Sound, Embodiment, and the Experience of Interactivity in Video Games & Virtual Environments
Lori Landay, Professor of Cultural Studies, Berklee College of Music
Society for Cinema and Media Studies Conference, March 2013, Chicago, IL
–DRAFT—3/3/13: warning: citations incomplete! Material cut from presentation shoved into endnotes!

This presentation is part of a larger project that explores perception and imagination in interactive media, and how they are created by sound, image, and action. It marks an early stage in extending my previous work on the virtual kinoeye in virtual worlds and video games to consider aural, haptic, and kinetic aspects of embodiment and interactivity.

Innovative sound can bridge the gaps between the physical body, represented body, and imagined body in video games.
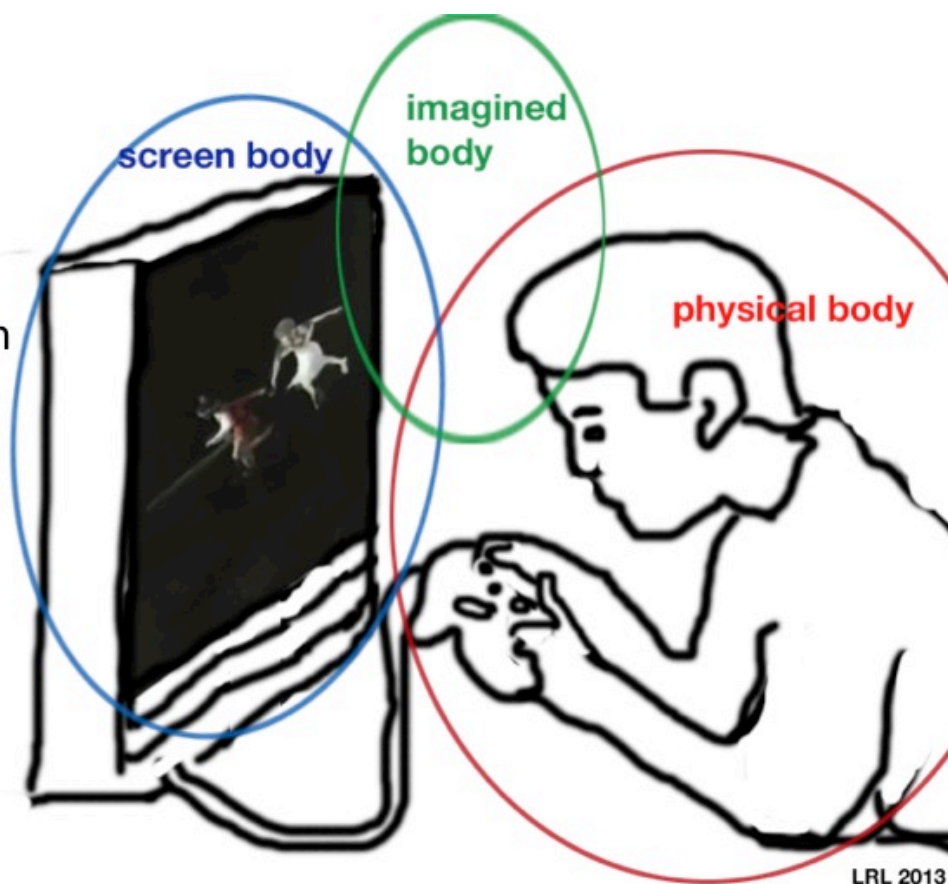
In particular, the argument is that innovative sound has the potential to bridge the gaps in experiencing embodiment caused by disconnections between the perceived body in physical, representational, and imagined contexts. To explore this, I draw on Walter Murch's spectrum of encoded and embodied sound, interviews with sound designers and composers, ways of thinking about the body from phenomenology, and analyze examples of the relationship between sound effects and music in video games. There's a QR code to a longer version of this material, with fuller attention to the theoretical model, literature review, and examples from more video games.

Let's start with a question philosopher Don Ihde poses when he asks his students to imagine then describe an activity they have never done. Often students choose jumping out of an airplane with a parachute, and as he works through the

distinctions between the possible senses of one's own body as the physical body, a first-person perspective of the perceiving body, which he calls the here-body, or the objectified over-there-body, the third-person view of one's own body, he asks, "Where does one feel the wind?" Ihde argues the full multidimensional sensory experience is in the embodied perspective.

Now we ask the same question of a video game experience, rather than something we imagine. We can't play a game right now, so we have to watch a recording of gameplay.

Do you feel the wind? Where? Why?
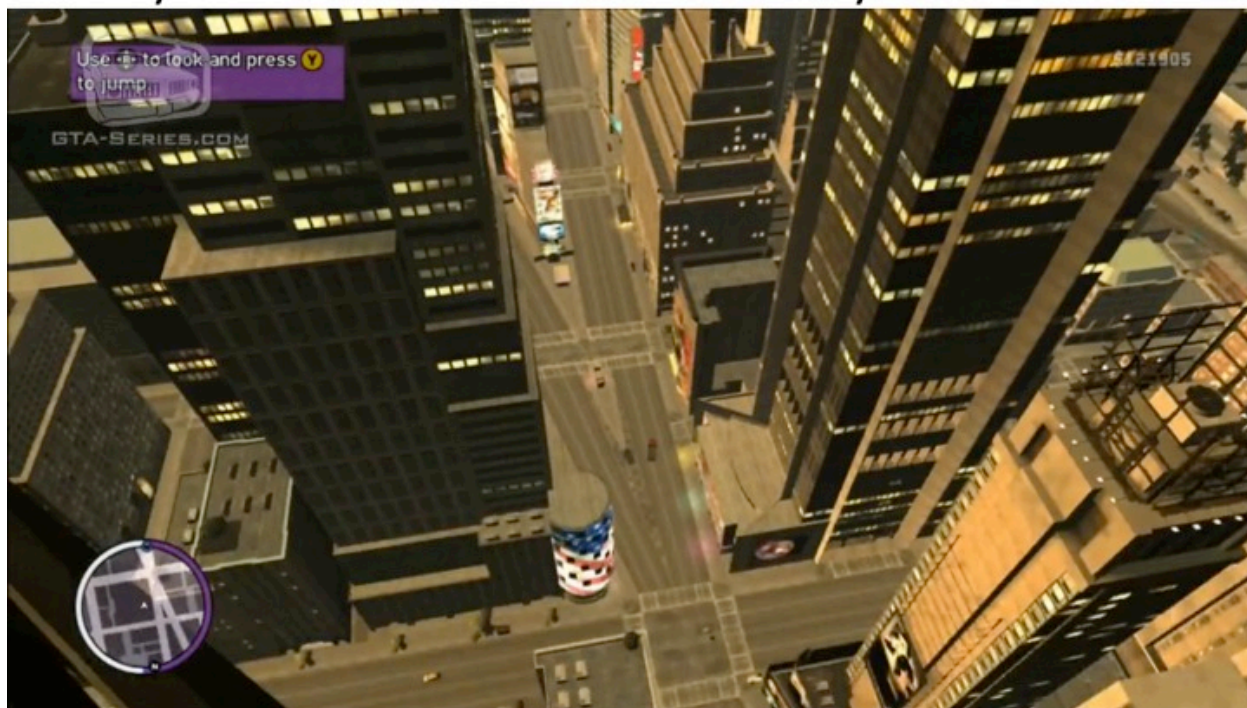
Grand Theft Auto IV: The Battle of Gay Tony
Rockstar North, 2009

Where do you feel the wind, if you do? There is no wind, although we can imagine the feeling of wind. There is image, and depiction of motion, but it is on the 2D screen in front of you. You can imagine the parachute jump, but the only sensory input you share with the embodiment position represented on the screen is

sound. You hear the wind, in 360 degrees, and 3dimensions, in your peripersonal space,[1] literally and figuratively between you and the jumper.

Let's experience another example, a very different game from *Grand Theft Auto*.

Do you feel the wind? Where? Why?

Flower, thatgamecompany, 2009

Flower, a game for Sony PlayStation3 so unusual that its creators "like to think of it as a video game version of a poem."[2] is perhaps more of a new media experience or interactive journey than a conventional video game. It's never clear who you are in this first person game; are you the wind? The embodiment of a flower petal? Something else? The visual, kinetic, and aural aspects are intertwined through interactivity and coherent meaning, with the motion-sensing gyroscopic Sixaxis PS3 controller changing the direction and speed of movement, which causes sound. PLAY FLOWER CLIP[3] [4]

Do you feel the wind? Do you feel you are the wind? There is the sound of not only the wind, but an intricate interactive sound design of the wind, and this brings me to the point of this argument. I imagine the feeling of the wind, I feel the wind in my imagined body, the one that is cruising through the grass, because *sound is always embodied*, always perceived in the here-body. As Frances Dyson argues in *Sounding new media: Immersion and embodiment in the arts and culture*, "Three-dimensional, interactive, and synthesthetic, perceived in the here and now of an embodied space, sound returns to the listener the very same qualities that media mediates: that feeling of being here now, of experiencing oneself as engulfed, enveloped, absorbed, enmeshed, in short, immersed in an environment. Sound surrounds." [5] Moreover, Dyson contends, sound has immediate physical and psychological effects because "to hear is also to be touched, physically and emotionally"—think of a high-pitched scream, or how low sounds vibrate in the body inducing panic. How does the sound in *Flower* touch us physically and emotionally?

Walter Murch's aural spectrum

Yellow

Sound Effects

Encoded-Embodied

Violet

Language

Encoded

Red
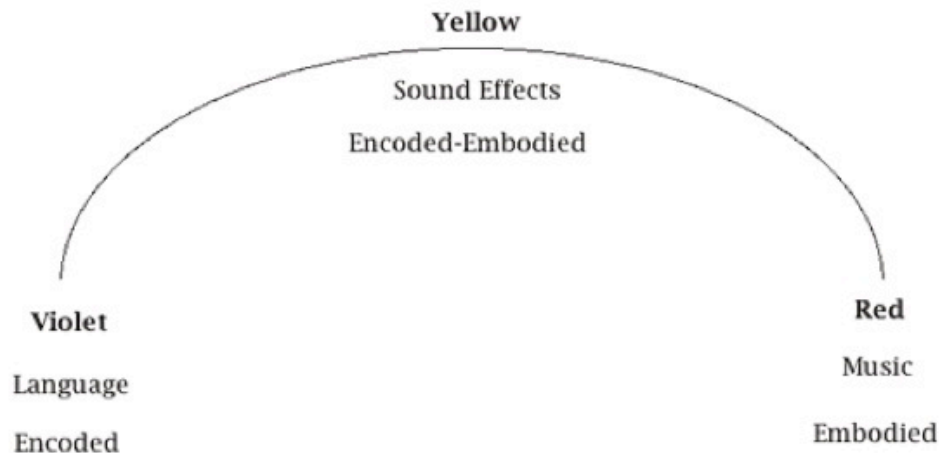
Music

Embodied

http://transom.org/?page_id=7006

One of the ways is by blurring the line between sound effects and music. Walter Murch describes a spectrum from "encoded sound" like language which has to be decoded and "embodied sound" like music which is experienced directly. Sound effects, Murch says, "fall mid-way, like 'sound centaurs,' they are half-language, half-music"[6]. In *Flower* and other games, when the conventional line between music and sound effects is blurred, when sound effects become part of the music, to be listened to directly, and music conveys information, both connected to the player's perception of interactivity, sound enables the player to create not only a series of sounds, but a sonic experience through his or her interactions, to, in the case of *Flower*, make music, which contributes to a sense of embodiment, because it is the experience you make.[7]

What has to be overcome to experience embodiment playing a video game? Current apparatus separates what happens in the virtual environment to the virtual character or avatar from the sensory input and the physical output of the player, and psychological and perceptual immersion are experienced despite the obvious fact that we are not really doing the action in the game, nor simply pretending or imagining.[8] Let's turn to an example of immersive new media art, *Osmose* by Char Davies for one way to join the gaps between the physical, virtual, and imagined sites of embodiment.[9]

unified sound, image, action in a virtual environment

Char Davies, *Osmose* (1994-96)

Play video of Char Davies' *Osmose* (1994-96)[10]

Davies explained, "in *Osmose*, sound and imagery have become one, amplifying each other." Davies's use of sound seems to me to sum up the holy grail of sound in video games and virtual environments: audio unified with the

visual and kinetic, generated in real-time by the participant's action in the environment, emotionally significant, and thematically coherent with the work as a whole.

What *Osmose* does with its head-mounted visual display is to overcome the gaps between the physical, virtual, and imagined bodies that persist in video games and much interactive media that use screens and controllers. There is no distinction between the 2D screen image and the player's immediate visual input, and the experience is controlled with breathing, not button-pressing. The sound dances between abstract and realistic, congruent with the visual, controlled by the body. [11].

Video games and other virtual environments experienced on screens and with typical game interfaces cannot rely on the kind of bodily interface that *Osmose* does for immersion.[12] Nevertheless, despite the limitations of interface and visual input, sound *could* and ideally does function in video games in the way it does in *Osmose*, because, as Mark Grimshaw explains, sound "exists and operates both in reality and in virtuality; it has a real volume and dimensionality that is a 3-dimensional representation of the 2-dimensional representation of the 3-dimensional world of the game."[13]
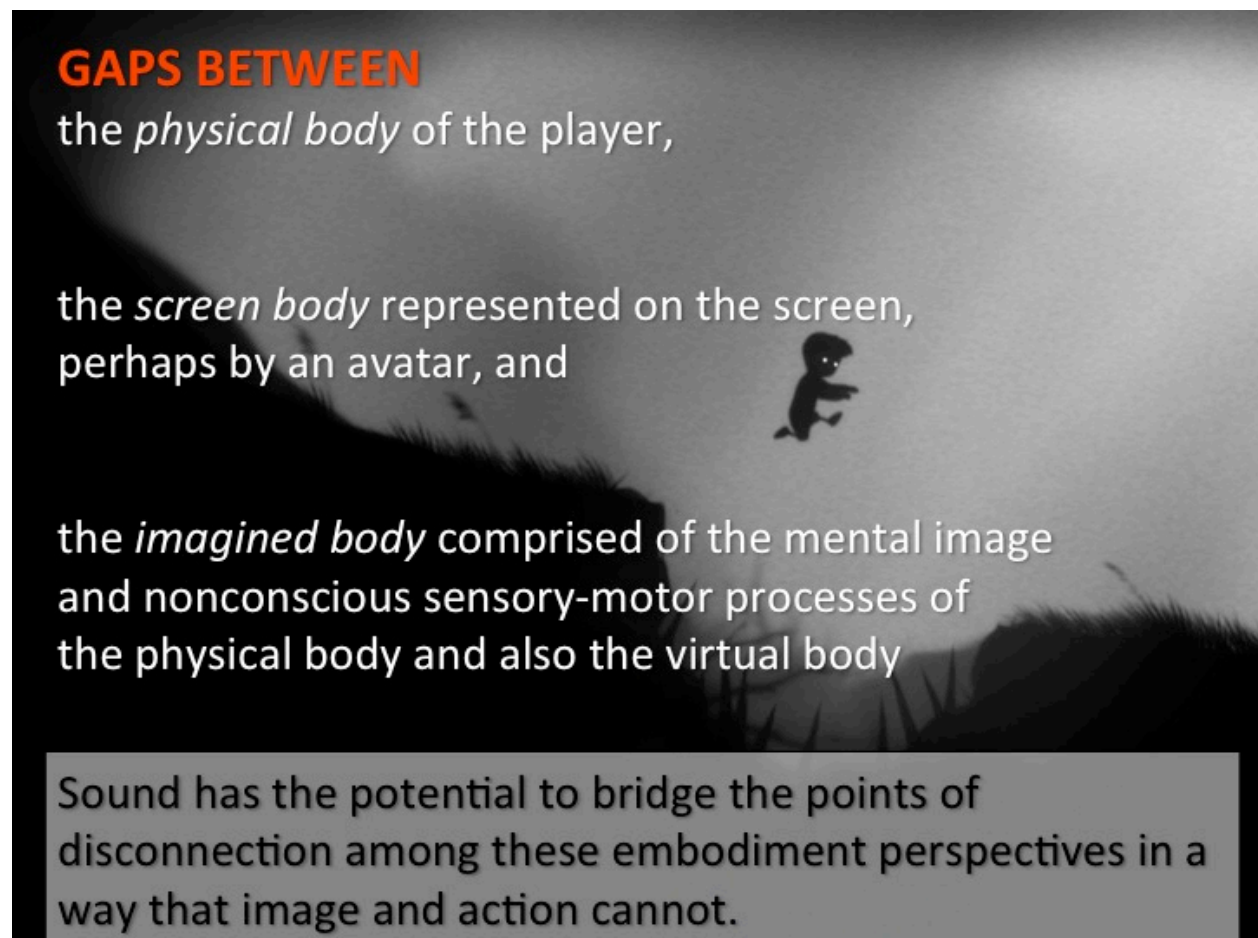
LIMBO, Playdead, 2010

An example of how innovative game audio can bridge gaps between the imagined body, the actual body, and the virtual body is *LIMBO,* in which sound and silence are used for narrative and emotional purposes to create a powerful sense of embodiment and heightened emotional response.

In the following clip, notice how although we are always visually distant from the boy, the sound is immediate, and changes according to circumstance as well as environmental realism. Instead of music throughout, there are times when there are only sound effects, like footsteps. *Limbo* sound designer Martin Stig Anderson used the prominence of certain sounds to indicate their importance in the game environment to the character, separating sound from visuals so that you hear "approaching obstacles and environments even before they're revealed visually,

and as you pass them they may be silenced entirely although they may still be in the frame thereby revealing new obstacles or environments to come.[14]" Anderson explains his strategy: "For me it has a much bigger psychological impact when you turn a naturalistic soundscape into abstraction by making your sound effects play as "music" rather than adding some traditional background music."[15]

 PLAY CLIP OF LIMBO

  Sound in LIMBO, therefore, moves sound effects closer to music on Murch's spectrum, into the orange range. The abstracted sound, more orange than yellow on Murch's spectrum, connects to the imagined body, and is affective.[16]

GAPS BETWEEN
the *physical body* of the player,

the *screen body* represented on the screen, perhaps by an avatar, and

the *imagined body* comprised of the mental image and nonconscious sensory-motor processes of the physical body and also the virtual body

Sound has the potential to bridge the points of disconnection among these embodiment perspectives in a way that image and action cannot.
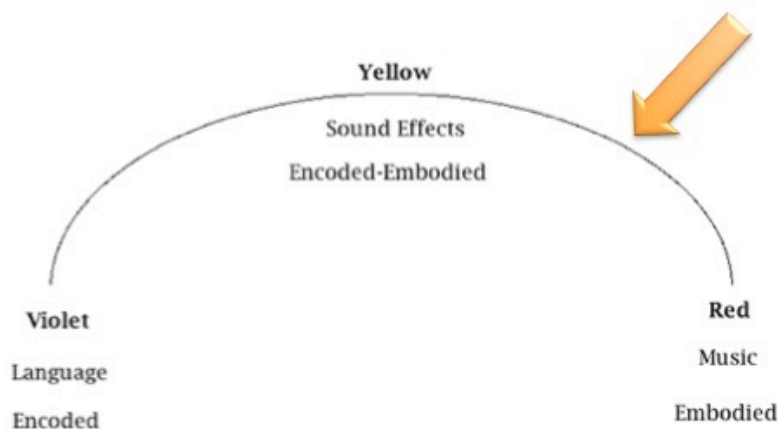
The visually-distant yet affective Limbo points out how the screen body is both the visual and kinetic representation on the screen and is also in the player's imagination; this is why I call it the screen body and not the virtual body, because

the virtual body is not only on the screen, but also in the imagination.[17] The representation on the screen is not really the over-there-body, and mind's version of the avatar is not exactly the imagined body of the player; we shift between these, and innovative uses of sound smooth the shifts, using aural input to connect the player to the physical and emotional experience of the avatar body, as it reinforces and amplifies visual and haptic aspects of interactivity. [18]

Sound effects & music in Portal 2

sound of soaring through the air after bouncing on "Aerial Faith Plate," a sound effect that can be better understood as a musical event than a diegetic sound

Yellow

Sound Effects

Encoded-Embodied

Violet

Language

Encoded

Red

Music

Embodied

Dan Bruno http://cruiseelroy.net/2011/06/portal-2-music/

To conclude, let's turn to Portal 2 a funny and clever puzzle game we play from the first-person perspective of Chell, but in an environment that gives us third-person glimpses of her through the portals we create in space. The unified sound design is characterized by sound effects and music that take on the characteristics of the other, some triggered interactively and others not, but both contributing to a sonic experience that does not recede into the background. Dan

Bruno, a composer who writes about sound in video games, uses Murch to describe how a game like Portal 2 "pushes sound effects and music towards each other." Here is an example:

PLAY AUDIO CLIP

The experience of Portal2 plunges you into an environment where polarities are not what they might seem, in all possible ways: up or down, in or out, behind and in front, of course, but also what is rational, emotional, abstract, concrete, useful, unnecessary, mechanical, organic, are all in flux. Chell is a center of calm, silence, action, and the sound design conveys this. Writer Geoff Keighley describes sound and music:

> The first *Portal* was renowned for its musical ending, and in *Portal 2* composer Mike Morasky wanted to up the ante with interactive music that would subtly evolve as players completed a puzzle. Run along orange speed paint and the music speeds up. Successfully jump across a ledge and the music shifts to let you know you're doing a good job. "The puzzles are thanking you for playing with them," is how Morasky puts it. "They love you."[19]

This is brilliant! The puzzles love you, and the way the game indicates its emotion is the sound. It is not surprising then, that this game would invest the nonplayer characters with something out of the ordinary, and here I'd like you to go with it as I consider speech as sound.

Speech as sound effects in Portal 2
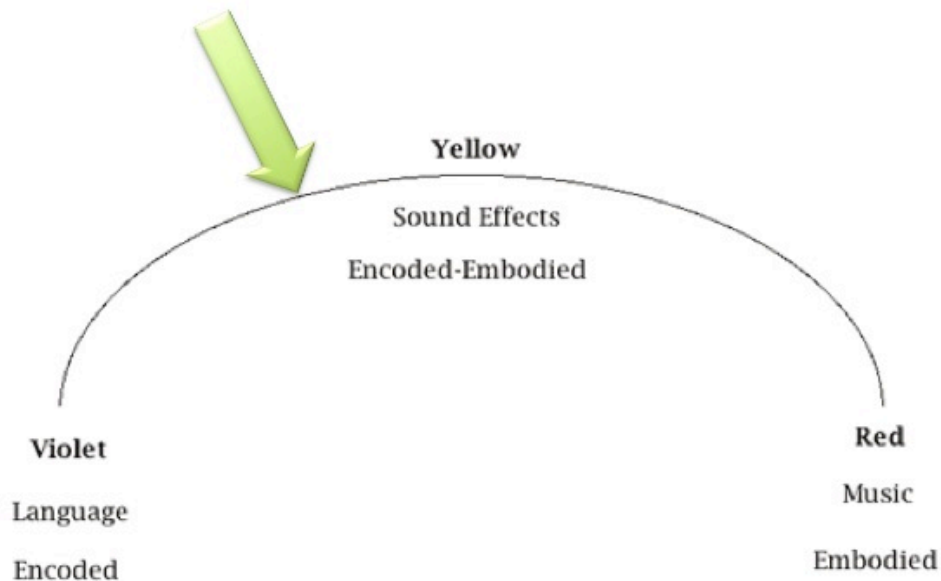
Murch writes, "To the degree that speech has music in it, its 'color' will drift toward the warmer (musical) end of the spectrum," and indeed it is the tone, the music, of the nonplayer characters that become part of the unified sonic experience. Brilliantly voiced by British actor Stephen Merchant, eyeball-robot Wheatley provides company, comedy, and exposition while the player solves the series of puzzles, but in a surprisingly natural, neurotic, and humanly-flawed way. Given that the protagonist, Chell, is silent, as so many characters in single-player games are, there is no interactive conversation between you/Chell and Wheatley, but he is so cleverly scripted, performed, and programmed that it feels like he is not only responding to your choices and outcomes, but interpellating you. Of course we decode the language that we hear, but it also pushes the encoded sound of the dialogue closer to the embodied/encoded hybrid of sound effects.

Portal 2, Valve, 2011

PLAY PORTAL CLIP

Wheatley is not the first character whose speech is not encoded sound but embodied sound; think of Mario and Luigi, or even of the princess and boy in *ICO*, who speak different made-up languages, with only Ico's words translated in subtitles.   But in Portal2, silence for Chell is more than conventional first-player silence; like the restraint in using music only, in Murch's terms, in the red part of the spectrum in favor of sound more in the orange part, Chell's lack of speech is meaningful in what it allows to take the prominent position in the sound design. As Portal2 head writer says about Chell, "She probably can talk, though. In our minds, she can talk. She's just pissed off the whole time and is refusing to dignify any of the things going by speaking about them."[20]

"Technology mirrors our desires; interactive technologies, in particular, reflect our desire to feel engaged."

David Rokeby, "Transforming Mirrors: Subjectivity and Control in Interactive Media," 1996

Artist David Rokeby commented, "Technology mirrors our desires; interactive technologies, in particular, reflect our desire to feel engaged." [21] Although embodiment is not the only goal of interactive media, it is often a way we measure engagement.

Sound, especially when music, sound effects, and even speech are interwoven innovatively in interactive media, is perhaps the most important of the triumvirate of sound-image-action because of its impact on embodiment. I suspect that when we overlay and intersect spectrums for image and action similar to Murch's, on which I'm working, we may have a very intriguing model for understanding embodiment and engagement in interactive media.

http://www.lorilanday.com/sound

OMITTED: *Bad Hotel, Urban Terror*[22], *EVE Online*[23]

---

[1] K. Collins (2013). *Playing with Sound: A Theory of Interacting with Sound and Music in Video Games,* Cambridge: MIT Press, p.42. I draw

[2] Kellee Santiago // President and co-founder of thatgamecompany http://blog.us.playstation.com/2009/01/19/stop-and-smell-the-flower-on-psn-february-12th/#comments

[3] thatgamecompany included composer Diamante in discussions with artists, illustrators, and level designers, so that in parts where traveling through a row of flowers causes a string of instrument sounds to intermingle with the music, Diamante contributed to deciding the location and color of the flowers because sound, image, and action are so closely connected in the game.

[4] Another example is how the sound of the wind, connected to the wind, gives hints for game objectives that are holistic in the game world. Let's listen to how the sound designer and composer discuss their work.
PLAY CLIP 2 OF FLOWER

[5] Dyson, F. (2009). Sounding new media: Immersion and embodiment in the arts and culture. Berkeley: University of California Press, 4.

[6] (qtd http://cruiseelroy.net/2011/06/portal-2-music/)

[7] The particular quality of this interactive feedback loop, of creating music through kinetic first-person movement in a sonically-rich environment, creates here a powerful sense of embodiment, on which the player can focus because they are not overly challenged by gameplay. This is different from music games, where the goal is to make music.
On a different spectrum, this time of interactivity, *Flower* allows the player to shape a musical experience, in a very different but analogous way that the iPad game *Bad Hotel* uses a generative music system so that player's actions and choices make music. The game doesn't sound the same twice because different game objects have different musical elements that combine in new ways, sonic equivalents of the structures the player builds to defend their hotel.[7]
BAD HOTEL CLIP

Bad Hotel, a game with a high degree of what some term open or complex interactivity in the player's ability to make sound has little embodiment. Of the three kinds of immersion categorized by researchers: *sensory immersion,* with the audiovisual (and sometimes propriocentric) input overriding actual world input, *challenge- based immersion* that requires motor or intellectual skills, and *imaginative immersion* by which people psychologically become involved in the characters, environment, and narrative of the experience,[7] Bad Hotel works primarily through challenge-based immersion and the music the player creates is part of the fun, but does not contribute to a strong experience of embodiment.

[8] Is it first-person perspective? An avatar? Sensory, or challenge-based, or imaginative immersion predominating over the others, or a magical balance of the three? Lots of interactivity? Story? A strong diegetic world? I would argue that all aspects of video game and interactive media are important, and can foster embodiment, but that sound plays a determining role.

[9] The instruments that limit image and action, as we can currently play video games, result in gaps, or breaches in experience. Many thinkers, in one way or another, have noted the different spaces and bodies that matter to the person engaged in virtual worlds, telematic interactions, and video games: the physical space where the computer or device and the physical body are; the screen space and whatever image-body is there, and then there is a third space, that of the imagination, which has been explored as a social "third place" in MMOs and the imaginary body.[9] Frank Biocca,[9] who researches human-computer interaction in virtual environments, concludes that embodiment is comprised of three bodies: "The objective body is the physical, observable, and measurable body of the user. The virtual body is the representation of the user's body inside the virtual environment. The body schema is the user's mental or internal representation of his or her body."
Biocca uses the term body schema to include both the non-conscious system of motor-capacities that enable us to move as well as what some distinguish as the body image, the perceptions and beliefs about one's own body.[9] The difference is a level of awareness. For my purposes, it is fine to put together the conscious and non-conscious aspects of the body in the mind, as they are mutually informing, and we can call that the imagined body, an umbrella for the perceptual, conceptual, emotional, and sensory-motor processes that occur in the mind.

[10] The artist of this immersive virtual environment, Char Davies, wrote:
The sound in *Osmose* is spatialised and interactive in real time, responding to changes in the immersant's location, direction and speed. My goal was to have sound that was neither literal nor illustrational, musical nor sound effect-like, but which was, in equivalence to the visual aesthetic, aurally ambiguous.. . . in *Osmose*, sound and imagery have become one, amplifying each other.[10]

[11] Margaret Morse, who unlike me, experienced Osmose personally, wrote:

While most of the virtual environments I have experienced use the metaphor of flying to move the visitor through virtual and symbolic space, Davies uses the metaphor of deep-sea diving in which we move between worlds organized largely in terms of up and down. We also hear sounds derived from sampling a single male and female voice. As we float through the dozen virtual world-spaces in the piece, we have as much access to the interior of transparent objects as to their exterior surface. . . . Davies herself and many visitors to the piece have had profound experiences related to the worlds themselves and the use of breath and balance to explore them. (Morse 1998, 208) **Virtualities: Television, Media Art, and Cyberculture**
By Margaret Morse
Bloomington, Indianapolis: Indiana University Press (1998)

[12] ; at best, the avatar is a visual stand-in for the immersant, and the player watches the avatar, or sees from its first-person perspective on he 2D screen. The interface is still primarily screen and hand controller, and even when there is a haptic, vocal, or kinetic element of control, there is a gap between the environment in which the game is played and the gamespace itself, a gap between the body on the screen (even if it is suggested in a first-person perspective) and the body of the player, that the game seeks to overcome through immersion and interaction.
[13] Mark Grimshaw, **SOUND AND IMMERSION IN THE FIRST-PERSON SHOOTER**
http://wlv.openrepository.com/wlv/bitstream/2436/35995/2/Grimshaw_CGAMES07.pdf
[14] "Besides contributing to the foreboding atmosphere of Limbo, such mixing minimizes the risk of making sounds become annoying to the player, simply because the sounds only play as long as they're important to the actual game." http://designingsound.org/2011/08/limbo-exclusive-interview-with-martin-stig-andersen/

[15] "Moreover, making your "music" emerge from the environment is likely to make the audience more forgiving towards it since they'll accept it as stemming, however abstractly, from the environment."
[16] Listen to the relationship between sound effects and music in this next clip of gameplay towards the end of the game.

PLAY CLIP 2

[17] The imagined body of the avatar in the player's mind is neither based on nor separate from the physical body of the player, and that is where game audio has its most impact, on the imagined experience of the avatar body.
[18] GO TO ICO SLIDE

Like LIMBO, *ICO's* sound design makes sound effects prominent over music to create the 3D environment.[18] The third dimensionality of the

game space is represented visually as well as aurally, and there are variations in scripted camera positions during gameplay as well as in the cut

scenes that guide the visual input as well as narrative. Here, sound effects position the player in the space, with emotional resonance, but it is the

music, cued to the princess, that supplies the affect.

Watch and listen.

PLAY ICO CLIP

We hear the music with special emphasis after we are used to listening to the sound effects so intently for information about the environment,

after listening to encoded sound, strengthening the meaning-making of the music.

[19] in his behind-the-scenes making of Portal 2 feature "The Final Hours of *Portal 2*"
[20] http://gameological.com/2012/07/interview-erik-wolpaw-portal-2-head-writer/
[21] ." ("Transforming Mirrors: Subjectivity and Control in Interactive Media" 1996 http://www.davidrokeby.com/mirrorsconclusion.html).
[22] Grimshaw uses first-person shooter games to show his point, explaining:
"the diegetic sonic world of the FPS game extends from the screen to physically encapsulate the player in the acoustic ecology's real resonating space. ...Thus, FPS game diegetic sounds extend the game environment from a flat, 2-dimensional screen to the 3-dimensionality of the external world. The player's proprioceptive sounds are replaced by the character's proprioceptive sounds and all other game world sounds envelop the player as part of the game's real resonating space. These sounds form part of not only the real resonating space but also the virtual resonating space of the game and thus help to immerse the player, both physically and mentally, in the FPS game acoustic ecology."
FPS SLIDE diagram [22]
Urban Terror clip: http://www.youtube.com/watch?v=N3tMh6nrGYw
By comparing the waveforms from an inactive and active player, we can see the difference that interaction makes. What we hear is a barrage of sound effects in the gunfight, far more direct sensory input than the first-person visual perspective gives us, and no music. The first-person perspective is intended to embody us, to connect the hand and gun on the screen to our physical body as extensions to another, imagined body.
SLIDE
But there are gaps, or breaches, in the experience, all the ways that the hand and gun behind the information also displayed on the screen in text and image, and my actions with keyboard and mouse control action on the screen, but my physical body is certainly not running or pointing.   I am disconnected from an embodied experience, perhaps because the game is more encoded than embodied, more like language than music.

[23] How does an avatar make a difference? There is more information, and sometimes more choice in camera control, from the third-person perspective, whether the follow camera position closely aligned with an avatar's physical point of view, or from another vantage point. One of the games I've been following is EVE Online, a MMPORG of particular interest to me because only recently have players been able to be represented by avatars who "walk in" space "stations," previously being represented by the ships they were in (other than in profile pictures). Although ships and other objects, especially vehicles, can be extensions of the body, the audiovisualkinentic rhetoric of the game did not seem to deliberately encourage a strong embodiment. In addition to limited experiments with avatars walking in stations, EVE seems to be adding to embodiment by strengthening the kinesonic aspects of the game, with more attention to soundscape. Most significantly for this discussion, the approach to game music shifted from an ingame jukebox to music connected to the context of place and action, some procedurally. Other changes in sound, according to language in the development blog, include "gratifying sounds" "to get that sense of accomplishment" after finishing a mission or training. It will be interesting to see how EVE Online develops, on its current platform and the new, connected PlayStation first-person shooter set on the planets the ships fly past, and the role that sound might play in its quest for greater embodiment.
http://community.eveonline.com/devblog.asp?a=blog&nbid=73569
EVE CLIP