMC. Through the collaborative efforts of the organs in this relationship, users will have more choices for how to interact with one another in virtual worlds, and can thus accomplish more diverse goals in a variety of contexts.

6. BIBLIOGRAPHY

- Bainbridge, William Sims. 2010a. "When Virtual Worlds Expand." Pp. 237-251 in *Online Worlds: Convergence of the Real and the Virtual*, edited by W. S. Bainbridge. New York: Springer.
- Bainbridge, William Sims. 2010b. *The Warcraft Civilization: Social Science in a Virtual World*. Cambridge, MA: MIT Press.
- Barnett, Jane and Mark Coulson. 2010. "Virtually Real: A Psychological Perspective on Massively Multiplayer Online Games." *Review of General Psychology* 14(2):167-179.
- Bartle, Richard. 2003. *Designing Virtual Worlds*. Indianapolis, IN: New Riders Publishing.
- Blizzard Entertainment. 2004. *World of Warcraft*. Platform: PC. Blizzard Entertainment.
- Blizzard Entertainment. 2010. "World of Warcraft Subscriber Base Reaches 12 Million Worldwide." [www.blizzard.com](http://us.blizzard.com/en-us/company/press/pressreleases.html?101007). Retrieved October 8, 2010 (http://us.blizzard.com/en-us/company/press/pressreleases.html?101007).
- Blumer, Herbert. 1969. *Symbolic Interactionism: Perspective and Method*. Berkeley, CA: University of California Press.
- Castronova, Edward. 2006. "On the Research Value of Large Games: Natural Experiments in Norrath and Camelot." *Games and Culture* 1(2): 163-186.
- Chen, Mark G. 2009. "Communication, Coordination, and Camaraderie in *World of Warcraft*." *Games and Culture* 4(1): 47-73.
- Daft, Richard L., and Robert H. Lengel. 1986. "Organizational Information Requirements, Media Richness and Structural Design." *Management Science* 32(5):554-571.
- Davis, Alanah, John Murphy, Dawn Owens, Deepak Khazanchi, and Ilze Zigurs. 2008. "Avatars, People, and Virtual Worlds: Foundations for Research in Metaverses." *Journal for the Association of Information Systems* 10(2):90-117.
- DeLuca, Dorrie, and Joseph S. Valacich. 2006. "Virtual Teams In and Out of Synchronicity." *Information Technology & People* 19(4):323-344.

- Dennis, Alan R., Robert M. Fuller, and Joseph S. Valacich. 2008. "Media, Tasks, and Communication Processes: A Theory of Media Synchronicity." *MIS Quarterly* 32(3):575-600.
- Dennis, Alan R., and Joseph S. Valacich. 1999. "Rethinking Media Richness: Towards a Theory of Media Synchronicity." Proceedings of the 32nd Hawaii International Conference on System Sciences, 1999 January 5-8, Maui, Hawaii, USA. Computer Society Press. 10 pages.
- Ducheneaut, Nicolas. 2010. "Massively Multiplayer Games as Living Laboratories: Opportunities and Pitfalls." Pp. 135-145 in *Online Worlds: Convergence of the Real and the Virtual*, edited by W.S. Bainbridge. London: Springer.
- Ducheneaut, Nicolas, and Robert J. Moore. 2004. "The Social Side of Gaming: A Study of Interaction Patterns in a Massively Multiplayer Online Game." Proceedings of the ACM Conference on Computer-Supported Cooperative Work (CSCW 2004), 2004 November 6-10, Chicago IL, USA. NY: ACM. 360-369.
- Ducheneaut, Nicolas, and Robert J. Moore. 2005. "More than Just XP: Learning Social Skills in Massively Multiplayer Online Games." *Interactive Technology and Smart Education* 2(2):89-100.
- Ducheneaut, Nicolas, Nick Yee, Eric Nickell, and Robert J. Moore. 2006. "Building an MMORPG with Mass Appeal: A Look at Gameplay in *World of Warcraft*." *Games and Culture* 1(4):281-317.
- Fine, Gary Alan. 1983. *Shared Fantasy: Role-Playing Games as Social Worlds*. Chicago: University of Chicago Press.
- Halloran, John, Yvonne Rogers and Geraldine Fitzpatrick. 2003. "From Text to Talk: Multiplayer Games and Voiceover IP." DiGRA Proceedings, Level Up Conference. 2003 November 4-6, University of Utrecht, The Netherlands. 130-142
- Hung, Aaron Chia-Yuan. 2009. "The Order of Play: Seeing, Teaching, and Learning Meaning in Video Games." DiGRA Proceedings, Breaking New Ground: Innovation in Games, Play, Practice and Theory. 2009 September 1-4, London, UK. 8 pages.
- Innocent, Troy and Stewart Haines. 2007. "Nonverbal Communication in Multiplayer Game Worlds." Proceedings of the 4th Australasian Conference on Interactive Entertainment. 2007 December 2-5, Melbourne, Australia. 9 pages.
- Jarvenpaa, Sirkka L., Kathleen Knoll, and Dorothy E. Leidner. 1998. "Is Anybody Out There?: Antecedents of Trust in Global Virtual Teams." *Journal of Management Information Systems* 14(4):29-64.
- Jørgensen, Kristine. 2008. "Audio and Gameplay: An Analysis of PvP Battlegrounds in *World of Warcraft*." *Game Studies* 8(2). Retrieved January 16, 2009 (<http://gamestudies.org/0802/articles/jorgensen>).
- Kahai, Surinder S., Elizabeth Carroll, and Rebecca Jestic. 2007. "Team Collaboration in Virtual Worlds." *The DATA BASE for Advances in Information Systems* 38(4):61-68.
- Keating, Elizabeth and Chiho Sunakawa. 2010. "Participation Cues: Coordinating Activity and Collaboration in Complex Online Gaming Worlds." *Language in Society* 39(3): 331-356.
- Klastrup, Lisbeth and Susana Tosca. 2009. "Because it Just Looks Cool! Fashion as Character Performance: The Case of WoW." *Journal of Virtual Worlds Research* 1(3). Retrieved July 10, 2010 (<http://journals.tdl.org/jvwr/article/view/305/427>).
- Kock, Ned. 2004. "The Psychobiological Model: Towards a New Theory of Computer-Mediated Communication Based on Darwinian Evolution." *Organization Science* 15(3):327-348.
- Kow, Yong Ming, and Bonnie Nardi. 2010. "Culture and Creativity: World of Warcraft Modding in China and the US." Pp. 21-41 in *Online Worlds: Convergence of the Real and the Virtual*, edited by W.S. Bainbridge. New York: Springer.
- Lewis, Chris, and Noah Wardrip-Fruin. 2010. "Mining Game Statistics from Web Services: A *World of Warcraft* Armory Case Study." Proceedings of the Fifth International Conference on the Foundations of Digital Games. 2010 June 19-21, Monterey, California, USA. 100-107.
- Manninen, Tony. 2001. "Virtual Team Interactions in Networked Multimedia Games – Case: "Counter-Strike" – Multi-player 3D Action Game." Proceedings of the PRESENCE2001 Conference. 2001 May 21-23, Philadelphia, Pennsylvania, USA. 9 pages.

- McCall, George J. 2003. "Interaction." Pp. 327-348 in *Handbook of Symbolic Interactionism*, edited by L.T. Reynolds and N.J. Herman-Kinney. Lanham, MD: AltaMira Press.
- Mead, George Herbert. 1934. *Mind, Self, & Society*. Chicago: University of Chicago Press.
- Miranda, Shaila M., and Carol S. Saunders. 2003. "The Social Construction of Meaning: An Alternative Perspective on Information Sharing." *Information Systems Research* 14(1):87-106.
- Montoya, Mitzi M., Anne P. Massey, and Nicholas S. Lockwood. 2011. "3D Collaborative Virtual Environments: Exploring the Link Between Collaborative Behaviors and Team Performance." *Decision Sciences* 42(2):451-476.
- Mortensen, Torill Elvira. 2006. "WoW is the New MUD: Social Gaming from Text to Video." *Games and Culture* 1(4): 397-413.
- Nardi, Bonnie A. 2010. *My Life as a Night Elf Priest*. Ann Arbor, MI: University of Michigan Press.
- Pearce, Celia. 2009. *Communities of Play*. Cambridge, MA: MIT Press.
- Prus, Robert. 1996. *Symbolic Interaction and Ethnographic Research: Intersubjectivity and the Study of Human Lived Experience*. Albany, NY: State University of New York Press.
- Reeves, Stuart, Barry Brown and Eric Laurier. 2009. "Experts at Play: Understanding Skilled Expertise." *Games and Culture* 4(3): 205-227.
- Salen, Katie, and Eric Zimmerman. 2004. *Rules of Play: Game Design Fundamentals*. Cambridge, MA: MIT Press.
- Sotamaa, Olli. 2010. "When the Game is not Enough: Motivations and Practices Among Computer Game Modding Culture." *Games and Culture* 5(3):239-255.
- Steinkuehler, Constance, and Dmitri Williams. 2006. "Where Everybody Knows Your (Screen) Name: Online Games as 'Third Places'." *Journal of Computer-Mediated Communication*, 11(4), article 1.
- Thornton, Russell, and Peter M. Nardi. 1975. "The Dynamics of Role Acquisition." *American Journal of Sociology* 80(4):870-885.
- Valve Software. 2000. *Counter-Strike*. Platform: PC. Vivendi Universal.
- Van Broek, Sigurd, Zhe Lou, and Marc Van den Broek. 2011. "Real-Time 3D Video Communication in 3D Virtual Worlds: Technical Realization of a New Communication Concept." *Bell Labs Technical Journal* 16(1):63-76.
- Walther, Joseph B. 2006. "Nonverbal Dynamics in Computer-Mediated Communication, or :(and the Net :(s with You, :) and You :) Alone." Pp. 461-479 in *The SAGE Handbook of Nonverbal Communication*, edited by V. Manusov and M. L. Patterson. Thousand Oaks, CA: SAGE Publications.
- Waskul, Dennis D., and Matt Lust. 2004. "Role-Playing and Playing Roles: The Person, Player, and Persona in Fantasy Role-Playing." *Symbolic Interaction* 27(3):333-356.
- Williams, Dmitri, Scott Caplan, and Li Xiong. 2007. "Can You Hear Me Now? The Impact of Voice in an Online Gaming Community." *Human Communication Research* 33(4): 427-449.
- Williams, Dmitri, Nicolas Ducheneaut, Li Xiong, Yuanyuan Zhang, Nick Yee, and Eric Nickell. 2006. "From Tree House to Barracks: The Social Life of Guilds in *World of Warcraft*." *Games and Culture* 1(4):338-361.

19

THE UNCANNY VALLEY AND NONVERBAL COMMUNICATION IN VIRTUAL CHARACTERS

By Angela Tinwell, Mark Grimshaw, and Debbie Abdel-Nabi

This chapter provides a case study of a research project investigating aspects of facial expression and viewer response to the ‘Uncanny Valley’ (Mori, 1970) in realistic, human-like, virtual characters. In humans, there exists a rich variety of information that can be obtained through nonverbal communication (NVC) (Ekman, 1965). However, empirical evidence collected as part of this research project, suggests that this information is lacking in animated realistic, human-like, virtual characters (Tinwell, Grimshaw, & Williams, 2010; Tinwell, Grimshaw, Abdel-Nabi, & Williams, 2011). Specifically, a perceived lack of emotional expressivity generated in the upper face region during speech was found to strongly influence perception of the Uncanny Valley in realistic, human-like, virtual characters (Tinwell, Grimshaw, Abdel-Nabi, et al., 2011). Building on empirical evidence provided so far in this research project, the consequences of a lack of NVC in the upper face region with regards to perception of the uncanny in characters is considered. New theories as to the cause of the uncanny phenomenon are put forward that go beyond Mori’s original theory based on a perception of spontaneous versus deliberate facial expression, the detection of transient microexpressions with suggestion of possible deceit, and a recognition of psychopathic behavior. This chapter also considers the limitations of the stimuli and methodology used in previous experiments, with suggestions made for future experiments. The implications of a lack of NVC in virtual worlds are discussed with possible caveats against the use of characters in virtual simulations used for assessment purposes in the real world.

1. INTRODUCTION

Viewer perception of a character’s emotive state is of ever-growing significance in virtual worlds. Developers of cinematic, interactive-drama, video games such as *Heavy Rain* (Quantic Dream) and *LA Noire* (Team Bondi, 2011) claim to have taken the next frontier in computer generated animation within gaming. Advanced motion capture technologies and performance captured animation allows for increasing sophistication of characters’ emotional expressions, gesture and movement. The video game *LA Noire* requires the player to interrogate both civilians and suspects to solve crimes in the story. In order to reveal new clues, it is important that players understand the facial expression made by character’s that they interview. In the crime thriller *Heavy Rain*, the player’s aim is to identify a serial killer. To do so the player must continuously seek feedback from the characters facial expression and emotive state to inform decisions as to their next actions in the game. Failure to do so may risk not solving the mystery or protagonist characters becoming victims of the killer.

Virtual worlds are also being implemented outside the remit of gaming, for use in training and testing simulations, where a clear understanding of a character’s emotive state is crucial to achieve a successful outcome for the viewer. Such technology is being introduced as an innovative method of assessing and improving people’s skills with the intention that participants will then apply those skills to similar scenarios

in the real world (ACTIVE Lab, 2011; Bergen, 2010; Bowers et al., 2010; MacDorman et al., 2010). As such, virtual characters are being used in roles once exclusively assigned to humans both in the public and private sectors. For example, a virtual character may be used to present moral or ethical dilemmas to trainees in the medical, military and legal professions (ACTIVE Lab, 2011; Bergen, 2010; Bowers et al., 2010; MacDorman et al., 2010). A virtual character may deliver a personal problem to a trainee doctor so that the trainee doctor's response to that problem may be assessed (MacDorman et al., 2010); or be featured as an injured civilian in a warfare scenario who requires help from a trainee soldier (Bowers et al., 2010).

Similarly, commercial companies are implementing virtual worlds for psychological assessment as part of the recruitment process to find the best possible employees (Bergen, 2010). Potential employees interact with virtual worlds that simulate 'real world' scenarios featuring annoyed customers or argumentative colleagues. Companies seem keen to explore this new method of assessing possible new recruits as it improves the value of their employee branding and raises awareness of the company as a forward thinking and advanced place to work. Handler, who is involved in the simulation development of virtual worlds at the human resources software company, Kanexa states, "It's the wow factor," as companies seek to be established as an exciting and desirable place to be (Bergen, 2010). In this way, virtual characters are used to test how effective potential new recruits may be in coping with the potential challenges of a role. However, recent evidence suggests that the Uncanny Valley phenomenon may act as a limitation to this type of endeavor and companies should be aware of what effect the uncanny may have on recruiting the right candidate for the job. This chapter investigates how a lack of Nonverbal Communication (NVC) in realistic, human-like, characters may affect perception of the uncanny and the potential consequences of interacting with such characters in virtual worlds, used not only for entertainment, but for training and assessment purposes.

ANGELA TINWELL ON HER METHODS

This chapter provides an overview of a current research project investigating the Uncanny Valley phenomenon in realistic, human-like virtual characters. The research methods used in this work include a retrospective of both empirical studies and philosophical writings on the Uncanny.

No other research has explored the notion that realistic, human-like, virtual characters are regarded less favorably due to a perceived diminished degree of responsiveness in facial expression, specifically, nonverbal communication (NVC) in the upper face region. So far, this research project has provided the first empirical evidence to test the Uncanny Valley phenomenon in the domain of animated video game characters with speech, as opposed to just still, unresponsive images, as used in previous studies. Based on the results of these experiments, a conceptual framework of the Uncanny Valley in virtual characters has been authored to allow developers to design either for or against the uncanny for antipathetic or empathetic-type characters.

This research is relevant to embodied conversational agents used in a wider context such as therapeutic and e-learning applications and has an outreach to the disciplines of psychology, social psychology, game studies, animation and graphics, and human computer interaction.

The first section provides a synopsis of how experience of the uncanny was first considered in psychological writings of the twentieth century, leading to Mori's hypothetical notion of the Uncanny Valley (1970). An account of previous research investigating the uncanny in synthetic agents is also given including: exploration of factors, which may exaggerate the uncanny and design guidelines published on how to control the uncanny in virtual characters. The origins and adaptive functions of facial expression in primates are discussed in section two. This serves as a rationale of why perception of the uncanny was stronger for some emotions than others when presented in characters with limited upper facial movement (Tinwell, Grimshaw, Abdel-Nabi, et al., 2011).

Section three provides a retrospective of the role of NVC in humans and the purpose and associated meaning of movement in the eyebrows, lids and forehead. The psychologist, Paul Ekman, has established a conceptual framework of facial action in humans used to create facial expressions. Based on Ekman's work, descriptions are provided for facial actions used in NVC including the differences between voluntary and involuntary facial movements and the facial muscles used in fabricated expressions. As well as the purpose of NVC, this section examines how a perceived lack of NVC in facial expression may exaggerate the uncanny for characters in virtual worlds. Importantly, consideration is also given to the potential confounding impact on the viewer if such facial actions are missing in virtual characters. A summary of the importance of the role of NVC in creating believable, realistic, human-like, characters and the antithetical consequences of aberrant upper facial movement is provided in the conclusion.

Finally, it should be noted that while the focus of this current research project explores the centrality of the role of a perceived lack of NVC in the upper face with perception of the uncanny in virtual characters, the authors acknowledge that these actions should not simply be studied in isolation from other facial and body movements. The stimuli used in experiments have included vocalizations and speech, yet head and body movements can also complement NVC in various social contexts. This chapter does disclose how NVC in the upper face may contribute to a more multi-dimensional model to measure the uncanny.

2. THE UNCANNY VALLEY

The subject of the uncanny was first introduced in 1906 by the psychologist, Jentsch, who likened the uncanny to a state of uncertainty as to whether an object was real or unreal or alive or dead. Jentsch gave examples of objects such as automaton or wax-work, human-like dolls that may elicit the uncanny effect. As a way towards understanding why some objects were not regarded as aesthetically pleasing to the extent of frightening or repulsing the viewer, Freud (1919) described the uncanny as a state of confusion that occurred as a seemingly familiar object behaved in a strange or unfamiliar way. Freud implied that the uncanny may exist as a revelation of what is normally withheld or hidden from others, for example, a revelation of a repressed emotion or thought, resulting in odd or disturbing behavior.

Building on this literature, a potential benchmark has now been set for designers in creating a virtual character that is believably (and authentically) human in their appearance and behavior. This benchmark is referred to as "The Uncanny Valley" and derived from observations made by the roboticist, Marashiro Mori (1970), when designing androids. Mori hypothesized that humans would be less accepting of synthetic agents as their human-likeness increased and created a graph to demonstrate this theory (see Figure 19-1).

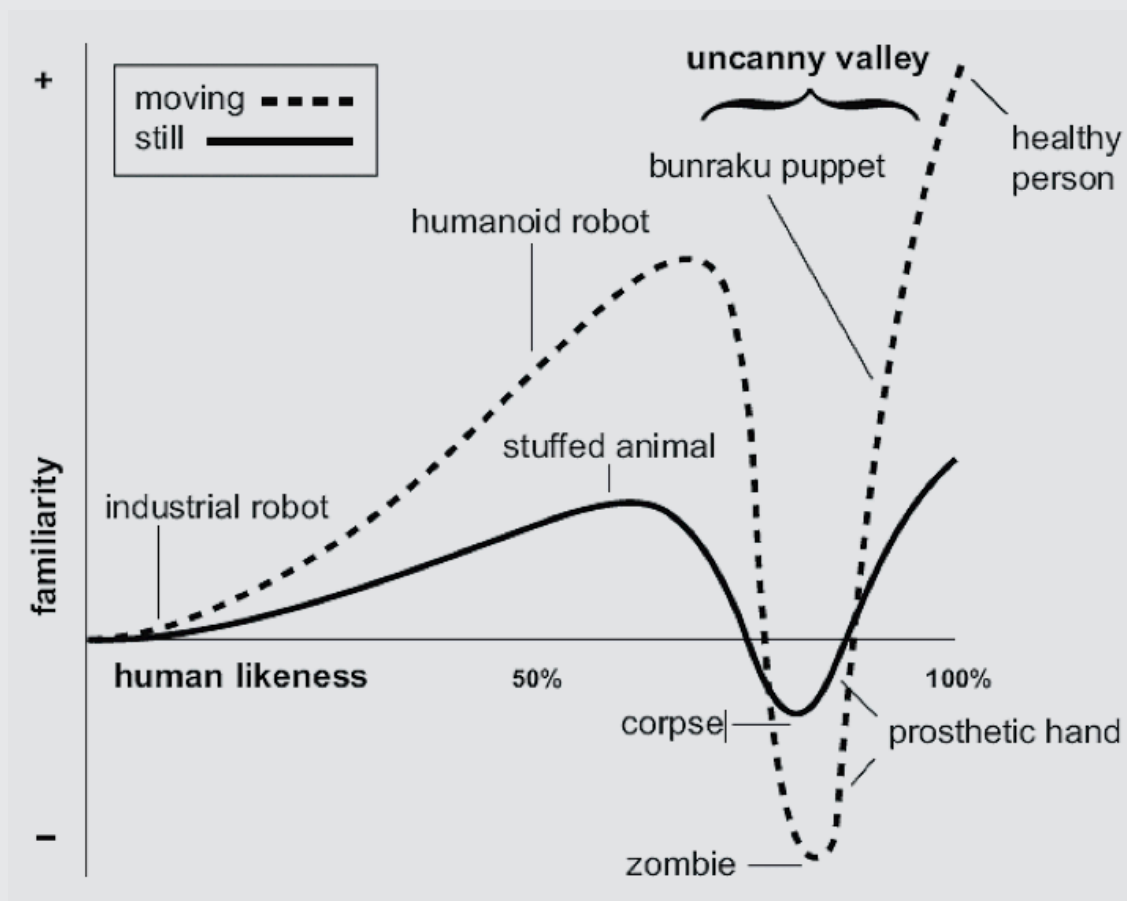


Figure 19-1: Mori's plot of perceived familiarity against human-likeness as the Uncanny Valley taken from a translation by MacDorman and Minato of Mori's 'The Uncanny Valley'

A valley shaped dip shows the negative affective response Mori observed in viewers towards android designs approaching fully, authentic human-likeness. Androids with a human-like appearance built an expectation from the viewer that their behavioral fidelity would match their human-like appearance. Androids that deviated from the human norm in their appearance and behavior repulsed the viewer, falling into the valley alongside zombies, corpses and prosthetic limbs. Mori (1970) recommended that for robot designers, it was best to aim for the first valley peak and not the second, developing humanoid robots with human-like traits, but avoiding android designs.

AUDITORY

With increasing realism achieved in virtual worlds, the Uncanny Valley phenomenon is frequently being associated with realistic, human-like, characters (see e.g. Geller, 2008; Pavlus, 2011; Plantec, 2007, 2008; Pollick, 2009; Stuart, 2007; Tinwell, Grimshaw, & Williams, 2011a). Viewers seem particularly discerning of characters' facial expression (Doerr, 2007; Gouskos, 2006; Thompson 2005). Critical of movements that appear odd or unnatural, parallels have been made between unsuccessful facial surgery in humans and virtual characters recognized as being uncanny. Thompson (2005) observed that realistic, human-like, comrade Marine characters featured in the video game *Quake 4*, "looked like the victims of thoroughly botched face lifts". Similarly, the admirably beautiful appearance of the heroine, Naomi (as modeled on the actress Anne Darrow), in the video game *King Kong* (Ubisoft Entertainment, 2005) was not enough to appease audiences and she was described as a monster. Facial expression that was regarded as stiff and distorted exaggerated the uncanny for this character, despite her aesthetically pleasing appearance.

In some ways, her avatar is an admirably good replica, with the requisite long blond hair and juicy voice-acting from Watts herself. But the problem begins when you look at her face – and the Corpse Bride stares back. The skin on virtual Naomi is oddly slack, as if it weren't quite connected to the musculature beneath; when she speaks, her lips move with a Frankensteinian stiffness. And those eyes! My god, they're like two portholes into a soulless howling electric universe. (Thompson, 2005)

With such criticism of characters designed to endear and not repel the viewer, designers required guidance as to how to control the uncanny in character design. Various factors have been found to influence perception of the uncanny. Importantly, humans can experience a less positive affective response towards androids and realistic, human-like, virtual characters when there is a perceived mis-match between a character's behavioral fidelity with their human-like appearance (Bartneck, Kanda, Ishiguro, & Hagita, 2009; Ho, MacDorman, & Pramono, 2008; Kanda, Hirano, Eaton, & Ishiguro, 2004; Vinayagamoorthy, Steed, & Slater, 2005). Viewers expect that, characteristics of speech (Tinwell et al., 2010; Tinwell, Grimshaw, & Williams, 2011b; Tinwell, Grimshaw, & Abdel Nabi, 2011; Mitchell et al., 2011), gestures and timing of movements (MacDorman, Coram, Ho & Patel, 2010; Ho et al., 2008; Minato, Shimda, Ishiguro, & Itakura, 2004), and a character's response to others and external events (MacDorman & Ishiguro, 2006) will match a character's human-like appearance. Failing this, the uncanny may be exaggerated for the character.

Despite the Uncanny Valley phenomenon often being regarded as a negative consequence, it can work to the advantage of characters within an appropriate setting and context. For example, robots designed with the intention to be unnerving and antipathetic characters, such as zombies within the horror game genre (MacDorman, 2006). Based on this analogy, a study was undertaken to investigate how cross-modal factors, such as motion and sound, may be manipulated to enhance the fear-factor in horror games (Tinwell, Grimshaw and Williams, 2010). One hundred participants rated 6, empathetic, realistic, human-like characters; 5 antipathetic, zombie characters; a Chatbot character with a stylized, human-like appearance; and a human.³⁶ The results indicated that a lack of human-likeness in a character's facial expression and speech exaggerated the uncanny. Specifically, perception of the uncanny was increased under the following conditions: an awareness of a lack of expressivity in the upper face region, the forehead area being of particular significance; a perceived over-exaggeration of mouth movement (articulatory motion) during speech; doubt in judgment as to whether the voice actually belonged to the character or not; and a perceived lack of synchronization between lip movement and speech (Tinwell et al., 2010).

Particular concern has been raised as to how the uncanny may have a potentially confounding impact on a participant's performance if presented with an uncanny character for assessment or training purposes (Bergen, 2010; MacDorman, et al., 2010). Doubt may be raised as to the validity and suitability of using virtual characters in such circumstances, given the negative effect associated with the uncanny. To eliminate the unintentional negative effect on those interacting with characters in virtual worlds (or to exaggerate the uncanny for those characters intended to be unnerving or frightening), designers should be aware of factors that may exaggerate or reduce the uncanny. Based on the body of research undertaken so far on the uncanny, this chapter investigates the relationship between NVC in virtual characters and the Uncanny Valley phenomenon. The next section examines more closely how a perceived lack of NVC in a character's facial expression, specifically in the upper face region, may be used to control the uncanny when designing emotive characters in virtual worlds.

³⁶Please see the original paper (Tinwell et al., 2010) for a full description of the stimuli used in this experiment.

3. FACIAL EXPRESSION OF EMOTION AND THE UNCANNY

It is well established that we rely on our innate ability to recognize and respond to others emotions as a primordial survival technique. The successful recognition of each type of the six universally recognized basic emotions, anger, disgust, fear, happiness, sadness and surprise, (Ekman & Friesen, 1978; Ekman, 1992a, 1992b) serves a different adaptive (survival or social interaction) function in humans and animals (Darwin, 1872; Ekman, 1979, 1992a, 1992b). Facial expressions evolved antecedently to help cope with fundamental life-tasks in our progenitors (Andrew, 1963; Darwin, 1872; Eibl-Eibesfeldt, 1970). The receiver can gain information about the sender's possible future behavior, or the event that may have elicited such an expression, and react aptly to that information (Ekman, 1979).

Importantly, the spoken language and movements of the body take second place to facial expression in human communication, "Words are not emotions, but representations of emotions" (Ekman, 2004, pg. 45). Evidence of this proposition has been found in rhesus monkeys (Izard, 1971). As more phylogenetically advanced mammals, the communication mechanisms of rhesus monkeys may be related to the evolutionary development of human communication (Darwin, 1872; Izard, 1971). An experiment was conducted with two groups of rhesus monkeys: In one group, the facial nerves of the monkeys had been cut so that they were unable to move facial muscles and create expressions; the other group had normal facial movement (Izard, 1971). When the groups merged, the monkeys with limited facial movement used full body gestures to communicate with other monkeys. Such gestures escalated to increased aggression and the attack of those monkeys with normal facial movement. Wary of their facial handicap, if the monkeys with limited facial movement felt fearful of others in the group, they had to resort to physical attack as a way to communicate that they felt threatened (Izard, 1971).

Similar patterns of behavior have been identified in members of the human population diagnosed with anti-social personality disorders (ASPD), otherwise known as psychopathy (Hart and Hare 1996; Herpertz et al. 2001; Lynam et al. 2011; Miller et al. 2011). Such people are more prone to violent outbursts as they lack emotions that would normally 'check' or regulate one's behaviour to prevent acting on impulse (Hart and Hare, 1996; Herpertz et al., 2001; Lynam et al. 2011). Hence, the patterns of behavior observed in our progenitors can still be observed in humans in those diagnosed with ASPD. It is imperative that humans can identify emotion promptly, to ensure minimal time is required to respond and act accordingly (Ekman, 1992a). Otherwise, the consequences of a delayed reaction, particularly to a perceived threat, may be life threatening.

Each emotion is perceived differently by humans, both cognitively and in terms of physical response (Darwin, 1872; Ekman, 1979, 1992a, 1992b; Johnson, Ekman, & Friesen, 1990). Humans react instinctively aversively to the emotions anger, disgust, fear, and sadness to avoid potential harm, threat or disease and the detection of such emotions in others may raise anxiety levels as a threat to one's own well-being. Evidence implies that there is distinct physiology (i.e. distinctive sympathetic autonomic nervous system (ANS) patterns) for separate emotions depending on the reaction required (Ekman 1992a, 1992b; Johnson, Ekman, & Friesen, 1990). Evidence of emotion-specific ANS activity has been found for anger, fear, disgust and sadness, but not so for happiness and surprise (Ekman 1992a, 1992b; Johnson, Ekman, & Friesen, 1990). An adaptive function for fear can be to run or flee from danger. When experiencing fear, evidence shows that blood rushes to the large skeletal muscles, to support such a reaction. Fighting may have been the adaptive action for anger, consistent with the movement of blood to the hands (Ekman 1992a).

Given that NVC in the face is the primary method of communicating the affective state of an individual, and the importance of recognizing and responding promptly to another's emotive state, Tinwell, Grimshaw, Abdel Nabi, et al., (2011) conducted an experiment to investigate whether inadequate movement in the upper face may have differential effects on perceived uncanniness depending on which emotion was being portrayed by a character. It was put forward that those survival-related emotions considered signals of a threat, harm or distress (including anger, fear, sadness and disgust (Ekman, 1979)) would be regarded as more uncanny in near human-like, virtual characters; especially when part of the facial expression was aberrant, impeding one's ability to recognize the emotional state of a character (Tinwell, Grimshaw, Abdel Nabi, et al., 2011). Emotions regarded as less important for survival such as happiness and surprise would be less noticeably strange or uncanny, even when the animation of facial features appeared odd or wrong to the viewer.

One hundred and sixteen participants rated perceived familiarity and human-likeness for head shots of a male human and a male virtual character ('Barney' from the video game Half-Life 2, (Valve, 2008)) who produced prosodically congruent utterances for the six basic emotions (in addition to a neutral expression). Two experimental conditions were created for the virtual character: a fully animated character (named *full*) and a partially animated character where movement above the lower eyelids was disabled (named *lack*) (see Figure 19-2). The results showed that obscuring the salience of an emotion by limiting NVC in the upper face did exaggerate the uncanny.



Figure 19-2: The three Conditions, Human, Full and Lack, expressing the emotion Anger.

Participants rated the virtual character as less familiar and human-like (more uncanny) than the human, but significantly more so when facial signals were removed from the upper face. As the authors predicted, the extent of this increased uncanniness varied depending upon which emotion was being portrayed. Yet, the results indicated that emotions with distinctly *different* adaptive functions (social vs. survival) were regarded as more tolerable than others with an ambiguity of facial expression. In the lack condition, perception of uncanniness was strongest for the emotions fear, sadness, disgust, and surprise. Tinwell, Grimshaw, Abdel Nabi, et al. (2011) postulated that as both fear and sadness can require only small facial movements, the character may have resembled a corpse-like state with no movement in the upper face, hence exaggerating the uncanny. It was suggested that as surprise is commonly mistaken with fear in humans (Ekman, 2003b), the same characteristics attributed to fear in the lack condition may be applicable to the emotion surprise. Despite disgust possibly serving as a warning signal to others to avoid a repugnant object or situation (Blair, 2003), a lack of detail in creating the folding and wrinkling of the skin in the upper nose (referred to as the "nose wrinkler" action (Ekman, 1992a, 1992b)) was regarded as a more viable explanation as to why disgust was rated significantly more uncanny with restricted upper facial movement.

As predicted, participants were less sensitive to the uncanny when the emotion happiness was presented in characters with a lack of upper facial movement. Yet, unpredictably, the uncanny was also less noticeable for anger in characters with restricted upper facial movement. This finding was accounted for in that participants were able to recognize anger and happiness due to action unit movement in the lower half of the face, articulation and prosody (Tinwell, Grimshaw, Abdel Nabi, et al., 2011). Surprisingly, the results also revealed that happiness was rated the least familiar and human-like (most uncanny) when presented in the fully animated character. This finding was unexpected due to happiness being a more positive emotion. Tinwell, Grimshaw, Abdel Nabi, et al. (2011) suggested that participants may have been suspicious of being presented with a “false smile” (Ekman & Friesen, 1982, pp. 244-248) thus, exaggerating the uncanny.

Based on these findings, it was recommended that designers make informed decisions as to how to control the perceived level of uncanniness in virtual characters. Strategic design modifications can be made when modeling the upper face region, bespoke to each different emotion. To reduce the uncanny, substantial amounts of time need not be invested to convey upper facial expressivity for anger and happiness. However, particular attention should be paid to facial expressivity in the upper face for the emotions, fear, sadness, disgust, and surprise. These guidelines may be reversed if the designer wished to exaggerate the uncanny. For example, characters may be perceived as stranger and less human-like with reduced upper facial animation when expressing the emotions fear and sadness (Tinwell, Grimshaw, Abdel Nabi, et al., 2011).

4. NONVERBAL COMMUNICATION (NVC)

Behavior (either conscious or subconscious) in the presence of others is instilled with meaning. The perceiver may obtain critical information as to the emotional state of a person by observing behavior of their face or body (Ekman, 1965, 1979, 2004; Ekman & Friesen, 1969, 1978). Nonverbal communication can be detected by: the intonation, speed, and volume of a person’s speech; body posture; gestures; and facial expression, that may otherwise render verbal communication ambiguous (Ekman, 1965). For example a narrowing of the eyes and a shaking fist show that a person is angry (Ekman, 2004). Empirical evidence shows that the perceiver can make accurate judgments about a person’s attitude, personality and emotive state (in keeping with independent assessment as to the personality traits and circumstance of an individual) from observing movements of their face and/or body.

In studies designed to determine what kinds of information could be derived from observing facial or body behavior we found that inferences about emotions, attitudes, interpersonal roles, and severity of pathology can be made by observers with no specialized training in emotion recognition. (Ekman and Friesen 1969, p.50)

So that the great number of human facial actions could be described, Ekman and Friesen (1978) devised a categorical scheme based on facial anatomy, called the Facial Action Coding System (FACS). FACS provided a method so that individual muscles or combinations of muscles used in creating facial expression could be recorded systematically (Ekman and Friesen, 1978; Ekman, 1979). The muscular activities used to generate changes in facial appearance for a particular emotive state have been assigned specific Action Units (AU). In the upper face, AU1, *Inner Brow Raiser*; AU2, *Outer Brow Raiser*, and AU4, *Brow Lowerer*, can work independently or together to achieve one of the seven observably distinct brow actions in humans across the six basic emotions (Ekman, 1979; Ekman and Friesen, 1978). When contracted, the facial muscle *Depressor Glabellae* (described as AU4) lowers the eyebrows resulting in a corrugated furrow between the brows, creating a frown expression. AU1 defines the changes in appearance when the inner section of the frontalis muscle is contracted (Ekman, 1979). Wrinkles appear (or deepen) in the centre of the forehead as a result of the inner corner of the eyebrow being raised. AU2 defines the visual changes that occur when the outer frontalis muscle is contracted. This action creates wrinkles in the lateral (outside) portion of the

forehead. As a signal of surprise and interest the AUs 1 and 2 work in cohesion to open the eye area and increase the visual field (Ekman, 1979; Ekman and Friesen, 1978).

Various quantitative studies have been undertaken to unravel the complex interrelationships between spontaneous and contrived facial movements (Duchenne, 1862; Ekman, 2003b; Ekman & Friesen, 1982). Ekman distinguished two different types of facial social signals to convey NVC in humans: *Emotional Expressions* and *Conversational Actions* (Ekman, 1979). Both emotional and conversational signals can occur during speech, but are related to different components of conversation. Conversational signals tend to be made voluntarily and may replicate, precede, occur simultaneously or in some cases, replace a spoken word. Emotional signals occur involuntarily and are used to support the semantic meaning of a word. The listener may also use emotional signals to convey that they understand, agree, or disagree, what the sender has said (Ekman, 1979).

In a way towards classifying the nonverbal signals presented during speech Ekman and Friesen (1969) included the two categories *Emblems and Illustrators*, both of which use movement in the upper part of the face. Brow lowering and raising can be used as Illustrators to provide additional emphasis for words or phrases during speech. A lowered brow (AU4), associated with more negative emotions such as fear, sadness, distress and anger may help accentuate a negative word, whereas a raised brow (AU1+2), associated with more positive emotions such as Happiness and Surprise may “baton-accent” a more positive word such as “easy, light, good, etc.” (Ekman, 2004, p.42). Ekman suggested that the brows are more frequently used as conversational signals than other facial actions, because they are contrastive (positive vs. negative) and easiest to perform (1979). Emblems can include movement of the hands, shoulders, head and face. For example, raised eyebrows typically demonstrate the emotion Surprise. Mostly intentional, Emblems are performed in the “presentation position” (Ekman, 2004, p.40) when facing people and a person is aware of presenting Emblems to others. Interestingly, some can occur unintentionally, as the person unwittingly divulges repressed or deliberately suppressed information. A person is unaware they have made such a gesture, akin to “a verbal slip of the tongue”, (Ekman, 2004, p.40).

In 1872, Darwin suggested that facial expression can give away a person’s true feelings, despite efforts to try to conceal or hide an emotion. People are also aware of being presented with a false expression, fabricated to try to convince them an emotion is actually felt (Darwin, 1872). Furthermore, people may be unable to mask those facial movements that are most difficult to make voluntarily (Darwin 1872, Ekman, 2003b; Ekman & Friesen, 1969b). For example, one may be able to control body movements, such as a clenched fist, in an attempt to hide that one is angry, but may not be able inhibit the momentarily passing of a frown expression. To investigate how facial expression may be used to detect possible deception, Ekman conducted experiments to establish which facial movements could not be made deliberately (Ekman, 2001, 2003b). Those movements that could not be made without the involuntary processes of spontaneous emotional response could then be recognized as reliable signals of a person’s emotional state (Ekman, 2001, 2003b; Ekman & Friesen, 1969b). For anger it was identified that AU24, *Lip Pressor*, that creates a tightening of the lips, could not be activated voluntarily. This finding may explain why the uncanny was less noticeable in virtual characters expressing the emotion anger when movement was disabled in the upper face as pressed lips clarified the authenticity of this emotion (Tinwell, Grimshaw, Abdel Nabi, et al., 2011). As well as a droopy mouth (activated by AU 15, *Lip Corner Depressor*), Sadness requires that AU1, the Inner Brow Raiser, be active for people to perceive this emotion as felt (Ekman, 2001, 2003b). As the uncanny was increased for those characters communicating sadness without upper facial movement (Tinwell, Grimshaw, Abdel Nabi, et al., 2011), this emotion may have been perceived as superficial without evidence of this facial action. The combination of AUs 1+2+4 cannot be voluntarily activated for the emotion fear (Ekman, 2001, 2003b). Ekman (1979) speculated that, as fear is considered an anticipatory response of distress, the origin of this facial movement may be explained in the following ways: AUs 1+2 provide greater visual input, thus increased attention; AU4 communicates that a distress experience is expected.

Without evidence of the reliable expressions (Ekman, 2001, 2003b; Ekman & Friesen, 1969b), one may question the authenticity of the portrayed emotion. If one attempts to communicate joy, when one is otherwise experiencing more negative emotions, the smile action that uses the *Zygomatic Major* muscle (referred to as AU12) will be used in conjunction with other facial expressions characteristic of fear, sadness or disgust (Darwin, 1872; Duchenne, 1862; Ekman, 2003b; Ekman & Friesen, 1982; Frank, Ekman, & Friesen, 1993). It is presumed a smile is false if the upper facial features show expressions associated with more negative emotions. If the upper face region for a realistic, human-like, virtual character is not modeled correctly, then expression in the lower face may not be sufficient to convince the viewer a positive emotion is felt. A reliable expression may be evident for fear in the upper face when a smile is shown in the lower face. Hence, the perceiver is presented with a potential leakage of emotion to portray concealed, negative emotion. This may have been the contributing factor as to why the uncanny was actually reduced for those realistic, human-like characters expressing happiness with no upper facial movement, when compared to fully animated characters (Tinwell, Grimshaw, Abdel Nabi, et al., 2011).

A perception of emotional hyporesponsiveness in virtual characters may evoke more sinister undertones as to what a character may be trying to conceal. As shown in our progenitors, monkeys with limited facial expression may become frustrated and physically attack other monkeys to make themselves understood (Izard, 1979). Similarly, humans with particular personality disorders may also be more likely to resort to more violent tactics in order to communicate their feelings (Hart & Hare, 1996; Herpertz et al., 2001). Empirical evidence has revealed that a salient trait in those members of the population who are notably affectively hyporesponsive, e.g. psychopaths, is a lack of fear response to aversive events (Herpertz et al., 2001). Specifically, this is communicated with an absence of movement in the upper facial region in response to fear-related stimuli. In those diagnosed with psychopathy, the startle reflex (evident with raised brows and a blink response) that includes the combined facial movement of AU 1+2+4 to demonstrate genuine fear, may not be evident when presented with aversive stimuli (Herpertz et al., 2001). Tinwell, Grimshaw, Abdel Nabi, et al., (2011) also found that perception of the uncanny was particularly strong in characters communicating the emotion fear with a lack of movement in the eyelids, brows and forehead. Hence, the uncanny may be evoked through viewer perception of a personality disorder bordering on that of psychopathy. What is repressed and now unconcealed (Freud, 1919) may have dangerous consequences for the viewer. Psychopaths may be more predisposed to violence as they fail to experience emotions that would otherwise inhibit acting on violent impulses (Hart & Hare, 1996; Herpertz et al., 2001). The viewer cannot be aware of that characters anticipated (potentially violent and threatening) behavior. Thus, perception of hypoemotionality in virtual characters may raise alarm and fear, as the viewer is wary of a heightened threat of attack, exaggerating the uncanny.

5. A LACK OF NVC IN VIRTUAL WORLDS

In 2006, Quantic Dream released a tech demo called “The Casting” for the much awaited video game *Heavy Rain* at the E3 in 2006. Modeled after film noir and pre-empted as a revolutionary, cinematic, interactive drama, the game was expected to set new heights in levels of player engagement and rapport with the characters and plot (Hoggins, 2010). Utilizing the ‘Vicon’ motion capture system to achieve the highest quality of precision in capturing not only full body movement, but small facial movement and expression (such as Emblems and Illustrators prevalent in facial NVC (Ekman and Friesen, 1969)), the developers intended that the viewer would be able to completely suspend disbelief with the virtual characters (Martin, 2007; Hoggins, 2010). Guillaume de Fondaumière, the co-founder of the game developer Quantic Dream, stated that “I can officially announce that there is no uncanny valley any more, not in real-time” (as quoted in Doerr, 2007). The Casting featured the empathetic character Mary Smith, providing a personal account as to the recent devastating news of her husband’s betrayal. It was intended that viewers felt sympathy and would empathize with Mary as she delivered her dramatic narrative; however, instead of eliciting empathy

from her audience, this character was simply mocked. Viewers claimed that Mary's emotion seemed awkward and wooden and that her facial expression did not match the emotive quality of her voice, thus reducing the overall believability for this character (Doerr, 2007; Gouskos, 2006). On closer inspection of Mary's facial expression, when expressing fear, whilst the lower facial features were fully animated as she spoke, there was little movement in the upper face. Viewers may otherwise have expected to see raised outer eyelids (AUs 1+2), with a furrowed brow (AU4) (Ekman, 1979; Ekman and Friesen, 1978) however there was insufficient movement to demonstrate these facial actions. This resulted in a lack of plausibility for this emotion. While the viewer could hear what she said, the character lacked believability as her facial expression did not suggest that this character experienced fear.

In 2009, the video game *Heavy Rain* was released. Whilst the graphical fidelity of the characters featured in the game had improved (such as the quality of their skin texture), their behavioral fidelity still did not match their realistic, human-like appearance. Again, aspects of the main characters' facial expression were regarded as odd or strange. A poignant scene takes place where the protagonist Ethan Mars loses his son in a busy shopping mall and hunts desperately to find him. As an interactive, drama game, eliciting emotion was one of the game's key targets, essential for viewers to participate fully with each evolving chapter (Hoggins, 2010). However, Ethan's facial expression was criticized as one of the reasons as to why this scene fell below expectations of fully engaging the viewer as the search progressed. With a demonstrable lack of NVC in this character's facial expression, the viewer could not empathize with Ethan's affective state and his apparent panic and fear at the trepidation of a loss of his son. As Ethan calls out his son's name the evident despair in his voice was not evident in his facial expression: there was a distinct lack of his brows being raised and wrinkles created as a result of this movement in the forehead, a typical facial action in humans when in despair (Ekman, 2001, 2003b); nor his brows being lowered to baton-accent and accentuate more negative words (Ekman, 2004). Quantic Dream intended that the viewer would experience a heightened emotive experience in playing the game based on the character's heightened emotive states. However, a perceived lack of NVC in the characters' facial expression reduced the overall believability and impact of this game for the player.

Perception of a lack of NVC in virtual characters leads to a state of confusion for the viewer. The viewer may risk receiving contradictory information from indiscriminate, contradictory facial signals raising confusion as to how to interpret the character. If one cannot identify the emotion for a character based on an incongruence between their upper and lower facial features (during or without speech), the viewer may be unsure as to the consequent actions of that character. As demonstrated with the empathetic character, Mary Smith, if there is possible doubt as to a character's emotive state, based on expectations of the given context within which the character is presented, the character may be perceived as less believable, less human-like and uncanny. The viewer is unsure of how to react to the given character or what their next intended actions will be. Moreover, the aesthetic consequences of an ambiguity of facial expression may have a more alarming response than one being confused as to a character's emotive state. This may make the viewer feel uncomfortable given that, under certain circumstances, a human may resort to more physical and aggressive means of behavior to communicate how they feel, if they perceive that they are not being understood (as evidenced in those with ASPD). Whilst this behavior works to the detriment of empathetic characters by increasing the uncanny, such characteristics may be beneficial for protagonist characters intended to be unnerving.

In the video game *LA Noire*, the player must identify a criminal by conducting face to face interviews with characters. As well as keeping pace with the story line and events, the player may use NVC as a way to detect potential suspects from the innocent. Those guilty culprits may exhibit a lack of the startle reflex in response to fear or generate a smile that may otherwise be regarded as false in a human, as a clue to the player that their behavior is strange or suspicious. When being interrogated, it may improve player engagement in a crime-thriller if those guilty characters may convey traits similar to those diagnosed with aspects of ASPD or psychopathy. Social interaction that is perceived as abnormal to the viewer raises alarm

that the character may be capable of atypical, anti-social behavior raising suspicion that they may have conducted the crime. Solving a mystery in this way, by observing subtle nuances in NVC provocative of lies of omission, would no doubt be a more rewarding feat for the player, rather than having to rely solely on more blatant clues such as more obvious lies of commission. A lack of facial expression may instill panic in the player that the character is not only untrustworthy, but capable of aggressive or threatening behavior. A hint of psychopathic tendencies in a character (due to a lack of upper facial movement) may raise suspicion that this character is capable of violent behavior to the extent of killing others with little or no remorse for their actions. Experiencing the uncanny may be used as a warning signal for the player to detect a killer within a crime game. However, such tactics would of course not be effective in helping the player eliminate criminals if all the characters were behaving in this way, guilty or not, as it seems is currently the case in such video games. So that the uncanny may be used to the advantage of the player (and the designer) the facial expression in those empathetic characters not intended to deceive, must be accurate to avoid confusion for the player. Acquiring a greater awareness in designers of how to manipulate NVC in facial expression may be necessary to achieve this sophisticated element of player-character interaction the game, alongside improved graphical realism achievable in virtual worlds.

The consequences of a lack of accurate modeling to communicate NVC in the upper facial region may have detrimental effects for those interacting with virtual characters for testing or training purposes. The participant may be less able to empathize with the character and relate to that character's emotional state in order to respond accordingly to the needs of the given dilemma, given the ambiguity of that character's perceived emotional state. As well as the spoken narrative, the participant may seek nonverbal facial cues such as Emblems and Illustrators (Ekman and Friesen, 1969) to understand a character's emotion. If a trainee doctor is presented with a virtual character who explains that they are feeling depressed, yet whose facial expression does not match the expected feelings associated with their given ethical dilemma (for example, despair, grief, fear) then the trainee doctor may be less sympathetic towards that character. The trainee doctor's response may be perceived by an assessor as unsympathetic towards the character. The candidate may fail the given task in the virtual world despite otherwise being measured as having a high level of empathy.

A trainee soldier may be less willing or delayed in their actions towards helping an injured citizen in a virtual world if there is doubt as to the actual emotion that character is attempting to convey. As a survival tactic, it is imperative that we can recognize and respond to other's emotions quickly and accordingly, to avoid potential harm. In a virtual warfare scenario, trainee soldiers must assess and respond promptly to the perceived emotional state of others, in this case to help establish them as friend or enemy (ACTIVE Lab, 2011). A delay in reaction times due to confusion as to the emotional state of a character (i.e. in deciding as to whether that character is a threat or not) may have a negative consequence for the trainee soldier. The trainee soldier may even be alarmed or put-off by the abhorrent, uncanny, facial expression of those characters he is supposed to help, resulting in a delay in his reaction times as he struggles in making an instant decision to help an injured citizen (or not). With such hesitation, the trainee soldier may become an easy target for other enemies featured in the virtual world or fail to react as quickly as they should, thus failing the given test. In this case, the uncanny may serve as a negative consequence for the participant as they fail to demonstrate their full capabilities and aptitude for performing well when under pressure.

6. LIMITATIONS AND FUTURE WORK

In experiments featured as part of this research project, characters have so far been tested in isolation. However, conversational and emotional signals are most likely to occur in verbal exchange between a sender and receiver (Ekman, 1965, 1979). For example, the brows can be employed as signals of "turn-taking" (Ekman, 1979, p. 186) when two people are engaged in conversation. The listener may use brow movements

to signal that have understood what the speaker has said (Ekman, 1979; Dittman, 1972). Without facial signals such as Emblems or Illustrators that help ease the flow of conversation, the conversation may appear staged or unnatural (Ekman, 1965). Future experiments should include role-playing between at least two realistic, human-like characters, with strategic design manipulations made to control facial movement in either, or both, the sender and receiver. The results may lead to a confounding detrimental effect on perception of the uncanny in other characters if they appear to react inappropriately to facial signals (or a lack of) made by the sender or vice-versa. Similarly, the uncanny may be exaggerated if inappropriate or lacking facial signals are made by the receiver in response to what another character has said. Such incongruence between the sender and receiver may increase the magnitude of perceived uncanniness in realistic, human-like characters with aberrant facial movement.

As stated previously in this chapter, NVC is not only limited to upper facial movement. During conversation people may make mouth movements as a signal that they are about to speak. The receiver may raise the corner of their mouth or part their lips signaling to the sender that they would like a turn in conversation (Ekman, 1979). Movements of the head and body, hand gestures, and direction of gaze can also provide conversational and emotional signals (see e.g. Ekman 1979, 2004; Ekman and Friesen, 1969, 1972; Johnson et al., 1975). Hence, future investigation into NVC and perception of the uncanny in virtual characters may be extended to other parts of the body including: mouth movements and animation in the lower face; head tilting; gaze direction; hand gestures; and the positioning and posture of the body.

7. CONCLUSION

The issues raised in this chapter suggest that the subtleties required to communicate accurate NVC in a character's facial expression may be difficult, if impossible, to replicate due to the current technical and practical limitations in the design and development process of virtual worlds. An inaccurate simulation of NVC in a realistic, human-like virtual character's facial expression may exaggerate perception of the uncanny by: preventing effective communication of the emotional state of the character; and implying ramifications of a personality disorder (bordering on psychopathic tendencies) in that character. Realistic, human-like, virtual characters are commonly being used in virtual simulations with objectives beyond that of purely entertainment purposes (ACTIVE Lab, 2011; Bergen, 2010; Bowers et al., 2010; MacDorman et al., 2010). The effect of experiencing the uncanny may have adverse consequences for those interacting with virtual characters used for training and assessment purposes. In the case of virtual worlds being used to represent real world scenarios for job applicants or trainee professionals, the implications of a lack of NVC in a character's facial expression may have a devastating effect on the outcome of how a participant performs on interacting with that character. If such confusion arises for the candidate, their ability to perform to their best possible ability may be impaired, despite their aptitude to fulfill the requirements of a given job. Interacting with characters judged as strange or uncanny may impair a user from learning new information from a character, or detract from their performance in a given test. A trainee soldier may hesitate in going to help an injured civilian in a virtual world if he cannot comprehend a clear signal of fear and distress from that character. A trainee medical student may fail to realize the mounting panic or fear in her virtual patient due to a lack of nonverbal facial cues in the virtual character and, as a result, fail to respond accordingly to the patient's needs. This may result in a user not accomplishing a task or unable to demonstrate particular skills if their performance is being monitored. It is hoped that continued experimental investigation into the importance of displaying NVC in facial expression and the cause of the uncanny in realistic, human-like characters will help establish a framework for when it is appropriate to use such characters in virtual simulations, and, importantly, when it is not.

Based on the above, it may be recommended that designers do not rely solely on the facial animation achieved by motion capture techniques to demonstrate emotion in their character designs. If the facial actions captured in the upper facial region are insufficient to demonstrate a given emotion (either during or without speech) models should be further manipulated in 3D animation software to ensure that sufficient detail is provided to communicate a given emotion effectively. Whilst this tactic may be beneficial for pre-recorded full motion video, for example, in cut-scenes or trailers featured in video games, footage rendered in real-time such as in-game play, presents a greater challenge for the designer. Thus, it may be proposed to actually reduce the graphical fidelity of realistic, human-like characters used in real-time gaming and other virtual simulations, so that the more basic and naïve facial expression, currently achievable for animation generated on the fly, matches the character's more basic, human-like appearance. Mori (1970) warned of the dangers of pursuing highly-realistic, synthetic agents, and empirical evidence has shown that the uncanny can be increased if there is a perceived mismatch (or imbalance) between a character's graphical fidelity and their expected behavior, based on their realistic appearance. In moving to heightened graphical realism in virtual worlds, a paradox occurs. Game developers and those creating virtual worlds to simulate real world scenarios for training and testing purposes are doing so in the pursuit of suspending disbelief for the viewer. However, this increased realism simply magnifies the apparent imperfections in the characters' facial expression and behavior. This puts the viewer off and reminds them that they are confronted with a man-made, synthetic agent, rather than being fully immersed within the game or scenario. Ironically, it seems by increasing the realism of human-like, virtual characters the character appears *less*, not more, believable for the viewer.

A lack of NVC in character's facial expression is just one example of how increased realism may impact on a viewer's engagement with a virtual world. There still remains unanswered questions as to the complexity of the cognitive and affective processes by which one interprets and responds to other's facial expression in humans. Until the complexities of emotion recognition are fully understood and developers are made fully aware of such processes (in addition to their advanced technical expertise), it may be that developers are simply setting themselves up for failure in trying to suspend disbelief for the viewer and simulate perception of emotion with increasing realism in virtual worlds. To improve upon this current circumstance, a greater synergy between the disciplines of psychology and animation may be required, rather than leaving all character development work to the hands of the developers. Combining knowledge gained from psychology, such as a typical profile of nonverbal responses in psychopathy, or how NVC may be relied upon in real-life scenarios to detect untruths, with a synthesis of research findings of perception of the uncanny in virtual characters may provide a new direction as to the cause of the uncanny. This is a matter for future discussion and debate.

8. BIBLIOGRAPHY

- Applied Cognition & Training in Immersive Virtual Environments (ACTIVE) Lab. (2011). <http://active.ist.ucf.edu/LinkClick.aspx?fileticket=YdDwMmTKYe4%3d&tabid=740> (retrieved May 27, 2011).
- Andrew, R.J., (1963). Evolution of Facial Expression, *Science*, 142, 1034–1041.
- Bailenson, J. N., Swinth, K. R., Hoyt, C. L., Persky, S., Dimov, A., & Blascovich, J. (2005). The independent and interactive effects of embodied-agent appearance and behavior on self-report, cognitive, and behavioral markers of copresence in immersive virtual environments. *Presence: Teleoperators and Virtual Environments*, 14(4), 379–393.
- Bartneck, C., Kanda, T., Ishiguro, H., & Hagita, N. (2009). My robotic doppelganger – A critical look at the Uncanny Valley theory. In *Proceedings of the 18th IEEE international symposium on robot and human interactive communication*, Japan, 2009 (pp. 269–276).

- Bartnek, C., Kulik, D., Croft E., & Zoghbi, S. (2009). Measurement instruments for the anthropomorphism, animacy, likeability, perceived intelligence, and perceived safety of robots. *International Journal of Social Robotics*, 1(1), 71-81.
- Blair, R. J. R. (2003). Facial expressions, their communicatory functions and neurocognitive substrates. *Philosophical Transactions of the Royal Society*, 358, 561–572.
- Brenton, H., Gillies, M., Ballin, D., & Chatting, D. (2005). The Uncanny Valley: Does it exist? *Paper presented at the HCI group annual conference: Animated characters interaction workshop, Napier University, Edinburgh, UK.*
- Darwin, C. (1965). *The expression of the emotions in man and animals*. Chicago: University of Chicago Press (original work published 1872).
- Dittman, A.T. (1972). Developmental factors in conversational behavior. *Journal of Communication*, 22, 404–23.
- Doerr, N. (2007). *Heavy Rain Devs have “Conquered” the Uncanny Valley*. *Joystiq*, <http://www.joystiq.com/2007/12/18/heavy-rain-devs-have-conquered-the-uncanny-valley/> (retrieved 30 August 2011).
- Duchenne, B. (1862). *Mecanisme de la physionomie humaine ou analyse electrophysiologique de l’expression des passions*. Paris: Bailliere.
- Eibl-Eibesfeldt, I. (1970). *Ethology, the biology of behavior*. New York: Holt, Rinehart and Winston.
- Ekman, P. (1965) Communication through nonverbal behavior: A source of information about an interpersonal relationship. In S. S. Tomkins, & C. E. Izard, (Eds.), *Affect, Cognition and Personality* (pp.389–442). New York: Springer.
- Ekman, P. (1979). About brows: emotional and conversational signals. In M.K. Von Cranach, Foppa, W. Lepenies, & D. Ploog, (Eds.), *Human ethology: claims and limits of a new discipline* (pp.169–202). New York: Cambridge University Press.
- Ekman, P. (1992a). An argument for basic emotions. *Cognition and Emotion*, 6(3/4), 169–200.
- Ekman, P. (1992b). Are there basic emotions? *Psychological Rev*, 99(3), 550–553.
- Ekman, P. 2001. *Telling Lies: Clues to Deceit in the Marketplace, Marriage, and Politics*, 3rd edit. New York: W.W. Norton.
- Ekman, P. (2003). *Emotions Revealed: Recognizing faces and feelings to communication and emotional life*. New York: Henry Holt.
- Ekman, P. (2003b). Darwin, deception, and facial expression. *Annals New York Academy of sciences*, 1000, 205–221.
- Ekman, P. (2004). Emotional and Conversational Nonverbal Signals. In M. Larrazabal, & L. Miranda (Eds.), *Language, Knowledge, and Representation* (pp. 39–50). Netherlands: Kluwer Academic Publishers.
- Ekman, P., & Friesen W. V. (1969). The repertoire of nonverbal behavior: categories, origins, usage, and coding. *Semiotica*, 1, 49–98.
- Ekman, P., & Friesen W. V. (1969b). Nonverbal leakage and clues to deception. *Psychiatry*, 32, 88–105.
- Ekman, P., & Friesen W. V. (1978). *Facial action coding system: A technique for the measurement of facial movement*. Palo Alto, CA: Consulting Psychologists Press.
- Ekman, P., & Friesen W. V. (1982) Felt, false and miserable smiles. *Journal of Nonverbal Behavior* 6(4), 238–252.
- Frank, M.G., P. Ekman, & W.V. Friesen. (1993). Behavioral markers and recognisability of the smile of enjoyment. *Journal of Personality and Social Psychology*. 64, 83–93.
- Freud, S. (1919). The uncanny. In J. Strachey & A. Freud (Eds.), *The standard edition of the complete psychological works of Sigmund Freud* (pp. 217–256). London: Hogarth Press.
- Geller, T. (2008). Overcoming the uncanny valley, *IEEE Computer Graphics and Applications*, 28(4), 11–17.
- Gouskos, C. (2006). *The depths of the Uncanny Valley*. *Gamespot*. <http://uk.gamespot.com/features/6153667/index.html> (retrieved July 1, 2010).

- Green, R. D., MacDorman, K. F., Ho, C. C., & Vasudevan, S. K. (2008). Sensitivity to the proportions of faces that vary in human likeness. *Computers in Human Behavior*, 24(5), 2456-2474.
- Hanson, D. (2006). Exploring the aesthetic range for humanoid robots. In *Proceedings of the ICCS/CogSci-2006 long symposium: Toward social mechanisms of android science, Canada, 2006* (pp. 16–20).
- Hart S.D. & Hare R.D. (1996). Psychopathy and risk assessment. *Curr Opin Psychiatry*. 9, 380–383.
- Ho, C.-C., MacDorman, K. F., & Pramono, Z. A. (2008). Human emotion and the Uncanny Valley. A GLM, MDS, and ISOMAP analysis of robot video ratings. In *Proceedings of the third ACM/IEEE international conference on human–robot interaction, Amsterdam, 2008* (pp. 169–176).
- Izard, C. E. (1971). *The face of emotion*. New York: Appleton.
- Jentsch, E. (1906). On the psychology of the uncanny. *Angelaki*, 2, 7–16 (R. Sellars, Trans.).
- Johnson, H.G., Ekman, P., & Friesen, W.V. (1990). Voluntary facial action generates emotion-specific autonomic nervous system activity. *Psychophysiology*, 27, 363–84.
- Kanda, T., Hirano, T., Eaton, D., & Ishiguro, H. (2004). Interactive robots as social partners and peer tutors for children: A field trial. *Human–Computer Interaction*, 19(1), 61–84.
- MacDorman, K. F. (2006). Subjective ratings of robot video clips for human likeness, familiarity, and eeriness: An exploration of the Uncanny Valley. In *Proceedings of the ICCS/CogSci-2006 long symposium: Toward social mechanisms of android science, Canada, 2006* (pp. 29–29).
- MacDorman, K. F., Coram, J.A., Ho, C.C. & Patel, H. (2010). Gender Differences in the pact of Presentational Factors in Human Character Animation on Decisions in Ethical Dilemmas. *Presence*, 19(3), 213–229.
- MacDorman, K. F., & Ishiguro, H. (2006). The uncanny advantage of using androids in cognitive and social science research. *Interaction Studies*, 7(3), 297–337.
- MacDorman, K. F., Green, R. D., Ho, C. C., & Koch, C. T. (2009). Too real for comfort? Uncanny responses to computer generated faces. *Computers in Human Behavior*, 25, 695–710.
- Martin, M. (2007). *Quantic Dream buys mo-cap tech for PS3 exclusive*, *Games Industry*, <http://www.gamesindustry.biz/articles/quantic-dream-buys-mo-cap-tech-for-ps3-exclusive> (retrieved July 20 2011).
- Minato, T., Shimda, M., Ishiguro, H., & Itakura, S. (2004). *Development of an android robot for studying human–robot interaction*. In R. Orchard, C. Yang, & M. Ali (Eds.), *Innovations in applied artificial intelligence* (pp. 424–434). Ottawa, Canada: Springer.
- Mitchell, W.T., Szerszen, K.A., Shirong Lu, A., Schermerhorn, P.W., Scheutz, M., & MacDorman, K.F. (2011). A mismatch in the human realism of face and voice produces an uncanny valley, *iPerception*, 2, 9–10.
- Mori, M. (1970). Bukimi no tani (the Uncanny Valley). *Energy*, 7, 33–35 (K. F. MacDorman, & T. Minato, Trans.).
- Pavlus, J. (2011). *Did the “Uncanny Valley” kill Disney’s CGI company?*, *Fast Company’s Co.Design*. <http://www.fastcodesign.com/1663530/did-the-uncanny-valley-kill-disneys-cgi-company#>, (retrieved May 27, 2011).
- Plantec, P. (2007). *Crossing the great Uncanny Valley*, *Animation World Network*, <http://vfxworld.com/?atype=articles&cid=3494>, (retrieved July 20 2011).
- Plantec, P. (2008). *Image metrics attempts to leap the Uncanny Valley*. *The digital eye*. <http://vfxworld.com/?atype=articles&cid=3723&page=1> (retrieved July 1, 2010).
- Pollick, F. (2009). In search of the Uncanny Valley. In P. Daras, & O. M. Ibarra (Eds.), *UC Media 2009. Lecture notes of the institute for computer sciences, social informatics and telecommunications engineering* (pp. 69–78). Venice, Italy: Springer.
- Herpertz, S. C., Werth, U., Lukas, G., Qunaibi, M., Schuerkens, A., Kunert, H. J., Freese, R., MD; Martin Flesch, M., Mueller-Isberner, R., Osterheider, M., & Sass, H. (2001). Emotion in criminal offenders with psychopathy and borderline personality disorder. *Archives of General Psychiatry*, 58, 737–745.

- Hoggins, T. (2010). *Heavy Rain Video Game Review*. *The Telegraph*, <http://www.telegraph.co.uk/technology/video-games/7196822/Heavy-Rain-video-game-review.html> (retrieved 30 August 2011).
- Schneider, E., Wang, Y., & Yang, S. (2007). Exploring the Uncanny Valley with Japanese video game characters. In *Proceedings of situated play, DiGRA 2007 conference, Tokyo, Japan* (pp. 546–549).
- Bowers, C., Smith, P.A., Cannon-Bowers, J., & Nicholson, D.M., (2008). Using virtual worlds to assist distributed teams. In P. Zemliansky, (Ed.), *The Handbook of Research on Virtual Workplaces and the New Nature of Business Practices* (pp. 408–423). Hershey, PA: IGI Publishing Group.
- Stuart, K. (2007). *Artists climb the uncanny valley*. *Games Blog*. <http://www.guardian.co.uk/technology/gamesblog/2007/aug/08/artistsclimb> (retrieved May 15, 2011).
- Thompson, C. (2005). *Monsters of Photorealism*, *Wired*, <http://www.wired.com/gaming/gamingreviews/commentary/games/2005/12/69739> (retrieved February 18, 2010).
- Tinwell, A. (2009). The Uncanny as usability obstacle. In A. A. Ozok & P. Zaphiris (Eds.), *Proceedings of the HCI international 2009: Online communities and social computing workshop* (pp. 622–631). San Diego, CA: Springer.
- Tinwell, A., & Grimshaw, M. (2009). Bridging the uncanny: An impossible traverse? In O. Sotamaa, A. Lugmayr, H. Franssila, P. Näränen, & J. Vanhala (Eds.), *Proceedings of the 13th international MindTrek conference. Everyday life in the ubiquitous era* (pp. 66–73). Tampere, Finland: ACM.
- Tinwell, A., Grimshaw, M., & Williams, A. (2010). Uncanny behaviour in survival horror games. *Journal of Gaming and Virtual Worlds*, 2(1), 3–25.
- Tinwell, A., Grimshaw, M., & Williams, A. (2011a). The uncanny wall. *International Journal of Arts and Technology*, 4(3), 326–341.
- Tinwell, A., Grimshaw, M., & Williams, A. (2011b). Uncanny speech. In M. Grimshaw (Ed.). *Game sound technology and player interaction: Concepts and developments* (pp. 213–234). Hershey, Pennsylvania: IGI Global.
- Tinwell, A., Grimshaw, M., Abdel Nabi, D., & Williams, A. (2011). Facial expression of emotion and perception of the Uncanny Valley in virtual characters. *Computers in Human Behavior*, 27(2), 741–749.
- Tinwell, A., Grimshaw, M., & Abdel Nabi, D. (2011). Effect of Emotion and Articulation of Speech on the Uncanny Valley in Virtual Characters. In S. D’Mello et al. (Eds.), *Proceedings of the ACII 2011, Part II, LNCS 6975* (pp. 557–566). Memphis, TN: Springer-Verlag.
- Tinwell, A. (2009). The Uncanny as usability obstacle. In A. A. Ozok & P. Zaphiris (Eds.), *Proceedings of the HCI international 2009: Online communities and social computing workshop* (pp. 622–631). San Diego, CA: Springer.
- Ubisoft Entertainment (2005). *King Kong (video game)*. Montreal, Canada: Ubisoft.
- Valve (2008). *Half-Life 2 (video game)*. Washington: Valve Corporation.
- Vinayagamoorthy, V., Steed, A., & Slater, M. (2005). Building characters: Lessons drawn from virtual environments. In *Proceedings of Toward social mechanisms of android science, COGSCI 2005 workshop, Stresa, Italy* (pp. 119–126).
- Von Bergen, J. M. (2010). *Queasy about avatars and hiring employees*. http://www.philly.com/philly/blogs/jobs/Queasy_about_avatars_and_hiring_employees_.html (retrieved July 20 2011).

SECTION VI

NEW DIRECTIONS FOR NVC IN VWs

20

THE FUTURE OF AVATAR EXPRESSION: BODY LANGUAGE EVOLVES ON THE INTERNET

By Jeffrey Ventrella

1. INTRODUCTION

When I use (or become) an avatar in a virtual world or game, I am given some degree of control over that avatar's visual manifestation. In a 3D world, I can navigate it around other avatars to generate proxemic interactions; a kind of signaling used by most animals. Some avatar systems allow animations, gestures, and facial expressions, in addition to customized static appearances. The avatar is the locus of *virtual body language*: I use it to represent my expressions, thoughts, emotions, and intentions. By generating real-time body language that is remote from my physical body, a new kind of mediated communication is enabled. This is *embodied communication*, but my real body is not present at the site of apprehension. We have only just begun to grasp the implications of this medium, and to study its potential uses.

Anthropologist Ray Birdwhistell, coined the term “kinesics”, to refer to the study of body language (Birdwhistell 1970). I propose the term “telekinesics” to denote the study of all emerging forms of body language on the internet, by adding the prefix, tele to Birdwhistell's, term *kinesics*. The word looks and sounds like “telekinesis”: the ability to cause movement at a distance through the mind alone. The term I propose is not about paranormal phenomena. “Telekinesics” may be defined as “the science of body language as conducted over remote distances via some medium, including the internet” (Ventrella, 2011). It appears that the term, “tele-kinesic” has been used by Mary D. Sheridan as a way to describe body language as experienced across some physical distance, in young children and babies (1977). Let us expand Sheridan's physical distance via digital technology: body language can be transmitted in real-time on a global scale.

This chapter takes a broad look at our various communications over the internet, and considers body language to be not just important for future communication technologies, but inevitable – as part of the natural course of evolution. In this chapter, I will be taking an imaginative look into future scenarios of remote nonverbal communication: how *telekinesics* will play out in future communication technology.

EVOLUTION

Human language is not so impressive when put in the context of billions of years of genetic evolution: DNA has an alphabet and a grammar. The language of DNA enabled the evolution of the biosphere, which gave birth to brains, culture, and natural language. The evolution of the spoken word enabled culture to be transmitted at a higher rate. When the written word evolved, speech was encoded in a more precise and portable form (though limiting). Encoded as text, speech leaped beyond the here and now, traveling vast spaces, beyond the pocket of air and light between groups of individuals. It became plastic and infinitely expressive. And it could replicate reliably— like DNA.

But natural language is not just about speaking mouths and listening ears (nor is it only about writing hands and reading eyes). There is a visual/physical dimension to natural language, which is in fact much older than speech, as earthly languages go. We generate this visual language with our bodies—our postures, gestures, facial expressions, and gaze.

Here's a prediction: similar to the way speech has been encoded as text, body language is currently undergoing a similar encoding process—as a kind of nonverbal text. Body language will become plastic – infinitely expressive. The internet is becoming more visual and interactive all the time. Anomalies of the written word, those “disembodied agencies” and “lettered selves” (Rotman 2008), have dominated society for so long. New texts and grammars of embodied communication are evolving in our midst, thanks to the internet, computer graphics, artificial intelligence, and new gestural interfaces. How is this evolution taking place? Well, avatars in virtual worlds are only one part of the story. Embodiment, and how it plays out in many facets of interface design, has been explored in depth by several authors, such as Dourish (2001).

BEYOND THE FACELESS INTERNET

People are increasingly using email, social networking services, texting, and virtual worlds to socialize. Business transactions and meetings are done increasingly online. What is happening as human communication shifts over to the internet? Answer: it is getting the cold shoulder. According to Sandy Pentland in *Honest Signals*, our communication technologies “treat people like cogs in an information processing machine” (Pentland, 2008). Without eye contact, gesture, and the nonverbal signals that we rely on in our personal and professional lives, we are at risk of losing our decision-making abilities, as well as the affective dimension that enriches our social lives. Text was invented for highly-crafted, premeditated use, in which a reader may not encounter the writer's words for a long time, perhaps several years. But writing and reading are marching toward each other, and they are approaching real-time conversation rates. This is why body language is trying to insert itself and evolve to compensate. In email and texting, emoticons and certain conversational dynamics have emerged spontaneously to help disambiguate conversational text, but they only help up to a point.

Meanwhile, other evolutionary trends continue as predicted. Telephony has finally acquired the video screen we had been dreaming about for so long. Applications like Skype are increasingly used as a way to include our faces and bodies in our remote communications. It is an obvious and effective way to create presence remotely and to broadcast one's honest signals. And it makes up for a great deal of the physical separation between people. However, video chat lacks the open-ended user-generated customization that virtual worlds afford, as well as the interpersonal geometry where mutual eye-contact, bodily gaze, and simulated physical touch can lend meaning and power to virtual embodiment. While virtual worlds may seem opaque and obstructive to my real body language, the technology is likely to advance in such a way as to allow my true expressions – and more importantly – ANY expressions – to be transmitted. Video chat is literal, but virtual worlds are imaginal.

A LANGUAGE OF REMOTE EMBODIMENT

What I am exploring here is a different kind of body language technology than video chat, and a new kind of *language* that comes with it – a language of remote embodied communication which is as plastic and expressive as the written word – perhaps even more so. Consider this: I could write a short story about myself in which I am a magnificent 100-armed creature that can generate symphonic sounds, dance, and express beautiful mathematical concepts – something a human could never do. I imagine myself as a professor in an online university, possessing this super-expressive body. At this time, I cannot *become* this creature or *perform* this act, in any meaningful or effective way. This is a brand of body language that has been the dream of Jaron Lanier (2006) for several decades now, referring to the experience of puppeteering

non-human animals, or generating super-human body language. For the kind of embodied communication I am talking about, I would have to *inhabit* a virtual creature and I would have to use advanced controls. This is an extreme example of a kind of body language that requires a special meta-humanoid avatar. But the principle still holds for realistic human-like avatars: In order for me to express with an avatar with the same degree of imagination and expressivity as I can with text, I need a combination of a sophisticated puppeteering control technology, and a language of embodied communication. The nuts and bolts of this language are currently being worked out by several researchers and practitioners in the field.

2. TRANS-WORLD BODY LANGUAGE

Avatars of the future may not occupy the same kinds of spaces that we associate them with today (i.e., single, monolithic virtual worlds like *Second Life*). Avatars may deconstruct into highly flexible and versatile manifestations of embodiment, able to occupy many social spaces and media. In fact, they may acquire the ability to shift their body language capabilities, chameleon-like, to take on forms that are appropriate to whatever media they occupy from one moment to the next. Consider the smiley :-). A semi-colon, an optional dash, and a right-round bracket. They come in many forms, but all serve as primordial avatar expressions. They live in text-land. When a user types a smiley into a chat window in some virtual worlds, the avatar will smile (in effect, creating two smiles: one typographical and one animated). If the user is driving an avatar in the form of a jellyfish, the smiley might cause its tentacles to flair slightly.

Here is a scenario I like to imagine: several open source software tools allowing the standard components of virtual worlds to be built (avatar navigation, avatar appearance customization, avatar expression puppeteering, terrains, user databases, 3D modeling and animation tools, renderers, physics engines, VOIP, gestural input, etc). These tools are built and integrated by enthusiasts, hobbyists, and professionals with personal interest. A series of virtual worlds and virtual world-like spaces emerge within a *non-corporate commons*, supported to some degree by several businesses (like the internet itself) but not owned by any single company. I bring this up in reference to the hardships and sorrow caused by major virtual world companies going belly-up: the rich virtual lives that communities had built over many years vanish into thin air. My own experience as the co-founder of There.com is a case in point: there were several difficult changes in management and disagreements among the business executives. (Fortunately – in some respects – I was not involved on that level). Ultimately, while some shareholders may encounter great disappointments when a virtual world business venture fails, the greatest loss, I believe, is with the communities that built the society; generated untold amounts of user-generated content; and lived important aspects of their lives in this world.

At the end of the day, a handful of executives with a new business plan can delete the virtual lives of thousands of people. They may wince, but money talks. I look forward to a time when major virtual worlds are simply a part of the *commons*.

Here's the key: woven into the fabric of the internet itself, these virtual worlds would have interfaces and APIs allowing users to move between worlds (because there would be no companies trying to keep all the users in their own worlds and out of other worlds). Now imagine that as users pass through this web of worlds, their avatar appearances and behaviors change according to the worlds they occupy. One world might be hyper-realistic (the avatar would use sophisticated AI, behavioral animation, and rendering to make the avatar appear uncannily real. Another world might be cartoon-like (the user's puppeteered body language would be converted into cartoon gestures and exaggerated expressions). Another world might render a microscopic primordial puddle; in this world, a user's green-blob-avatar could twitch and roll, and perhaps wave its flagella as a result of the same puppeteering UI.

Although these worlds may be vastly different, I propose one thing remain constant: a body language alphabet/grammar. I imagine three components: (1) a set of input scenarios, (2) a body language alphabet/grammar, and (3) the kaleidoscope of possible worlds in which avatars can manifest. I am proposing a standard body language scheme that provides universal translations across various input scenarios at the one end, and across various worlds at the other end. Even if we just do this as a thought experiment, we will find ourselves asking provocative questions about the nature of embodied communication. Whether or not this system deliberately gets implemented or specified by a standards committee, such a body language alphabet will eventually evolve on its own, due to the persistence of human imagination, our need to connect socially, and the ubiquity of the internet. Perhaps letting it evolve on its own is the best way. And perhaps I am wrong to want to force it to be a universal scheme. After all, it is a language I'm talking about.

3. THE PROBLEM OF INPUT

This chapter looks at the topic of avatar expression from the standpoint of my own real world experiences designing and engineering avatar systems for *There.com* and *Second Life*. In these worlds, avatars are controlled via keyboard and mouse. These input devices were initially invented for things like word processing and desktop computing – not for avatar expression. When we talk about controlling avatars in virtual worlds, we are often heavily steeped in the context of *typewriter ergonomics*. Mobile computing, touch screens, and gestural inputs are helping to open up the thoughtspace of avatar control, not because these technologies are new (e.g., sophisticated console game character navigation has been advancing for decades), but because a great variety of input is now commercially available and more accessible for widespread experimentation. Escaping the confines of typewriter ergonomics helps us think about the possibilities – although I believe that many people are jumping too far ahead and forgetting some important factors regarding mediated communication.

Many people see the Kinect as a harbinger, pointing the way towards the ultimate avatar expression input device, along with other motion-capture systems that can read our motions, our facial expressions, even track our eyes. The holy grail in this case is natural, hands-free, gesturing and facial-expressing. The thesis I would like to put forward is that such completely transparent expressing will *not* become the primary mode of avatar control in the future – even though it may find its way into scientific research and some high-end virtual reality entertainment media. Three reasons: (1) it requires too much physical work and continuous attention from the user, (2) it is not straightforward how to throttle it down for asynchronous modalities, or partial-attention uses, and (3) people take comfort in using physical devices as instruments with which to give and receive information.

Puppeteering controllers are not necessarily obstructions – they can become techno-companions (like the smart phone, iPod, or laptop). I would even refer to age-old accoutrements like fans and Rosary beads). These and many other physical artifacts are used as vehicles for communicating, expressing, signifying, prayer, and other language-like activities. Humans evolved using artifacts - tools that serve both mechanical needs and language needs. This is clearly up for debate, and many would argue that our ultimate desire is to communicate with each other in the most natural, hands-free way possible, and that it shouldn't matter if we are in cities on the opposite sides of the globe. My contention is that we still live and breathe among physical people throughout most of our lives. We need ways to have all modes of communication with our colleagues and loved-ones, and each of these modes should be experienced as such. Total naturalness and transparency in all natural language activities (proximal as well as remote) would only result in cultural schizophrenia and chaos.

So then, what does the future hold for avatar expression interfaces? What ways will people puppeteer their virtual body language in the future? *Many* and *All!*

4. FROM GESTURE TO SPEECH, AND THEN BACK TO GESTURE

Among the many ways that people will be able to generate remote body language, here is one possible way that doesn't require cumbersome input devices: *they could use their voice*.

Gestural Theory (Hewes 1973) claims that speech emerged from the communicative energy of the body: from iconic, visible gestures. It has been found that the regions of the brain responsible for mouth and hand movements are sitting right next to each other. It is well established that sign language is just as sophisticated as any spoken language in terms of grammar. And sign language communicators appear to use the same language areas of the brain. It is believed that gesturing is not only used as a way to illustrate words; it is actually an integral part of thought process. We gesture in *order to think*. These ideas are supported by many researchers and authors, including McNeill (1996) and Iverson (2000).

Now imagine an inverse of the gesture-to-speech vector of Gestural Theory I am talking about reconstituting gesture virtually, using *speech*. This vector reversal would take inspiration from the long evolution of neural mechanisms that tie gesture to speech.

Gesticulation is the physical-visual counterpart to vocal energy. We gesticulate when we speak—moving our hands, head, eyebrows, and other body parts...and it's mostly unconscious. Humans are both verbally-oriented and visually-oriented. Our brains expect to process visible body movement when we hear spoken words, even if the speaker is not visible. In fact, neurologist Terrence Deacon suggests that the brain's auditory processing of speech is based on the gestural production of sounds and the original vocal energy, as referenced from the body, rather than being based on the acoustical attributes of the sounds (Deacon 1997). It is almost as if our brains had evolved to understand visible gesture, and then had re-adapted for auditory gesture, on the path to processing spoken words.

Now consider gesticulation and its intimate relationship with audible speech as the basis for an avatar animation algorithm, which I call "voice-triggered gesticulation". This is the basis for some avatar animation techniques that process voice signals (diPaola, 2008), (Ventrella, 2011). In addition to mere gesticulation (nonsemantic speech energy made visual), we can glean an unlimited depth of meaning. With speech-to-text software becoming more sophisticated, and natural language processing continuing to advance, it is more feasible to extrapolate meaning and context from words, phrases, and sentences – not to mention the many cues from prosody: voice intonation, pauses, etc. These can become triggers for any number of gestural emblems, postures, stances, facial expressions, lip-sync, and modulated gesticulations.

WATCH WHAT I DO

What if avatars had mirror neurons? (Mirror neurons are fired when a person does an act, sees another person do the same act, or even *thinks* about someone doing the act. Mirror neurons are likely to be important for developing a *theory of mind*). Here's a possible future scenario to consider: You put on some motion capture markers. Alternatively, motion-detection systems like the Kinect could be used. Your head and hand motions would be picked up in real-time as you go about your virtual world activities. Now, you select an option on the avatar control panel called "watch me and learn". Then, you engage in natural conversations with other people, making sure to let it all hang out, letting your head and hands move freely, etc. While you speak, the sounds you generate are processed as text, perhaps being tagged with various nonverbal triggers to create a script, analogous to the Behavior Markup Language (Kopp et al., 2006).

The motions you make are stored in a database with the associated text. As these motions and associated texts are collected, your avatar starts building a vocabulary of identifiable gestures, postures, and so-on. It starts associating those patterns with specific words and phrases that you spoke while generating the body language. Instead of using Good Old-fashioned AI, a form of avatar body language intelligence is built up over time – from experience.

Researchers have developed several techniques and experiments whereby gestural data is collected from performers, or even from analyzing video segments, and then later reconstituted in virtual humans to accompany the production of verbal language. This research shows promise for becoming the basis for effective auto-body language schemes. For instance, hand gestures researched by Neff and colleagues (2008), indicate how the unique gestures of an individual can be captured and then reconstructed in a form that can be played back with modifications to accompany the original spoken word, or associated text. The buildup of gestures doesn't have to stop at the nuances of speech itself. The system could take into account environmental situations (to whom you are talking, where you are in the world, what events have recently happened, etc.) As the database resolves into focus, a “nonverbal idiolect” emerges in your avatar's memory banks. (While “dialect” refers to the unique language variety of a people, “idiolect” refers to the language style of a single individual. I use “nonverbal idiolect” to refer to an individual's unique body language).

Continuing on with this scenario: imagine that you return to the virtual world, and—without requiring any motion detection—you see that your avatar has begun to take on some of your unique body movements. At any time, you would be able to switch on the motion detection apparatus, and train your avatar more. This would be an iterative scheme: there would be no final state of correct behavior; no ultimate conversational style that your avatar would converge upon. You would always be able to re-train your avatar with new kinds of body language, thereby causing it to unlearn previous body language behaviors, or at least have those subsumed by more recent body language. You could have an interface that allows your avatar to “wear” different versions of your own nonverbal idiolect. There could be another option that allows your avatar to pick up the nonverbal idiolect of the avatar you are talking to (whose nonverbal idiolect may have been learned by another avatar, and so-on: a *hall-of-mirror-neurons*). Perhaps there could be a “chameleon” slider on the interface that determines how easily and quickly your avatar takes on new nonverbal idiolects. There may be intellectual property concerns with having one's personal movement signature used (or “stolen”) by someone else. This kind of blurring between property and personhood has already been expounded upon by some pundits (de Andrade, 2009).

Once recorded, the nonverbal idiolects of many kinds of people could be uploaded into a library. Judith Donath of MIT has developed this idea as well (e.g. 1997, 1999). Just as easily as buying avatar clothes or body types, one could buy the behavioral personalities of people whose nonverbal idiolects have been stored. Nonverbal idiolects would become virtual goods.

5. CONCLUSIONS

This chapter explores the future of nonverbal communication on the internet, with current avatars in social virtual worlds as a reference point. The approach here is to consider both *evolution* and *design* as critical factors as to how this future will come about. Evolution is considered necessary because we are talking about language, not just about engineering tools. Living languages must evolve – and as the embodied languages of avatars in virtual worlds evolve, we must continually refine our software interfaces and algorithms to accommodate this evolution. The large-scale evolution of human language is seen as a continuation of the ancient communicative buzz of the biosphere, over the last few billion years. Our brains and bodies are the product of this evolution, and even as our lives become more digital and more virtualized, we will always

rely on the grounding metaphors that are baked into our brains. But at the same time, the plasticity of our brains (and our culture) will allow us to speak posthuman body languages – involving a hundred arms, animated skin, or the ability to morph our bodies to mimic objects that we are thinking about – what Lanier calls “postsymbolic communication” (Lanier, 2010).

Consider how people naturally create their own grammars and vocabularies, often subverting the features provided to them by designers, in order to communicate pragmatically. One example is in *World of Warcraft*: people use jumping up and down in specific ways to signify excitement, impatience, and other emotions. Another example: I once heard of an early virtual world with very limited avatar controls; to engage in the sex act, one user’s avatar would be made to lie down, while the other user’s avatar would be placed on top, and switch rapidly between standing and sitting.

Natural languages evolve through iterated use by many people in real-world situations. Natural languages cannot be designed. But that doesn’t mean that as a designer of user interfaces and virtual worlds, I should just throw up my hands and give up! Virtual worlds are after all, engineered artifacts. The conclusion is that designers of communication technologies can (and must) be fully aware of the dynamics of language evolution. We cannot predict where this evolution will take us. We can guide it to some degree. It is really a balancing act: setting down the tools and their constraints, watching how people use them, and refining those tools (or discarding them). Perhaps the best approach is to create meta-tools that allow users (the agents of language genesis) to build their own tools. Understanding the principle of affordances is critical to good design. We have billions of years of evolution and a whole planet of continual biosemiotic messaging to use as our guide. Any new communication technology that is not understood (or understandable) within a framework of *affordances*...is suspect.

ARTIFACTS OF EXPRESSION

Given our human bodies, versatile as they are, we cannot easily map our emotions, expressions and intentions onto a virtual body, especially given current awkward input devices derived from people sitting in chairs in front of typewriters. As gesture-recognition systems and touch-screens become more advanced, we will be able to articulate more flavors of body language input. But even as our real-time spontaneous gestures make their way into the virtual world, the ultimate goal, as I see it, is not complete transparent natural language, but rather, an evolving set of tools for multimodal communication. This chapter proposes that online body language of the future will most often be a mediated form of expression, and some aspects of the technology will remain opaque (like musical instruments). The key is refining these tools, keeping up with cultural evolution, and articulating a *body language text* that enables new forms of communication that are just as expressive as the written word, if not more so.

6. REFERENCES

- Birdwhistell, R. 1970. *Kinesics and Context*. University of Pennsylvania Press, Philadelphia.
- de Andrade, N.N.G. 2009. “Striking a Balance between Property and Personality: The Case of the Avatars”. *Journal of Virtual Worlds Research*, Volume 1, Number 3. <http://journals.tdl.org/jvwr/article/viewArticle/362>.
- Deacon, T. 1997. *The Symbolic Species: The Co-evolution of Language and the Brain*. New York, Norton.
- DiPaola S, Turner J. 2008. “Authoring the Intimate Self: Identity, Expression and Role-playing within a Pioneering Virtual Community”. *Journal of the Canadian Games Studies Association*, Vol 1, #3
- Dourish, P. *Where the Action Is: The Foundations of Embodied Interaction*. 2001. MIT Press.

- Hewes, Gordon W. 1973 . “Primate communication and the gestural origins of language”.
Current Anthropology 14:5-24. 1973.
- Iverson, J. M., Tencer, H. L., Lany, J. and Goldin-Meadow, S. 2000. “The Relation Between Gesture and Speech in Congenitally Blind and Sighted Language-Learners.”
Journal of Nonverbal Behavior. vol 24. Number 2. Springer.
- Kopp, S., Krenn, B., Marsella, S., Marshall, A. N., Pelachaud, C., Pirker, H., Thórisson, K. R., and Vilhjálms H. 2006. “Towards a Common Framework for Multimodal Generation: The Behavior Markup Language”. *Lecture Notes in Computer Science*. Springer Volume 4133.
- Lanier, J. 2006. “What cephalopods can teach us about language”. *Discover*. April 2, 2006.
- Lanier, J. 2010. *You Are Not a Gadget – A Manifesto*. Vintage Books.
- McNeill, D. 1996. *Hand and Mind: What Gestures Reveal About Thought*. University of Chicago Press.
- Neff, M. Kipp, M., Albrecht., I., and Seidel, H. 2008. “Gesture Modeling and Animation Based on Probabilistic Recreation of Speaker Style”. *ACM Transactions on Graphics*. vol. 27. Issue 1, March.
- Pentland, Alex (Sandy). 2008. *Honest Signals*. MIT Press.
- Rotman, B. 2008. *Becoming Beside Ourselves: The Alphabet, Ghosts, and Distributed Human Being*. Duke University Press.
- Sheridan, M.D. 1977. “Observations on the development of spontaneous tele-kinesic communication in babies and young children”. *Child Care Health Development*. May 1977. (Vol. 3, Issue 3, Pages 189-99) .
- Ventrella, J. *Virtual Body Language*. ETC Press, 2011.

21

CHALLENGES AND OPPORTUNITIES FOR THE ONGOING STUDY OF NONVERBAL COMMUNICATION IN VIRTUAL WORLDS

By Theresa Jean Tanenbaum, Magy Seif El-Nasr, and Michael Nixon

As virtual worlds grow more sophisticated, and as users grow more adept at navigating and interacting within them, we anticipate the range and sophistication of nonverbal communication within these spaces to grow as well. The variety of research into nonverbal behaviors contained within this book provides us with important insight as we attempt to look into the future and identify open problems and future challenges.

As is evident by the diversity of perspectives represented in this book, we are operating within a profoundly interdisciplinary field, with representation in sociology, ethnography, the digital humanities, the performing arts, computer science, artificial intelligence, and design (to name just a few of the disciplines represented herein). In this book we have celebrated this diversity by inviting authors from different disciplines to share their work on NVC. We regard this as the beginning of a conversation and a first step toward collectively moving forward within this emerging field. While the chapters in this book may hail from disciplines which are traditionally separate we believe that there are many productive overlaps and interrelationships between them. It is at these points of intersection between these different methodologies and perspectives that we see the road forward for the field. In this final section we synthesize the lessons learned within this book and offer some future directions and opportunities for NVC research in virtual worlds at the intersection of these disciplines and perspectives.

1. FACTORS IN THE EVOLUTION OF NVC FOR VWS

Since we started this project in 2010, many changes have already occurred in commercial virtual worlds. Indeed, the danger of any serious large scale research into a subject that is moving at the speed of digital culture is that by the time you have finished your work, the thing you are studying might no longer exist, or might have transformed so as to be unrecognizable. The chapters in this book, thus, represent a snapshot of research into nonverbal communications in virtual worlds across a variety of fields and perspectives. In collecting this material we have attempted to synthesize insights into the broader phenomena at work within human communication in virtual environments. It has become apparent to us that there are at least four factors at work that play into how people can and will communicate in virtual worlds: *Infrastructures, Ubiquity, Literacy, and Interfaces*. These may be witnessed across the material in this book and are especially important when trying to imagine the future of nonverbal communication in virtual worlds. The divisions between these themes are not always clean: they reflect the messiness and interconnectedness of the interdisciplinary research and perspectives that have informed them.

FOUR THEMES FOR THINKING ABOUT THE FUTURE

Infrastructures, Architectures, Access, and Bandwidth

One theme that has arisen within the book is the ways in which the infrastructures of virtual worlds implicate their users and designers in meaningful ways. Jennifer Martin makes a strong case in Chapter 17 for the impact of designed systems such as avatar marketplaces as key determinators of social and communicative possibilities in virtual worlds. Jacquelyn Morie's chapter on avatar appearance (Chapter 7) and Elisabeth LaPensée and Jason Edward Lewis's chapter on First Nations NVC (Chapter 8) are also clearly situated within these architected techno-social affordances. Jeffrey Ventrella (Chapters 6, 14, and 20), Jim Parker (Chapter 10), Ben Unterman and Jeremy Owen Turner (Chapter 12), and Hannes Högni Vilhjálmsson (Chapter 15) all also bring a designerly perspective to their contributions, and each of their chapters address the ways in which the technological capabilities of the systems in which they work both affords and constrains how they design for communication. These constraints include aspects of our hardware infrastructure like the availability of internet bandwidth and the expected processing power of servers and clients in our computer networks, as well as software constraints rooted in the development of architectures for communications in virtual worlds.

It is clear to us that the first major factor that impacts the ongoing evolution of NVC in VWs is the ongoing development of these core infrastructures. As bandwidth, processing cycles, and storage become cheaper and more abundant, the basic challenges faced by the first generation of virtual world designers are transforming. The question of how to create a graphical virtual environment populated by remotely operated avatars is no longer interesting. Instead, we are beginning to ask questions about what we most want out of these virtual environments, and how to better design for the types of participation that we imagine could take place in them. More importantly, we are learning about how people actually use virtual worlds, by observing their practices in the "wild".

A central issue here becomes one of access: who is using virtual worlds, and who is excluded? As the economics of creating and maintaining virtual worlds are transformed by changes in the underlying technologies, how does this impact the accessibility of these environments? This issue of access also plays into two of our other factors: *Literacy*, and *Ubiquity*.

Ubiquity and Context

Computing and computational practices are growing less and less attached to the desktop: an evolution that we believe will have significant as-yet-unexplored impacts on the evolution of NVC in digital environments. With only a few exceptions, the chapters in this book all envision the users of virtual worlds as people seated at a traditional desktop or laptop computer, in part because up until recently there were not all that many alternatives. However with the recent explosion of tablet computing and high-powered smartphones, the possibilities for more seamless interplay between virtual and physical worlds are growing, potentially leading to new opportunities for nonverbal communication.

The *TimeTraveller*[™] ARG described by Elisabeth LaPensée and Jason Edward Lewis in Chapter 8 hints at a possible future for these types of interwoven realities, in which artifacts from the fictional world are spread out through the world in which we live. Similarly, Jeffrey Ventrella calls for the development of a new "trans-world body language" (in Chapter 20) that can accommodate the translation of our avatars across a wide range of contexts and situations.

As computing becomes more ubiquitous, several important things appear to be happening. First, screens are multiplying throughout our physical environment. The range of possible windows into our virtual worlds is propagating at a rate that was unimaginable even 10 years ago. The spread of screens into our material world can be seen as a vector of infection by which our virtual lives distribute them into our

physical spaces: it means that where once we might have had to choose between virtual embodiment or physical embodiment (marked by sitting down at a desktop computer) now we must instead start learning to negotiate both of these conditions at the same time. Another significant transformation that comes with ubiquity is that the input space for virtual worlds is expanding beyond the keyboard, mouse, camera, and microphone. The presence and availability of accelerometers, heart rate sensors, skin conductivity sensors, and electrodes for simple brainwave sensing has increased by an order of magnitude in recent years, and these sensing systems are going to continue to propagate through our lives. The significance of passive sensing technology to nonverbal communication within virtual environments cannot be overstated. As has been pointed out throughout this book, one of the biggest challenges facing the deployment of NVC for avatars is that most nonverbal behaviors are *unconscious* or exist before and outside of language. The more accustomed we become to the presence of sensing devices that can be used to triangulate our emotional states, and map the positions of our bodies in space and time, the more we can translate our unconscious body languages into our digital communications.

As suggested above, an important aspect of ubiquity and context is access: as we develop more pervasive means of getting our information into virtual worlds and back out again, the lines between the material world and the virtual world become less distinct. One positive outcome of this is that some of the longstanding barriers to participation in a virtual public become less of an issue. In this sense, virtual worlds are following a similar trajectory to other technologies of mass communication as they transition out of the hands of a minority of “elite” and “privileged” users and into a public space. We believe that these transforming contexts of use for virtual embodiment will be a major factor in our remaining two themes: *Literacy and Interfaces*.

Literacy and Acculturation

Our third theme deals with how people develop the literacies needed to communicate successfully within virtual worlds. Winograd and Flores refer to this as “communicative competence”: the ability to express oneself, and take responsibility for the range of possible commitments and interpretations that that expression brings into the world (Winograd & Flores, 1986). As Hannes Högni Vilhjálmsón points out in Chapter 15, most users struggle to feel in control of a conversation when shifting between textual communication and avatar puppeteering tasks. This is a significant challenge for NVC in virtual worlds, and it is as much about the literacies, cultural expectations, and core competences of interactors as it is about the design of the system. Even experienced users can feel lost at sea when entering a new virtual world for the first time. It takes a long time to develop enough familiarity within any given virtual world to feel comfortable and competent communicating, in part because there is a substantial learning curve around the interface and in part because each virtual world has its own set of social and cultural conventions that must be learned. Jeffrey Ventrella argues in Chapter 20 that it is necessary to develop a more articulated and specific grammar of gesture and interaction for virtual worlds: that body language in virtual worlds must be treated as a distinct linguistic system that must be learned in order to communicate within virtual worlds. Understanding how to design for this type of interactional grammar, and how to understand emergent conventions that arise from use in virtual worlds is critical to the evolution of human communication in these spaces.

One way that we can develop a deeper understanding of the literacies involved in NVC for VWs is to draw on other systems of knowing and bodily communication to help bridge the gap. Leslie Bishko and Michael Neff’s chapters on applying knowledge from dance and theater to animation and avatars (Chapters 9, 11 and 13) demonstrate how these existing domains of bodily knowledge can inform a deep understanding of how to design, perform, and understand nonverbal behaviors with avatars. Angela Tinwell et al.’s chapter on the Uncanny Valley (Chapter 19) does similar work by unpacking how our experiences of faces and gaze in the physical world inform our experiences of them in digital settings.

Interfaces for NVC

All of our themes converge at the level of the interface: the point of contact between the user and the virtual world. As new hardware and software platforms evolve, new interfaces become possible including gesture recognition, bespoke puppeteering interfaces, multi-touch screens, and passive sensing technology. These interfaces are spreading into our physical environments at an unprecedented rate, greatly diversifying the possibilities for interacting with virtual worlds, but requiring the evolution of an ever shifting set of new literacies and cultural practices. In this book we have seen this theme manifest across many chapters. Elena Erbiceanu and her colleagues describe a puppeteering interface that draws on the expertise of professional puppeteers to control the behavior of virtual agents (Chapter 16). David Kirschner and Patrick Williams also take a close look at the interface; however they focus on the practices of expert players as they customize their interfaces for specific tasks (Chapter 18). In both cases, highly specialized interfaces are required to translate expertise into action within the virtual worlds, and performing these actions takes on an almost virtuoso quality.

The tendency of interfaces in virtual worlds has been to diverge: each system has its own set of conventions, affordances, and constraints. Different virtual worlds automate different levels of NVC, with different levels of success. In the case of many nonverbal behaviors, an ideal interface is one in which the user doesn't have to stop the flow of communication in order to adapt the actions of her avatar. In other cases, it might be desirable to provide the player with extremely granular control over the avatar's movements. Finding the balance between automatic behaviors and intentional ones is perhaps one of the central challenges faced by designers of virtual characters as the medium evolves.

2. FUTURE DIRECTIONS FOR RESEARCH

As the study of nonverbal communication in virtual worlds moves forward there are some important questions that researchers will need to grapple with. Some of these questions are already being asked, and we are well on our way to putting some of the early questions to rest for good. For instance, the most fundamental question of early research in the field: "is nonverbal communication possible in virtual worlds?" is essentially played-out. In this section we list some of the focalizing questions that we anticipate becoming relevant as the field advances. Some of these questions are already being addressed within ongoing research, while others remain largely unexamined. Many of these questions are still very basic, while others suggest more challenging and complicated research trajectories.

We propose to separate questions about NVC in virtual worlds into three broad categories: designing virtual worlds; understanding how people communicate in virtual worlds; and nonverbal communication in a hybrid reality.

DESIGNING VIRTUAL WORLDS

Much of the current discourse on virtual worlds is explicitly about improving the design of these environments. Thus, many of the questions about NVC in virtual worlds are about how to make it work better. Unlike much of the early design literature, that was focused on questions of "how?" (How do we make avatars? How do we implement gestures? How do we balance client-server relationships in our network?) these questions are about what type of experience is desirable for interactors. Some are quite simplistic, and it is these that have seen the most attention from the research community.

- What nonverbal signals lend themselves to implementation within virtual worlds?
- What nonverbal signals are necessary for rich communication within virtual worlds?
- Which signals may be safely ignored (or de-prioritized)?
- What is the right balance of intention and automation in social signaling behavior?
- How should control of social signals be divided between the player and the system?
- What interface best affords a rich NVC interaction?
- How visible/invisible should this interface be?
- How do pervasive sensing technologies change the design of NVC for virtual worlds?
- How will increased access to virtual worlds transform how we design for them?

NONVERBAL COMMUNICATION WITHIN VIRTUAL WORLDS

As much as we might design for specific social signaling behaviors, in practice we can seldom anticipate the ways in which people in virtual worlds will develop their own particular ways of communicating. Our second set of questions is concerned with how interactors within virtual worlds use the limitations and affordances of these environments to create their own social signaling behaviors that are not necessarily dependent on the physical world as a point of reference.

- What types of idiosyncratic NVC behavior arise within different worlds?
- How aware of their use of nonverbal signals are users of virtual worlds?
- In what contexts do people use more or less nonverbal cues in virtual worlds?
- What nonverbal behaviors are unique to digitally mediated environments?
- What is the relationship between interactional skill and NVC in virtual worlds?
- How do users subvert the underlying systems of virtual worlds for their own communicative ends?
- What literacies do users develop over time in virtual worlds?
- To what extent are those literacies transferrable from one virtual world to another?
- How do other communication modalities like voice chat, and video chat affect the types of NVC that people use in virtual worlds?

NONVERBAL COMMUNICATION IN A HYBRID REALITY

Virtual worlds are often viewed as a perfect laboratory for phenomena that are hard to study in the physical world. Yee et al write:

“If people behave according to the same social rules in both physical and virtual worlds even though the mode of movement and navigation is entirely different (i.e., using keyboard and mouse as opposed to bodies and legs), then this means it is possible to study social interaction in virtual environments and generalize them to social interaction in the real world” (Yee, Bailenson, Urbanek, Chang, & Merget, 2007).

Whether or not virtual worlds can be used as a controlled space for investigating real world phenomena is not an argument we intend to engage in here, however it does give rise to some important questions that we expect to continue to see as the field evolves. Central to these questions is how our nonverbal communications will evolve to respond to a growing overlap between the physical and the virtual world. We include here some questions pointing toward important and potentially challenging issues around ethics, identity, and value systems that research into this domain could fruitfully explore.

- What can NVC in virtual worlds tell us about human social behavior in general?
- Do cultural norms in NVC from the physical world (such as proxemic spacing behaviors) translate to consistent behaviors in virtual worlds?
- Do indicators of intimacy and relationships translate between virtual worlds and the physical world?
- What is the relationship between physical world identity and virtual identity, and which social norms dominate when these two are not in agreement? (When crossing gender in a virtual world, do things like proxemic spacing behavior and gaze change to suit the avatar gender, or the player's gender?)
- Do nonverbal behaviors from virtual worlds translate back out into the physical world?
- How does the increasing ubiquity of virtual worlds transform or augment physical NVC?
- What physical behavior characteristics can be sensed and translated into NVC in virtual worlds?
- What should be sensed and communicated, and what should be left alone?
- What are the ethical implications of a mixed physical and virtual world?
- To what extent should our virtual selves literally reflect our physical realities?
- To what extent should we be free to shape our identities, and our social signals in virtual worlds independent of physical (or biological) limitations in the material world such as gender, ethnicity, and able-bodiedness?
- Inversely, to what extent do our virtual embodiments need to reflect our material realities, and who is entitled to know the personal details of our physical identities?

As virtual worlds continue to evolve, we anticipate many more questions being added to these lists. What is clear is that there is no small amount of work remaining to be done on this topic.

3. FINAL THOUGHTS

The study of virtual worlds is moving out of a period where the discourse was dominated by the sheer novelty of these environments and into a period where we must treat them as fully realized social spaces. With the maturing of the field, comes a need to understand virtual environments as spaces with unique social affordances and as spaces that are closely related to the physical world on which they are modeled. The tradition of research into nonverbal communication provides us with a powerful set of analytical and descriptive tools for evaluating social interactions within virtual spaces, however in order to fully understand these phenomena, we need to incorporate a more sophisticated understanding of nonverbal communication into our research. In this book, we have collected perspectives from a wide range of practices, including professional designers, artists, theater practitioners, animators, and multi-disciplinary scholars. Looking at this body of work, we have been able to identify some critical themes and questions for the evolving study of nonverbal behavior in both virtual worlds and the physical world. It is our hope that with a richer understanding of NVC, future work in virtual worlds can begin to address some of the gaps we have identified and begin answering some of the questions that we have asked. We regard this work as the beginning of a conversation in the field, and believe that as virtual worlds become more pervasively spread throughout our experience of the physical world these issues and questions will become increasingly important.

4. BIBLIOGRAPHY

- Winograd, T., & Flores, F. (1986). *Understanding Computers and Cognition: A New Foundation for Design*. Norwood, New Jersey: Ablex Publishing Corporation.
- Yee, N., Bailenson, J. N., Urbanek, M., Chang, F., & 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.

