Making Gamers Cry: Mirror Neurons and Embodied Interaction with Game Sound

Karen Collins
University of Waterloo
200 University Ave W
Waterloo, ON
+001 (519) 888-4567 x38326
collinsk@uwaterloo.ca

ABSTRACT
In this paper, I draw on an embodied cognition approach to describe how sound mediates our identification with and empathy for video game characters. This identification is discussed in terms of mirror neurons and body schema, drawing on theoretical and empirical research to explore ways in which identity is created from our embodied interaction with sound. I conclude by suggesting ways in which sound designers and composers can use this information to create more empathy and identification between players and their game characters.

Categories and Subject Descriptors
J.5 [**Computer Applications**]: Social and Behavioral Sciences – *Psychology*.

General Terms
Design, Human Factors, Theory

Keywords
games, sound, embodiment, gesture, mirror neurons, empathy.

1. INTRODUCTION
Can a videogame make us cry? Enter this question into a search engine and thousands of pages will result: having games that make us cry is considered by many designers to be the "holy grail" of game design. Quantic Dream's David Cage for instance said of *Heavy Rain* (Sony 2010), "Our starting point… was to say look, all the games around here are just based on adrenaline… It's just on excitement and frustration and competition—all these simple and primal emotions. What about all the more complex emotions like empathy, sadness, happiness?… Can we make you cry?"[50] Rhianna Pratchett, writer of the game series *Mirror's Edge* (EA 2008) and *Overlord* (Codemasters 2007) argues against the desire to make gamers cry, suggesting that, "As an industry we get very, very obsessed with making people cry."[23] Designer and author Richard Rouse even gave a talk entitled, "Five Ways A Video Game Can Make You Cry" at the Game Developers' Conference in 2010.[42] Actually making people cry is not directly the point of the issue, of course. At stake are two related questions: 1) can we create emotional depth in games, and 2) *how* do we do that: what techniques or tools are necessary to create the empathy needed to reach such emotional depth? In this paper, I suggest that gesture-based, embodied interaction with music and sound in games opens up possibilities for enhancing just such an emotional involvement in the game.

AM'11, September 7–9, 2011, Coimbra, Portugal.

Philosopher Don Ihde explores three distinct ways of understanding embodiment.[27] Ihde first makes a distinction between the historical approach to the phenomenological body (the sensory, material body of spatial orientation, physical perception and emotion as described by Merleau-Ponty [36]) and the cultural body (the socially, politically and culturally constructed body of postmodern philosophy, as described by Foucault[20]). But Ihde claims these understandings of the body fail to account for technology: the phenomenological body will be unaffected by technology, whereas the cultural body exaggerates technology's impact. Ihde then describes a third body that brings together elements of the first two, what he calls the *technological body*, which is characterized by our interactions with technology (by which he means everything from primitive to high-tech tools). In this way, the human body (both material and cultural) is shaped by and can be extended through the use of tools (technologies), what Ihde calls "extended embodiment", an "instrument-mediated experience in which the instrument is taken into one's experience of bodily engaging the world."[27]

Here, I explore the possibilities of the gestural control of video games as such an extension of the body, and in particular discuss the ways that gestural interaction with sound and music can be exploited to increase empathy and emotion in narrative games. I do not wish to suggest that this approach is the only way to generate empathy or identification with the character (for another perspective see Gee[22]), but rather that this is *one* way to create empathy. Although this paper is theoretical, I provide evidence from both theoretical and empirical research from areas of embodied cognition theory, neurobiology, psychology, philosophy and musicology. Drawing these together, I then suggest some directions for ways that we can better implement sound in games to improve the player's experience.

2. EMBODIED COGNITION AND THE MIRROR NEURON SYSTEM
Embodied cognition holds that our understanding of the world is shaped by our ability to physically interact with it. In this way, embodied cognition is complementary to phenomenology, but whereas phenomenology is grounded in philosophy, embodied cognition is grounded in psychological theory and cognitive science. According to embodied cognition theory, our cognitive processes use re-activations of sensorimotor states from our past experiences to understand the world. In other words, our knowledge is tied to the original neural state that occurred in our brain when we initially acquired the information. Our cognition is therefore "embodied" because it is inextricably tied to our

sensorimotor experience; our perception is always coupled with a mental re-enactment of our physical, embodied experience.

Scientific evidence that supports the embodied cognition approach can be found in the recent research into the mirror neuron system in the brain, through which a three-way mapping between sound, image and action occurs. Research has shown that the same group of neurons will fire when either performing or observing an action, and so these are referred to as "mirror neurons". Put differently, the neuroactivity in our brain responds as if we are performing an action that we are not performing ourselves, but are rather witnessing (through vision or sound). Our emotional and neurophysiological state can be directly affected by what we see: for instance, if we see pain or fear in someone else, we understand this in terms of our own psycho-physiological experience of similar pain or fear.[40] For example, neurons that normally fire when a patient is pricked with a needle will also fire when the patient watches *another* patient being pricked.[41] Mirror neurons are therefore closely tied to—and possibly responsible for—our experience of empathy. Neuroscientist V.S. Ramachandran believes that mirror neurons dissolve the barrier between self and others, and so refers to them, with humor, as "Gandhi Neurons".[41]

Keysers et al. describe research into monkey mirror neurons that they conducted, in which they found that the same neurons fired whether an action is performed, seen or *heard*.[29] These audiovisual mirror neurons respond as if we are experiencing the cause behind the event, when only the sound of the action is presented.[31] In other words, when the monkey hears the sound, the brain responds as if it is also seeing and experiencing the action creating the sound. Lahav et al. [32] tested mirror neurons in relation to musical sounds, and found that when participants were trained on a piece of piano music, the sound-action mirror neurons became active: When they listened to music they had not played, the neurons did not become active, indicating our own motor experience of the gestures involved are critical to understanding the sound of actions. Indeed, in other cases mirror neuronal activation has been shown to be stronger when we have undertaken the action ourselves.[10] In this sense, we can understand not only the physical action, but also the *intentions* that underlie that action.[21]

Morrison and Ziemke argue that such phenomena can function not just with witnessing other humans, but also with virtual characters, and thus we "can facilitate a user's identification with the character's 'body' as well as provide the groundwork for empathy."[38] There is, in other words, a cognitive multimodal mapping between visual, motor and auditory representations that is closely integrated to our own feelings of empathy, even with virtual characters, due in part to our tendency towards anthropomorphism.

2.1 Mirror Neurons and Anthropomorphism

As shown above, mirror neurons provide evidence that we experience what we see/hear in our own egocentric terms. Mirror neurons are still very much under-explored, however, and it is somewhat unclear if we experience the same empathic and egocentric understanding of other creatures or inanimate objects. Some evidence would suggest that we do. It is our natural tendency to imbue inanimate objects with life, what is sometimes referred to as the *pathetic fallacy* or anthropomorphism. Human beings have a predisposition to anthropomorphize objects—particularly those objects with which we interact regularly, that have human-like physical characteristics (especially neotonous characteristics), or that behave in "an apparently complex or intentional manner."[19] Interactivity in particular contributes to anthropomorphism, notably through receiving feedback on our own actions. Explains Morse, "[A computer's] very capacity to give feedback and the immediacy of its response lends the quality of 'person' to what is a computational tool. This responsiveness allows it (and the virtual entities it displays) to pose or function as subjects—however partial, quasi, imaginary, and virtual—who are involved in the interactive exchange."[39] Our tendency to anthropomorphize interactive objects would suggest that we would anthropomorphize characters in video games more so than characters in film or animation, for example.

What is particularly interesting about this tendency is that we do not need the objects to visually resemble an animate being in terms of its physical characteristics (beyond movement). Morrison and Ziemke describe that, "it does not take very much for humans to anthropomorphize even simple animate agents or to make personality trait attributions to geometrical shapes or point-light figures …[we] need very little provocation to interpret a triangle as 'chasing' a square or to think that the triangle is 'mean' and the square is 'frightened'."[38] Tremoulet and Feldman used a single moving object and found that animacy was attributed to the shapes when we could explain the *intent* of the movement (such as "reacting" to other shapes), illustrating that it is not just motion, but our attribution of cause or intent to that motion that leads to anthropomorphism, which they refer to as the *intentionality hypothesis*.[46] Moreover, Cooley demonstrated with her experiments of animated walking triangles that sound can especially help to anthropomorphize an inanimate object, and give it a sense of personality.[14] So sound, intentional movement, and interactivity are all critical to our anthropomorphizing tendency.

If we interpret what we see/hear in animated sequences in terms of intentionality, and our neuronal activity is tied to understanding the intentionality behind the actions of others, it may be the case that when we interpret animated sequences we sense that intentionality, because our own motor-neuronal activity is activated in the same way that it would be if we were witnessing another (live) person. That is to say, we may interpret video game characters in anthropomorphic terms, and this anthropomorphism may help us to create an empathic relationship to those characters. Game players may flinch, twitch, or otherwise indicate that the game world is "bleeding into" the real world. This suggests that we experience empathy with game characters in a similar fashion to our empathic response to the actions of other people. The photos of video game players by Robbie Cooper and Shauna Frischkorn, Beate Geissler, Oliver Sann and others illustrate the engagement of the physical body with the virtual space, for example, and anybody who has watched a child play a game in front of a television will be able to testify to the leaning with racecars or characters, particularly in moments of stress, or jumping the controller with the character. The research suggests that interactivity, movement and sound all help to forge this empathic experience.

If we accept that we can see/hear/feel the experience of game characters through our own egocentric mirror neuronal system, then there are several steps that designers can take to enhance this empathic experience. In particular, we can use the knowledge of our visualizing causality and intentionality of sounds to better express the character's emotional state, so that the sound/music is more likely to impact upon our own emotional state.

3. MUSIC, GESTURE AND SOUND IN GAMES

The concept that we mentally mimic the causality behind sounds we hear is not new to musicologists. Research has shown that we can recognize and feel the emotion conveyed by a performer when we listen to music, and this also results in physiological changes (blood pressure, heart rate, galvanic skin response).[8] An embodied cognition approach as to why this occurs suggests that we understand human-made sounds (including those of playing a musical instrument) in terms of our own experience of making similar sounds and movements. Our mental re-creation of the music causes a neuronal and motor-sensory response in ourselves that mimics the performer, and thus we are able to interpret the emotional inflections through our re-creation of the action. We therefore give meaning to sound in terms of emulated actions, or corporeal articulations.[33] Put differently, we mentally and sometimes physically imitate the expressiveness of the action behind the sound, based on our "prior embodied experience of sound production."[15] Winters describes, "The mimetic hypothesis might also provide an explanation for why we might find ourselves unconsciously 'imitating' the emotion seemingly being expressed, in addition to any willing participation in a game of make-believe" [49].

Electronically generated or synthesized sounds and music remove this corporeal connection to causality. Issues of "liveness" frequently arise in discussions of electronic music (see e.g. [26]), for instance, and we might extend these to a broader concept of "a-liveness". For example, the use of a laptop as an instrument is criticized as "not live", since "the body of the musician is not directly and causally acting on an object physically to create a sound. The causal relationship between what we hear coming out of the speakers and the body of the performer is broken."[16] In short, with our mirror-neuronal response to interactive sound, liveness is tied in the minds of many to a causal link between the performer's action and the instrument's sonic response. Croft elaborates, "we expect a sound proportionate to the energetic characteristics of the performer's action… to have a more or less transparent relation to the properties of the sounding body we see before us."[16] Calling on Barthes' *Grain of the Voice,*[2] Croft points to imperfections, and the sounding body of the instrument inherent in acoustic music as having implications for the meaning of liveness and by extension our empathic response. It is (at least in part) the human gesture in sound and music that creates the emotional mimicry and connectedness between listener and creator. As discussed above in relation to mirror neurons, we mentally re-create (visually and motorically) what we hear, and we hear in terms of intentionality and causality—including emotional intent—and thus we empathize with the originator behind the sound.

As with electronic/digital music, early video games also removed the gesture from our interaction with sound in the games. In a majority of cases the gestures we make with our input devices are still largely unrelated to the sonic output. A historical overview of the "push-button" control in games and a critique of their lack of gestural ability has been previously taken up by Griffin, who writes, "The dependency of video game control on the button reflects a disregard for the body's abilities. By relying on this artifact of automation, the video game medium must adopt a cognition-centric approach to interaction—giving up the pleasures and benefits of physical involvement."[24] Gestural control over sounds can provide us with useful feedback: the strength of our gesture, temporality, proximity, angle, and so on. With non-gestural game controllers, this sonic feedback is decoupled from the player's gesture. In other words, we are merely pressing a button, and it does not matter how hard we press that button nor how long, or in which direction we hold the controller; we will still hear the same sonic feedback. In a study of musical instruments, Bongers found that such decoupling of sound source from gestural control results in a loss of useful feedback to the performer (feedback intrinsic to acoustic instruments), and thus leads to a feeling of restricted, if not loss of, control.[5] With the loss of gesture, there is a loss of emotional articulation inherent in gesturally produced sound, and through loss of control a lessening of identification with the character.

The introduction of the recent gesture-controlled input devices such as the Wiimote, the Xbox Kinect, as well as controllers adapted from real-world objects, such as the musical instruments of *Guitar Hero* (Harmonix 2005) and *Rock Band* (Harmonix 2007), have ushered in a new wave of opportunities for coupling the actions of the player to sound events. Gestural input devices are (arguably) easier and more intuitive to use, judging by the wide success of the Wii: but does this sonic connection to our own body also make it easier for us to create an embodied, cognitive connection to the characters? Bianchi-Berthouze, Kim and Patel found that physical body movement on the part of game players increased their level of engagement and modified the ways in which they became engaged in the game.[4] By using gestural controllers (in this case, musical controllers like the guitar of *Guitar Hero*), and thus inducing body movement, players felt more of a sense of presence in the game-world and had altered their affective state. The authors suggest that these effects are directly the result of the controller's ability to enable roleplay: "The players appeared to quickly enter in the role suggested by the game, here, a musician, and started to perform task related motions that were not required by the game itself." Viewed in other terms, mimicry leads to a greater engagement, attachment, identification and thus empathy. Bodily movement in general—particularly the mimicry of posture and gesture—is closely tied to emotion in a bi-directional manner: emotions generate movement and movement generates emotions.[3] For this reason, when we inhibit the physical movement of the game player, this interferes with the emotional experience.[8] Ekman et al. suggest an elaboration on this through their study of what they termed "emotional contagion", in which we mirror the facial expressions on others that we see, and thus elicit the proper autonomic nervous system's response to the perceived emotion (that is, we feel empathy).[19] In this way, we similarly may be able to "catch" a character's emotional state.[25]

With gestural input devices, either the player determines the length of the sound output being played (assuming the sound is mapped to the player's input), or the sound is only mapped to the start of the event (i.e. triggered at the start of the input gesture), which means that the length of the gesture may be incongruent with sonic response (the length of the sound). We can refer to the congruity between the player's gesture and the sound as *kinaesonic congruity*. Kinaesonic (kinaesthetic + sonic) refers to "the physicalization of sound or the mapping of sound to bodily movements."[50] When sounds are kinaesonically incongruent, how does this impact our embodied understanding of that sound?

In kinaesonic congruence, we have a sonic reaction that matches the action that our body is making, thus embedding our own expressiveness into the game through the character. Even in cases where we do not have kinaesonic congruence with sound, however, we still receive *some* of that sonic response to our

action. Although our physical body did not kinaesonically create that sound, the mimetic hypothesis (and mirror neuronal research) suggests that we may still *feel* that it did through our own subsequent mental/corporeal imitation. If we hear sound in terms of our own embodied experience of that sound, then when we hear those action sounds in games, even though we did not kinaesonically create them, *we hear them as if we did*. In other words, we have a direct, embodied interaction with the sounds that we evoke and hear in games, and coupled with our physical and/or kinaesonic-congruent action, these sounds (and thus the game character) can become an extension of the self.

4. THE TECHNOLOGICAL BODY

I would now like to suggest a second related means by which we may experience empathy with the game character, which is through the extended, technological body. Scientific evidence shows that the areas of the brain related to sensorimotor activity will treat a tool such as a game controller as an extension of the hand or arm.[38] With this view of tools, it is easy to conceive of the game controller as becoming an extension of the body—we do not view the controller as *part* of our body, but we can experience the virtual world *through* the controller. In other words, we may extend our bodily representation without altering our body schema. This extensibility of tools is not a new conception, of course: Merleau-Ponty for instance talk at length about how tools become an extension of the self.[36] A tool frequently used in a sense becomes an extension of our self because we no longer focus on it as a part of our experience: it facilitates a feeling of non-mediation between us and the (in this case virtual) world. McLuhan wrote extensively on the concept of technologically mediated extensions of mind and/or body, arguing that any media form or technology can become an extension of our senses and therefore ourselves, with all media altering our cognitive (and by extension social) organization, comparing new technologies "to what happens when a new note is added to a melody."[35] In *The Global Village*, McLuhan describes that "Communication media of the future will accentuate the extensions of our nervous systems, which can be made disembodied and totally collective."[34] Ihde expresses a similar concept in his discussion of "embodiment relations" where "the experience of one's body image is not fixed but malleably extendable and/or reducible in terms of the material or technological mediations that may be embodied."[26]

The rubber hand illusion illustrates how multimodal sensory perception can influence our sense of self through a physical extension, in this case a rubber arm. In this series of experiments, Botvinick and Cohen found that by stroking a rubber hand in view of a participant while simultaneously stroking the participant's real hand (hidden from view); this led to the participant's taking "ownership" over the rubber hand, to the extent that they could "feel" through the rubber hand.[6] Although these experiments have not used auditory feedback as a tested condition, they illustrate the point that multimodal feedback impacts the ways in which we understand our sense of self. We could, in other words, make the conceptual leap to say that using image, sound and haptic feedback with a virtual arm or character, as in a video game, could lead to a similarly altered (extended) sense of self.

What the rubber hand illusion illustrates is that we have a *body schema* into which we may assign non-corporeal and corporeal objects.[9] Our body schema can extend beyond our physical body to incorporate non-corporeal objects. However, to the extent that we incorporate such objects into our body schema, these must be representational of our own body: If we replace the rubber hand with a wooden stick, the effect fails.[9] De Preester and Tsakiris make a distinction, therefore, between *incorporation* and *extension*. The rubber hand is *incorporated* into our body schema, and becomes a physical part of our body to the extent that our brain takes ownership over it, and our body schema is altered. *Extension*, on the other hand, is the feeling of non-mediation between ourselves and the world through the use of a tool or technology.[17] Incorporation and extension are closely related—indeed de Preester and Tsakiris point out Merleau-Ponty's conflation of the terms in his discussion of the blind man's cane, in which the cane becomes a tool through which the blind man senses the world.[36] McLuhan likewise conflates extension with incorporation in his theories of media.[35] There is an important distinction in the sense that with incorporation the object (e.g. the cane) becomes a part of (is incorporated into) our existing body schema; with extension the object (cane) is an expansion of that body schema into what we might refer to as the peripersonal space. Peripersonal space is an intermediary space between our body (personal space) and what we see as the external environment (extra-personal space).[9]

Cardinali, Brozzoli and Farne argue that auditory information falls into the peripersonal space, and therefore is an extension, rather than becoming part of our body schema (an incorporation).[9] In this way, it could be said that in games, sound extends our sense of self beyond our physical body and into the intermediary space between ourselves and the virtual world. Sounds that we make—including in the virtual world—become a sensory extension of our self into that virtual world. The auditory realm of games thus becomes an extension of the self, a technological body through which we sense the game-world.

An important consideration of this sonic involvement in the game is the idea of self-produced sound. Self-produced sound can be defined as sound produced by one's own body or movement.[1] Film sound theorist Michel Chion [11] proposes the term "ergo-audition" to refer to how we hear ourselves taking an action, noting that, "this concept extends the concept of feedback to incorporate the subjective experience of one's impact on the world." Our connection to the sound, in other words, is arguably much stronger (than externally produced sounds) when it comes to self-produced sounds, because of the embodied connection between self and the sound. In other words, we have a physical experience of sounds that we produce ourselves that is different from other sounds in our environment (see, e.g. [43]). Knoblich and Flach describe that the action system of the brain is "responsible for creating an immediate sense of self by determining whether certain sensations and perceptions are the result of one's own actions."[30] Put simply, then, if we undertake an action and have an immediate sonic response, even if that sonic response takes place in a virtual world, because it also takes place in the real world (our peripersonal space), we can integrate that sound as an extension of our body schema.

The elasticity and mutability of our body schema is readily apparent. As Ihde comments: "We are our bodies—but in that very basic notion one also discovers that our bodies have an amazing plasticity and polymorphism that is often brought out precisely in our relations with technologies. We are bodies in technologies."[28]

The ease with which we might adjust our body schema through our sensory perception illustrates just how possible it may be that we can identify with a game character (who is an extension of our self through sound). In other words, it is (in part) through sound that we extend our self and become the character (or the character

becomes us). The embodied cognitive connection to sound is vital in our extension of our self, as an extension of our body-in-technology. In this sense, it is not the *controller* through which our body is extended, it is through the *game character*. The character is the tool through which we experience the virtual world—through which we bump into walls, get shot, dig holes and talk to other characters. This is most apparent with games that use the Kinect, and thus eliminate a handheld controller altogether.

To summarize, then, not only may we experience the game world in egocentric terms due to our audiovisual mirror neurons, but we may experience the game character as an extension of our self, since we are primarily responsible for the character's sounds, thus they may facilitate an extension of our body schema to incorporate that peripersonal auditory space between the virtual and real world.

5. PUTTING IT INTO PRACTICE

I have suggested two significant related ways in which empathy might be created in video games, and I have discussed how sound contributes to these phenomena, but I have not yet illustrated how we might use this information in terms of the creation and implementation of game sound to better facilitate empathy in narrative games. I will now explore a few examples of how we might use the theoretical and empirical research presented above in real-world video game sound design and composition.

5.1 Music in Games

Music is one of the key elements that drives emotion in media. Countless studies have illustrated how music influences the interpretation of emotion and intent in films, for instance (see e.g.[12]). As discussed above, humans can distinguish the emotional expressions of the performer in live music. While this mimetic element of music is not by any means the only method by which emotion is articulated, it could be suggested that, particularly in scenes of high emotion, music in games could be more effective in communicating emotion if recorded with live performers.

While there are certainly reasons not to use live performers (budgets are the primary reason, but electronic music is sometimes aesthetically desired), much of the more "cinematic" style narrative video games draw on film music traditions, and a single live instrumentalist recorded over a synthesized score can go a long way towards "humanizing" a score, while still keeping costs low. For example, I may have a synthesized orchestral score, but perhaps the lead violin may play a particularly prominent motif that expresses the emotional state of my character: using a live violinist in the recording who understands the motivations of the character in the way an actor might, would help to add that additional expressive quality to the music and drive empathy in that scene. In addition, many sequencers have "humanizer" or randomizer plug-ins that produce small timing errors and inflections. While these are (arguably) not as effective as a live performer, understanding the instrument being emulated and the expressive qualities of playing those instruments can go a long way to incorporating such effects into the score, and increasing empathy. I do not wish to suggest here that this is the only way to induce emotion in music, rather that it may be one more way to add an expressive quality.

Perhaps the most effective way to implement the ideas in this paper are to allow the expressiveness of the player into games, particularly music-based games. As described above, it is useful to start with a live recording of a band/instrument in music-based games, to capture the emotion of the players into the original recording. But what may be particularly useful is to allow the player to attempt to emulate those expressive qualities. It is this expressiveness for instance that characterizes much air guitar.[33] However, the current music game controllers are limited in their expressive abilities: we can lift the neck of the guitar in *Guitar Hero* to gain "star power,"[37] and occasionally employ the whammy bar, but we cannot make use of the nuances of vibrato, palm mutes, bending, slides or harmonics, and the strength of the button-presses makes no difference to the notes produced. The loss of the expressive elements of gesture, therefore, lessens the emotional involvement of the player. Imperfections and inconsistencies that signify liveness and personality in an acoustic recording are penalized in the game as they are interpreted as errors. By allowing the player to add these expressive elements through controller design, we can increase the involvement of the player in the game, the feeling of connectivity to the original performance and, through getting them to emulate the emotions, could even alter their emotional state and empathy for the character.

Similarly, we may better use the technologies available in gestural devices (either through motion/image capture like the Kinect or remotes like the Wii controller) to allow for the expressive control of self-produced sound effects in games. In this way, we encourage the technological extension of the body while at the same time reinforcing the empathic, egocentric connection to the game character. In particular, through tying the gesture of the player directly to the sound (discussed above as kinaesonic congruence), the player has a more realistic extension through the game character into the virtual world.

5.2 Synthesized Sound Effects in Games

As with synthesized orchestras and electronic music, synthesized sound effects in games lack an embodied connection to any causality. The consequences of gesture-based input for sound designers have been interesting, for with the timing and articulatory, energetic qualities of a player's gesture remaining unknown in advance of an action, the sound designer is suddenly faced with unpredictable sound length and strength requirements. The response of some sound designers has been to return to sound synthesis methods. While it was common to synthesize sounds in the early years of video games, by about 1990 this practice had been largely abandoned, at least in terms of console games, in favor of sampled sounds.[13] It was felt that sampled sounds were more realistic, and that this realism would drive game sales.

Synthesis solves some of the problems of temporality, expressiveness and unpredictability tied to gestural input, but how effective such sounds are in provoking the same kind of neuronal response seen with recorded (sampled) sound remains unknown. As illustrated, if sounds appear to have been generated by a "real" object, we can infer intentionality and emotion behind those sounds. Experiments in quality ratings of different synthesis methods vary in response and have been undertaken primarily without the context of image, e.g.[7] It is certainly likely that it is more difficult to imply intentionality and emotion through synthesized sounds. It may be worth experimenting with a combination of sample and synthesized sound in sound effects: We typically identify sounds by the attack portion of the sound envelope.[47] By retaining a sample of the attack but using synthesis to extend the signal (through the decay, sustain and release), we might therefore solve some of the potential issues with realism or fidelity of synthesized sound effects while allowing for more extended or shortened temporality to adjust to

the player's gesture. Real-time DSP effects are already used to adjust the sound games during run-time, but temporality remains a particularly tricky issue with which to contend.

Another way to overcome this realism issue is through contextualization. Many sounds in media are not realistic in the sense that they may be chosen for exaggeration, metaphor, and so on. For example, the sound of bones breaking in film is often a recording of a fresh piece of celery snapping. The context of hearing the bones break is enough for us to mentally "feel" the break in egocentric terms, even if what we actually hear is celery. Moreover, this remains effective with or without the visual of the bone breaking to reinforce the feeling. Such a visceral, corporeal use of sound can be particularly useful in *acousmatic* (off-screen) sound. Much can be implied through sound, since as shown we always hear those sounds in a frame of reference of our own previous experience with similar sounds. Thus hearing bones snap off-screen can be as disturbing (sometimes more so) than actually seeing it happen (I say more so because this is particularly the case when the visual action is mimicked, thus potentially disrupting the verisimilitude of the scene). As others have argued, this is a particularly fruitful technique in the horror genre.[44] While the acousmatic mentally draws our attention off-screen in cinema, in games we can "de-acousmatize" sounds by actually seeking out their source through altering the camera angle or moving the player-character.[45] We may, in other words, seek out the cause of the sound, which can be particularly useful as a gameplay device. Tinwell and Grimshaw for instance use the example of a little girl crying in *Left 4 Dead* (Valve 2008), which appeals to our protective (empathic) side, but when we attend to the sound, we discover that she is in fact a witch to be avoided.[44]

5.3 Dialogue in Games

Dialogue is also an important means to convey emotion in games. As with sound effects and music, we experience vocal sound in terms of our own embodied experience of similar sounds: we mentally mimic the voice in our own body. The consequence of this is that we need professional voice actors in games, as the loss of emotional inflection will significantly affect our ability to *feel* the emotion and thus develop empathy. The writers and directors of game dialogue likewise need to be aware of the various means of drawing the emotion out of the actors. For this reason, it is also sometimes germane to not use voices at all, because if the mental mimesis fails to map onto the actual voice, then the result may be less immersive than if there had been silence. In other words, if the voice in our head that we ascribe to our character (perhaps our own voice, perhaps an affected voice) is incongruent with the voice we do hear in the game, we lose that empathic connection. Kenny Young described his approach to the dialogue in *Little Big Planet* (Media Molecule 2008) for instance as follows:

"I think the sound design decision not to litter the Sackfolk with inane voice samples contributed significantly towards [winning an] award. This is something which some sound designers find hard to resist—there are a couple of 3rd party trailers and adverts out there where some [idiot] has added chipmunk voices to the Sackfolk to make them sound comic and cute… not having a voice allows players to feel that their Sackboy, which they lovingly dress, customise and emote with, is theirs. This sense of ownership would be hard to achieve if the character you were controlling had a mind of their own, voice and language being the most personal way of communication and expression. Which is why when Sackboy does speak it is with the voice of his player, his lips moving to match those of his puppet master."[51]

Young makes an important point here, which is that it is, in part, the role of the sound team to make the player more fully engaged with their character. Young argues that another approach to player-character dialogue is to eliminate the voice of the player's character, since the player has already read the words on-screen, and thus already "heard" (and re-enacted) the voice in their heads. In this way, we are aiding the creation of the extended body. If we are to become our character, shouldn't they sound like us, or at least how we imagine they should sound? The voice is so intimately personal, that the "wrong" voice can destroy our illusion of the character. Young continues:

"*Dragon Age* [EA 2009]... uses a different system whereby you are presented with several verbatim options for what your character could say and, then, as soon as you click on one of these phrases it is as if your character has already said it and you immediately hear and see the other party's response. This works beautifully for several reasons: Having read all the options, considered whether it fits with the character you have established and any potential outcomes, there is no need for you to hear your character speak this information out loud again (a trap fallen in to by earlier games, such as Ion Storm's *Deus Ex* [Eidos 2000]) because you've already just 'heard' it in your head when reading it. And so, the act of clicking replaces the act of speaking....by not hearing a prescribed character voice which takes them out of the experience the player is empowered to *fully inhabit their character*" [51: my emphasis].

6. SUMMARY: GESTURAL SOUND AS EMPATHY DEVICE

In summary, I have presented here two significant means through which scientific research (mirror neurons and body schema) and philosophical theory (the technological body) blend together to suggest ways in which the player may be more empathic towards their game character: bodily extension and gestural interaction. My focus has been on how an embodied interaction with sound in particular can be used to forge this relationship between the player and the character, and I have suggested several ways in which we might employ this knowledge in game sound design to enhance the ability of the player to understand and experience game characters in egocentric terms. These can be summarized into several suggestions as follows:

1. Allow for more expressivity in music-based games.

2. Experiment with means to better employ sound effects and synthesis to tie the sounds to the player's gestural input.

3. Have a clear tie between sound and (emotionally inflected) causality—either through live performance/sampling or through contextualization.

4. Explore means by which a player can identify with their character through voice (either absence of voice, use of the player's voice, or ways to encourage role play through voice).

The suggestions here are by no means exclusive. There are certainly other ways that we can improve the auditory experience. For example, if self-produced sound aids our immersion and identification with the character, the sudden jumps to cinematic sequences found in many games surely disrupts this subjectivity, as we are suddenly prevented from having the media respond to our actions. Eliminating cinematics and finding another means to convey game information may better immerse the player and continue the extended self. In addition, I have not touched on spatial sound here. I would suggest that encouraging the use of surround sound speakers would likely increase the effects

described here, since we are then even more immersed in the sound, our own body in the middle of the peripersonal auditory space.

I will conclude with a philosophical point: If game technology has reached the point where game designers can make us cry, then it follows that they can also make us angry, aggressive or happy. Evidence from studies of *Second Life* (Linden Labs 2003) suggests that qualities that we take on in the virtual world spill into our conduct in the real world. Bailenson has shown that after only 90 seconds, the interactions that take place virtually can elicit real behavioural changes in our physical self.[18] It will be the responsibility of designers in the future to take into account this ability to generate real change in the world through games.

7. ACKNOWLEDGEMENTS

The author would like to acknowledge the ongoing support of the Social Sciences and Humanities Research Council of Canada. Elements of this paper were inspired by conversations with Neil Randall, Chris Salter and Colin Milburn.

8. REFERENCES

[1] Ballas, J. Self-produced sound: Tightly Binding Haptics and Audio. in *Haptics and Audio Interaction Design 2007*, 1–8.

[2] Barthes, R. 1977. *Image, Music, Text.* Hill and Wang, New York.

[3] Benyon, D., Höök, K. and Nigay, L. 2010. Spaces of Interaction. In *Proceedings of the ACM-BCS Visions of Computer Science Conference*, Edinburgh, UK, April 2010.

[4] Bianchi-Berthouze, N., Kim, W.W. and Darshak, P. 2007. Does body movement engage you more in digital game play? And Why? In *Proceedings of the 2nd International Conference on Affective Computing and Intelligent Interaction*, ACII '07, (2007), *LNCS* 4738. Springer-Verlag, Berlin, 102–113. DOI=10.1007/978-3-540-74889-2_10

[5] Bongers, B. 2000. Physical Interfaces in the Electronic Arts: Interaction Theory and Interfacing Techniques for Real-Time Performance. In *Trends in Gestural Control of Music*. M. M. Wanderley and M. Battier (eds.) IRCAM–Centre, Paris, 41–70.

[6] Botvinick, M and Cohen, J. 1998. Rubber hands 'feel' touch that eyes see. *Nature* 391 (Feb. 1998), 756. DOI= 10.1038/35784

[7] Böttcher, N. and Serafin, S. 2009. Design and evaluation of physically inspired models of sound effects in computer games. In *Proceedings of the Audio Engineering Society, AES 35th International Conference*, London, UK, February 2009, 11–13.

[8] Bresin, R. and Friberg, A. 2001. Expressive Musical Icons. In *Proceedings of the 2001 International Conference on Auditory Display*, Espoo, Finland, Hiipakka, J., Zakarov, N., & Takala, T. (Eds.). 141–143.

[9] Cardinali, L., Brozzoli, C., and Farnè, A. 2009. Peripersonal Space and Body Schema: Two Labels for the Same Concept? *Brain Topogr,* 21, 252–260. DOI=10.1007/s10548-009-0092-7

[10] Chandler, J., and Schwarz, N. 2010. Use does not wear ragged the fabric of friendship: Thinking of objects as alive makes people less willing to replace them. *J. Consum Psychol,* 20 (2010), 138–145, 139. DOI=10.1016/j.jcps.2009.12.008

[11] Chion, M. 1998. *Le Son*. Paris: Editions nation.

[12] Cohen, A. J. 1999. The functions of music in multimedia: A cognitive approach. In S. W. Yi (Ed). *Music, Mind, and Science*. Seoul National University Press, Seoul, Korea, 53–69.

[13] Collins, K. 2008. *Game Sound: An Introduction to the History, Theory and Practice of Video Game Music and Sound Design.* The MIT Press, Cambridge.

[14] Cooley, M. 1998. Sound + Image in Design. in *Proceedings of ICAD '98.* Glasgow, UK.

[15] Cox, A. 2001. The Mimetic Hypothesis and Embodied Musical Meaning. *Music Sci,* 5, 2 (Fall 2001)195–212, 195. DOI= 10.1353/not.2002.0116

[16] Croft, J. 2007. Theses on liveness. *Organ Sound,* 12,1 (2007), 59–66, 61. DOI=10.1017/S1355771807001604

[17] De Preester, H. and Tsakiris, M. 2009. Body-extension versus body-incorporation: Is there a need for a body-model? *Phenom Cogn Sci,* 8, 3 (2000) 307–319. DOI=10.1007/s11097-009-9121-y

[18] Dell, K. 2008. How Second life Affects Real Life. *Time*, Monday May 12, 2008. http://www.time.com/time/health/article/0,8599,1739601,00.html

[19] Ekman, P., Friesen, W.V., and Levenson, R. W. 1990. Voluntary Facial Action Generates Emotion-Specific Autonomic Nervous System Activity. *Psychophysiology*, 27,4 (Jul 1990) 363–384. DOI=10.1111/j.1469-8986.1990.tb02330.x

[20] Foucault, M. 1982. The Subject and Power. *Crit Inquiry*, 8,4 (1982) 777–795.

[21] Freedberg, D. and Gallese, V. 2007. Motion, Emotion and Empathy in Esthetic Experience. *Trends Cogn Sci,* 11, 5 (2007) 197–203. DOI=10.1016/j.tics.2007.02.003

[22] Gee, J. P. 2004. What video games have to teach us about learning and literacy. Basingstoke, UK: Palgrave Macmillan.

[23] Gilbert, B. 2009. GDC09: Rhianna Pratchett says games should forget about making people cry. *Joystiq*. March 27 2009. http://www.joystiq.com/2009/03/27/gdc09-rhianna-pratchett-says-games-should-forget-about-making-p/

[24] Griffin, S. N. 2002. Push. Play: An Examination of the Gameplay Button. *Loading* 2, 2 (2008) http://journals.sfu.ca/loading/index.php/loading/issue/view/3.

[25] Hatfield, E. Cacioppo, J.T. and Rapson, R.L. 1994. *Emotional contagion*. Cambridge University Press, New York.

[26] Ihde, D. 1979. *Technics and Praxis*. Dordrecht, Holland, D. Reidel Publishing Group, 508.

[27] Ihde, D. 1986. *Consequences of Phenomenology*. New York State University of New York Press, New York, 141.

[28] Ihde, D. 2002. *Bodies in Technology*. University of Minnesota Press, Minneapolis, 138.

[29] Keysers, C., Kohler, E., Umiltà, M. A., Nanetti, L., Fogassi, L. and Gallese, V. 2003. Audiovisual mirror neurons and action recognition. *Exp Brain Res,* 153, 4 (2003) 628–636. DOI= 10.1007/s00221-003-1603-5

[30] Knoblich, G., and Flach, R. 2003. Action identity: Evidence from self-recognition, prediction, and coordination

Conscious Cogn, 12 (2003) 620–632. DOI=10.1016/S1053-8100(03)00070-9

[31] Kohler, E., Keysers, C., Alessandra, M., Umilta, L.F., Gallese, V. and Rinolatti, G. 2002. Hearing Sounds, Understanding Actions: Action Representation in Mirror Neurons. *Science*, 297, 5582 (2002) 846–848. DOI=10.1126/science.1070311

[32] Lahav, A., Saltzman, E., and Schlaug, G. 2007. Action Representation of Sound: Audiomotor Recognition Network While Listening to Newly Acquired. *J Neurosci,* 27, 2 (Jan. 2007) 308–314. DOI=10.1523/JNEUROSCI.4822-06.2007

[33] Leman, M. 2008. *Embodied Music Cognition and Mediation Technology*. MIT Press, Cambridge, MA.

[34] McLuhan, M., and Powers, B. R. 1989. *The Global Village*. Oxford University Press, Oxford.

[35] McLuhan, M. 1962.*The Gutenberg Galaxy*. University of Toronto Press, Toronto, 83.

[36] Merleau-Ponty, M. 1998. *Phenomenology of Perception*. Routledge, London, 41.

[37] Miller, K. 2009. Schizophonic Performance: *Guitar Hero, Rock Band*, and Virtual Virtuosity. *Journal of the Society for American Music,* 3, 4 (2009) 395–429. DOI=10.1017/S1752196309990666

[38] Morrison, I. and Ziemke, T. 2005. Empathy with Computer Game Characters: A Cognitive Neuroscience Perspective. in *AISB'05: Proceedings of the Joint Symposium on Virtual Social Agents*. AISB, UK, 73–79.

[39] Morse, M. 1998. *Virtualities. Television, Media Art and Cyberculture*. Indiana University Press, Bloomington.

[40] Niedenthal, P.M. 2007. Embodying Emotion. *Science,* 316 (2007) 1002–1005. DOI= 10.1126/science.1136930

[41] Ramachandran, V.S. 2009. The neurons that shaped civilization. TEDIndia, November 2009. http://www.ted.com/talks/vs_ramachandran_the_neurons_that_shaped_civilization.html

[42] Rouse, R. 2011. GDC 2011 Preview: Richard Rouse III. *Next-Gen*. http://www.next-gen.biz/features/gdc-2011-preview-richard-rouse-iii

[43] Rochat, P. Early Development of the Ecological Self. In *The Self in Infancy*, P. Rochat (ed.), 53–71. Elsevier, Amsterdam, 1995.

[44] Tinwell, A., Grimshaw, M., and Williams, A. 2011. Uncanny speech . In Grimshaw, M. (Ed.), *Game Sound Technology and Player Interaction: Concepts and Developments*. Hershey, PA: IGI Global.

[45] Todd, N. and McAngus, P. 1995. The Kinematics of Musical Expression. *J Acoust Soc Am,* 97, 3 (1995) 1940–49. DOI= 10.1121/1.412067

[46] Tremoulet, P. D., and Feldman, J. 2006. The influence of spatial context and the role of intentionality in the interpretation of animacy from motion. *Percept Psychophys*, 68, 6 (2006) 1047–1058.

[47] Truax, B. 1996. Soundscape, Acoustic Communication and Environmental Sound Composition. *Contemporary Music Review*, 15,1 (1996) 49–65. DOI=10.1080/07494469608629688

[48] Wilson-Bokowiec, J. and Bokowiec, M. A. 2006. Kinaesonics: The intertwining relationship of body and sound. *Contemporary Music Review*, 25, 1 (2006) 47–57. DOI= 10.1080/07494460600647436

[49] Winters, B. 2008. Corporeality, Musical Heartbeats, and Cinematic Emotion. *Music, Sound, and the Moving Image*, 2, 1 (2008) 3–25, 12.

[50] Yin-Poole, W. 2009. Muzyka: Dragon Age and ME 2 will make Gamers Cry. *Videogamer.com* http://www.videogamer.com/news/muzyka_dragon_age_and_me_2_will_make_gamers_cry.html

[51] Young, K. Voice in Bioware's 'Dragon Age: Origins'. *Sound spam blog*, Sunday, 16 May 2010, http://soundspam.blogspot.com/2010/05/voice-in-biowares-dragon-age-origins.html