Rock Band: A Case Study in the Design of Embodied Interface Experience

Joshua Tanenbaum
Simon Fraser University
250 13450 102nd Avenue
Surrey, BC V3T 0A3 Canada
1-778-858-4113
joshuat@sfu.ca

Jim Bizzocchi
Simon Fraser University
250 13450 102nd Avenue
Surrey, BC V3T 0A3 Canada
1-778-782-7437
jimbiz@sfu.ca

Abstract

There has been a recent surge of novel interface devices available for home gaming systems. With the rise in popularity of games like Guitar Hero and consoles such as Nintendo's Wii comes new opportunities for game design at the interface level. In this paper we propose three interrelated dimensions for the analysis of embodied and gestural game interface hardware devices. We demonstrate how gestural and embodied interactions can be understood as ludic, kinesthetic and narrative experiences. We ground this discussion in a close analysis of the interface affordances of the game Rock Band and demonstrate how these three dimensions allow us to understand more clearly the place of the interface in the design and the experience of games.

Categories and Subject Descriptors

H.5.2 [**Information Interfaces and Presentation**]: User Interfaces - *Input devices and strategies, Interaction styles, Theory and Methods, Haptic I/O, User-centered design.*

General Terms

Design, Human Factors, Theory

Keywords

Game Interface, Embodiment, Narrative, Ludology, Novel Interaction, Physical Interaction Design

1. Introduction

The lead singer grips her microphone in sweat drenched hands, eyes closed. There is a perfect moment of anticipation before the song, when the crowd of spectators goes quiet and the air is charged with electricity. Behind her, the drummer counts off four quick beats, clicking his sticks together in the air. Cymbal crash, and the guitar and bass both growl, and the audience roars back. The lead singer can feel the rhythm of the music vibrating through the floorboards in time with the pounding of the kick drum. Next to her, the guitar player is "pogo-ing" up and down in time with the music, his long hair flailing around his head. The song segues out of its intro and the singer belts out the first lyrics as the crowd goes wild. She wields her voice with carefully crafted precision – an instrument that she has honed as finely as any of her bandmates – and tonight the music rewards her by coming easily. A grin breaks across her face; she whips her head back, and dances in time with the music as the crowd joins in.

This story is remarkable, not because it is about a singer in a rock'n'roll band, but because it is about a group of gamers playing a video game with their friends at a party. The game is Rock Band, and the experience described is one of a new generation of embodied and gestural game interactions that have become very popular.

As digital games have evolved, so have the interface devices that we use to play them. For many years, the video game arcade was the best place for players to enjoy a particular type of specialized game interface - those that offered a robust, embodied experience that physically immersed them in the world of the game. In Namco's Alpine Racer (1992) players maneuvered a downhill skier by shifting and tilting their body, with their feet strapped into a pair of free-floating "skis"; Atari's Road Burners (1999) simulated motorcycle racing by placing players on a motorcycle shaped controller which they could steer by turning the handlebars or by tilting the bike from side to side.

Interfaces such as the "light gun" and the driving wheel moved a version of this embodied experience out of the arcade and into the home, with many examples of each available for home gaming consoles and computer systems. Recently, however, the use of embodied game hardware devices in home systems has exploded, led by the unprecedented popularity of Nintendo's Wii – a gaming console that relies on a combination of accelerometer data and camera vision to track positions of controllers in each hand. The Wii is by no means the first home system to use gesture as input – the Sony Eyetoy uses image tracking for gestural control; the Sega Activator used IR sensing; and the venerable Nintendo Power Glove used ultrasonic pulses for positioning. Nonetheless, it is the first such generalizable gesture-friendly device to achieve such widespread popularity.

Concurrent with the appearance of the Wii, another new interface phenomenon has emerged as a significant force in the home gaming landscape: the music simulation game. Rock Band (Harmonix, 2007) is one of the latest members of this new species. Popularized by the game Guitar Hero, the genre is notable for the use of unique and specialized game controllers, much like its sister genres: the rhythm game (as exemplified by games such as Dance Dance Revolution and Pump It Up, Exceed) and the Karaoke Game (as exemplified by games such as SingStar and Karaoke Revolution). Rock Band combines the vocal interface of the Karaoke Game with a variation of the guitar controller introduced in Guitar Hero and provides a new interface device in the form of an electronic drum controller. [Figure 1]

Rock Band's three controllers represent a range of the possibilities for new gestural and embodied game interface devices. To unpack the operations of these three novel interfaces we propose an analysis along three related dimensions: the ludic, the kinesthetic and the narrative. We argue that these three dimensions allow us to understand more clearly the place of the interface in the design and the experience of Rock Band.

Figure 1: The Three Rock Band Controllers

2. Ludic Interaction

We use the term "ludic" from game studies as a shorthand for describing the aspects of game experience concerned with the pleasures of tactical and strategic gameplay in general, and with the notion of winning and losing in particular. The pleasures of ludic performance have often been linked to Csikszentmihalyi's notion of Flow, which proposes that pleasurable experiences often exist in a balanced state between boredom and frustration [Csikszentmihalyi 1990]. It is generally accepted that interactions at the ludic level must maintain a balance between being too simple, and thus boring, or too challenging, and thus frustrating. This balanced state is known as the Flow Channel [Csikszentmihalyi 1990]. Using Csikszentmihalyi's concept as a starting point, we propose two distinct dimensions to more clearly articulate the understanding of ludic interface operation: *ludic efficiency* and *granularity*.

2.1 Ludic Efficiency

We define *ludic efficiency* here as: *the extent to which an interface device eases or hinders the player's attempt to perform any given operation within the game*. Ludic efficiency explicitly remaps Csikszentmihalyi's notion of *challenge* onto a continuum. The most efficient interface is the one which is the *easiest* and most *simplistic,* the one which affords the shortest interactional path to success. A ludically inefficient interface is one that is challenging, and actively *problematizes* the interaction needed to succeed. To a certain extent, all game interfaces are at least slightly problematic, and need to be, in order to support a satisfying experience.

We use the term *ludic optimization* to refer to the process of balancing these two sides of the ludic efficiency spectrum in order to provide a satisfying game experience. We can imagine the most efficient interface as a simple button, which – when pressed – automatically wins the game. This game would not be very much fun, but the interface would be highly efficient. This button is an extreme example of the *simplistic* side of the spectrum. In contrast, we can imagine a game where the interface actively frustrates the interactions of the player by randomly changing the mapping of the controls, without cause. This game would also not be very much fun, due to the highly inefficient operations of the interface. This interface is an extreme example of the *problematic* side of the spectrum. Based on the flow channel's balance between boredom and frustration, there is an *optimal* position for an interface that balances the simplistic against the problematic.

In many cases the design goal of the interface may be *easing ludic efficiency*. If other game variables are sufficiently challenging, the optimal interface would be an efficient one. However there are also examples of game interfaces where the interface is designed to *problematize ludic efficiency*. In the popular Soul Calibur games, ludic success relies on memorizing and executing complicated combinations of buttons and joystick movements. One of the central pleasures of this game is mastering a highly *problematic* interface. An example of a fighting game that uses more *simplistic* interface to achieve similar results is Capcom vs. SNK 2 EO, which allowed players to map these complicated operations to positions on a single analog joystick. Finding a balance between an interface which is too *simplistic* and an interface which is too *problematic* is a careful design choice, and one which is often central to the strategic experience of play, as it provides one of the sets of constitutive constraints that bound the game experience.

The interface of Rock Band is particularly interesting in terms of its design stance towards *ludic efficiency*. The *efficiency* of Rock Band's interface varies with the controller chosen[1]. The microphone, for instance, provides a direct audio interface between the voice of the player and the game software, and does little to mediate that interaction (feedback is provided via an arrow that indicates whether the sung pitch is higher or lower than the expected pitch in real time). In this case, the game operation required for ludic success is perfectly aligned with the activity supported by the interface. Not only is the microphone a very simplistic interface, it is also a uniquely *specialized* one; the microphone could not be used as a controller for any game that does not rely on audio signals as input. We describe this type of controller as highly constrained in that it cannot be used for a more general purpose; however these constraints are *optimizing constraints* in that they uniquely afford the desired interaction. In the case of vocal interactions with the game, the challenge is mediated by two factors: the vocal capabilities of the player, and the extent to which the system forgives mistakes.

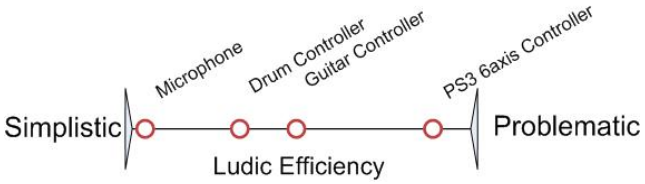

Figure 2: Ludic Efficiency Scale for Rock Band Controllers

The drum and guitar controllers are also specialized to the interactions demanded by the game, and are subsequently more *efficient* than a more general purpose interface, such as the standard Playstation or X-Box controller. [Figure 2]

In particular, the physical design of both interfaces supports player success by mapping hardware positions and colors directly to visuals on the screen. The red, yellow, blue and green drum pads correspond to red, yellow, blue, and green "notes" on the screen, and are spatially laid-out to correspond with the screen based feedback [Figure 3]. Ludic success with the drums is a function of hitting the correct pad when prompted by the screen interface. The guitar is similarly designed, however, ludic success requires the player to select the right "chord button" with her left hand, and trigger it at the appropriate time by activating

[1] While all three controllers share the same game narrative—one of playing rock music—it should be noted that the play itself across each of the three "instruments" is highly individuated; the vocalist and the drummer are playing very different games, even within the same shared experience. As such, it seems fair to treat each of these interactions as three distinct games that happen to overlap the same content domain. This being said, the game mechanics of the guitar and the drum controllers are more similar to each-other than they are to the microphone controller.

Figure 3: The software and hardware interface in action.

the "strum" button with her right hand. Both interfaces, however, are more problematic than the microphone, in that both require a set of *interface specific learned skills* in order to succeed at the game. It might be argued that the guitar is slightly more problematic, due the increased complexity of the interaction. It is of particular interest to note that each of the interfaces in Rock Band finds a different point of *ludic optimization* on the spectrum between simplistic and problematic. In evaluating their relative efficiencies, we do not apply any value judgment to the interfaces; we simply observe how the various solutions across all three interfaces yield three distinct play experiences. The source of challenge in these experiences shifts progressively away from manipulation of an abstract interface device, and closer to the performance of a literal activity as the interface becomes more *simplistic.* We return to this notion of literal and abstract interactions in Section 4.

2.2 Granularity

Granularity is another ludic interface dimension we find useful, particularly in relationship to predictability. Interface granularity in Rock Band is primarily related to the game's *temporal resolution.*

Temporal resolution is the number of times that the system samples the interface state within a given period of time. We see a connection between the resolution of the system and perceptual capabilities of the player. We posit that systems whose granularity exceeds a resolution which the player can perceive may be experienced as unpredictable, while systems whose granularity is below a "perceptual threshold" for a given player may be experienced as unresponsive or "clunky". One common complaint that players have in games is that there is a "lag" between their actions and those occurring on screen. This is a case of low temporal resolution interfering with the game experience.

In the case of all three of the interface devices provided with Rock Band, input is measured at a very high temporal resolution. Vocal input appears to be continuous, and only the physical limitations of the button pushing appear to restrict the granularity of the drum and guitar controllers. The high temporal granularity of the system introduces some challenges of *predictability*, in that the controllers are able to measure nuances not immediately apparent to the players.

An example of this is the vocal controller at the higher difficulty settings. Unlike the other controllers, which are modeling discreet and binary controller states ("Did the player push the right button at the right time?"), the vocal controller measures a continuous analog audio signal. The vocal system must assess whether or not the player is singing at the right time, whether or not she is singing the correct pitch, and in the case of atonal spoken word parts, whether she is enunciating the correct phonemes. While it is possible to scale the difficulty of the songs on the other controllers by simply introducing more complex patterns of button pushing, in order to make the vocals more difficult the game must increase the temporal resolution, narrowing the zone within which it considers a given audio input correct. At the highest levels this requires a degree of control over vocal intonation that is not perceptible to most players, and which is difficult to consistently reproduce. In this case the high granularity of the system, while predictable from the perspective of the system's technical design, is experienced as unpredictable when experienced by any but the most trained vocal performers.

3. Kinesthetic Interactions

Beyond concerns of winning and losing, there is a physical pleasure associated with engaging in Rock Band with the entire body. Traditional video game interactions take the player outside of an awareness of the body. A player using a control pad is often only aware of her body when something happens to disrupt the experience, such as when she contracts the repetitive stress injury known colloquially as "Gamer's Thumb". By contrast, the interface of Rock Band leverages the player's sense of bodily perception, and incorporates it within the play experience in a meaningful way. In this section we will look at this notion of embodiment in some detail, and consider how systems for evaluating physical gesture and movement in other domains, such as theater and dance, may be used to understand the interactions in Rock Band.

3.1 Embodiment

Embodiment is a widely used term which commonly refers to an awareness of one's bodily condition, or presence. Rohrer terms the notion of bodily awareness *phenomenological embodiment,* and also describes a form of embodiment that arises from the particular *socio-cultural situation* in which it is located [Rohrer 2007]. We connect this to notions of context and the concept of *affordance.*

3.1.1 Phenomenological Embodiment

Phenomenological embodiment refers to our own conscious awareness of the role of our bodies in shaping our identities and experiences [Rohrer 2007]. Rock Band leverages this awareness both implicitly and explicitly. The guitar controller contains a "tilt sensor" which responds when the guitar neck is rocked upwards. This interaction requires the player to drop one shoulder down, while raising her fretting hand upward. This twists the player's torso and shifts her center of gravity. The result of this simple interface action is that the player "rocks" her body into a stereotypical guitar pose, assuming the postures of hundreds of lead guitar players throughout the ages. In doing so, the player activates a special ability in the game called "Overdrive" which results in a temporarily higher score, as well as a more enthusiastic performance from the player's virtual on-screen counterpart. For drummers, the interface relies on continuous and explicit bodily awareness in order to succeed, due to the demands on the player's coordination of arm and foot motions.

In addition to the explicitly embodied interactions found in the design of the interface, there is an implicit focus on the player's body that emerges from the game's emphasis on musical

performance. While it is possible to play Rock Band as a purely strategic exercise, one of the central pleasures of the game is bodily engaging in the fantasy of being a Rock Star. This often means moving in rhythm to the music, dancing with the microphone, "pogo-ing" up and down while playing guitar, twirling the drum sticks, and generally performing the role of a rock star. All of these behaviors, while not explicitly required in order to succeed at the game, are afforded and encouraged by the design of the interface and an implicit aspect of the bodily experience of the game. As we will see in the Section 4, there is a close relationship between these embodied activities and the game narrative.

3.1.2 Embodied Context

Another facet of embodiment that is at work in Rock Band illustrates Rohrer's notion of the *socio-cultural situation*. In his survey of embodiment he describes how "habitual interactions with the environment" shape our "embodied mind", writing:

> "The experiential worlds with which we interact are more than simply physical; we are born into social and cultural milieus which transcend our individual bodies in time. Tools are an excellent example of the elements of our physical world that come to us already shaped by socio-cultural forces which predate each individual's body, if not the human body in general."[Rohrer 2007]

Tools are shaped by an exchange of the affordances of the task to which the tool is to be applied and the affordances of the body which will be using the tool. This notion of affordances is one which was first described in 1977 by J.J. Gibson [Gibson 1977], and articulated in terms of design more recently by Donald Norman [Norman 1988]. Norman describes the relationship between tools and the body in terms of a three part conceptual model comprised of the notions of *affordance, constraints* and *mappings:*

> "Consider a pair of scissors: even if you have never seen or used them before, you can see that the number of possible actions is limited. The holes are clearly there to put something into and the only logical things that will fit are fingers. The holes are affordances: they allow the fingers to be inserted. The sizes of the holes provide *constraints* to limit the possible number of fingers: the big hole suggests several finger, the small hole only one. The mapping between holes and fingers – the set of possible operations – is suggested and constrained by the holes." [Norman 1988].

Similarly, the different interfaces of Rock Band provide distinct affordances, constraints, and mappings. The microphone affords being held in the hand, and constrains game interactions to vocalizations (and occasionally hitting the mic with the palm of the hand to simulate hitting a tambourine, or hand clapping.) The mapping between these two presents no ambiguity about how the microphone is to be used, but this conceptual model only works due to the pre-existing socio-cultural models that have shaped the evolution of the microphone into a recognizable object. Similarly, the guitar controller is the sum of its affordances and constraints, resulting in a conceptual model of use that invokes many of the bodily experiences of actual guitar playing. The drum controller has some of the clearest constraints in that it must be played sitting down, using both hands and feet.

Beyond the conceptual models of the interface as tools, there are even broader contextual implications embedded within the Rock Band interface. Embodied interaction theorist Paul Dourish regards context as arising from a social situation, asking "how and why, in the course of their interactions, do people achieve and maintain a mutual understanding of the context for their actions?"[Dourish 2004] This notion of context as socially constructed provides us with a useful perspective for understanding the setting in which Rock Band's interactions occur.

Although it is possible to enjoy the game alone, there is an implicit social experience built into the design of Rock Band's gameplay and narrative. Unlike Guitar Hero, which glorifies the lead guitar player as an icon, Rock Band pays homage to the band as a collaborative vehicle. When one player fails at a song, another player is able to "rescue" him. The ensemble nature of the controllers reinforces this, as does the narrative framing of the game experience. Rock Band also remediates the rock concert experience from both sides, creating an environment that allows non-players to engage as spectators, and encourages players to *perform* their parts for their audience, as well as for the game score. This social setting draws the game experience out of what is happening on the screen, and causes it to infuse the relationships of all of the bodies in the room.

3.2 Viewpoints

Theater and dance have long been concerned with the ways in which meaning emerges from the movements and positions of the human body. *Viewpoints* is a framework from dance and theater designed to interrogate and design bodily actions for performative purposes [Bogart and Landau 2005]. These same tools may be brought to bear on the kinesthetic experiences that arise as a result of the affordances of gestural and embodied game interface devices. The viewpoints framework is split into two categories – space and time – containing nine distinct viewpoints. The Viewpoints of Time include: Tempo, Duration, Kinesthetic Response, and Repetition. The Viewpoints of Space include: Shape, Gesture, Architecture, Spatial Relationship, and Topography. We will not explore all nine of these viewpoints in depth in this paper, but we will consider several of them in context of the interactions afforded by Rock Band.

3.2.1 Viewpoints of Time

Rock Band belongs to a genre commonly known as "rhythm games" and so the viewpoints of *tempo*, *duration,* and *repetition* are central to the ludic play of the game. In Viewpoints, there are two distinct forms of repetition: *internal* and *external*. *Internal repetition* is used to describe the repetition of movements within the human body, such as heartbeat or breathing pattern; *external repetition* is used to refer to repetition of movements and tempos outside the body [Bogart and Landau 2005]. These are most apparent in the play afforded by the Drum controller, which requires both *internal repetition* (players must create and maintain an internal model of the song's rhythm) and *external repetition* (players must follow visual and auditory clues on the screen) in order to succeed. *Kinesthetic response* is defined as "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."[Bogart and Landau 2005] This is observable across all Rock Band interactions as players move their bodies in response to the music, but is most afforded by the vocal interface, which leaves the player free to move her body without sacrificing ludic success. In our own play we have noticed a correlation between dramatic performance, ludic success and kinesthetic response: as we have grown more skilled in the ludic aspects of the game we are more likely to "rock out" narratively and kinesthetically.

3.2.2 Viewpoints of Space

Shape is "the contour or outline the body makes in space."[Bogart and Landau 2005] In regards to game interface we can consider shape as the posture which the device places the player in. This corresponds closely to Norman's notion of *constraints*. *Gesture* is "shape with a beginning, middle and end" [Bogart and Landau 2005]. Norman's notion of mapping includes a "set of possible operations", which closely corresponds to this notion of gesture. In our discussion of the embodied affordances of Rock Band's interface we have already considered how the design of the devices places the player into a particular *shape* and affords specific *gestures*. As we will discuss in the section 4, these embodied interactions invoke a narrative of rock and roll performance that goes beyond the constraints of ludic play. This means that even if the player is playing with the minimum amount of kinesthetic attachment to the experience, the interface has placed her bodily into the role of the Guitar player through its physical constraints and ludic requirements. The remaining three viewpoints of space—*Architecture, Spatial Relationship,* and *Topography*—deal with the context in which the play occurs. *Architecture* deals with the design of the physical environment in which the performance occurs; *spatial relationship* addresses the relationship between the bodies of the performers, and *topography* describes the pattern described by a performer's movement through space [Bogart and Landau 2005]. While the experience of Rock Band is constrained somewhat by architectural and spatial factors, it does not engage the player in her environment as much as it might. Specifically, the game affords little awareness of the *topography* or stage picture painted through play. However, it is possible to alter the context of play in ways that draw on these three viewpoints. It is not uncommon to see performances of Rock Band at gaming conventions and expos, and in these cases the a*rchitecture, spatial relationship,* and *topography* of the play become significant aspects of the game's appeal. By placing the game on a stage with lights and effects and by encouraging the players to be aware of their role as performers for a real audience, it is possible to transform Rock Band from a game about rock music to a *simulation* of a rock concert. Even inside more humble settings, this phenomenon can be observed, with players rocking out for friends and family at social gatherings, as described in section 3.1.2.

As a final observation on the place of the kinesthetic in our overall evaluation of Rock Band, it is interesting that the temporal kinesthetic perspective correlates strongly with the ludic. This is not surprising for a game where the winning and losing are dependent on timing. The spatial kinesthetic perspective, on the other hand, correlates more with the narrative - for reasons which our discussion in the next section should make clear.

4. Narrative Interactions

The role of narrative and story within the larger game experience is an ongoing, and sometimes controversial, issue in the area of games studies. Luckily, the overblown conflict between "narratologists" and "ludologists" has subsided, and researchers are concentrating their efforts on developing a more focused understanding of how narrative can inflect and enhance the experience of a game. Bizzocchi maintains that an emphasis on the overall narrative arc, although critical to the understanding of story in media such as film or the novel, is of limited utility in understanding interactive media such as electronic games. He argues that the examination of "narrative parameters" such as storyworld, character, and emotion is more useful in this regard [Bizzocchi 2007]. Rock Band is not an experience with the tightly controlled narrative arc of a film, but there is a looser central narrative in the progression of the player's character from unknown musician to rock superstar. As the band moves through the game, the player unlocks progressively higher-profile concert venues, and is awarded with the trappings of rock and roll success: cooler clothes and cooler instruments. The game invites players to live out the positive aspects of the rock'n roll fantasy, without invoking the darker side that is often a part of such rise-to-stardom narratives.

One of the parameters Bizzocchi cites is the "narrativized interface", which involves the incorporation of narrative elements and sensibilities within the design of the game interface [Bizzocchi and Woodbury 2003]. We have already seen how the embodied aspects of the interface can kinesthetically invoke certain aspects of narrative experience. In this section we will take a more focused and explicit look at the role of narrative within the game's interface.

4.1 The Narrativized Interface: Iconic and Functional

Bizzocchi identified two broad strategies for narrativizing the interface. The first is modifying the look of the interface to reflect narrative themes such as character. The second is to modify the functionality of the interface so that the player's interactions were shaped and channeled in ways that expressed and reinforced narrative concepts [Bizzocchi 2003]. Lin extended this look at the narrativized interface, starting with a formal classification of interface structure, and then identifying six distinct strategies used by game designers to incorporate narrative within the design and the experience of the interface [Lin 2007] .

Lin and Bizzocchi agree that two possibilities for interface narrativization are the look of the interface ("iconic" narrativity) and the actual operation of the interface ("functional" narrativity). Both narrativization strategies can be seen in Rock Band, with differences in emphasis depending on the instrument. For example, the drum and the guitar interfaces are both narrativized at the iconic level - that is, each of these interfaces resemble the instruments they represent. However, because the interactions with the Rock Band drums more directly resemble the interactions with an actual drum kit, the drums represent a greater degree of functional narrativisation than the guitar.

4.2 Metaphorical vs. Literal Interactions

Functional narrativization of an embodied and gestural interface can be situated on an axis from the literal to the metaphoric. A more literal functional narrativization within an embodied interface will afford and channel physical interactions that more directly reproduce the movement, gesture and functionality of its real world referent. A more metaphoric functional narrativization will afford and channel physical interactions that evoke (rather than reproduce) a real world referent. [Figure 4]

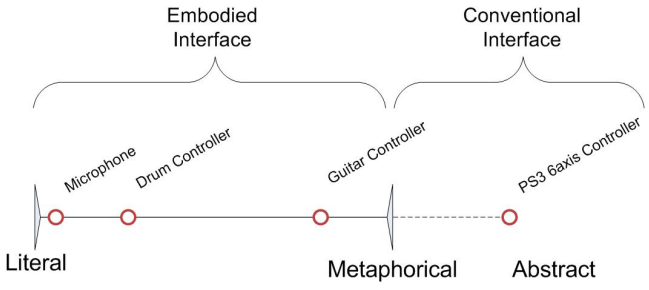

Figure 4: Metaphorical & Literal Functional Narrativization

The drum controller more accurately affords the functionality of its reference instrument than the guitar controller, so the latter is more metaphorical and the former is more literal. However, the microphone voice controller is even more literal than either. The player performing the vocal part of the song does not play at singing the music, but must instead actually sing the song in order to succeed. Unlike the other interactions in Rock Band, which represent points embedded within this continuum, the vocal performances require that players engage in the represented activity completely and authentically. This is a rare example of an almost completely literal narrativized interface design.

In contrast to all three of the Rock Band interface devices, the conventional generic input device, such as the PS3 6-axis controller, affords little if any narrativized gestural or embodied interaction. These standardized controllers fall off the metaphoric end of the scale—any relationship they have to narrativized action is purely abstract. [see Figure 4]

Metaphorically or literally, each of the three Rock Band interfaces offers the user a narratively-rich channel for interactive experience. This robust narrative connection is one of the most potent attributes of the interface and the game. Each individual interface allows us to *play* and to *be* a character whose role we understand, admire, and covet—a rock and roll performer. The combination allows us to enter a richly simulated storyworld in the socially rewarding context of a live band. Narrative and gameplay are united in a mutually reinforcing act of performance.

5. Conclusions

In this paper we have presented three separate dimensions for the analysis of interface in the game Rock Band: the Ludic, the Kinesthetic, and the Narrative. While these separate dimensions have allowed us to explicitly address specific phenomenon within the game experience, we do not claim that these vectors of analysis are truly isolated from each other. Over the course of this analysis we have encountered a number of issues that have manifested across all three vectors. We view the places where these categories overlap and merge with each other as indicators of larger, more universal phenomena. These junctures throw into relief the broader issues at play within the realm of embodied and gestural game interactions. We see connections across our categories in the dynamics of *immersion* within the game. Another common theme that runs through our analysis is the understanding of how embodied interfaces leverage a player's *prior knowledge*. Finally, we would like to discuss the relationship between the concept of *transparent interface* and our three dimensions.

5.1 Three Types of Immersion

Immersion is often positioned as the "holy grail" of interactive experience. It has been described as "the sensation of being surrounded by a completely other reality […] that takes over all of our attention, our whole perceptual apparatus" [Murray 1997]. As a phenomenon, immersion is complex and difficult to quantify. Ermi and Mäyrä parse immersion into three distinct categories: imaginative immersion, challenge-based immersion, and sensory immersion [Ermi and Mäyrä 2005].

Ermi and Mäyrä's conception of immersion echoes the formulations of earlier media theorists. Their "imaginative immersion" corresponds to Coleridge's notion of the "willing suspension of disbelief" [Coleridge 1952 c1906]. Murray extends Coleridge's notion and suggests that in interactive environments the interactor "actively creates belief" in the fictional world [Murray 1997]. Ermi and Mäyrä explicitly ground their "challenge-based immersion" in Csikszentmihalyi's notion of Flow, which we discussed in section 2 [Csikszentmihalyi 1990]. "Sensory immersion" is consistent with McLuhan's claim that there is a strong connection between the human sensorium and the mediated experience [McLuhan 1998].

We see a strong correlation between our three dimensions of interaction and Ermi and Mäyrä's three-fold conception of immersion. In exploring the operations of the interface at the ludic level, we described several design metrics which were rooted in dynamics of challenge and flow. In considering the ways in which narrative played out at the interface level, we saw several techniques that supported the imagined experience of character and storyworld. Finally, in our discussion of the kinesthetic aspects of the Rock Band interface, we described the manner in which the player is immersed in the sensory and embodied experience of the game.

The effectiveness of the Rock Band interfaces lies in the fact that they support immersion across all three of Ermi and Mäyrä's directions. By optimizing the ludic efficiency, extending the game sensorium into the bodily context of the player, and transforming the experience into an embodied representation of narrative these interfaces provide a complex gateway for multivalent immersion.

5.2 Prior Knowledge

One widely held belief about embodied game controllers is that they are more "intuitive", or more "transparent", than a traditional game controller. We will unpack the notion of transparency in further detail below. For the moment we'd like to propose that one reason for the pervasiveness of this belief is due to the extent to which the interface leverages *prior knowledge* and experience into the gameplay, and consequentially how the interface allows successful gameplay to be transposed back into real world capabilities.

In Rock Band, we find a clear example of this in the social conventions surrounding the use of the microphone controller. When we introduced the game at a social function, the first thing that became apparent was that the person with the microphone was responsible for selecting the song for the group to perform. This was due to a number of observable factors, but the most significant was that the vocalist needed familiarity with the song in order to succeed at the game, whereas the guitar player and drummer did not. It was also observed that the vocal controller quickly became the subject of the most social pressure, due in part to its reliance on actual singing ability. In Rock Band, a skilled vocalist who is familiar with the songs is able to succeed much more immediately than an inexperienced or unskilled vocalist.

In this sense, the interface allows the player to translate prior knowledge directly into ludic success. Conversely, we suspect that learning how to play the game better could result in the player attaining a greater degree of "real world" vocal skill. To a certain extent, the narrative of the lead singer as band "front-man" is invoked by the simple fact that the vocal part relies on a greater degree of *prior knowledge.* This is congruent with the idea that the vocal controller is more narratively *literal* than the other interface devices.

By comparison, we see a far more tenuous connection between the act of playing guitar, and the interactions afforded by the guitar controller in Rock Band. The relationship between prior guitar playing experience and success at playing the guitar controller is in fact the subject of some debate. Arsenault describes the "simulational fidelity" of Guitar Hero's (and by extension Rock Band's) guitar controller as favoring *breadth* over *depth*. He describes in detail the ways in which the guitar

controller fails to simulate activities required by an actual guitar, concluding that the controller broadly simulates the core tasks of harmony, melody, and rhythm, but in an abstracted sense [Arsenault 2008]. MTV performed an informal study in which it asked a band of professional rock musicians to play their corresponding instruments in order to explore contentions that musical skill could translate into successful play [John 2007]. The results suggested that actual guitar ability confounded success in the game.

> "There's only five buttons on this guitar, where on the real guitar there's tons of frets. So you have to really pay attention to where all the dots are going and what you're supposed to be playing than how the song [actually goes]."[John 2007]

While this is only a single case, it does indicate that understanding how the real-world playing of the song "actually goes" can interfere with success in the game-world playing of the song. This supports our notion that the guitar interaction in Rock Band is *problematized* by prior knowledge, rather than *simplified* by it. Unlike the vocal controller, the operation of the guitar controller is a specialized and learned skill that does not map directly to any prior life experience and so while the design of the hardware interface supports success in the game, it also requires an extensive learning period in order to be mastered. By extension, it is fair to assume that skill acquired through the use of the guitar controller does not translate back to skill at playing an actual guitar.

One common assertion within the gaming community is that drumming in Rock Band can train players to drum in real life. In the aforementioned MTV study, this was at least partially borne out, as is evidenced in the comments of the drummer playing Rock Band Drums:

> "He also thought playing the "Rock Band" drum kit was more authentic to real drumming when it was played on expert mode. "The easier [the song is in the game], it takes away what's actually in the song," he explained. "But if you actually know how the drums go in your head and you're reading the notes [on the screen], it doesn't really make sense in the tune. I guess the harder [the difficulty] is, the more real the song is [to actually playing it]." [John 2007]

It is especially notable that the drummer claimed to find the "difficult" level songs to closely resemble actual drumming. It is evident from this comment that the higher difficulty setting allowed him to leverage his own prior knowledge into a higher score.

The drum controller's interaction paradigm draws on the skills needed by actual drummers (keeping time, maintaining multiple beats simultaneously, coordinating hands and feet), but operates via a simplified abstracted interface. The perception of the drum controller as a "bridge" between the ludic activity and the actual activity is a valid one. Much like the guitar controller, the drum set consists of a set of color-coded buttons, which correspond to a stream of color-coded instructions that scroll toward the player on the screen. Like the guitar controller, mastery of the drum kit requires extensive practice; however, like the voice controller it is possible for players with prior drumming experience to bypass a large portion of the learning curve. It seems reasonable to suggest that this learning curve operates both ways, although more research would be needed to confirm that Rock Band could act as a training aid for real drumming.

When someone says an interface is intuitive, we suggest that what they are actually talking about is the extent to which the interface leverages *prior knowledge,* however, the discussion above includes an implicit warning against interfaces that bear a *resemblance* to an actual activity, but whose operation may be confounded by *prior knowledge*. In the next section we will look at the similar, but distinct notion of *transparency* in interface.

5.3 Transparency

One common claim made about gestural and embodied interfaces is that they are *transparent*. Like claims that an interface is *intuitive,* it also operates on a relationship between the expectations of the user and operations of the interface. For the most part, this notion of transparency in gestural and embodied interfaces has remained unexplored. In this section we will discuss some common theories of transparency from HCI and explore how they can be understood in terms of our three dimensions.

Transparency is a condition that is commonly associated with the idea of the user's awareness of the interface disappearing, resulting in an "unmediated" relationship between the interactor and the activity being performed, as in Bolter and Grusin's notion of *transparent immediacy* [Bolter and Grusin 1999]. Research on transparency in interaction suggests that transparency is a contingent phenomenon, emerging from a context of use. Bardram and Bertelsen write that first "transparent interaction is not a property of the interface by itself, but a quality of the use activity; and second, that transparent interaction is developed by the user during interaction."[Bardram and Bertelsen 1995] They use Activity Theory to analyze the conditions which support the development of transparent interaction, concluding that it is impossible to identify absolute "transparent features" which need to be included in any interaction. However, they do recommend possible options which may help when designing for transparency. One is the notion of supporting *development in use*: the design of interactions that automatically train the interactor in the intended purposes of the system. Another is the notion of *initial familiarity*, which draws on metaphors from the users' prior experience in order to initiate deeper exploration and learning of the system [Bardram and Bertelsen 1995]. These conditions suggest that a transparent interface is one which affords the type of interactions required for success within the system, while leveraging the prior experience of the user in order to bridge the gap between her existing capabilities and those needed to succeed within the system. For these reasons, claims about interfaces that are intuitive are bound up in issues of *prior knowledge* and *transparency.*

It is possible to see the conditions for transparency throughout our earlier analysis of Rock Band. We argued that the vocal interface leverages a player's prior knowledge into success in the game, and supports learning real world skills through the acquisition of game skills. In this sense the vocal controller for Rock Band is highly transparent in that is draws on *initial familiarity*, while the highly specific design of the controller affords the single interaction needed for ludic success: vocalization. The same can be said of the drum controller. The guitar does not draw on *initial familiarity* at the same level; however it ably supports *development in use,* by *affording* the interactions needed to succeed, *constraining* the ways in which the interactor uses it to those which are ludically appropriate and narratively salient, and ultimately *mapping* to a set of operations that support the learning of the interface. At the same time, the guitar controller also shows how *prior knowledge* can interfere with the learning of an interface's operation if that knowledge fails to properly align with the interactions actually afforded by that interface.

6. Future Work

We have presented three different perspectives for the analysis of embodied and gestural interface devices, and demonstrated their utility through a close analysis of the game Rock Band. We have connected these perspectives to larger themes that we believe further inform the understanding of the experiences afforded by these devices. The next phase of our research is to determine the reach of our three perspectives in the analysis of a wider range of embodied and gestural game interface devices. A related question is whether these analytical tools have further utility as *design* guidelines. We hope this is the case, because we feel that these are important questions. We believe that embodied interface devices such as those used in Rock Band and the Wii will continue to grow in popularity and in their cumulative impact on user experience. The true test of our scholarship will be in its ability to help in both the analysis and in the evolution of this new generation of interface devices.

7. Acknowledgements

This research has been undertaken as part of the Creativity Assistive Tools for Games Project, funded by Canadian Heritage.

Additional financial support has been provided by the Province of British Columbia through the Ministry of Advanced Education in the form of a Pacific Century Graduate Scholarship.

8. References

ARSENAULT D. Guitar Hero: "Not like playing guitar at all"? *Loading... 1* (2).

BARDRAM J.E. AND BERTELSEN, O.W. Supporting the development of transparent interaction. in *Human-Computer Interaction*, Springer Berlin / Heidelberg, 1995, 79-90.

BIZZOCCHI J. *Ceremony of Innocence* and the Subversion of Interface: Cursor Transformation as a Narrative Device *DAC 2003 : Digital Arts and Culture*, Melbourne, Australia, 2003.

BIZZOCCHI J. Games and Narrative: An Analytical Framework. *Loading - The Journal of the Canadian Games Studies Association*, *1* (1). 5-10.

BIZZOCCHI J. AND WOODBURY, R. A Case Study in the Design of Interactive Narrative: the Subersion of the Interface. *Simulation and Gamin*, *34* (4).

BOGART A. AND LANDAU, T. *The Viewpoints Book: A Practical Guide to Viewpoints and Composition*. Theatre Communications Group, New York, 2005.

BOLTER J.D. AND GRUSIN, R. *Immediacy, Hypermediacy, and Remediaton*. The MIT PRess, Cambridge, Mass, USA, 1999.

Coleridge S.T. *Biographia Literaria*. E.P. Dutton, London & New York, 1952 c1906.

CSIKSZENTMIHALYI M. *Flow: The Psychology of Optimal Experience*. Harper Perennial, New York, 1990.

DOURISH P. What we talk about when we talk about context. *Personal and Ubiquitous Computing*, *8* (1). 19-30.

ERMI L. AND MÄYRÄ, F., Fundamental Components of the Gameplay Experience: Analysing Immersion. in *Digital Games Research Association (DiGRA) 2005*, (Vancouver, British Columbia, Canada, 2005), Digital Games Research Association.

GIBSON J.J. The Theory of Affordances. in Shaw, R.E. and Bransford, J. eds. *Perceiving, Acting and Knowing*, Lawrence Erlbaum Associates, Hillsdale, N.J., 1977.

JOHN T. Is "Rock Band" Drumming Just Like The Real Thing? Cartel Tests The Theory, MTV Multiplayer, 2007.http://multiplayerblog.mtv.com/2007/10/17/is-rock-band-drumming-just-like-the-real-thing-cartel-tests-the-theory/

LIN B. Narrative Interface Design: The Use of Interface Elements to Enhance the Narrative Experience in Videogames *School of Interactive Arts and Technology*, Simon Fraser University, Surrey, BC, 2007, 152.

MCLUHAN M. The Playboy Interview. in McLuhan, E. and Zingrone, F. eds. *The Essential McLuhan*, Harper Collins, New York, New York, 1998, 233-269.

MURRAY J. *Hamlet on the Holodeck: the future of narrative in cyberspace*. The MIT Press, Cambridge, Massachusetts, 1997.

NORMAN D.A. *The Design of Everyday Things*. Doubelday/Currency, New York, New York, 1988.

ROHRER T. The Body in Space: Dimensions of Embodiment. *Body, Language and Mind*, *2*.