

SECTION 6

.....
**THE PRACTITIONER'S
POINT OF VIEW**
.....



CHAPTER 28

LET'S MIX IT UP

Interviews Exploring the Practical and Technical Challenges of Interactive Mixing in Games

HELEN MITCHELL

GAME audio has come a long way since the simple electronic beeps of the early 1970s, when significant technical constraints governed the scope of creative possibilities. Recent years have witnessed technological advancements on an unprecedented scale; no sooner is one technology introduced than it is superseded by another, boasting a range of new refinements and enhanced performance. Those working creatively within the sphere of audio for games have been faced with a growing array of creative possibilities resulting from significant technological advancements and changing game platforms, fuelled by increased market demand. While these are all positive and exciting developments, they are driven by a rate of change that allows little time for the absorption and exploration of these developments and the new creative opportunities that they offer.

Certain parallels might be drawn between this history and the early days of sound film. The development of film was also driven by the introduction and progress of technology and by market pressures (Lastra 2000). Following the success of the Warner Brothers film *The Jazz Singer* (1927), the film studios realized that synchronized sound would be central to the future of the film industry. During subsequent years the studios sought to answer box office demand for sound films by producing films that showcased the new technology, such as musicals. However, to remain competitive the studios had to find more subtle ways of exploring the possibilities of sound film while maintaining tight production schedules. Consequently, films became vehicles for the creative exploration of new and emergent technologies, resulting in a fast succession of film firsts. Some of these film firsts proved more successful than others and would be imitated in future films; less successful approaches were discarded. Over many years of trial and error, a sophisticated and standardized approach to sound for film emerged.

In the same way that early sound films explored new creative possibilities, modern videogames offer the potential to explore new creative approaches as technical constraints are gradually eroded. There is one essential difference however; film sound

emerged with few points of reference other than radio and stage performance and so had to forge a new path: games, on the other hand, emerged into a world familiar with a variety of audiovisual media. It is not surprising therefore, that film sound should become just one point of reference for the game industry, particularly since the tastes and expectations of potential gamers and game developers have been shaped by current conventions of sound for film and television. Game sound designer Jerry Ibbotson, for instance, states that he has often been asked for specific film sound styles in his work for game developers, and traces this trend to the formative influence of film and television from an early age:

Games have definitely been influenced by film...it's subconscious because you've grown up watching films in the cinema and on television; you know what things are supposed to sound like. You know what guns are supposed to sound like from watching films... [but] they don't actually sound like that in real life (we're based opposite a firing range and guns don't sound like that at all)... Everything you think you know about things sounding is based on film, so even subconsciously it's carried through into what you do...the designer and the producer will tell you how they want it to sound, and nine times out of ten...they will use sound analogies like, "can you make it sound like the drop-ship from *Alien*" or "something from *The Fast and the Furious*" [or] "that has to sound Hollywood"... What they mean is that it's got to be like... watching a film, so it runs through everything.

The desire to achieve the Hollywood sound as expressed by some game developers is reflected in much of the literature relating to game audio today. Even the briefest foray into game sound literature (academic and practice based) soon reveals references to film sound and to examples of best practice in particular. Without question the film industry's experience of sound, acquired over many years, represents an invaluable resource within various spheres of creative endeavor, but within the context of game audio there is one essential flaw in the film sound model: linearity.

Dynamic mixing, defined as the "real-time manipulation of audio replay parameters and inline effects treatments for organizational, technical, or creative purposes" (Broomhall 2010) distinguishes game sound from film sound. With the exception of some linear sequences such as cut scenes, sound in games is dynamic and changes in response to the interaction of the player(s). By contrast, film sound is mixed to footage, which once the final cut has been made will not change. The linearity of film sound allows tremendous creative and technical control over every aspect of the final mix. Such levels of control become quite challenging within the context of dynamic mixing, in which the mixing console is replaced by computer code that governs the ways in which sounds are triggered and controls changing audio parameters in real time. This chapter seeks to explore some of the practical and creative challenges of dynamic mixing for videogames, focusing on sound effects in particular. The discussion will be based on a series of interviews by the author and Jamie Maxwell in 2010 with leading audio specialists who are actively involved in the creation and standardization of audio for games.

28.1 DEFINING A “GOOD” GAME MIX

Before outlining some of the challenges of dynamic mixing for games, it is necessary to consider what is generally meant by “a good game mix” in the industry. Nick Peck defines this as “a mix in which the appropriate layers of material are the ones which are the most audible to the listener at the particular time that they need to be,” and suggests that this definition is equally applicable to film and television mixes. In refining the definition still further it is possible to identify three general classifications of sound that are present in most game and film mixes: music, dialog, and sound effects. These three basic categories may be further deconstructed into subdivisions that give some indication of the ways in which audio functions and is applied. Peck identifies five categories of sound:¹ ambience or background sounds, namely, “location-based material that is used to set the scene and give a context to the narrative”; Foley sounds, defined as “footsteps, cloth rustling, prop manipulation, and other sounds made by people”; hard sound effects, or “primary sounds that are often signature elements of the game”; music, which conveys the emotional content of the material, telling the viewer “what to feel”; and finally, “the most important layer,” dialog, the function of which is to “tell the story.” Of paramount importance is the ability to hear and understand what people are saying, although, as John Broomhall points out, one should not discount the “storytelling power of sound.”

Having outlined the individual layers making up the mix, it is necessary to consider what is meant by the word “appropriate” in this context. This word is the linchpin of the definition and also the most problematic element of it since it suggests a subjective assessment. If dialog is to be heard clearly throughout game play then it would be logical to assume that dialog should always take precedence when balancing different elements within the mix. However, imagine a scenario in which a military commander is shouting instructions to his men during a combat sequence. A huge explosion takes place, deafening all those close to it. In such circumstances it could be argued that it would be more *appropriate* to duck the dialog below the sound of the explosion or to filter out the high frequencies of the dialog to mimic the effects of temporary hearing damage, making the game feel more real and believable but potentially compromising the intelligibility of the dialog. Alternatively, the clarity of the dialog might be preserved by ensuring that the frequency range of the explosion remains well below that of the dialog, possibly dropping the volume of the explosion in relation to the voice as well. This approach would prioritize the dialog at the expense of other sonic elements within the mix. Broomhall points out that some game developers today have taken this approach one stage further by introducing the concept of “critical dialog.” He defines this as “dialog which must be heard come what may” because it provides important information and feedback to the player. He suggests that the technical treatment and handling of “critical dialog” differs from approaches applied to other types of in-game dialog. Perhaps the treatment of dialog in games seems an obvious example, but achieving a good game mix involves making many decisions relating to the prioritization and technical treatment of sound but

on a much greater scale, involving multiple audio layers responding fluidly to interactive game play.

There are, however, some generally accepted and objective hallmarks of a good game mix, hallmarks that might indicate that the right and most “appropriate” decisions have been made. Ibbotson suggests that a good game mix is one that “you shouldn’t really notice.” A mix that brings together individual, high-quality “elements” in a “balanced” way, preventing what might otherwise become “a complete cacophony” while avoiding “irritating” repetitions of individual elements and providing informative “feedback.” A good mix therefore is as much about omission as inclusion. The result however, according to Ibbotson, is a mix no one notices; one that draws the player into the game. In this regard Peck suggests that a good game and film mix share the same goal:

The best soundtrack is the one that completely melts away in the perception of the viewer so that they’re not thinking about the sound; they’re letting the sound wash over them as they are immersing themselves in the experience and they are engaging in the story itself. . . . to that end the aims of a good film mix and a good game mix should really be the same. . . . The aim of both of them is to engage the player, engage the viewer; to take them away from their life for a little while. . . [placing] them into the middle of another story.

John Broomhall makes a further addition to this appraisal of a good game mix, suggesting that there may be moments when it might be appropriate for a game mix to draw attention to itself by highlighting particular sounds that facilitate game play by, for example, helping a player to navigate within the game world or to “understand their next objective,” and so on.

In summarizing the overriding objective of a good mix, we might therefore define this as the creation of a more “interesting and entertaining gaming experience,” one that enables players to “suspend [their] disbelief” as they become immersed and absorbed in game play. The extent to which this is achieved must surely be the quality benchmark for dynamic mixing in games today.

28.2 CREATIVE AND TECHNICAL CHALLENGES: INTEGRATION

For those working toward the achievement of successful game mixes there are many creative and technical challenges. One of the key challenges concerns the integration of sound into the game mix itself. The extent to which sound designers are involved in this process can vary significantly. An in-house sound designer may have considerable input when determining how the finished sounds might be integrated into the mix and processed during game play. Having in-house status provides opportunities to facilitate creative collaboration. In addition, being onsite as the game is developed provides the

potential to become involved in the development of the audio much earlier in the production process. Clearly when sound design work is outsourced, creative collaboration can become more difficult and time constraints can be intensified as production schedules slip; something an in-house sound designer would be better positioned to monitor and potentially counteract.

It should be noted that neither of these models offers standardized approaches in terms of integrating sound into games. In some cases sound designers might be given the opportunity to work with audio programmers and other in-house personnel, directly influencing the ways in which their audio samples are integrated and processed within the game. In others they will simply be asked to deliver audio assets and will have little or no involvement in the integration process. Ibbotson's experience of working on games in which the sound design is outsourced, testifies to these differences of approach:

We're working on a PS3 game at the moment with someone in-house who we ship the samples to;... [he asks] our opinion, but he's the one who is laying... [the audio assets] out. . . . In some projects we've worked on, we've had completely bespoke tools. In *Broken Sword: Angel of Death* we had tools where you could load up every level and move in through the level, placing sound emitters where we wanted. . . . Likewise we have worked on games where we've had a lot of control but then it's been the world builders instead who are placing the sounds and, when you see it, you're banging your head against a wall saying "No, no, no, I want that moving!" They don't necessarily understand because, even though they've placed it, they haven't even had headphones on or played the game. They're just placing little dots where they think a sound should be.

Thankfully Broomhall suggests that the experience Ibbotson describes is not as common as it used to be, although he does acknowledge that it is still largely true of "hand-held" games. On the other hand, Ibbotson's experience of nonstandardized approaches and integration is echoed by that of Peck:

[The] type of integration process you have... to get your sounds working in the game... changes on a game-to-game basis. There are many games that use middleware solutions... in which, if there is enough time, the sound designers can... sculpt and change some of the interactive parameters of the sounds as well as the sounds themselves. However there are many other situations... in which... the sound designer spots the game, figures out what it is they want to happen, and then generate(s) the assets. [The sound designer]... hand(s) off the sound effects to the client who takes them... to an engineer, who wires them up to the game itself. In reality that often happens, particularly when you're operating in a consulting... realm... without very much input from the sound designer after the fact.

Clearly these differences of approach can have a profound effect on the success or failure of the audio mix. Distancing the sound designer from the integration process and from decisions relating to the audio mix and real-time processing are less likely to yield satisfactory results. However, it is not simply a case of collaboration; successful audio

integration hinges upon the expertise of the audio programmers and the dynamics between individuals, whether sound designers, audio programmers, audio directors, audio editors, producers, or game designers, as Ibbotson describes:

We've worked on games with award-winning sound and a large part of that is down to the sound coder. I sometimes think that producers don't realize just how massively important a sound coder is. When you look at really successful games, one of the reasons why they sound so good is because they have dedicated audio programmers who really know how to get stuff to work, and when it works on a project it's normally because they are on the same wavelength as you. They are just as creative as any of the artists or graphics people.

According to Nick Peck, the best audio programmers are musical and have an inherent interest in sound. However, he also suggests that the context in which personnel such as programmers and audio directors work is also central to the success or failure of the finished audio:

There is a need for game designers to be able to have trust in their audio directors; to be able to let them go and do what it is that they think is best for the game. The best game audio programmers that I have worked with are the ones that are also musicians and that love manipulating sound on their own at night when they're not doing it during the day for a job. The communication with them becomes effortless. There's no question... that audio directors often face an uphill battle in terms of being able to educate the... people around them as to the importance of sound... at LucasArts sound was always considered to be of utmost importance because that was the aesthetic that was laid down by George Lucas across his entire company.

Encouraging a creative culture in which audio programmers and other personnel recognize the potential power of sound and value creative exploration as well as the expertise of sound specialists, whether in-house or out of house, offers a number of distinct advantages during the integration process. Integrating sounds into a game mix involves making a number of decisions regarding the prioritization and processing of certain sounds, as outlined earlier. It could be argued that distancing sound specialists from these core decisions limits the scope to use the full potential of sound within games. John Broomhall suggests that, in his experience, most game developers now have in-house sound integrators, even when the creation of some sounds is outsourced. Developments of this type are already helping to resolve some of the issues surrounding the integration of audio into games.

Conversely, one might assume that audio samples might be integrated within the game and mixed or processed according to principles of physics without reference to sound specialists. Individual volume levels might be governed by distance measurements for example. Reverberation times might be set according to typical acoustic principles and absorption rates. Sound-emitting objects traveling at speed might be subject to pitch processing to mimic the Doppler effect. Frequency filtering might be linked to changes in distance, obstruction, occlusion, and so on. Such attempts to recreate the

real sound world in miniature are relatively standard but do not fully address the sonic needs of games today, perhaps because very few games truly represent reality. Nick Peck asserts that “we need to understand that a video game is not the real world,” and goes on to suggest that “while they’re rooted in reality,” this is supplemented with “artistic license,” the object of which is “to heighten the dramatic aspect . . . of the story.” He adds that “it is more important for a game to be able to succeed aesthetically than to succeed realistically.” He illustrates his point by referencing “phenomenal examples of sound design” in films such as *Apocalypse Now*, *The Conversation*, and *Saving Private Ryan*. With reference to game sound, he rightly points out that the emphasis on scientific reality could yield less than entertaining results; for example, “it would be a mighty boring outer space science fiction game if we tried to mimic the real world since there is no sound in a vacuum.”

Just as representations of sounds in space are not scientifically accurate, so it might be argued that the very act of recording sounds and diffusing them between stereo or surround speakers creates an artifice, a removal from the real world, altering the ways in which the human brain perceives those sounds. As anyone who has ever done any recording could testify, the results achieved from the placement of a single microphone in a room tend not to recreate the auditory reality of standing in the same position and simply listening. The effect and filter of human perception may not operate in the same ways when applied to recorded sounds as opposed to real-world sounds. For example, John Broomhall highlights the ability of the human brain to filter out certain sounds, and points to the use of Foley sounds in the game *Operation Flashpoint* as an illustration of one way in which this facility might be recreated in games. The Foley sounds are kept very low within the mix unless there is a dramatic reason for highlighting them; for example, if it becomes important for the game character to move stealthily, the volume of the footsteps increases, adding to the dramatic tension of the game play and focusing the attention of the player more directly on their movement within the game environment. If scientific modeling governs the ways in which sounds are integrated within game play without taking into account the effect of human perception upon those sounds then the model is incomplete and therefore flawed.

28.3 CREATIVE AND TECHNICAL CHALLENGES: NONLITERAL SOUND

If the scientific model is incomplete or inadequate when applied within the unreal realms of fantasy or hyperexaggerated reality, it must be supplemented by something else, namely nonliteral sound. The requirement to consider nonliteral approaches to sound, mixing, and processing makes the involvement of sound designers and other sound specialists desirable much earlier in the development of the game and particularly throughout the integration process. It could be argued that if this approach were

standardized, the potential quality of interactive game audio, and its ability to entertain, manipulate, and immerse players, would be significantly enhanced.

It should be pointed out that an exploration of past and current game titles would reveal many examples of nonliteral sound successfully integrated into dynamic game play. However if one considers the use of nonliteral sound within the context of pioneering films such as *Apocalypse Now*, it seems logical to consider whether similar approaches might be possible when translated from the linear sphere to the nonlinear. For example, in thinking about one function of sound in *Apocalypse Now*, namely the subtle heralding of future events, we might consider to what extent future events might be anticipated or heralded in game play when those events have not yet been determined by the player. One application of this technique in games would be the introduction of dark or menacing sounds in the mix to suggest danger, implying, almost subliminally to the player, that if they were to pursue their current course of action the outcome would not be good. For example, as a player approaches a door, a sound might be added to the mix, or the real-time processing of the existing game sounds might be manipulated, indicating danger behind the door. Alternatively, sounds might be pulled from the mix, as if to focus the attention, creating a sense of apprehension and internalizing the action. To use a real example; in an unreleased game project John Broomhall worked on some years ago, nonliteral sounds and music were linked to game states such as “exploring, enemy not present,” “enemy present not sighted,” “enemy sighted,” and “combat.” When “exploring,” the music and sound combined to create unease, suggesting the enemy might be present; when the enemy was present but not sighted, changes in the music, such as greater use of dissonance and altered instrumentation, suggested greater danger. Additional sounds, such as whispering, were also added to the audio mix to intensify the sense of growing danger and fear. Once the enemy was sighted, the music and sound combined to add even greater tension and drama, anticipating the combat to follow. Interestingly once the combat game state was triggered, all music was pulled from the mix, allowing the focus to transfer to the sounds of combat, effectively shifting the focus of the player from the tension of anticipation to the reality of combat, while also preventing a cacophony of competing sounds. In this way the music and sound combined to “lead the player by the nose to the point of combat,” and once there, the sounds of combat held the player in the moment, right in the heart of the action. Tying the design of the music and sound to specific game states as the dynamic mix responded in real time to changing game conditions helped to herald danger, and direct game play.

In a more recent example, the audio team for the game *Dead Space* created a game object called the “fear emitter,”² which attaches to other game objects, reacting dynamically in response to the location of the avatar. As the avatar approaches the fear emitter, the scaled response of the emitter adjusts the mix and the real-time processing in order to ratchet up the sense of tension and fear. By manipulating the emotional response of the player as they move within the game environment it becomes possible to manipulate and direct the player by heralding possible dangers. The player’s response to these object-focused “fear emitters” is intensified by a dynamically changing RAM-based “creepy ambi patch,”³ designed to recreate something of the player’s internal and

subconscious mind. The sound samples making up this patch may not correspond to a visible onscreen source since they represent the internalized world of the mind, or “in your head sounds.” The use of “fear emitters” and the “creepy ambi patch” also help to create a more engaging gaming experience, one that offers other layers of interactive control and emotional manipulation.

However, such approaches raise other questions, such as to what extent should the mix seek to help the player by warning them of danger, or should such sounds actively mislead players to make game play more challenging? Jerry Ibbotson suggests that the use of sound to mislead or confuse would be appropriate in some games, and highlights *Bioshock* as an example of one that draws near to this approach, singling out the game as significant for its use of sound as “motif” and for its focus on sound design:

If you think of the start of *Bioshock*...you have no idea what's going...in minutes it...runs into double figures before you have any clue what's going on... Sound design is such a bandied-around phrase, often it's just sound effects, but then sometimes it just steps over into something else. The whole game had... [a] motif of sound.

It is interesting to consider to what extent such approaches might work during protracted game play, since players might inevitably become acclimatized to such effects and therefore anaesthetized to them. Answers to such questions will always be subjective, but the early and active involvement of sound specialists in seeking to address these issues, given their knowledge of the scope and potential of sound, should be at the forefront of game sound development today. Unfortunately, all such explorations and innovations take time and within the pressurized environment of game development, time is a rare luxury.

28.4 CREATIVE AND TECHNICAL CHALLENGES: PRODUCTION SCHEDULES

Game production schedules, unlike film schedules, tend not to incorporate an official postproduction stage, although John Broomhall suggests this is now being addressed through the work of people such as Rob Bridgett and Marty O'Donnell. Production schedules for those involved in the creation and integration of game sounds are extremely tight and this can have a significant impact upon the quality of the finished sound and the mix, according to Peck:

It's an invariable aspect of game production that audio always comes at the end of the line. Often what ends up happening is that the people that are upstream from the audio... will continue tweaking and changing their materials past the point where the audio people need to be doing their work, but the deadline, the ship date itself,

doesn't slip on the other end. As a result, the audio people are often pinched... there's a time crunch right at the very end of development when they're trying... to get their work in. So as a result, during the time when the sound designers and audio directors should be mixing their game, they may very well be altering sounds to be able to fit revised animations or fielding change requests from their creative directors and other people on the project; the result being that there's less time to mix than you would ever like... There's no question about the fact that having less resources in terms of less people... [and]... less time to be able to work on the game, means that the mix, which is... the very last step in the production chain of audio, is going to be affected.

Peck asserts that allowing more time for the creation and integration of audio by bringing in audio personnel right at the beginning of game development yields the most successful results:

The games that succeed the most as far as audio goes, are the ones in which audio is brought in at the very beginning of the process just as in film. If audio is left until the end... you will often have a "slapdash," stress[ful]... race... to shovel sound effects in as quickly as possible. Sadly this happens all too often, however there are... many superb exceptions... particularly in the case of some of the larger videogame developers.

Sometimes scheduling pressures are intensified by the availability of key personnel, such as audio programmers; as Ibbotson reflects, "sometimes you work on a game with a coder who's doing many other things and has got Thursday alone to do the sound." Broomhall suggests that situations of this type are less common today but does acknowledge that leaving the mixing to the end of the process is "problematic." He points out that there might be a place for mixing "as the game progresses through production milestones," so avoiding problems of this type. Scheduling pressures compound the creative and technical challenges of audio creation and dynamic mixing. In addition, the challenges outlined so far are also impacted by the fact that the games industry lacks standardization, which of itself presents something of a challenge for audio personnel.

28.5 STANDARDIZATION: TOOLS

Steve Martz likens the games industry to the "Wild, Wild West." It is an interesting analogy, which conjures up images both good and bad (and perhaps ugly) in equal measure. Given the proliferation of game genres and platforms, coupled with the fast pace of technological change, it is easy to see why standardization has not been a priority. For some the lack of standardization provides greater creative flexibility and more opportunities for experimentation when developing and mixing sound. Jerry Ibbotson likens "game sound people" to problem solvers; "we are inherently tinkerers... people in games development are always... trying to find different ways of doing things. Even if

the first method works . . . we'll look at another method." He contrasts this with the more set approaches of colleagues in film and television, but makes no judgment between game and film industry methods.

However, when it comes to mixing, the film model does have real appeal, whether or not it is achievable in a nonlinear context. Peck suggests that "when you can get to the point when you can mix a game [in] the way that you can mix a film, on a digital console, then I will say that we're finally getting somewhere." A point that is echoed by John Broomhall, who argues for accessible technical facilities that enable "the sort of freedom a film mixer/sound editor would have to manage and exploit dramatic [and] narrative . . . opportunities." Goals such as these have encouraged the development of various middleware tools for the development and integration of game sound.

Recent years have seen the move towards greater standardization of these tools as they have been developed by companies such as Audiokinetic (Wwise) and Firelight Technologies (Fmod). Middleware solutions have had a significant and positive impact and are addressing some of the creative and technical challenges facing sound designers and other audio personnel. Nick Peck states that "Wwise and Fmod . . . allow you to be able to spend more time working on fine-tuning your sounds and less time having to . . . shovel the sounds into the game." He goes on to suggest that these tools have "allowed mixes to sound better."

The contribution of middleware tools cannot be overestimated. However, as middleware providers operate in a competitive market place, it is not surprising that different game developers use different middleware tools. Sound designers have to work with different middleware tools and also with custom software tools in some cases. Ibbotson's experience reflects this:

Sometimes coders will design tools for you. For example, we've got vehicle tools; if you want an engine to backfire, it will dip everything else down underneath it, because you haven't got limitless headroom . . . So you've got fairly complicated . . . [things] going on in real time using these tools. The tools used will vary from developer to developer of course; some will just ship software in and other times they'll get things written for that particular project.

If one compares this with the more standardized preference for ProTools software found elsewhere in the film and music industries, the contrast appears quite marked. The use of custom tools would be unimaginable in this context. Even within the context of game audio, the development of custom tools has to be carefully considered, as Michael Kelly points out:

I think designing and supporting custom tools for one title is expensive. To make it cost effective, it's likely you would want to reuse it for many titles. However, even very similar titles (such as sequels) will have different requirements, so rework will be necessary. Further, technology changes so fast that even the most flexible tool will date if it is customized to a particular technology level. My view would be that bespoke technology appears to become unsustainable for anything but the largest

developers. The analogy in the music industry would be if each record label were to write their own mixing software and build their own mixing desks. As technology gets more complicated, I expect we'll see more from off-the-shelf tools which people can easily adapt to their own requirements; plug in their own extensions, or integrate with their own custom tools by exporting and importing files between them.

Adaptation of software to suit individual requirements might offer a move away from custom software to "generalized frameworks for software which are heavily customizable to bespoke requirements," but such a development would require agreement regarding "plug-in frameworks and...standardized data formats so the plug-ins...could communicate with each other." Clearly anything that simplifies the process of creating and integrating game sound would be a step forward, as Peck reflects:

I'm a huge fan of simplicity; I want to put the power into the hands of the creative audio professionals...What is still to be achieved is the notion of making it easier...to be able to get sounds into the game interactive environment quickly and simply and to be able to mould and manipulate those materials with the minimal amount of hassle and number tweaking.

One of the most significant steps taken towards achieving this was made by the MIDI Manufacturers Association in October 2001, when they published plans for an interactive eXtensible Music Format (iXMF):

Interactive XMF is intended to be an open-standard, cross-platform means for audio artists to bundle audio content files with general information and audio implementation instructions. Initially, it will be highly beneficial for the game industry, and since it may be used in any interactive audio application a potential to expand or create markets in other areas exists. This new file format will put artistic control into the hands of the artists, keep programmers from having to make artistic decisions, eliminate rework for porting to new platforms, and reduce production time, cost, and stress. (MIDI Manufacturers Association 2010)

The file format is being developed by the iXMF Working Group (IXWG) of the Interactive Audio Special Interest Group and work is ongoing. Just as MIDI revolutionized music making when it was first introduced, a universally recognized and interactive file format would be likely to have a similar impact and would offer a number of advantages, as Michael Kelly points out:

In terms of more linear audio tools (like ProTools) there's a clear advantage to game developers if pro-audio (linear) tool developers were able to provide data in a format that was game-ready. Currently many developers must implement their own conversion tools, or worse still reinvent very similar tools which can never do the job as well as off-the-shelf tools which can afford investment in advanced interfaces. Of course, linear tools don't directly meet the interactive needs of game studios, but if pro-audio tools have a means to better fit into the game audio pipeline this has to be a good thing for everyone.

Imagine a situation in which sound designers could create and mix their sounds in ProTools before saving their work to a standardized, interactive file format for use with any interactive software, middleware, or hardware, perhaps including an embedded scripting layer holding detailed information for implementation and integration by any game engine. This would give tremendous creative freedom to sound designers and other audio personnel, offering them more direct control over the integration of audio into the mix. Implementing this level of control and access to a scripting layer would represent a challenge, as Kelly describes:

The requirement for a scripting language, and ultimately the tool that makes this visible to the sound designer, is that it shouldn't limit the imagination of the sound designers. One way or another it should allow them to derive mix values from whatever they want, whether this is something high level like the number of players on screen, the weather in the game, or something low level like the current gain value of a compressor on a parallel bus. Similarly it should not limit what they can use to mix, whether it is volume, pitch or center frequency of a filter, or another scripted variable that links to any parameter on any other voice or bus. When you have flexibility like that, the team can decide what mixing system meets their specific game requirements. Such flexibility does come at a cost, but the proportional cost is lower and lower all the time. Ultimately it should be an agreement between the sound designer and the programmer . . . Traditionally it was dictated by the programmer; most likely because it was dictated by the hardware. Since the hardware restrictions are reducing, it is important that tools offer this flexibility and therefore it is important that any scripting language supports that, both for now and in preparation for the future.

If such tools were available, it is likely that the general public would also be interested in exploring the possibilities offered by such standardized interactive formats, something that Kelly recognizes: "I think there is a need for tools the public can get their hands on to create interactive audio; it is not just the games industry who need them." Above all he stresses the need for flexibility: "standardized file formats and standardized connectivity (such as plug-ins systems) between interactive and noninteractive tools mean that the user can choose the right combination of off-the-shelf and bespoke tools for their needs—that is what we lack at present." With reference to plug-ins, Kelly suggests that the range of available plug-ins is limited but adds that third-party plug-ins are gaining some ground in the games industry:

we're beginning to see some use of third party plug-ins in the games industry. It will be interesting to see how it pans out. If the aim is flexibility and customization then plug-ins are definitely useful. For a developer who works on a cross-platform title it is more difficult because not all of the plug-ins will be available on all platforms. There may be technical reasons why a plug-in cannot be supported on all platforms, or more commercial reasons. Whilst these constraints exist (and I expect they will for some time to come), the tools can help by finding the best way to author data by making the best of each platform (and available plug-ins) but prevent the need to duplicate design work for different platforms with different abilities.

All of these developments would be to the advantage of game developers and content creators, but the same might not be said of middleware providers, who have invested time and money in the development of the tools they offer; standardization would inevitably involve redevelopment and could provide competition from creators of linear audio tools. Steve Martz suggests that there might also be some resistance to a standardized game data format on the part of publishers, developers, and some larger studios:

They spend a considerable amount of their internal resources customizing... game engines to accommodate new features and advancements for the next title as well as streamlining project workflow. [That is] not to say that... resistance is unwarranted. They do what works best for their production pipeline;... third-party software or [standardized game data] format(s) may or may not be able to meet their needs in a timely manner.

For those more loosely associated with game development, the advantages of an open standard would be significant, but for now, despite various proposals and the establishment of working groups, Steve Martz notes that “their progress in creating a ubiquitous standard has achieved limited success. They may be successful at some point, but just not now.”

28.5.1 Toward Standardization

Questions of standardization apply not only to tools for interactive audio and interactive data formats but also to technical specifications for game audio; a full discussion of which would be beyond the scope of this chapter. However, if we consider one example, namely reference levels, questions of standardization soon become apparent. In a paper written for the Audio Engineering Society’s 35th International Conference in 2009, Masataka Nakahara, Kazutaka Someya, and Steven Martz surveyed differences in playback levels for in-game sound, game start-up sounds, and reference levels of game production studios. In the opening statement of the paper the authors state that:

The audio experience can vary greatly when playing a video game. After establishing a playback level for a reproduction device, the game may still be too loud or too soft. Average reproduction levels can fluctuate from title to title and even within a game. It is important for game production studios to carefully consider how they configure their systems and the effects it may have on consistency in their games. (Nakahara, Someya, and Martz 2009).

In addition to suggesting useful reference levels for games based on average figures resulting from the survey, the authors of the paper conclude that “there are large differences among production companies, and... that the level of start-up sounds which are implemented into game consoles are approximately 10dB lower than the level of in-game sounds in opening cut-scenes” (Nakahara, Someya, and Martz 2009). If one compares such differences with the standardized specifications found in the film industry,

the contrast appears quite marked. The work of the Audio Engineering Society and its Technical Committee on Audio for Games seeks to address issues of this type while furthering educational goals. Conferences, such as those organized by the AES and also the games industry itself, also have a role to play in establishing greater standardization, while also sharing best practice and promoting education.

The lack of standardization in relation to various technical specifications has a direct impact on mixing for games. Consider one example: game audio, unlike film audio, is mixed primarily for home media systems and mobile media devices. The speakers of these systems vary hugely in quality and specification, ranging from the cheapest in-built speakers through to high-end, audiophile systems designed for the well-heeled connoisseur. While equipment bearing the THX logo indicates a certain standard of maximized performance for audiovisual components and systems based on standards and parameters adapted from the cinematic context, not all manufacturers follow THX guidelines. Despite some general improvements in the quality of cheaper speakers for the home, it could be argued that the growing market for surround-sound systems has proved somewhat counterproductive and has not facilitated quality enhancement. For those on a budget, the money that would once have secured a reasonably decent pair of stereo speakers is now expected to secure anything up to seven speakers plus a sub-woofer. In some cases surround-sound systems are simply thrown in as an extra feature accompanying the purchase of large-screen television sets, arguably resulting in a dilution of the quality of these systems. Even the best surround-sound systems may not be calibrated correctly for the domestic spaces in which they are placed. Consequently, those mixing audio in preparation for release to the home entertainment and mobile device markets must consider the variable quality and specifications of this equipment. Jerry Ibbotson accommodates such considerations in the final stage of the mix process:

One of the key problems people have is that if they have a 5.1 setup, the majority of them will be setup incorrectly. They've probably got the sub up way too loud, drowning everything else out and many people won't know what the center speaker is for... With games that are played in the home it is grimly inevitable that someone has their speaker system setup oddly... With a game you always try and work on it in a studio environment and have that benchmark, but then you need to know what it's going to sound like in a domestic environment. When we worked on a PSP game we had a PSP development kit in the studio. At first we had the sound coming out of the studio speakers but as we got further towards the end of the project we were unplugging it from the mixer in the studio and either plugging in the worst headphones we could find or just playing it through the inbuilt speakers, so you know what it's going to sound like to the punter on the bus. For us, the logical process would be to work on the samples and the overall implementation at the highest possible quality you can manage, and then towards the end, to get the overall mix right, listen to it on what the consumer would listen on. Most home systems are pretty good these days and will pick up on most things but on a handheld like the PSP with tiny speakers, some things will just get completely lost. We have done stuff in the past, say a menu sound, where we have added a frequency element on top so that you'll be able to hear it better through smaller speakers.

Broomhall also suggests that great care is taken within the industry to consider different home setups. For example, depending on the playback configuration selected, adjustments can be made to the settings of multiband compressors strapped across the final mix, optimizing the sound for a particular setup. He also points to a growing trend towards the use of external, premium studios in which game teams set up their game development equipment and “tweak their mix in a high-end reference listening environment, possibly with a consulting mix engineer on hand to advise.” It is clear that those within the games industry are concerned about the integrity of dynamic game mixes when transferred to the home environment and it is also clear that steps are being taken within the industry to move towards greater standardization in terms of tools and technical specifications. Just as standardization developed over many years within the film industry, it appears as if the games industry is at the start of a similar journey towards greater standardization.

The limitations of game audio technology, which were once so apparent within the games industry, have faded into the background, being replaced by rapidly evolving and unique creative and technical challenges. In addressing these challenges lies real potential for innovation and the exploration of “the power of ideas” (Broomhall 2010). The rich resource offered by the best examples of film audio and practice provides a valuable source of inspiration, but the interactive nature of game audio opens up new avenues waiting to be explored by the power of the imagination.

28.6 CONCLUSIONS

From very simple beginnings, game sound has progressed rapidly and continues to evolve. Film sound has provided an important referential resource in the development of game sound and continues to be of relevance moving forward into what promises to be a future of further rapid technological change and innovation. However, despite some common ground, there is one crucial point of divergence at which the paths of film and games separate, namely interactivity.

The creative and technical challenges presented by interactivity, as outlined in this chapter, fall into a number of key areas: the integration and mixing of sound within dynamic, real-time environments and the involvement of audio creators in this process (determining the prioritization, placement, real-time processing, and functionality of audio); and the impact of production schedules and associated budget constraints, coupled with nonstandardized approaches, tools, and audio specifications.

What might be the outcomes if these challenges are successfully tackled? It seems likely that as the challenges of integration and dynamic mixing are gradually eroded, audio creators will have far greater input in determining the ways in which sound functions and is prioritized within games. If this were to occur we might expect to see a shift away from literal, “see it, hear it” approaches, toward nonliteral explorations of the “storytelling power of sound.” Within this context, the old adage “sound keeps it real, and

music tells you what to feel” is likely to be challenged as the full potential and functionality of sound are explored within games, matching the very best examples of nonliteral sound on film but, more importantly, exploring the unique possibilities of sound in interactive environments.

Ironically such nonliteral explorations of functionality and precedence within game mixes might create games that feel far more real and immersive. For example, the prioritization of dialog in games mixes today can conjure up the impression of disembodied voices floating above the action, recreating anything *but* reality. The primary reason for the development of this approach is the need to communicate information to the player to facilitate game play. Perhaps this is appropriate in the first levels of a game, or in games for children, but should dialog be prioritized in this way as a rule of thumb? Might there be other, more subtle ways of communicating information to the player through sound rather than dialog? Might there be a case for depriving the player of information, or even misleading them, in order for game play to become more real or challenging?

If production schedules become a little more accommodating and consideration is given to the use of audio right from the beginning of the game development process, there should be more time for experimentation of this sort, posing and addressing questions such as these. Technological advances, in terms of integration, development, and so on, might “free up” additional time within production schedules, allowing more time for creative experimentation while also providing audio creators with easier ways of integrating audio assets, enabling them to try out different ideas and approaches. If, in addition, a dedicated postproduction phase of development becomes standardized we should also see improvements in sound production and mastering.

In terms of answering the technological challenges of dynamic mixing and games development more specifically, we should expect to see further developments in middleware tools and dedicated game development technologies. However, given technological trends already outlined, we are also likely to witness developments in “off the shelf” and customizable software and plug-ins. If the move toward standardization progresses, the features of standardized connectivity and multiplatform use will become more important, as will standardized data formats and audio specifications. Developments such as these would certainly be of interest to the general public and to providers of pro-audio (linear) tools, and perhaps to the film industry as well. At this point we might witness a reversal in the tide of influence as the film industry adopts technologies originally developed for games. More specifically, technological advances developed for the purposes of integration, dynamic mixing, and real-time processing, might be of use when producing, pre-dubbing and mixing audio stems for film.

In addition to technological advances supporting the development of games, we should expect to see further developments in gaming hardware and user interfaces. It seems likely that the human body itself and associated physiological or psychophysiological responses might well become the focus of interface development. Nintendo’s ongoing development of a “vitality sensor” (designed to measure pulse response) and Microsoft’s proposed “Joule” heart-rate monitor for the Kinect Play Fit, might be indicative of future trends. A patent application for biometric controls (designed to measure

galvanic skin resistance, electrocardio data, and electromuscular data), filed by Sony in 2010, would suggest that the company is at least thinking along similar lines. The introduction of interfaces with biosensors of this type could be quite significant for developments in dynamic mixing. For example, galvanic skin response, heart rate, respiration, and so on, might be linked to real-time processing and changes in the layering of the mix; changes designed to intensify physiological and also emotional responses. Technological advances such as these will no doubt raise new technical and creative challenges, not to mention ethical and medical considerations. As the games industry addresses these challenges in the interests of interactive and immersive entertainment, outcomes might result that could be applied in other fields, such as medicine and music therapy, perhaps offering new approaches in the treatment of specific physical and mental health conditions.

To date, innovation and a pioneering spirit have characterized the games industry. The technical and creative challenges of interactivity have fuelled rapid and diverse developments. If moves toward standardization progress, with all the potential benefits this might afford, it will be interesting to see whether these traits continue to flourish. A standardized games industry might not nurture innovation, experimentation, and diversity to the same degree. The exploratory zeal of the games industry might be stifled in less fertile, homogenized ground.

Moving into the future, the advantages of standardization might well outweigh concerns of this kind, paralleling the approach taken by the film industry in its own historical development. But do the parallels end there? It seems likely that the influence of the film industry will lessen in coming years. As the games industry grapples with the challenges of interactivity, the relevance of linear film as a referential resource will continue to diminish. From its first tottering steps the games industry continues to mature towards greater independence. Like a stropky teenager with a growing sense of their own identity, the games industry will become more self-sufficient and the balance of influence will shift full circle.

ACKNOWLEDGMENTS

The author wishes to thank the contributors, all of whom gave so generously of their time, expertise and contacts in the preparatory interviews for this paper:

John Broomhall (2010): Broomhall Projects Limited; audio director and consultant (see <http://www.johnbroomhall.co.uk> for further details), BAFTA Video Games Committee and Chair of BAFTA Interactive Entertainment Audio Awards, Columnist for *Develop Magazine*. Previous roles: formerly Head of Audio (and Game Producer) at MicroProse, Spectrum Holobyte, Hasbro Interactive, and Infogrames UK Internal Studios. Interviewed February 2010.

Jerry Ibbotson (2010): founder and Creative Manager of Media Mill (a game audio production company founded in 2000; see <http://www.mediamill.co.uk/> for further details). Writer and reviewer for *Audio Media* magazine. Guest lecturer in sound design

at Hull University, York St. John University, and York University. Consultant and beta tester for Adobe Systems. Previously a BBC radio journalist and freelance radio producer. Interviewed February 2010.

Michael Kelly (2010): Senior Audio Engineer, SCEE R and D, Sony Computer Entertainment Europe and Co-chair of the AES Technical Committee on Audio for Games. Previous roles: Audio Research Engineer at Creative; Lead Sound Designer at Revolution Software. Interviewed by telephone and email, July and August 2010.

Steve Martz (2010): Senior Design Engineer THX Ltd and Co-chair of the AES Technical Committee on Audio for Games. Interviewed by telephone and email, June to August 2010.

Jamie Maxwell: Research Assistant.

Nick Peck (2010): Perceptive Sound Design (owner). Previous roles include: Audio Director, Activision; Sound Supervisor, LucasArts Entertainment Company Games; Game Sound Supervisor, Skywalker Sound (see <http://www.underthetree.com/perceptivesound/> for further details). Interviewed August 2010.

NOTES

1. Academic literature tends towards identifying a greater number of audio types or classifications. These are sometimes linked to terminology more commonly associated with film studies, such as diegetic and nondiegetic sound or music. See Collins 2008, 125–7.
2. For further details, see Original Sound Version 2010.
3. See Original Sound Version 2010.

REFERENCES

- Collins, Karen. 2008. *Game Sound: An Introduction to the History, Theory and Practice of Video Game Music and Sound Design*. Cambridge, MA: MIT Press.
- Lastra, James. 2000. *Sound Technology and the American Cinema: Perception, Representation, Modernity*. New York: Columbia University Press
- Nakahara, Masataka, Kazutaka Someya, and Steven P. Martz. 2009. An Analysis of Differences in Playback Levels Among In-game Sound Levels, Game Start-up Sounds, and Reference Levels of Game Production Studios. *Proceedings of the 35th AES International Conference on Audio for Games*. New York: Audio Engineering Society.
- MIDI Manufacturers Association. 2010. *The Interactive Audio Special Interest Group and the iXMF Working Group (IXWG)*. <http://www.iasig.org/wg/ixwg/>.
- Original Sound Version. 2010. *Don Veca Interview: Dead Space Audio*, <http://www.original-soundversion.com/?p=693>.