This paper was presented at the 56th Conference of the Audio Engineering Society, as paper number 3-2. The full published version can be found at http://www.aes.org/e-lib/browse.cfm?elib=17591

EXTREME NINJAS USE WINDOWS, NOT DOORS: ADDRESSING VIDEO GAME FIDELITY THROUGH LUDO-NARRATIVE MUSIC IN THE STEALTH GENRE

Richard Stevens[1], Dave Raybould[2], and Danny McDermott[3]

[1]*Leeds Beckett University, Faculty of Arts, Environment & Technology, Leeds, UK*
r.c.stevens@leedsbeckett.ac.uk
[2] *Leeds Beckett University, Faculty of Arts, Environment & Technology, Leeds, UK*
d.raybould@leedsbeckett.ac.uk
[3] *Leeds Beckett University, Faculty of Arts, Environment & Technology, Leeds, UK*
d.mcdermott67272@leedsbeckett.ac.uk

A significant factor in the aesthetics of video games is the need to compensate for a lack of, or poor fidelity of, sensory information that would be present in the physical world. Although dialogue, sound and music do play a ludic role, by providing information to compensate for this, in general there remains an over reliance on visual UI (User Interface) which has to fight for attention within an already overwhelmed sensory channel. Through a methodical analysis of the functions of audio in the stealth genre this paper identifies the limitations of current binary threshold approaches to audio feedback and puts forward music as a potential vehicle for providing richer data to the player. Music is accepted as a continuous audio presence and is able to provide information to help to prevent player failure, rather than sound effects or dialogue which often serve simply as a notification of failure.

INTRODUCTION

In video games the soundtrack is often relied upon to provide instruction, notification, feedback and orientation to the player, in addition to maintaining the more familiar narrative roles established in film [1]. Music and sound must provide this vital gameplay information to allow players to fulfill their intrinsic need for mastery [2] and for them to maintain a motivational state of flow (defined by Csíkszentmihályi as emerging when a person's skills are finely balanced against the challenges they face [3]). These 'Ludic' roles of audio [4] are of particular importance in the stealth genre where communication of states such as visibility, AI alertness[1] and proximity are crucial in supporting the strategic advantage of remaining undetected.

Through the analysis of 11 stealth games this study identifies a range of issues with existing practice and examines the potential for music to ease the pressures on the visual communication channel, and avoid the pitfalls of dialogue, by conveying information in new ways. The sample set was derived from games appearing more than twice in the results from the internet search query "best stealth games", with platform specific results and anecdotal forum discussions being discounted. The resulting set comprised: A - Assassins Creed III [6], B - Batman: Arkham City [7], C - Deus Ex: Human Revolution [8], D – Dishonored [9], E – Hitman: Absolution [10], F - Mark of the Ninja [11], G - Metal Gear Solid 4 [12], H - Sly Cooper: Honor Among Thieves [13], I - Splinter Cell: Conviction [14], J - Tenchu: Shadow Assassins [15], K - Thief: Deadly Shadows [16].

The paper focuses on three specific aspects of key importance to the player; the location of the non-player characters (NPCs), whether the player avatar is visible to the NPCs, and the state of alertness of the NPCs, in other words:
- Where are they?
- Can they see me?
- Are they actively pursuing me?

1 WHERE ARE THEY?

1.1 Why We Need to Communicate NPC Location

Upon analysis it can be seen that the majority of the ludic or UI (User Interface) components that are a familiar part

[1] AI or Artificial Intelligence refers to the behaviour of the Non Player Characters that result from game design algorithms [5].

of the aesthetics of many games are the result of attempts to compensate for a lack of fidelity of information that one might expect in the physical world. For example in order to avoid detection it is crucial that the player is aware of the location of the NPCs, but the lack of peripheral vision implicit in the limited field of view (FOV) of games[2] [17] not only impairs spatial navigation in which it plays such an important role [18], but also denies us the heightened automatic alert response to stimuli in the peripheral zone [19]. Given the limited FOV players are more heavily reliant upon spatialized sound to identify the location of potentially threatening NPCs. Such "signal listening" [20] is acknowledged in four of the games (B, D, E, F[3]) which offer a specific 'focussed listening' mechanic to alter the audio mix, more clearly isolating sounds of particular importance. Batman's 'Detective' mode attenuates and filters the ambient sounds. When peering through a keyhole in Dishonored or leaning against a door in Mark of the Ninja the audio in the adjacent room becomes focussed by an increase in volume, and Hitman's 'Instinct' mode filters out music and ambience to emphasise NPC dialogue.

The use of ambient NPC dialogue (present in 91% of the games) is a good example of the conflicts that can arise between audio's narrative and functional 'fit' [21] since its functional role (or ludic role), of alerting the player to the location of the NPCs, demands that it be a frequent and continuous presence. This undermines its narrative role, of making the world seem real and alive, through the repetition that consequently arises[4]. Even if one allowed for the production costs in creating and recording enough variety to convincingly cover many hours of gameplay, and the associated overhead in terms of disk storage, such ambient dialogue would still not be a good compensation for the lack of peripheral vision when locating NPCs.

Orientation through stereophonic listening, or "navigational listening" may be "fundamental to the players relationship to the game and other players" [22] but its accuracy is poor. The variety of playback hardware is highly variable [23] and whilst the poor utility of stereo speakers on a television 6 feet from the sofa for spatialization is obvious, it should be appreciated that even a sophisticated audio setup such as a 7.1 system does not come close to the fidelity of a real world soundfield. Given a theoretically ideal number of speakers (a practical impossibility since even a simple soundfield sphere of 40cm diameter would require 25 separate speakers [26]), the players ability to listen for enemy location would remain compromised since the dynamic range of a 16bit playback system is limited [27] and games tend not to use full dynamic range anyway, given the unpredictable and highly variable noisefloor of the playback environment [28][5].

Given unpredictable variables and limitations in fidelity, the usual careful considerations of the environment and apparatus required for making sure that sonic information can be detected [29] are impossible and therefore game designers have had to implement compensatory alternatives to a 'naturalistic' soundscape.

1.2 Visual UI for NPC Location

91% of the games in the sample set have some kind of visual UI for alerting the player to the location of NPCs. 82% indicate their actual location[6], and 18% use a visual icon to indicate that an NPC is within a given proximity to the player. (See Fig. 1)

Figure 1: Examples of visual UI to indicate NPC location. From top left : Hitman – Radar; Splinter Cell - Tagging; Dishonored - See through walls; Sly Cooper - Proximity icon

[2] Even in third person view the camera sticks close to the players position, limiting the extended FOV that might be available with a more distant model. The 2D game 'Mark of the Ninja' is an exception to the otherwise 3D perspective games in the study.

[3] When appropriate the games will be referred to in the main text according to their designated letters, as described in Footnote 1.

[4] The slightly questionable logic of guards constantly talking to themselves aside, when your game dialogue (Dishonored's "Should we gather for whiskey and cigars tonight?") is specifically commented on by reviewers [24], or becomes an internet meme (such as Skyrim's "arrow to the knee" [25]) it is probably an indication that player immersion in the convincing 'reality' of a game world has been broken.

[5] These differences are highlighted in the average loudness recommendations of Sony Audio Standards working group which differentiates between audio for home (-23 LUFS), or portable (-18 LUFS) consoles [30].

[6] 45% use a Radar or Map, 36% have a mechanic to allow the player to see NPCs through walls or obstructions, 18% allow you to 'tag' NPCs with icons and some games use multiple mechanisms.

Once a player is discovered and attacked 64% have an additional visual indicator to inform the player about the direction from which the attack is coming (see Fig. 2).

Figure 2: Examples of visual UI indicating the AI state and direction of NPCs. From top left: Hitman, MGS, Splinter Cell, Dishonored, Deus Ex, Mark of the Ninja

Such state and direction indicators (SADIs) are so commonplace in 1st and 3rd person shooters (the more iconic one from Mark of the Ninja (see Fig. 2) being the exception here for this 2D game) as to become an almost unquestioned 'grammar' of games, but they are indispensable only because of the limited FOV, poor spatialization of sound, and the limited dynamic range necessitated by the unpredictability of playback environments.

The proliferation of orientational UI in games indicates the degree to which the visual and audio cues we might ordinarily take for granted in the physical world to identify the location of another person must be compensated for. Although there remains some debate as to whether such visual UI impacts on the players sense of immersion [31], it is worthwhile examining how we might better use audio to convey this information in new ways, since it allows the player to remain visually focussed on their task and evidence shows that communication is more effective (and therefore task performance improves) when it takes place through multiple perceptual channels [32].

The issues around using dialogue, and the limitations of using sounds produced by NPCs (for example the localization of footsteps being affected by poor spatialization and limited dynamic range) means that other, more artificial audio 'UI' solutions may be required. UI type sounds to repel or attract the player when threats or items of interest fall within a given proximity to the player avatar are not uncommon in other games, such as the ghostly radio crackle of Silent Hill [33] which alerts the player to the proximity of the undead [34] or the accelerando of the scanner beep as the aliens draw near in Alien vs Predator [35]. Within the sample set the 'heart' item in Dishonored beats faster to indicate the players distance from secret 'Runes' hidden around the city. To be more widely applicable in games (the above examples require a narrative justification in their respective worlds for their existence) what is required is a fairly continuous sound whose presence would be readily accepted by players, criteria that are met by music.

Although spatialized 'diegetic' music is sometimes used in games to indicate the approximate location of enemies (since they are likely to be in a similar location to the sources of such music, such as radios), the use of a military drummer to accompany the pursuing band of troops in Assassins Creed 3 (a 'diegetic' sound source spatialized within the world that serves to indicate both the troop's spatial location through panning, and proximity through volume), is unusual, and justified by the games specific historical context. It is an exception to the more general position of music in games, which is to maintain the 'non-diegetic' position of film music, outside of the narrative world[7]. But if non-spatialized music cannot function to indicate the geographical location of NPCs then perhaps the presence of, or volume of, the music could function to indicate the player proximity to NPCs. There are examples of this in the games analysed; the dissonant strings in Dishonored alert the player to the diseased rats that might attack them[8] and the cartoon-like creeping bassline footsteps of Sly Cooper indicate when the player's avatar has become dangerously close to the enemy NPCs. Given our heightened tolerance of, or indeed expectation of, repetition in music (when compared to our tolerance of exact repetition of dialogue) these instances avoid the Ludo-Narrative conflict brought about by using repetitive ambient dialogue to aid localization, as they can simultaneously serve the ludic function, of indicating proximity, and the narrative function, of mirroring the dramatic tension of the situation.

1.3 Music for Proximity: Continuous Variables Through Parallel Forms

Given research that indicates sound's effectiveness in conveying dynamic variables [36], it is perhaps surprising that its use in our sample set remains that of a threshold mechanic, switching on/off at a given distance between the player and NPC. In these cases it is most likely that the primary purpose of the music is a narrative or emotional one (since a pack of rats tends to make little sound in reality, the dissonant strings were required as a 'danger' metaphor; the cartoon-like aesthetic of Sly Cooper is emphasised by the 'mickey-mousing' [37] of the characters movement) but there is potential here for music

[7] Used here for immediacy, these terms are in fact highly questionable and problematic for games [38].

[8] Given that this mechanic does not apply to the main NPCs we have not included it in our summary statistics shown on page 6.

to play a more critically important role in conveying proximity, and the richer informational stream of continuous, rather than threshold, notification would be beneficial.

The use of parallel forms of music (sometimes referred to as vertical re-orchestration [39] is common in games and involves multiple synchronised channels or 'stems' of a single musical arrangement that are able to be attenuated by variables in the game. This allows the music to vary in arrangement (and typically in intensity) in response to game events whilst maintaining a musically coherent sense of structure and timing. The distance, 'D', between the player and the NPC could be scaled to an appropriate value to enable control of the volume of music stem 03 (see Fig. 3), communicating the ludically useful information about the players proximity to danger whilst also narrativizing [34] the events through an increase in musical tension. Multiple NPCs would be handled by simply placing a limit on the maximum value of the volume of the music stem.

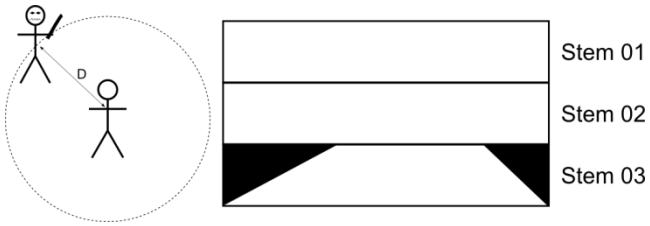

Figure 3: Distance between player and NPC (D) X Volume of Stem 03 = Proximity

In this section we have outlined why it might be worth considering the untapped potential for music in conveying information about NPC location ('Where are they?') given the issues with fidelity, the ludo-narrative conflict that arises through using ambient dialogue, and the possibly limited application of sound as a UI given a perceived need for narrative justification. We have also seen how parallel forms of interactive music are well suited to providing congruent emotional/narrative tension whilst supplying this ludic information, and that it is able to do so in a continuous and subtle, rather than threshold/binary and simplistic way, through simple volume changes in synchronised parallel stems

2 CAN THEY SEE ME?

2.1 Why We Need to Communicate Player Visibility

As human beings we are very aware of the position and location of our bodies through the related senses of proprioception and kinaesthesia [40], and the shapes our bodies can form are highly adaptable. As a consequence we are adept at using objects for hiding behind or in, and have a good sense about whether or not we are visible to others. In games our proprioceptive sense is missing (since we are not our avatar), our input mechanisms are limited [41] (typically via a gamepad), the adaptation of our avatar's positions are restricted by pre-set animations (e.g. stand, crouch or lie prone (See Fig. 4)) and the avatar's interactions with objects and the environment are often unsatisfactory (with many characters still typically being comprised of several rectangular boxes for collision detection purposes [42]).

Figure 4: Character animation fidelity in MGS4. The character can only stand, crouch, or lie prone.

Although a 3rd person perspective (available in 73% of the games analyzed) helps a little in terms of compensating for the lack of proprioception (whilst unfortunately making line of sight harder to evaluate) these factors (together with already established limited FOV) make it extremely difficult for the player to form a critical judgment about whether, when falling into a passing NPCs line of sight, they would be seen or not.

2.2 Visual UI for Visibility

In another example of how games need to compensate for a lack of sensory fidelity, 82% of the games within the study utilize a visual UI indication of player visibility. Visually these range from the use of icons (A, D, E, G, J) (see Fig. 5) and visual shading (I, F) (see Fig. 9 below) to Thief's 'Light Meter' (see Fig. 6).

Figure 5: Use of icons to indicate the visibility of the player character in Tenchu and Assassins Creed 3

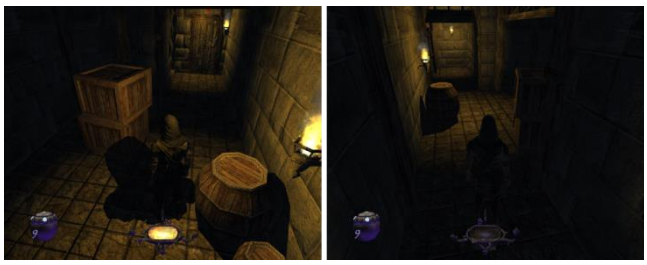

Figure 6: Use of 'light gem' in Thief (at the base of the screen in the centre)

If we look historically for a moment at the development of visual UI for visibility in a specific game series, Splinter Cell, we can see that developers appear to be continuing to struggle with communicating this effectively whilst competing with other aspects of the game for visual

attention. In the first three Splinter Cell games (Splinter Cell [43], Pandora Tomorrow [44], and Chaos Theory [45]) a light meter icon (see Fig. 7) was used (placed in the bottom left corner of the screen). Splinter Cell: Double Agent [46] then placed a 'traffic light' type meter onto the back of the avatar (see Fig. 8). Splinter Cell: Conviction used whole screen colour desaturation to indicate visibility (see Fig. 9) yet in the forthcoming Splinter Cell Blacklist [47] they have returned to the 'traffic light' meter.

Figure 7: Light meter icon in Splinter Cell

Figure 8: Avatar-based light icon on the character's back (Green, Amber, Red) in Splinter Cell: Double Agent

Figure 9: Colour desaturation in Splinter Cell: Conviction

2.3 Music for Visibility: Continuous Variables through Parallel Forms

Music is used to aid awareness of visibility in three of the games (A, E, H) where it simply attenuates to indicate when a player is hidden (alternatively F, G, J have sound notifications for 'hidden') but no games use sound or music for a more continuous notification of the player's visibility. Since the amount of light falling on a player is a continuous variable that could be analogous to visibility this could be used, again within a parallel music form, to indicate a player's visibility. The amount of light falling on the player, 'L', can be scaled appropriately to control the volume of stem 02 (see Fig. 10), communicating the ludically useful information about the player's visibility whilst again narrativizing the events through an increase in musical tension.

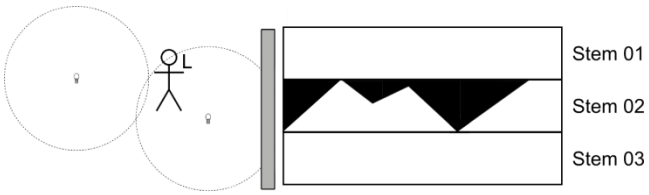

Figure 10: Light falling on player[9] (L) X Volume of Stem 02 = Visibility

2.4 'Seen', 'Spotted' and the Threshold Problem

Given the attentional demands on the visual system during gameplay it is not surprising that these visual mechanisms appear not to be enough to protect the player from frequent catastrophic failure and so 91% of the games analyzed have also implemented a 'seen' phase to act as a failsafe. This phase (based on the slightly unrealistic premise that although the NPCs might 'see' you, they will take a moment to process this information before actually 'spotting' you and becoming alert) appears to be an attempt to introduce some more increments, or a more analogue approach to the binary unseen/spotted paradigm. This momentary 'what was that?' window allows the player to move position and remain 'unseen' without consequence but is unconvincing in terms of realistic NPC behaviour.

So far we have looked at two instances where it might be preferable to have continuous information rather than a notification of thresholds (for proximity to NPCs, and the visibility of player) but the difficulty with threshold notification is thrown into sharper relief by this mechanic of 'seen' and 'spotted'. There's a greater use of sound here (45% for notification of 'seen') but although sound is effective, in that it does not rely upon visual attention to alert us [48], relying on the reactions of others (NPCs) only gives the player an indication of their visibility when they have already been seen. In other words, the problem with giving a notification of thresholds is that you only receive the notification once you have passed the threshold, i.e. it is a notification of failure, which then requires the artificial tolerance of a 'seen' phase to allow for recovery.

All of the games that have a 'seen' phase, bar one (Hitman: Absolution), use dialogue as a UI notification. As discussed above, using dialogue to supply information that is important ludically is problematic since the inherently repetitive nature of games means that a sense of narrative reality is quickly broken (through different characters in different locations throughout the game all using the same

[9] Potentially derived from the "number of bones on the player that can be illuminated" [49]

sound file exclamations[10]). The one exception to the use of repeated dialogue to indicate a 'seen' threshold is Hitman: Absolution, where the use of a continuous audio feedback mechanism (the gradual opening of a filter on a noise-like source) to indicate that the player is 'about to be spotted' allows it to dispense with the binary notions of 'unseen' or 'seen' altogether. Alerting the player to the fact that they are about to be noticed, the sound's transformation over time provides the player with the information required to avoid failure (indicating their time window to move back into cover) rather than the threshold approach which simply notifies them of their failure.

2.5 Music for 'About to be Seen'

The more continuous or 'analogue' design of the Hitman system suggests that again, parallel musical forms would be suited to conveying the player's visibility in a more continuous way, and would also allow a more narratively convincing solution to the rather incredulity stretching, 'seen' mechanic without impacting on visual attention. When the player falls within the line of sight of an NPC and has the potential to be spotted, the volume of Stem 01 begins to gradually increase (see Fig. 11) as the likelihood of being spotted increases over the time that the player remains visible.

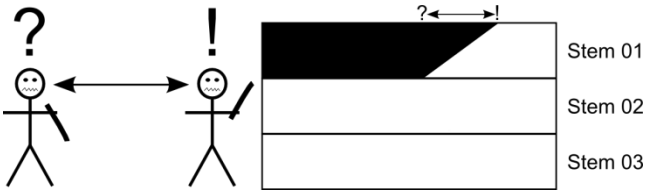

Figure 11: When 'visible' and falling into the line of sight of an NPC, Stem 01 begins to gradually increment in volume to indicate that you are 'about to be seen!'.

By using parallel stems for visibility, we can provide the player with richer continuous information, whilst avoiding visual overload and attention problems. This also avoids the threshold problem that has led to an artificial 'seen' phase whilst narrativizing the tension that you are 'about to be seen'. Using music to convey information in a new way may help games to avoid the current approaches in some games which appear to neither provide sufficient information to avoid being seen, nor a realistic reaction when you are.

3 ARE THEY ACTIVELY PURSUING ME?

3.1 Why We Need to Communicate AI State

A third critical piece of gameplay information a player must know within the stealth genre is whether the NPCs are actively pursuing them (and they no longer have the cushion of a 'seen' phase) or have returned to their default patrolling state (we will refer to this as the NPCs 'AI state'). In the physical world a person's body language or facial expression would provide sufficient cues for determining their emotional state, but within games the memory and processing confines mean that there are a limited set of animations available for a character's movement and posture [50], their facial animations are typically restricted to basic expressions [51]. The screen resolution also means that despite claims of 'photo-realism' [52] even a sophisticated facial expression is lacking in appropriate fidelity once any distance has been established between the player and that character.

Figure 12: Illustration of lack of fidelity in facial expression when viewed over distance

3.2 Visual UI for AI State

To compensate for these issues, 73% of games in the sample have some kind of visual UI to indicate AI state. In addition to UI overlays such as text (see Fig. 13) or icons (see Fig. 14) 36% of the games use emoticons, a rather blunt instrument in terms of expressing emotion that utilise icons placed above the NPCs head (see Fig. 15).

Figure 13: The use of text and colour within a graphical overlay in Deus Ex to indicate AI State as Neutral, Alarmed or Hostile.

Figure 14: The use of a colour changing icon in Tenchu: Shadow Assassins to indicate AI state.

[10] From Dishonored: "I'll find you", "Oh, I'll find you", "I'll find you, you hagfish".

Figure 15: The use of emoticons in Dishonored, Assassins Creed 3 and Metal Gear Solid 4 to indicate AI state.

In contrast to the previous examples (See Fig. 16), where it is perhaps underutilised, music is used ubiquitously, with only one exception (Thief).

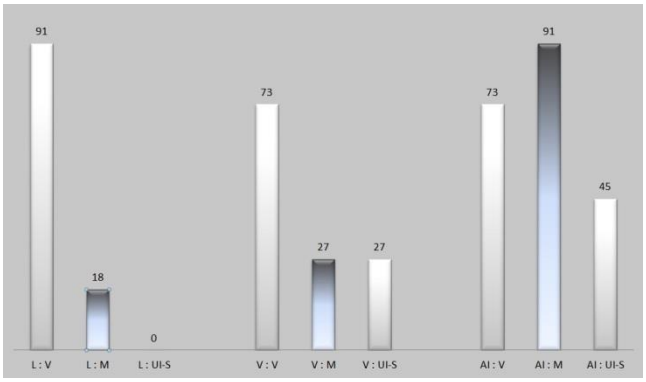

Figure 16: Summary of notifications by type: L – Location and/or proximity of NPCs, V – Visibility of player character, AI – AI state of NPCs. (V:Visual, M:Music, UI-S:UI Sound)

3.3 Ludo-Narrative Conflict in Music

It is perhaps logical to use music for these climactic circumstances as it can increase arousal [53], 'dynamizing' the gameplay experience to make it more thrilling [54]. However, unlike our previous examples, the ludic function of indicating the AI State frequently conflicts with the music's narrative or emotional role.

Any 'UI' must be consistent in order that its message be learnt and understood [17] and must provide immediate feedback [55]. These are the very characteristics that undermine music's narrative/emotional effectiveness, since in order to be ludically effective it must be repetitive (consistent), and it must disregard musical structures by reacting instantaneously (providing immediate feedback), with sometimes jarring musical effect [56].

Within the parallel forms discussed so far these are less of an issue since the 'consistent' message is abstracted to that of the presence of a musical stem, not necessarily its specific content, and since the stem is in synchrony with the underlying musical structure in many cases it can enter almost immediately without appearing to be musically clumsy. Players typically empathise strongly with, and share the goals of, their characters [21] and therefore designers should want their actions and mastery to be rewarded musically and major gameplay changes, such as entering combat or defeating enemies, necessitate a heightened drama and emotion more commonly created by transitional musical forms (sometimes referred to as horizontal resequencing [39]) where a change is made from one music state, or piece, to another.

Parallel forms may be very effective in communicating information to support mastery (and therefore help maintain a 'flow' state), but to enhance a feeling of 'fiero' (a sense of triumph in overcoming major obstacles [57]) we require more whole-scale changes. Many of the strong emotions evoked by music are associated with the creation of, and confirmation and violation of, expectancy [58], and a sense of closure (identified as another motivating factor behind play [59]) created by a musical cadence [53]. These powerful musical effects require time-based structures in which to operate, an idea seemingly in opposition to the interactivity of games, and to the autonomy of players to instigate actions or events at any time.

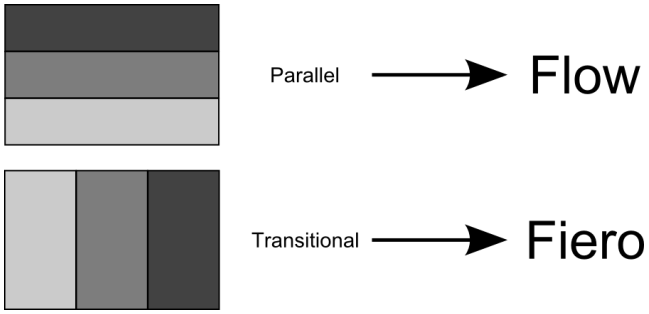

Figure 17: Musical forms and emotional affect.

These dual aims, of informing the player immediately about the AI state, and of providing the most musically pleasurable structure (in order to heighten the players sense of achievement) seem irreconcilable until we view the problem through a different design paradigm, one very rarely used in current game development practice.

3.4 Arbitrary Thresholds Offer Opportunity Windows for Integrated Design

If one considers 'Interactivity' as a reciprocal, two way process [60] it can be seen that there is little that is truly interactive about most music in video games, instead it is 'reactive', (simply responding to instructions from the game state).

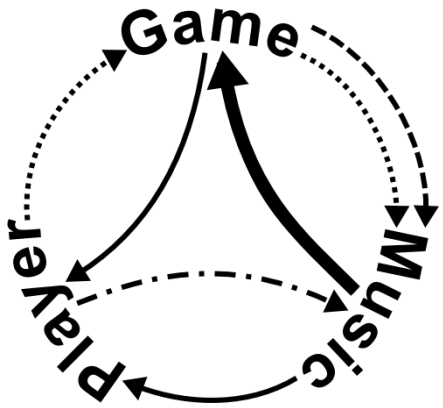

Figure 18: Illustration of 'connections' between game, player and music [61]

The diagram above represents the typical position of music in games, as a passive receiver of instruction from the game state. Sometimes the game state alters due to actions from the player ('Reactive' music shown by the dotted line), sometimes the game state changes autonomously ('Adaptive' music indicated by the dashed line), but rarely does the game take input as to the current music state (as indicated by the bold line from Music to Game) [61]. If we are to begin to address the Ludo-Narrative conflict inherent in music's use in games, and better utilise music's role as a powerful emotional reward, then the game needs to accept musical data as an input to its decision making structures; i.e. become a truly 'interactive' system. We have looked in the previous sections at how threshold information might be better communicated as continuous data, and our solution below suggests that the presence of any arbitrary threshold might also provide us with the opportunity for a more integrated design approach.

3.4.1 AI Stand-down

It is in the nature of most games, and stealth games in particular, that after a given amount of time the NPCs give up looking for you as a player, or 'stand down' (all games bar one[11] in the sample used this used this mechanic). The decision as to how long the NPCs will continue their search is no doubt the result of considered thought as to the balancing of the game but we would suggest that a few seconds one way or another would not impact significantly on gameplay. A general 'stand down' time could be specified, not as a threshold after which the AI stands down immediately, but rather as a window, during which the AI may stand down, only actually doing so at the next available musical point (typically the next bar or measure), therefore providing an opportunity for the player's successful evasion to be greeted with a more satisfying musical reward, instead of a musical fade-out (a technique whose musical effect is described by Clint Bajakian, senior music manager at Sony, as 'disappointing' [62]).

The top diagram (1) in Fig. 19 below illustrates the current 'reactive' approach (as found in C, D and G) where the intense music for the AI active pursuit state fades in on AI alert, and crossfades back to the ambient music state at a given time later when the AI stand down (indicated by the ¡ symbol). (The vertical lines illustrate the position of musical barlines or measures). The centre diagram (2) illustrates a slightly more musical solution (found in games A and B) where the AI stand down, but the music does not react immediately, instead waiting until the next available measure to transition more satisfactorily into the ambient musical state. The lower diagram (3) illustrates the crucial difference for a truly 'interactive' approach. Around the time that the game state wants the AI to stand down a 'window of opportunity' is opened, input is sought from the music engine as to its next bar/measure instance, and this feeds back into the game state to actually define when the stand down will occur, this event now coinciding with a musically appropriate time, and musically pleasing transition.

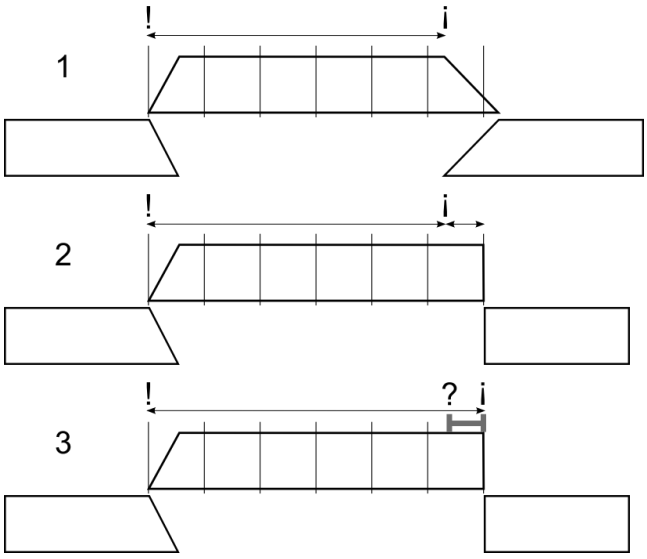

Figure 19: Illustration of different musical behaviours in relation to AI alert (!) and stand down (¡) states: 1. Reactive 2. Reactive but Musical 3. Interactive

This may appear to be some effort to simply acknowledge the perhaps undramatic event of an enemy giving up pursuit (as typically indicated by the weary mutterings of the NPCs, "Probably rats" in Dishonoured) but for many games the triumph of the player in eliminating all enemies within a given area is responded to in exactly the same way, since (using the logic of the game state, rather than the dramatic state) there are no longer any enemies present, therefore the pursuit phase must be over. Six games out of the ten that use music in the sample set use the same musical response for 'stand down' and 'all AI killed', with the players achievement being undermined by an inappropriate, or at least anticlimactic, musical response (for example in the musical fadeouts of Deus EX and Dishonored).

3.4.2 AI Killed

A similar interactive system for the more dramatic 'fiero' moment of triumph over an enemy seems at first glance equally insurmountable, since the game is balanced so that the enemy may take 8 blows with a sword, or 6 shots of a gun to die. But again the amount of damage taken by an NPC before death is an arbitrary threshold, and again there is an argument to be made that using an opportunity

[11] In Splinter Cell: Conviction once the player has been spotted the NPCs continue to search diligently for an infinite amount of time.

window around this threshold, that looks to the music state to determine the timing of the actual death event, will result in the powerful "heightened emotional impact provided by the close synchronisation of musical and visual events" familiar from film [56].

There is of course a potential threat to the player's sense of consistency in the game, however the idea of dynamically altering the game 'balance' in order to heighten the enjoyment of the player is not new [63]. There are many examples where the game conditions are altered dynamically, by changing damage multipliers (as in the X blows of a sword / shots from a gun example we give above) [64], or through the 'rubber banding'[12] that is pervasive in racing games. There is further research to be done in identifying whether dynamic changes to thresholds, based on a current music state, have a negative effect on player experience, but the potential positive emotional result of the synchrony between game action and music structures is an opportunity that should be explored.

3.4.3 Conceptual System

Although the production hurdles to implementing a more integrated design approach to music in games may be numerous [61] the actual implementation is generally straightforward. All good Ninjas know that it is more effective to use windows (of opportunity), than doors (thresholds) so instead of triggering the music directly (and the music responding reactively to events), the game 'call' opens a 'gate', thus creating a window of opportunity. The gate receives impulses from the music system that indicate musical transition points, and with the gate now open the next potential 'juncture' [1] (or transition point) will pass through the gate, triggering the event (e.g. NPC Death) and musical transition simultaneously, at a musically appropriate time.

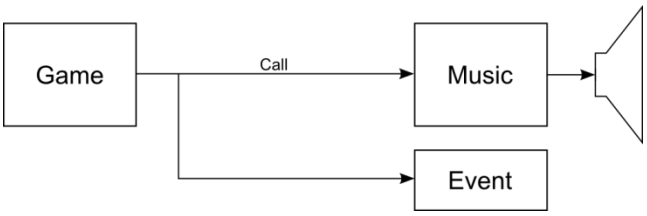

Figure 20: Illustration of a 'reactive' system

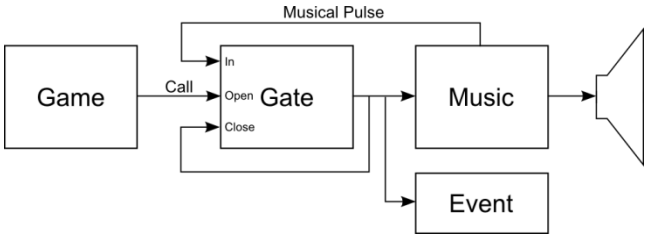

Figure 21: Illustration of an 'interactive' system

4 CONCLUSIONS

Our ability to monitor and process multiple auditory data sets [66] means that, in the absence of a "crazy high level of fidelity" to aid player's perception of their environment or NPC state [67], music could serve an important compensatory role in interaction design for video games that might allow us to move away from repetitive dialogue and an over-reliance on visual UI. Some of the ideas outlined above are already found in several games (music for proximity), and others have a potential application that appears to be as yet underutilised (we have found no examples in either our sample set or more widely of music for 'visibility' or 'about to be seen').

Continuous notifications via parallel musical forms for the players proximity to NPCs, visibility, and 'about to be seen' state may have the potential to extend periods spent in the immersive state of flow by providing the player with the information they need in order to avoid failure (not just threshold information that serves simply as a notification of failure[13]), and the use of truly 'interactive' forms could heighten a players sense of triumph (fiero) by synchronizing game events with musical structures. Testing these hypotheses empirically will form the basis of future research.

5 REFERENCES

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

[2] Rigby, S. and Ryan, R. 2010. *Glued to Games: How Video Games Draw Us in and Hold Us Spellbound*. Praeger Publishers Inc., Santa Barbara, CA.

[3] Csíkszentmihályi, M. and Csíkszentmihályi, I. S. 1992. *Optimal Experience: Psychological Studies*

[12] Rubber banding in racing games refers to the technique of artificially adjusting the speed of the opponents in order to maintain a sense of challenge for the player [65].

[13] We have focussed on a particular implementation technique for simplicity and clarity but would highlight that the game variables could be applied to many other parameters of the music besides the volume of a particular music stem.

of Flow in Consciousness. Cambridge University Press, Cambridge, MA.

[4] van Elferen, I. 2011. ¡Un Forastero! Issues of Virtuality and Diegesis in Videogame Music. *Music and the Moving Image, 4, 2*, 30-39.

[5] Mateas, M. 2003. Expressive AI: games and artificial intelligence. In *Proceedings of International DiGRA Conference*. November 4-6, Utrecht, Netherlands.

[6] Ubisoft. 2012. *Assassins Creed III* [Game].

[7] Rocksteady Studios. 2011. *Batman: Arkham City* [Game].

[8] Eidos Montreal. 2011. *Deus Ex: Human Revolution* [Game].

[9] Arkane Studios. 2012. *Dishonored* [Game].

[10] IO Interactive. 2012. *Hitman: Absolution* [Game].

[11] Klei Entertainment. 2012. *Mark of the Ninja* [Game].

[12] Kojima Productions. 2008. *Metal Gear Solid 4: Guns of the Patriots* [Game].

[13] Sucker Punch Productions. 2005. *Sly 3: Honor Among Thieves* [Game].

[14] Ubisoft Montreal. 2010. *Tom Clancy's Splinter Cell: Conviction* [Game].

[15] Acquire. 1998. *Tenchu: Stealth Assassins* [Game].

[16] Ion Storm Inc. 2004. *Thief: Deadly Shadows* [Game].

[17] Grimshaw, M. 2010. Player Relationships as Mediated Through Sound in Immersive Multi-player Computer Games. *Revista Comunicar, 17, 34*, 73–80.

[18] Yamamoto, N. and Philbeck, J. 2013. Peripheral Vision Benefits Spatial Learning by Guiding Eye Movements. *Memory & Cognition, 41, 1*, 109-121.

[19] Bayle, D. J., Henaff, M. A. and Krolak-Salmon, P. 2009. Unconsciously Perceived Fear in Peripheral Vision Alerts the Limbic System: A MEG Study. *PLoS ONE, 4, 12,* e8207.

[20] Huron, D. 2013. In *Playing with Sound: A Theory of Interacting with Sound and Music in Video Games.* K. Collins. MIT Press, Cambridge, MA.

[21] Ekman, I. and Lankoski, P. 2009. Hair-raising entertainment: Emotions, Sound, and Structure in Silent Hill 2 and Fatal Frame. In *Horror Video Games: Essays on the Fusion of Fear and Play*. B. Perron, Ed. McFarland & Co Inc., Jefferson, NC.

[22] Grimshaw, M. 2008. Autopoiesis and Sonic Immersion: Modeling Sound-Based Player Relationships as a Self-Organizing System. In *6th Annual International Conference in Computer Game Design and Technology*. November 12-13, Manchester, UK.

[23] Stockburger, A. 2003. The Game Environment from an Auditive Perspective. In *Level Up: Proceedings from DiGRA 2003*. November 4-6, Utrecht, Netherlands.

[24] Senior, T. 2012. *The Best NPC Barks of the Year 2012: Dishonored*. [Online] http://www.pcgamer.com/2012/12/26/the-best-npc-barks-of-the-year-2012-dishonored/ [Cited: 02/05/2013].

[25] Totilo, S. 2012. *How They Came up with Skyrim's 'Arrow in the Knee' Line*. [Online] http://kotaku.com/5886648/how-they-came-up-with-skyrims-arrow-in-the-knee-line [Cited: 02/05/2013].

[26] Ward, D. B. and Abhayapala, T. D. 2001. Reproduction of a Plane-Wave Sound Field Using an Array of Loudspeakers. In *Speech and Audio Processing, IEEE Transactions on, 9, 6,* 697–707.

[27] Pohlman, K. C. 2011. *Principles of Digital Audio,* 6th ed. McGraw-Hill, New York, NY.

[28] Hays, T. 2011. Dynamic Rand and Mix Levels: Where We're at in 2011. Presented at *Game Developers Conference*. February 28-March 4, San Francisco, CA.

[29] Walker, B. N. and Nees, M. A. 2012. Theory of Sonification. In *The Sonification Handbook*. T. Hermann, A. Hunt and J. Neuhoff, Ed. Logos Publishing House, Berlin, Germany.

[30] Sony Worldwide Studios Audio Standards Working Group. 2012. *Average Loudness and Peak Levels of Audio Content on Sony Computer Entertainment Platforms*. Sony Computer Entertainment, TR No.

ASWG-ROO1. [Online] http://gameaudiopodcast.com/ASWG-R001.pdf [Cited: 29/04/2013].

[31] Jørgensen, K. 2012. Between the Game System and the Fictional World: A Study of Computer Game Interfaces. *Games and Culture, 7, 2*, 142–163.

[32] Brewster, S. A., Wright, P. C. and Edwards, A. D. N. 1993. An Evaluation of Earcons for use in Auditory Human-Computer Interfaces. In *Conference on Human factors in Computing Systems*. April 24-29, Amsterdam, Netherlands.

[33] Konami Computer Entertainment Tokyo. 1999. *Silent Hill* [Game].

[34] Whalen, Z. 2004. Play Along – An Approach to Videogame Music. *Game Studies, 4, 1*. [Online] http://www.gamestudies.org/0401/whalen [Cited: 02/05/2013].

[35] Rebellion. 2010. *Aliens vs. Predator* [Game].

[36] Brewster, S. 1994. *Providing a Structure Method for Integrating Non-Speech Audio into Human-Computer Interfaces*. [PhD Thesis] York University, York, UK

[37] Chion, M. 1994. *Audio-Vision: Sound on Screen*. Columbia University Press, New York, NY.

[38] Jørgensen, K. 2010. Time for New Terminology? Diegetic and Non-Diegetic Sounds in Computer Games Revisited. In *Game Sound Technology and Player Interaction: Concepts and Developments*. M. Grimshaw, Ed. ICI Global, Hershey, PA.

[39] McAlpine, K. B., Bett, M. and Scanlan, J. 2009. Approaches to Creating Real-Time Adaptive Music in Interactive Entertainment: A Musical Perspective. In *AES 35th International Conference: Audio for Games*. February 11-13, London, UK.

[40] Stillman, B. C. 2002. Making Sense of Proprioception: The Meaning of Proprioception, Kinaesthesia and Related Terms. *Physiotherapy*, *88, 11*, 667–676.

[41] Gerling, K. M., Klauser, M. and Niesenhaus, J. 2011. Measuring the impact of controllers on player experience in FPS games. In *Proceedings of the 15th International Academic MindTrek Conference: Envisioning Future Media Enviornments (MindTrek '11)*. September 29-30, Tampere, Finland.

[42] Suaib, N. M., Bade, A., Mohamad, D. and Sulaiman, H. A. 2009. Bounding Volume Hierarchy for Avatar Collision Detection: Design Considerations. In *The 5th Postrgraduate Annual Research Seminar (PARS '09) UTM*. June 15-18, Johor, Malaysia.

[43] Ubisoft Montreal. 2002. *Tom Clancy's Splinter Cell* [Game].

[44] Ubisoft Shanghai. 2004. *Tom Clancy's Splinter Cell: Pandora Tomorrow* [Game].

[45] Ubisoft Montreal. 2005. *Tom Clancy's Splinter Cell: Chaos Theory* [Game].

[46] Ubisoft Shanghai. 2006. *Tom Clancy's Splinter Cell: Double Agent* [Game].

[47] Ubisoft Toronto. 2013. *Tom Clancy's Splinter Cell: Blacklist* [Game].

[48] Hatch, W. and Pirhonen, A. 2011. Designing Alarm Sounds for the Control of a Hydraulic Platform. In *Proceedings of the 6th Audio Mostly Conference: A Conference on Interaction with Sound*. September 7-9, Coimbra, Portugal.

[49] Redding, P. 2012. In *Redding's Recommendation – The Splinter Cell: Blacklist Interview*. D. Hindes [Online] http://sneakybastards.net/theobserver/blacklist-interview/ [Cited: 02/05/2013].

[50] Edsall, J. 2003. *Animation Blending: Achieving Inverse Kinematics and More*. [Online] http://www.gamasutra.com/view/feature/3456/animation_blending_achieving_.php [Cited: 02/05/2013].

[51] Tice, S. E. and Lander, J. 2000. Character Animation Engines for Interactive Game and Web Applications using Hardware Assisted Functionality. In *International Workshop on Human Modelling and Animation*. June 28-29, Seoul, Korea.

[52] Smut, A. 2005. Are Video Games Art? *Contemporary Aesthetics, 3*. [Online] http://www.contempaesthetics.org/newvolume/pages/article.php?articleID=299 [Cited: 30/04/2013].

[53] Juslin, P. N. and Sloboda, J. 2011. *Handbook of Music and Emotion: Theory, Research, Applications*. Reprint. Oxford University Press, Oxford, UK.

[54] van Tol, R. and Huiberts, S. 2008. *IEZA: A Framework for Game Audio.* [Online] http://www.gamasutra.com/view/feature/3509/ieza_a_framework_for_game_audio.php [Cited:02/05/2013].

[55] Laitinen, S. 2008 Usability and Playability Expert Evaluation. In *Game Usability: Advancing the Player Experience*. K. Isbister and N. Schaffer, Ed. Morgan Kaufmann, Burlington, MA.

[56] Munday, R. 2007. Music in Video Games. In *Music, Sound and Multimedia: From the Live to the Virtual*. J. Sexton, Ed. Edinburgh University Press, Edinburgh, UK.

[57] Hazlett, R. L. 2006. Measuring Emotional Valence During Interactive Experiences: Boys at Video Game Play. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 1023–1026.

[58] Sloboda, J. A. 1991. Music Structure and Emotional Response: Some Empirical Findings. *Psychology of Music, 19, 2*, 110–120.

[59] Lindley, C. A., Nacke, L. and Sennersten, C. C. 2008. Dissecting Play - Investigating the Cognitive and Emotional Motivations and Affects of Computer Gameplay. *13th International Conference on Computer Games (CGAMES '08)*. October 24-26, Dundee, UK.

[60] McQuail, D. 2005. *McQuail's Mass Communication Theory*. 5th ed. Sage Publications Ltd., Thousand Oaks, CA.

[61] Stevens, R. and Raybould, D. 2013. Designing a Game for Music. In *Oxford Handbook of Interactive Audio*. K. Collins, B. Kapralos and H. Tessler, Ed. Oxford University Press, Oxford, UK.

[62] Bajakian, C. 2013, Audio Bootcamp: Producing music for AAA Video Games. Presented at *Game Developers Conference*. March 25,-29 San Francisco, CA.

[63] Hunicke, R. and Chapman, V. 2004. AI for Dynamic Difficulty Adjustment in Games. In: *Challenges in Game Artificial Intelligence AAAI Workshop*. July 25-26, San Jose, CA.

[64] Marinello, S. 2013. AAA Level Design in a Day Bootcamp: Architecting a Multiverse - Cooperative Design in Dead Space 3. Presented at *Game Developers Conference*. March 25-29, San Francisco, CA.

[65] Jimenez, E. 2009. The Pure Advantage: Advanced Racing Game AI. [Online] http://www.gamasutra.com/view/feature/3920/the_pure_advantage_advanced_.php [Cited: 02/05/2013].

[66] Fitch, W. T. and Kramer, G. 1994. Sonifying the Body Electric: Superiority of an Auditory Display Over a Visual Display in a Complex, Multivariate System. In *Auditory Display: Sonification, Audification and Auditory Interfaces*. G. Kramer, Ed. Addison-Wesley, Reading, MA.

[67] Colantonio, R. 2012. In *Trade Secrets: Harvey Smith & Raphael.* D. Hindes, [Online] http://sneakybastards.net/shadesofgrey/trade-secrets-arkane/ [Cited: 02/05/2013].