A preliminary experiment to assess the fear value of preselected sound parameters in a survival horror game

Garner, Tom Alexander; Grimshaw, Mark Nicholas; Abdel Nabi, Debbie

Published in:
AM '10 The 5th Audio Mostly Conference

DOI (link to publication from Publisher):
10.1145/1859799.1859809

Publication date:
2010

Document Version
Early version, also known as pre-print

A Preliminary Experiment to Assess the Fear Value of Preselected Sound Parameters in a Survival Horror Game

Tom Garner
University of Bolton
School of Business & Creative Technologies
Deane Rd
BOLTON
United Kingdom
++44 1204 900600
garner8484@hotmail.com

Mark Grimshaw
University of Bolton
School of Business & Creative Technologies
Deane Rd
BOLTON
United Kingdom
++44 1204 903083
mng1@bolton.ac.uk

Debbie Abdel Nabi
University of Bolton
School of Health & Social Sciences
Deane Rd
BOLTON
United Kingdom
++44 1204 900600
a.nabi@bolton.ac.uk

ABSTRACT
This paper presents an experiment testing which sound parameters, in a survival horror game context, most warrant further investigation as a means to control the level of fear in such games. The experiment is part of a long-term study ultimately designed to support the development of a biofeedback procedural audio engine for computer games. By this means, it is hoped to provide an enhanced gaming experience whereby sound synthesis and audio processing is conducted in real-time according to the player's affect responses and emotional state. Results indicate that coarse manipulation of audio parameters has the potential to influence the intensity of the player's fear response whilst playing a survival horror game. Evidence is also presented that supports the integration of event logging and real-time participant vocal response into an experimental design to gather unbiased, quantitative data that can be associated with qualitative emotional response.

Categories and Subject Descriptors
H.5.1 [**Information Interfaces and Presentation**]: audio input/output.

General Terms
Human Factors, Measurement, Theory.

Keywords
Emotion, affect, perception, fear.

1. INTRODUCTION

1.1 What Do We Mean by Fear?

We describe an experiment that is part of an ongoing study on enhancing the perception of fear in survival horror games through the medium of sound. The aims of the study are to aid game sound designers in manipulating the player's perception of fear and, ultimately, to design a game procedural audio system that is capable of tracking the player's emotion valencies and intensities through biofeedback and that responds by adjusting pertinent parameters of audio samples and real-time synthesized sound in order to decrease or increase the level of fear. Such a system could theoretically be used to manipulate other emotions but we chose fear as the initial paradigm because of its close association to the survival horror game that is the genre most prevalent on the game engine we are using.

Before designing the experiment detailed here, we examined the concept of fear and perused literature on the relationship between sound and fear and related affect and emotion. This enabled us to do a preliminary selection of sound properties that were likely to be of most use in achieving our goal. Quite apart from properties of sound such as timbre, intensity and envelope, there are many other factors that can affect the player's perceptions such as culture and experience, the cross-modal relationship to vision and, in this case, the game's context. In this conception of meaning in sound, we follow the hermeneutic approach to language analysis espoused by, among others, Gadamer [1]. To investigate all of these is a lifetime's work and thus we limited our study to sifting through available evidence for the most potentially useful candidates in the context of the survival horror game.

The concept of fear, in fictional media such as computer games, may be broken down into the concepts of horror and terror which Varma [2] describes respectively as the "sickening realisation" and the "awful apprehension". We argue further, in our chosen context, that horror and terror cannot be present without the perception of a significant threat to the player's character. In immersive environments, players identify with their avatar (particularly in first-person games) in large part through the game's manipulation of emotion [3, 4]. Thus, a threat to the player's character becomes a threat to the player. Suspense and shock, both reliant upon time, play their part; horror is usually combined with shock and terror with suspense. In the former, we refer to any experience causing a pre-cognitive, involuntary fear reaction (recoil, scream, for example)

whilst the latter is shaped over time, typically growing in intensity. Often they are sequenced – suspense-terror preceding shock-horror – but they might also operate independently; part of the effectiveness of suspense-terror comes in holding back the climactic shock-horror. Once again, threat, or its perception, must be present for fear to be experienced – looking at an horrific image does not necessarily provoke fear. Finally, fear is often viewed as a negative experience but a number of recent authors describe it as a positive experience itself or one that leads to experiences such as pleasure [4] or excitement [5] and this suggestion echoes the thoughts of earlier authorities: Aristotle's concept of *catharsis*, de Quincey's characterization of serial murderers as sublime artists and Kant's identification of a sublime aesthetic of objects and actions more mundanely viewed as horrific and macabre. Indeed, such positive experiences, as Perron contends, may well provide the incentive to place 'oneself' in harm's way through playing a survival horror game.

1.2 Quantitative and Qualitative Properties of Sound

Sound is a critical component to consider when developing emotionality as it is directly associated with the user's experience of emotions [3, 12]. Parker and Heerema [13] suggest that sound carries more emotional content than any other part of a game. Grimshaw et al. [18] discovered that players felt significant decreases in immersion and gameplay comfort when audio was removed from gameplay; an assertion also made by Jørgensen [19] who, via observations and conversations with players, revealed that an absence of sound caused a reduction in engagement such that "the fictional world seems to disappear and that the game is reduced to rules and game mechanics". Foley sound design supports the emotionality of sound effects in creating both fantastic and everyday worlds. Ekman [11] describes how "often non-realistic sounds are purposefully used to make the action sound better". She exemplifies this process as "walking on cornstarch sounds much 'more real' on film than the actual sounds of walking on snow". Shilling et al. [3] quote industry professionals: "A game or a simulation without an enriched sound environment is emotionally dead and lifeless", implying that sound effects must be analysed in terms of their emotional qualities so that they may be implemented in a way that will maximise the audience's sensory experience

If we agree that sounds must be manipulated to maximise emotionality, it is reasonable to assume that specific game genres require specific audio 'emotioneering' [20]. Therefore the survival horror genre, most commonly associated with the emotion of fear, would require emotion-based sound design that strived to evoke fear [21].

As mentioned in section 1.1, there are many properties of sound that could be investigated as to their fear-inducing potential. Some are quantitative in that they can be objectively measured and applied to synthesis and audio processing whereas others are more qualitative; perception of sound meaning and import are influenced by factors such as culture, experience, context and expectation. We briefly survey the literature on the affective properties of sound in both classes particularly with reference to discomforting properties, the context of games or both.

Slaney [6] concedes that the dynamic characteristics of sound make it difficult to analyse using objective acoustical measurements. Nevertheless, several approaches have been documented that identify quantifiable sonic parameters that can be associated to a sound's emotionality. Cho et al. [7] provided evidence that pressure level, loudness and sharpness of a sound can directly affect emotional valence and intensity. Loudness and sharpness are admittedly, perceptual, psychoacoustic properties, however, using a model outlined by Zwicker and Fastl [8], such properties can still be measured to provide objective values. Moncrieff et al. [9] reference attack-decay-sustain-release (ADSR) as a quantifiable sound energy parameter showing a significant association between ADSR and specific emotional responses. Bach [10] documents the concept of increasing intensity as a measurable audio property that is psychoacoustic in nature via its intrinsic nature as a warning cue while signal to noise ratio [11] can also affect a sound's emotional impact because of ease of cognitive processing. Periodicity, tempo and rhythm have the potential to elicit substantial affect through audio-physiological effects such as entrainment wherein, according to Alves and Roque [12], a rhythmic simulation of a heartbeat, steadily increasing in tempo, has the potential to induce an increase in the heart rate of the listener. Parker and Heerema [13] suggest that an evolutionary survival instinct exists today that encourages humans to associate low-pitched sounds (growls and rumbles) with predators and consequentially experience fear in response to such a stimulus.

Reverberation is one regularly implemented effect that can affect a player's perception of the game environment [14]. An important function of the audio effect *delay* is to provide architectural and material information regarding the listener's environment: long reverberations and delays suggest reflective spaces that are large in comparison to the listener who can be made to feel quite small and lonely through this technique. Winer [15] documents how the application of frequency manipulation or equalisation (EQ) affects a sound's emotionality and aesthetic. Localization of a virtual object, although currently limited in terms of game implementation, has significant emotion-related potential [14, 16, 17]. The doppler-effect can also be measured objectively and manipulated to further create a more realistic illusion of position, direction and speed. Compression and normalisation techniques are used regularly across a multitude of audio applications; whilst their primary function is to limit erroneous sound pressure levels and create a more uniform audio stream, manipulation of such parameters creates noticeable differences to a sound's psychoacoustic properties and therefore begs investigation as an *emotioneering* parameter.

Section 1.1 detailed the characteristics of the sub-categories of fear: horror (associated with shock/surprise) and terror (suspense, anxiety and threat). Established literature describes implementation of this knowledge via a number of audio design techniques. Breinbjerg [22] posits that intentional ambiguity of a sound's source and location is critical to building suspense and terror, arguing that "[k]nowing that something is happening around the corner, without knowing precisely what it is, is most frightening". Breinbjerg also suggests that a "lo-fi" audio soundscape consisting of many interfering sounds can increase disorientation and decrease the player's perceived coping

ability. Kromand [21] exemplifies this by describing the implementation of sensory fillers (sounds irrelevant to gameplay) that nevertheless resemble sounds relevant to gameplay. This practice dissolves the barrier between diegetic and non-diegetic sound, consequently encouraging the player to cautiously treat every sound as a threat harbinger; suspense is characterized (in this context) as a more prolonged, less intense feeling of terror. Kromand [21] suggests that this can be achieved via a system of audio 'warning' cues that steadily reveal localization and movement information. He argues that the consequentially slow rising of intensity, plus no clear indication of when the inevitable shock will occur, manifests as suspense for the player. Parker and Heerema [13] propose that acousmatic sounds perceived as threatening increase the sensation of terror: "A prey animal that can only hear the predator is in an unknown amount of trouble, and it pays to believe the worst". Schwarz and Winkelman [23] argue that positive value judgements of audio strongly correlate with the ease with which they can be processed. Inverting this argument supports the notion that a sound that is difficult to identify, localize, and/or apply semantic meaning to evokes negative judgements.

Shock-horror requires a different approach. Despite Alfred Hitchcock's famous objection to shock (often referred to as 'cheap' and 'simplistic') it remains a hallmark of the survival horror game genre. The most frightening part of the original *Resident Evil* (Capcom, 1996) is arguably the shocking moment when two mutant dogs jump through a window to attack the player's avatar. Xu et al. [24] state that an audio shock is most effective when it is preceded by silence – a technique utilized in the aforementioned example. Cho et al. [7] insist that acoustical properties of audio (specifically intense loudness and sharpness) can produce quantitative increases in negative emotional valence. These sonic characteristics are typically descriptive of audio designed to shock. Kromand [21] details a deceptive technique that can be arguably associated to shock. This technique first establishes a sonic convention that aids player survival (Kromand uses the radio from *Silent Hill 2* [Konami, 2001] as an example) then intentionally defies this convention and morphs the semantic meaning of the sound from supportive to antagonistic. Cox [25] tested various sounds assumed to be 'disgusting' and 'horrible' in nature suggesting that (mainly as a result of cultural factors) individual sounds can have distinctly different levels of perceived disgust. There appears to be a fine line between the disgusting and the horrific and, although Cox suggests that a sound can be exclusively either, it seems reasonable to assume that perceived disgust will impact upon an overall sensation of horror when combined with a perceived threat. Parker and Heerema [13] describe third-person audio cues as distinctly horrific in nature. They use a human scream as an example, asserting that "[a]s humans we tend to react with emotional similarity when we hear such sound, not in sympathy so much as in fear of whatever is inflicting pain or fear on the other". In this example the sound is not only shocking due to its sudden, sharp and intense acoustic quality, but also horrific in that it implies the presence of a horrific creature and/or act.

2. METHODOLOGY

2.1 Preliminary Testing

Table 1 represents a range of sound properties and effects organized according to their objective/subjective parameters, their potential (based on the literature review) for inducing different types of fear and the ability for a game procedural audio engine to manipulate. Several preliminary trials were conducted using the same game level and selection of sounds used for the experiment described below. These preliminary trials utilized the same procedure that is outlined in section 2.6 below and similar equipment was used but data was collected entirely using self report. In these trials, participants were asked to complete a modified version of the questionnaire that was to be used in the main experiment. 20 individuals participated and 3D positioning, distortion, chorus/modulation, equalization, loudness, reverberation, stereo panning, ADSR, dissonance, and pitch were selected as treatments. Each treatment was applied to a separate sound and players compared the original to the treatment once in each trial. Mean participant results revealed, 3D positioning (particularly sound coming from a sharp left or right), pitch (particularly high pitched sound) and loudness (specifically greater relative loudness) to be notably effective in increasing participants' perceived intensity ratings. These three treatments were consequently selected for the experiment and are shown in red in Table 1.

Table 1. Potential affective properties of sound

	Semantic Association	Control via software DSP effects	Control via sound selection / position / timing / combinations
Horror-Shock	Contact Character Disgust Dynamics Function Material Shape Size Uncanny Geographic location	3D Positioning ADSR (loop/ambient sound) Chorus/modulation Dissonance Distortion Equalisation Loudness Pitch Reverb Simulated binaural Stereo panning	Acoustic spectrum ADSR (one shot) Sharpness Sound on sound
Terror-Suspense	Contact Character Disgust Dynamics Function Material Shape Size Temporal data Uncanny Geographic location	3D Positioning ADSR (loop/ambient sound) Chorus/modulation Delay Dissonance Distortion Doppler effect Equalisation Loudness Pitch Mix/Key-sound clarity Reverb Simulated binaural Stereo panning	Acoustic spectrum ADSR (one shot) Complexity Entrainment MMN Periodicity Rhythm Tempo Sharpness Sound on sound

2.2 Preparation of Sounds

The 5 sounds utilized in the experiment are all taken from the source engine originally created for Half-Life 2 (Valve, 2004). In its untreated state, each sound is presented as a single monophonic channel. In addition to the 5 test sounds, avatar footstep and vegetation rustling sounds can also be heard during gameplay.

2.3 Game Level Design

A bespoke game level was judged to be the most appropriate choice of presentation medium for the sounds. Because the specific interest of this research is to develop the audio within a survival horror computer game, contextualization is therefore the key to producing results with (virtual) real-world validity. Whilst this method could allow several non-sonic variables (particularly audio/visual synchresis and gameplay-related emotional experience) to impact upon the results, it should be acknowledged that any correlations/data patterns drawn from this experimentation must be observed within the context of a computer game as this is the only environment in which the research aims to apply gained knowledge.

The custom level was built using the unmodified *Cry Engine 2* (Crytek, 2007) game engine and sandbox level editor. Although the game engine supports third person and first person perspective play, research suggests that a first person display can increase the sense of urgency and immersion [26]. The avatar is not completely absent however, and (in traditional First-Person Shooter (FPS) style) visible forearms, hands and a pistol are outstretched into the virtual world. The level was non-linear with no suggested direction and could be completed by reaching one of three evacuation points.

Figure 1. Overhead view of level.

To achieve the desired aesthetic and encourage any negative player valence to be fear-related, certain survival horror conventions were utilized, including a night-time setting and a dense forrest environment. Near-zero visibility without the aid of a flashlight restricts the field of vision [27], creating large volumes of "blind space" [28]. In keeping with not only survival horror convention but also traditional FPS formats, the player was pursued during the level by an unknown creature which facilitates the *hunter & hunted* principle [29]. This creature was, however, only implied through the narrative in the level introduction and the sounds heard during gameplay.

Control layout was addressed to support gameplay accessibility and increase the chance of participants using tacit knowledge to control their avatar and keeping their focus more explicitly on the sound, graphics and atmosphere [30]. The default controls followed the standard setup found on most FPS games and the participant was given the opportunity to customize the controls before playing. The audio for the level differs depending on the level type. Type 1 used untreated audio whilst type 2, 3 and 4 used treated audio (pitch shift, 3D and loudness respectively). All level types housed the same group of 5 source sounds (a distant zombie call, a near-by twig snap, a woman's scream, a monster's attack scream and a sudden distorted monster scream) activated by a series of proximity triggers built in concentric circles. Regardless of gameplay, a minimum of 5 seconds of silence was guaranteed between sounds. All sound points were fixed and always produced the same sound (not accounting for treatment variations). All sounds within a specified type were treated with equal parameter settings of the same DSP process). This treatment was one of the following: 3D (Binaural processing placing the sound to the right side of the player), loudness (an intensity increase of 25dB), or pitch (300 cents rise in pitch compared to the untreated sound). Given that this is a preliminary experiment to assess which factors, easily processed by a game audio engine, might be most emotive in the survival horror game context, such differences were designed to be noticed without being too obvious – the fine-tuning is for later experiments.

2.4 Environment and Game Equipment

The game level ran on a bespoke 32-bit PC with *Windows Vista* (Service Pack 2) operating system, *AMD Phenom 2* (3.2GHz) quad core processor, 8GB RAM, ATI *Radeon* 4850 (1.5GB) GPU. At time of writing, this is a mid-level gaming specification PC able to run most new release games at medium/high settings. The PC monitor was a *LG*, 22" LCD screen, supporting the game level's 1920x1080 (full HD) graphics resolution. This configuration was designed to resemble a typical consumer home setup that was powerful enough to run a game representative of current gaming technology, whilst avoiding an elite specification that would be likely to exclude the majority of the casual gaming community. The testing was executed in a small studio space, providing natural light and a glass partition window through which participants could be observed without disruption.

The sound was processed and reproduced via an *Asus Xonar* 7.1 sound card and *Tritton AX Pro* 7.1 headphones. It has been suggested that the choice of headphones or speakers could be a significant contextual variable [31] particularly in terms of localization and immersion [14] and impact. In a comparable study, Murphy and Pitt [32] show a preference towards headphone use, arguing that it "…enables the designer to incorporate more complex sound objects whose subtleties will not be lost due to background noise, speaker cross-talk, etc". We agree that headphones produce a more immersive experience, and the commercial availability of a range of headphones (many specifically designed for computer games) suggests that headphone use is common in a player's 'natural environment'.

2.5 Participants

Similar experiments in related fields of study reveal a large range of participant numbers, with smaller numbers ranging from 15-25 and larger numbers reaching 100. Although practical constraints for this experiment set the participant number at 12, a number of relevant published experiments reveal that statistical significance is possible with relatively small participant numbers [33, 34, 35]. The 12 participants each experienced a different order of the 4 level variations (untreated

audio, pitch shift, 3D surround and loudness increase). This structure was implemented to reduce order effects which have been identified as a further possible cause of bias [34]. All participants were students or recent university graduates aged between 18 and 49, 9 male and 3 female. Participants were asked for their gender, age, ethnic background, game playing experience and if they suffered from any visual or hearing impairments.

2.6 Procedure

Before playing, each participant was given a brief detailing the exact procedure along with game instructions and control information. Participants were aware that they needed to rate the emotional impact of a sound, but not that fear (or negative valence) was under investigation. Participants were required to provide their own single word descriptors to illustrate the emotion they perceived, thereby not biasing subjective response towards fear. The game level took between 50 and 140 seconds to complete and each participant played 3 variations which, including the brief and debrief time, set the total typical completion time at 10 minutes for 1 audio property. Testing four separate treatments in a single sitting would take approximately 55 minutes (allowing 5 minute breaks between each treatment test). The debriefing questionnaire required immediate response after each play-through, followed by a more detailed set of questions to be answered after the last level was completed.

2.7 Data Collection

Moffat and Kiegler [33] argue that, although "physiological measurements … can be valuable in helping to read the emotional state of game players", the links between emotion and physiological response are currently unreliable and psychophysiological data collection alone cannot provide a complete account of a participant's emotional state. An overview of psychophysiology suggests that quantitative response measurement (heart-rate, galvanic skin response, electromyography, etc.) is capable of providing accurate emotional valence and intensity data [36] but cannot distinguish between different emotional states of the same valence. Research has attempted to counter this problem via near-simultaneous collection and correlation of objective physiological response and subjective player responses [10, 18].

Cacioppo [36] admits that "specific types of measurement of different physiological responses…are not by themselves reliable indicators of well-characterized feelings", suggesting that empirical data must be cross-examined alongside additional data sources. To provide supporting data, direct participant opinions were collected using a real-time vocal response system. A software based digital audio workstation (Pro Tools LE 8) synchronized to the game engine recorded participants' vocalized input via the integrated headset microphone whilst they played the game. The initial brief requests that the player rates the 'emotional impact' of each sound heard using a specified scale and then communicates that score vocally. During gameplay (all types) a visual prompt [1-2-3-4-5] appeared on the screen for 2 seconds immediately after a key-sound was triggered. The headphone setup (section 2.4) recorded the vocal responses via a microphone integrated into the headset. Audio data was recorded as a separate channel and synchronized to a video recording of the in-game performance. The exact time of the vocal response was recorded as text data in the game event log. The rationale for this approach comes from 2 concepts: memory and flow. Rugg and Petre [37] argue that a participant's explicit knowledge regarding information they have recently received is stored in short-term memory (STM), requiring rehearsal or meaningful association to migrate towards long-term memory (LTM). Whilst it may be a fair assumption that an intense emotional response could facilitate such a memory translation there are no guarantees, and the possibility that participants could forget a number of the sounds by the end of the level presents a genuine risk. Conversely, requesting that the participant break from the game to respond immediately to each stimulus severely diminishes the potential for flow. Flow is central to attention and consequently a break in flow negates immersion [38, 39]. Because the intention is to evaluate sounds' emotive potential within a game environment it is vital, in order to achieve contextual validity, that the player feels they are playing a game. Real-time vocalization is an attempt to find a middle-ground between the two extremes. The number of audio samples used obeys *Miller's law* [40] and subjects were asked to rate each sound using a 5 point scale (1=least emotive, 5=most emotive), in keeping with recommendations suggested by recent research [41] and experimentation [17]: subjects spoke or shouted the appropriate number while playing in response to the visual prompt [1-2-3-4-5].

Freeze, fight and flight are response actions associated with fear-inducing stimuli [42]. Perron [43] argues that such response actions can be applied to the experience of fear in a computer game. Reversing this, an analysis of player action and performance might reveal insight into their emotional state. Wolfson and Case [44] suggest that emotional arousal can greatly impact upon performance: "[W]hen highly aroused, people tend to be faster but less accurate, and they focus mainly on the most salient aspects of a task". Currently there is no existing framework correlating player performance and fear-related arousal; a broad analysis of performance may help further support the other data (user-input, psychophysiological response) and provide a launching pad for further study within this specific area. To this purpose, *Fraps* (version 3.2, 2010) real-time video capture software was implemented to provide a complete visual recording of each participant's actions within the game.

The debrief questionnaire requested only explicit knowledge from participants and was in 3 sections. Section A required participants to provide individual words that they felt reflected the atmosphere of the game and the emotional content of the sounds and then to rate the perceived 'scariness' and difficulty of the level overall using 5 point ordinal scales. This section was answered immediately after the player had completed each game level and was repeated for each play through. Section B requested each participant to rank the 3 levels in order of perceived 'scariness' and provide quality control information regarding sense of immersion, flow and general game experience using the same 5 point scale system. Section C asked the participants to state how often they played computer games, if they suffered from any visual or hearing impairment, their age, gender, nationality and country of origin. Subjects' participation and the data collection were conducted in accordance to the University's Research Ethics Framework.

3. RESULTS

Each participant (N=12) completed the four game segments featuring the four alternative sound treatments, completing questionnaire sections between levels and provided demographic and further data in the debrief at the end of the session. All participants were White British, 9 male and 3 female, with ages ranging from 18 to 55. During the debrief participants rated the immersive quality of their overall experience and how disruptive to flow providing the real-time audio responses were (**table 2**).

Table 2. Means/Standard deviations of participant ratings of immersion and flow disruption

	N	Range	Min.	Max.	Mean	SD
Immersion	12	2.00	2.00	4.00	3.25	.621
Disruption	12	2.00	1.00	3.00	2.33	.7785

To test the statistical significance of the results, one-way repeated measures ANOVA was employed via PASW/SPSS (v.18), with sound treatment type as the within subject factor for the dependent variable data collected. The dependent variable classes tested were completion time and 3 measurements associated with player in-game use of the run function (total activation number / total time run was activated / mean run time).

Table 3. Means (and standard deviations) of DV measurements for the four sound modalities

	Sound treatment factors			
DV measures	**Original**	**Pitch**	**Surround**	**Loudness**
Completion time	244.348 (184.823)	136.735 (89.453)	217.997 (182.061)	178.137 (114.879)
No. of RUN activations	2.08 (1.929)	2.67 (3.393)	1.42 (2.021)	2.17 (2.823)
Total RUN time	21.606 (23.894)	29.969 (34.836)	23.856 (32.757)	23.41 (32.621)
Mean RUN time	13.518 (19.97)	15.324 (20.206)	15.028 (24.56)	17.744 (26.563)

Mauchley's test of the DV measurements completion time (χ^2 (5) = 7.157, p > .05) and total run time (χ^2 (5) = 9.529, p > .05) signified that the assumption of sphericity had been met. However, no. of run activations (χ^2 (5) = 13.108, p < .05) and mean run time (χ^2 (5) = 123.203, p < .05) indicated a violation. The Greenhouse-Geisser estimates of sphericity were employed to correct the degrees of freedom (ε = .671, ε = .598 respectively). One-way repeated measures ANOVA revealed no statistically significant difference between the four sound treatment factors when measuring no. of run activations ($F_{2.041, 22.15}$ = .954, p > .05), total run time ($F_{3, 33}$ = .953, p > .05) or mean run time ($F_{1.794, 19.736}$ = .608, p > .05). Completion time did indicate significance ($F_{3, 33}$ = 2.888, p < .05) revealing that the objective DV measurement of level completion time was significantly affected by the different sound treatment factors. Measurements associated with the run function were tested further to assess correlation between use of the run function and RTI. The Spearman's Rank correlation (two-tailed) was selected to test the relationship between these variables ($R_{(RTI, number of run activations)}$ = -.483, p < .001, $R_{(RTI, total run time)}$ = -.634, p < .001, $R_{(RTI, mean run time)}$ = -.55, p < .001) indicating a moderate/strong *RTI-total time* correlation, a weak *RTI-number of runs* and a moderate *RTI-mean run time* correlation. The debrief questionnaire also revealed variation in PC player experience and confidence (PEC) *prior to testing* between participants. These grouped differences were evenly distributed and ranked (25%, R = 1-4) across the total sample. Multivariate variance analysis (MANOVA) assessed the significance of variation between PEC groups when measuring completion time and RTI ($F_{(3, 44)}$ = 20.616, p < .001). Bonferroni post hoc revealed specific significant difference lay between PEC levels 1 and 3 ($\bar{x}_1 - \bar{x}_3$ = 247.831, p < .001), 1 and 4 ($\bar{x}_1 - \bar{x}_4$ = 233.606, p < .001), 2 and 3 ($\bar{x}_2 - \bar{x}_3$ = 216.728, p < .001) and between levels 2 and 4 ($\bar{x}_2 - \bar{x}_4$ = 202.503, p < .001) with completion time the dependent variable. Comparable results were obtained in post hoc with RTI the dependent variable ($\bar{x}_1 - \bar{x}_3$ = 1.55, p < .001), ($\bar{x}_1 - \bar{x}_4$ = 1.8, p < .001), ($\bar{x}_2 - \bar{x}_3$ = 1.217, p < .001), ($\bar{x}_2 - \bar{x}_4$ = 1.467, p < .001).

Comparisons between real-time intensity averages and debrief intensity ratings averaged across all treatment types revealed a moderate/strong correlation via Spearman's rank coefficient ($R_{(RTI, debrief intensity)}$ = .694, p < .001), suggesting that although players reported similarly both during and after playing the game, the 30% margin of error arguably supports the use of vocal intensity responses collected during the game. The final analysis tests searched for difference (measured in RTI) between the game level source-sounds (**table 4**). A series of Freidman's tests of variance for repeated measures were employed to identify significant difference between source-sounds and search for any interaction between source-sound factors and sound treatment factors.

Table 4. Results of Friedman's test series

Sound Name	$\bar{x}$ RTI Untreated	$\bar{x}$ RTI Pitch	$\bar{x}$ RTI 3D	$\bar{x}$ RTI Loud
Zombie Call	2.75	2.63	2.25	1.81
Twig Snap	1.38	1.56	1.69	1.69
Woman Scream	2.69	3.31	3.44	3.38
Monster Attack	3.75	3.38	3.31	3.69
Intense Scream	4.44	4.13	4.31	4.44

Results revealed significant difference between source sounds across all treatments: untreated (χ^2 (4) = 20.593, p < .001), pitch (χ^2 (4) = 16.964, p < .01), 3D (χ^2 (4) = 16.176, p < .01), loudness (χ^2 (4) = 22.545, p < .001) and when tested across all treatments (χ^2 (4) = 71.891, p < .001). As with earlier tests, no significant difference was identified between treatment type (χ^2 (3) = 2.123, p > .05).

4. CONCLUSIONS AND FUTURE WORK

Whilst this paper has identified a great number of audio parameters that have the potential to affect player intensity response, the scope has currently limited this study to 3 sonic parameters. The results obtained present no significantly conclusive evidence to support the hypothesis that pitch alteration, decibel level or binaurally processed panning techniques affect player intensity response. One initial possible explanation is the set levels for the 3 treatments were too conservative to trigger significantly different intensity readings and that a more focussed analysis on an individual parameter across a greater range of treatment values would be of profit to further investigations. Within the remit of a preliminary test, such lack of significant conclusiveness is not unexpected.

Incorporation of real-time audio responses and game engine event logging allowed an accurate synchronisation of several data sets and provided an opportunity to analyse data at (and around) specific points of gameplay. Testing sound treatment modalities via various measurements of in-game avatar *running* action revealed no significant difference between groups for any of the 3 measures (number of run activations, total time spent with run engaged, average length of run time). Several possible reasons lie in alternative player motives for activating the run function (accelerating the player once the level exit has been visually identified in order to vary gameplay when the player experiences frustration or boredom) and inexperienced player difficulties in coordinating both run and movement controls simultaneously. Event logging provided a data filtration system, only recording data if certain criteria were met. Parameters of this logic-gate were set to record run associated data only between activation of a key sound and for the subsequent 5 seconds of gameplay (the logic suggesting that any run activity occurring during this time would be more likely to be a response to the key sound). This system unfortunately did not reveal any further insight due to too little data recorded (under such stringent filters) to perform any statistical analysis, suggesting that the run function was not used as an evasive player movement in reaction to any key sound. The Spearman rank correlation test provided further analysis data regarding possible relationships between the *emotional intensity ratings* (RTI) and the 3 *run-function* measurements. Results identified a strong link between the participants' RTI and *total run time*, and a moderate relationship between RTI and *mean run time*; suggesting that analysis of the run function may have future potential as an objective measurement.

Results of the data analysis also suggest that other critical factors were affecting the dependent variables, most notably player experience and confidence (PEC) prior to experimentation. As we previously asserted, fear cannot be experienced without a genuine perception of threat and that perception can alter in both presence and intensity depending on the appraised severity of threat, and the individual's capacity to overcome that threat. It could therefore be further asserted that high levels of experience (knowledge of game conventions, a variety of fear induction tactics, etc.) coupled with confidence and adept skill in game controls is very likely to reduce the threat severity and increase the coping ability. The test results presented above support this, presenting a highly significant negative correlation between PEC and RTI. At both a qualitative and quantitative level, the testing revealed that players with very little gaming experience were more likely to struggle to reach the level exit when they were feeling more intense fear, to an extent that they reported frustration and dislike towards the game. Quantitative data analysis supported this finding, revealing a significant correlation between RTI and completion time. The exact nature of completion time as a variable becomes a matter of opinion, with a logical assertion being that because sensation of fear-intensity positively correlates with completion time, which variable is the cause and which is the effect may fluctuate throughout the game. Taking longer to reach the level exit may increase negative emotional valence (worry, fear, frustration) which in turn creates a feedback loop as increased negative valence causes the player to lose their way or begin travelling in circles (as we observed in several of the gameplay capture videos). It could also be asserted that completion time impacts upon the potential of a sound to evoke increased fear response. Existing literature has suggested that forewarning cues that denote a frightening event is imminent (such as extended periods of unexpected silence) may significantly increase the impact of a sudden sound [27] and increased completion time dictates a greater mean time between key sound events. Such an aspect of sound design shows great potential to manipulate player fear response and is certainly a candidate for further study.

The nature of the experimental design afforded the opportunity to run statistical analysis of variance between each of the 5 alternative sounds heard in every play-through. Because each player reported an emotional intensity rating for each sound for four repetitions (across the 4 treatment types) order effects can arguably be dismissed alongside audio/visual interaction effects due to the repetitive and low visibility graphic environment. Results posited that despite contaminating variables (alternative audio treatments, PEC) a significant difference in RTI existed between groups and an ordinal rank and specific mean differences between each sound was also identified. Such findings provide a strong argument for the value of game sound in manipulation of player emotional states and calls for the continuation of this research line of enquiry to establish exactly what sonic differences between these sounds caused these significant and substantial variances in fear-response intensity.

The results of this paper confirm that differences between sound parameters can affect the degree of intensity an individual experiences whilst playing a survival horror game. The specific measures tested (3D, Loudness and Pitch), although not statistically significant, do reveal potential if subject to greater parameter extremes. Although not formally analysed via statistical data, qualitative interpretation posits periodicity (specifically, the length of silence experienced before a key sound) as a good candidate for individual study. The statistically

significant difference between source-sounds identifies timbre and attack (ADSR) as strong potential candidates for further study. Future experimentation will explore each of the above sound parameters individually, assessing each parameter across a detailed number of measures rather than a single measured difference between the treatment group and the control. Psychophysiological measurements to increase objectivity will also be integrated into all further experiments.

It could be posited that these results support the notion that the experience of fear is (at least within the confines of this context) a complex matrix of interacting variables. Whilst real-time intensity response and event logging have proved substantially valuable the exact execution of this approach requires minor alteration. Real-time audio responses collected during periods of silence in the game in addition to immediately after each key-sound could provide a more detailed account of player emotion at the risk of further interrupting immersion and flow. The substantial interference effect of player experience and confidence prior to testing is acknowledged and future experiments should endeavour to focus testing upon a single PEC group or balance the groups and increase the sample number to compensate.

5. REFERENCES

[1] Gadamer, H-G. 1976. *Philosophical Hermeneutics* (trans. D.E. Linge). University of California Press, Berkeley.

[2] Varma, D. 1966. *The Gothic Flame*. Russell and Russell, New York.

[3] Shilling, R., Zyda, M., and Wardynski, E. C. 2002. Introducing emotion into military simulation and videogame design: America's Army: Operations and VIRTE. In *Game On*, London.

[4] Perron, B. 2005. Coming to Play at Frightening Yourself: Welcome to the World of Horror Video Games. In *Aesthetics of Play* (Bergen, Norway, October 14-15, 2005).

[5] Svendson, L. 2008. *A Philosophy of Fear.* Reaktion books.

[6] Slaney, M. 2002. Semantic–Audio Retrieval. In *IEEE International Conference on Acoustics, Speech and Signal Processing* (Orlando, Florida, May 13-17, 2002).

[7] Cho, J., Yi, E., and Cho, G. 2001. Physiological responses evoked by fabric sounds and related mechanical and acoustical properties. *Textile Research Journal*, 71, 12, 1068-1073.

[8] Zwicker, E. and Fastl, H. 1999. *Psychoacoustics – Facts and Models*, 2nd edition, Springer-Verlag.

[9] Moncrieff S.,Venkatesh S., and Dorai C. 2001. Affect computing in film through sound energy dynamics International Multimedia Conference. In *Proceedings of the ninth ACM international conference on Multimedia*, Vol. 9, 525-527.

[10] Bach, D. et al. 2009. Looming sounds as warning signals: The function of motion cues, International. *Journal of Psychophysiology*, 74, 1, 28-33.

[11] Ekman, I. 2008. Psychologically Motivated Techniques for Emotional Sound in Computer Games. In *Audio Mostly 2008* (Piteå, Sweden).

[12] Alves, V. and Roque, L. 2009. A Proposal of Soundscape Design Guidelines for User Experience Enrichment. In *Audio Mostly 2009* (Glasgow).

[13] Parker J. and Heerema J. 2007. Audio Interaction in Computer Mediated Games. *International Journal of Computer Games Technology*, Vol. 2008, Article ID 178923.

[14] Grimshaw, M. 2007. The Resonating spaces of first-person shooter games. In *Proceedings of the 5th International Conference on Game Design and Technology* (Liverpool, November 14-15, 2007).

[15] Winer. 1979. The Art of Equalisation, *Popular Electronics Magazine*, http://www.ethanwiner.com/equalizers.html.

[16] Sonnadaraa, R, Alainb C, and Trainora L. 2006. Effects of spatial separation and stimulus probability on the event-related potentials elicited by occasional changes in sound location. *Brain Research*, 1071, 175-185.

[17] Steele, D.L. and Chon, S.H. 2007. A Perceptual Study of Sound Annoyance. In *Audio Mostly 2007,* 9-24.

[18] Grimshaw, M., Lindley, C.A., and Nacke, L. 2008. Sound and Immersion in the First-Person Shooter: Mixed Measurement of the Player's Sonic Experience. In *Audio Mostly 2008* (Piteå, Sweden).

[19] Jørgensen, K. 2006. On the Functional Aspects of Computer Game Audio. In *Audio-Mostly 2006* (Piteå, Sweden).

[20] Freeman, D. 2003. *Creating Emotion in Games.* New Riders Games.

[21] Kromand, D. 2008. Sound and the diegesis in survival-horror games. In *Audio Mostly 2008* (Piteå, Sweden).

[22] Breinbjerg, M. 2005. The Aesthetic Experience of Sound – staging of Auditory Spaces in 3D computer games. In *Aesthetics of Play* (Bergen, Norway, October 14-15). http://www.aestheticsofplay.org/breinbjerg.php.

[23] Reber, R., Schwarz, N., and Winkielman, P. 2004. Processing Fluency and Aesthetic Pleasure: Is Beauty in the Perceiver's Processing Experience? *Personality and Social Psychology Review*, 8, 4, 364-382.

[24] Xu, M.,Chia, L., and Jin, J. 2005. Affective content analysis in comedy and horror videos by audio emotional event detection. In IEEE International Conference on *Multimedia and Expo.*

[25] Cox, T. 2007. Bad vibes: an investigation into the worst sounds in the world. In *19th ICA Madrid, PPA-09-003*.

[26] Calleja, G. 2007. Revising immersion: A conceptual model for the analysis of digital game involvement. In *Situated Play* (University of Tokyo, Japan, September 24-28, 2007).

[27] Perron, B. 2004. Sign of a Threat: The Effects of Warning Systems in Survival Horror Games In *COSIGN 2004*. http://www.cosignconference.org/cosign2004/papers/Perron.pdf.

[28] Bonitzer, P. 1982. *Le champ aveugle. Essais sur le cinéma*, Cahiers du cinéma and Gallimard, Paris.

[29] Grimshaw, M., and Schott, G..2008. A Conceptual Framework for the Analysis of First-Person Shooter Audio and its Potential Use for Game Engine. *International Journal of Computer Games Technology*, Vol. 2008, Article ID 720280.

[30] Cunningham, S., Grout, V., and Hebblewhite, R. 2006. Computer Game Audio: The Unappreciated Scholar of the

Half-Life Generation. In *Audio-Mostly 2006* (Piteå, Sweden).

[31] Cox, T. 2008. Scraping sounds and disgusting noises. *Applied Acoustics*, 69, 12, 1195-1204.

[32] Murphy, D., and Pitt, I. 2001. Spatial sound enhancing virtual story telling. *Lecture Notes in Computer Science*, 2197, 20-29.

[33] Moffat, D.C. and Kiegler, K. 2006. Investigating the effects of music on emotions in games. In *Audio-Mostly 2006* (Piteå, Sweden).

[34] Nacke, L. and Lindley, C. 2008. Boredom, Immersion, Flow - A Pilot Study Investigating Player Experience. In *IADIS Gaming 2008: Design for Engaging Experience and Social Interaction.*

[35] Ullsperger P., Erdmann U., Freude G., and Dehoff W. 2007. When sound and picture do not fit: Mismatch negativity and sensory interaction. In *International Journal of Psychophysiology*, 59, 3-7.

[36] Cacioppo, J.T., Tassinary, L.G., and Bernston, G.G. 2000. *The Handbook on Psychophysiology.* Cambridge University Press.

[37] Rugg, G. and Petre, M. 2007. *A Gentle Guide to Research Methods.* McGraw-Hill Education, Open University Press.

[38] Csíkszentmihályi, M. 1990. *Flow: The Psychology of Optimal Experience.* Harper and Row, New York.

[39] Brown, E. and Cairns, P. 2004. A Grounded Investigation of Game Immersion. In *Conference on Human Factors in Computing Systems,* 1297-1300.

[40] Miller, G.A. 1956. The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological Review*, 63, 2, 81-97.

[41] Gillham, B. 2000. *Developing a Questionnaire.* Continuum, London.

[42] Bracha, S.H. 2004. Freeze, flight, fight, fright, faint: Adaptationist perspectives on the acute stress response spectrum. *CNS Spectrums,* 9, 9, 679-685.

[43] Perron, B. 2005. A Cognitive Psychological Approach to Gameplay Emotions. In *DiGRA 2005: Changing Views – Worlds in Play.*

[44] Case, G. and Wolfson, S. 2000. The effects of sound and colour on responses to a computer game. *Interacting With Computers*, 13, 2, 183-192.