Gabriel Sauvage

UFR Ingémédia, Pro TSI license, NTS course

University of Toulon

Professional License Dissertation

The relationship between sound and the state of Flow in the gaming experience

Work carried out under the direction of Jean-Michel DENIZART

Month and year of defense: July 2021

Gabriel Sauvage

UFR Ingémédia, Pro TSI license, NTS course

University of Toulon

Professional license dissertation

The relationship between sound and the state of Flow in the gaming experience

Work carried out under the direction of Jean-Michel DENIZART

Month and year of defense: July 2021

The relationship between sound and the state of Flow in the gaming experience

Abstract (2000 characters)

This thesis focuses on the relationship between sound and the state of Flow within a video game experience. The “flow”, a psychological state theorized by Mihály Csikszentmihályi in 1975 then explored in particular in his book “Flow: The Psychology of Optimal Experience”, would constitute according to the author “the optimal experience”, that which a human being seeks at through any practice or activity whatsoever. The frequent achievement of this state would then be directly correlated to the happiness and psychic well-being of an individual.

In this thesis, I explore how sound designers contribute to recreating this state of flow in the gaming experience, through its sound direction. First, I make an introduction to flow and undertake to define this state. Then, I endeavor to underline the distinction between “immersion” and “flow”, two different states and yet interconnected in their functioning and conditions of appearance. Following this, I explain how the sound aspect of a game constitutes the pillar of the appearance of the state of flow, and how the integration of sound design in the game design process, as well as certain uses of the characteristics of a sound, make it possible to facilitate the appearance of the state of flow. I then talk about the intrinsic relationship between “game feel”, sound and flow. Then, I address the question of the need for sound interactivity vis-à-vis the appearance of the state of flow, via (in particular) the use of music, the interpretation of game data in real time, the arrangement of sound materials in the form of proactive and reactive systems, and finally budgetary issues related to the design of these systems. I then explore the different dangers of breaking the Flow related to the sound direction, and conclude with the following question: "Without the image, is the sound enough to reach the flow?", before proceeding to a conclusion .

Keywords: Flow, Video Game, Gaming Experience, Sound Design, Composition, Sound Design, Interactivity, Sound Interactivity, Game Feel, Wwise

Thanks

Thanks to Arnaud Roy, Masami Komuro, Joonas Turner, Christophe Héral, Paul Weir and finally Bjorn Jacobsen, for having, in turn, advised, guided and provided solid resources that greatly contributed to the conception of this dissertation. Thanks also to Jean-Michel Denizart, my thesis director, who helped me design the common thread of my interviews with the members of the professional body mentioned above, and guided me throughout this work.

Contents

Preamble

1. Flow, introduction and definition

2. Distinction between "Immersion" and "Flow"

3. Sound, pillar of Flow

A - The technical and human issues related to the realization of the sound design of a video game
B - The sound direction conceived as an integral part of the game design
C - The use of sound characteristics as a vector of flow

4. *Game Feel*, Flow, Sound: an intrinsic relationship

A - Definition of "Game Feel"
B - Sound design at the center of the game feel
C - Identification with the avatar as a motor of flow

5. The implementation and design of sound interactivity, cement of Flow

A - The implementation of fluid and natural transitions
B - The use of interactive music as a vector of flow
C - The interpretation of game data as a driver of interactivity
D - The need to design a modular and flexible implementation
E - Optimization of the sound implementation and its systems
F - The budget, an obstacle to complex sound integration depending on the scale of the productions
G - The decorrelation between complexity and efficiency of a sound implementation

6. The Dangers of Flow Breakage Related to Sound Direction

A - Abuse of sound stimuli and fear of heights (talking about silent hill)
B - The question of digital latency, separating the input from its sound response
C - Ignorance of psychoacoustic phenomena
D - The concept of nuisance score and the unit of measurement "NUFS"
E - Issues of decorrelation between a sound event and its intention

7. Without the image, is the sound enough to reach and maintain the state of Flow?

Conclusion

Introduction

This thesis is a work carried out within the framework of a professional license at the University of Toulon, within the New Technologies of Sound course. At a time when the attention potential of the average individual is in free fall, and when any type of media struggles to capture and maintain the interest of the viewer/reader/player, it is appropriate to take an interest in the phenomena and tools to preserve it. Since the beginnings of cinema and television broadcasting, we have seen that the image-sound ratio was one of these tools. But naturally, the video game stands out here by its interactive dimension, allowing to adopt a point of view and a posture impossible in any other form of media. At a time when the video game has just celebrated its 46th birthday (if we trace the origin of its democratization with the sale of the Atari Pong in 1975), it remains despite everything a very recent medium. The questions of perception of a video game work in the player, as well as the mechanisms put in place to arouse his interest in the latter, constitute as many subjects filled with gray areas, waiting to be clarified. To understand the relationship between the player and the video game, it is then necessary to go back to what binds humans to entertainment, until coming to question the sensory experience that it constitutes.

Through this dissertation and the answers of professionals who have been questioned on the subject, we will try to understand the phenomenon of flow, its application in a video game experience and how the use of sound can facilitate, even encourage its appearance. First, we will focus on the definition of flow as such, and its application in the field of video games; secondly, we will try to discern the state of Flow from Immersion; then we will explore the different specificities, methods and dangers related to the use of sound with a view to creating and maintaining the state of flow, before continuing with various case studies for illustrative purposes. The opinions and ideas advanced in this work will be, except quotation, mine, and will engage only its author.

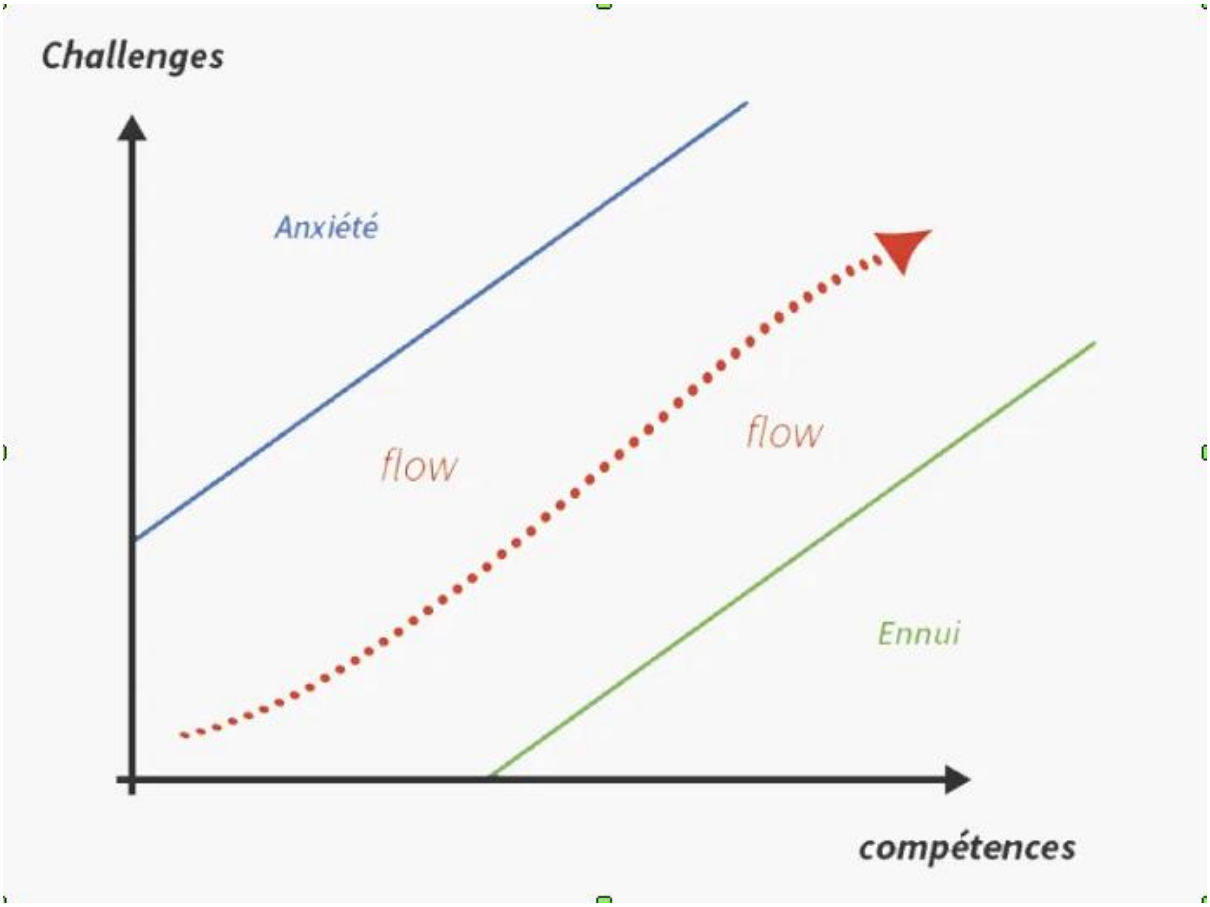
1. Flow, introduction and definition

A property like that of hysteresis could be compared to the state of flow: hysteresis, from the Greek ὕστερος/ *hústeros* (meaning "after", "later"), is the property of a system whose evolution would not follow not the same path depending on whether an external cause increases or decreases. Many game designers use the term “flow” to designate a flow of information conveyed as a process of communication to the player, to which the latter would respond in one way or another. However, this conception comes up against what actually constitutes the flow, or rather “the state of flow”. Theorized and named by Mihály Csikszentmihályi in 1975, the concept of "flow" is characterized as a state made up of 8 specific components (although the presence of all of these components is not required for a state of flow to be experienced) :

- A challenge that requires one or more skills.
- The fusion between action and awareness of it.
- The clarity of the current objective.
- A direct feeling of feedback.
- Intense concentration on the task being performed.
- The feeling of permanent control.
- The loss of awareness of the Self*.
- And finally, a distortion of temporal perception.

The notion of challenge and performance is therefore intrinsically linked to the appearance of this phenomenon: if the difficulty of a task undertaken is too great in relation to the skills of the subject, the latter will be confronted with a feeling of anxiety; conversely, if this difficulty is too low, boredom will stalk the subject. According to Mihaly, the state of absolute flow therefore requires a balance between the skills of the subject and the difficulty of the task in progress in order to appear.

** "Self", or "Me", theorized by Freud as the visible, conscious part of an individual; the one to which he assimilates and defines himself.*



However, a certain tolerance exists, even if it is difficult to measure due to the plurality of player profiles: approaching more or less boredom or anxiety without reaching it (in the pictorial context of this graph) does not cause the flow to disappear; on the other hand, it increases the risk of reducing its intensity, or even making it disappear. Thus, for Mihaly, the state of Flow constitutes “the optimal experience”, that which a human being seeks through any practice or activity whatsoever. The frequent achievement of this state would therefore, according to him, be directly correlated to the happiness and psychic well-being of an individual.

If this state of Flow therefore appears commonly during actions performed in real life, it must be artificially recreated in the context of a video game work. Since each player's ability to concentrate over an extended period is different for each profile, and any gaming session is necessarily destined to end, creating and maintaining the state of flow for as long as possible, on the other hand, allows the potential of pleasure and investment that can be felt by the player during the video game experience. In Jenova Chen's thesis called "Flow in Games" (transposition of Csikszentmihalyi's work to the field of video games, and which will be cited several times during this work), the author describes 3 conditions allowing the appearance of the state of Flow within a video game:

1. First, the game must be inherently worthwhile, and the player must try it of their own free will.

2. The game should provide a balanced challenge, appropriate to the player's abilities, that allows them to dive deep into the experience.

3. The player should feel a sense of personal control through their experience.

Still according to Chen, the result of these 3 elements would cause the player to lose the notion of time and mindfulness (or loss of Self, mentioned above). Over time, the concept of adaptive difficulty (or more commonly known as DDA, for Dynamic Difficulty Adjustment) ended up appearing, consisting in calibrating the difficulty of the experience based on a measurement of the player's performance and skills. . However, continues the author, the use of adaptive difficulty "would greatly diminish the control of game designers over the experience", making the use of this system a bit trivial and rather rare in the current video game spectrum. Thus, an adaptive difficulty system alone could not be the solution to the appearance of the state of flow, where a system based on the main elements of a game (the feeling of feedback, the game design, etc.) would be more direct and useful in order to obtain the desired effect.

However, if creating flow within an experience is a challenge in itself, retaining this phenomenon in the player is all the more important: in an era where the player is sometimes encouraged to remain "captive" of he same title lasting several hundred hours, it is not uncommon for high-budget games, and in particular open worlds, to become less and less finishable for the majority of consumers. If the latter can tolerate a slight frustration or moment of hesitation, the key to an optimal experience would therefore be to keep the subject in the "Flow zone". However, each profile is different, and has its own area depending on its skills. "In order to produce content that can be appreciated by a majority, it would then be necessary to extend the panel of potential experiences present in a video game, in order to fit in with the different player profiles" (Jenova Chen). Open worlds, by their intrinsic design, as well as via their gameplay loop, generally offer a varied choice of "games in the game" (capture of territories, puzzles, and other ancillary activities), allowing the player to adapt his experience according to their needs and desires.

One of the processes making it possible to anticipate or adapt game systems to the player according to his actions consists in measuring the latter in the form of data, and in predicting reactions or modifications in play according to the variation of these. . For example: depending on the player's life bar, the music could increase or decrease, in order to offer more impactful feedback than a simple visual change on an interface. However, more complex data (such as the notion of challenge perceived by the

player), is subject to interpretation, and therefore potentially in contradiction with the real user experience. This therefore limits the use of this data as a source of gameplay alteration or feedback. The author therefore concludes with this: “The only solution is to include choices in the gameplay, let the player treat these choices as part of the experience and possibly ignore them. Then his decisions will become intuitive and reflect his true desires.” Thus, apart from low-level data analysis, creating Flow via game mechanics would amount to creating a multitude of systems, responding or not to each other, while empirically testing on a panel of users more or less exhaustive the relevance of these systems. Another solution, this time post-development, would be to statistically interpret certain user data by the developer, in order to synthesize them into statistics that can be used to readjust certain mechanics, or be taken into account for the next title produced. Even if this process is based on more concrete data rooted in reality, it is still anchored in a more or less subjective approach, and therefore cannot be sufficient on its own to arrange mechanisms to achieve a state of flow in the most optimal way possible for the different target(s) of player(s) and player(s).

However, if the design of the mechanics and the choices of game design constitute a capital part in the appearance of the flow, the sound direction is not left out: "The sound being a material very easily analyzed by the brain, it quickly conveys a notion of quality and listening comfort. It therefore constitutes the first sense vector of flow within a video game: it envelops the player in a cocoon, a reassuring framework, which allows him to let go.", we said Masami Komuro. The sound would then be inseparable from the forgetfulness of the Self (one of the major components of the state of Flow), because it would give the player the opportunity to escape from his consciousness, while keeping physical control of his person. "We could also establish a parallel between the frequency composition of a sound and its ability to immerse the player in this so-called cocoon, depending on the level of comfort and quality perceived by the player vis-à-vis this sound." he continues. Thus, the sound would, among other things, plunge the player into a state conducive to the appearance of flow. However, it is then necessary to distinguish the very appreciation of a sound (whether it is pleasant or unpleasant) from the state of comfort in which it plunges the subject: a game like Doom Eternal comprising aggressive industrial sounds, can sometimes offer a panel of very unpleasant sound elements to the player's hearing, while placing him in a state of comfort conducive to the appearance of Flow due to its qualities (rich textures, worked and balanced despite the type of sounds offered). The frequency spectrum (and therefore, the level of pleasantness) of a sound signal would therefore be more correlated with a scenario goal than with an aid or an obstacle with respect to the state of Flow.

We can then assume that sound, by its physical nature and much more direct in terms of transmission to the human brain, would be a more effective means of inducing the state of flow in the player's experience, compared to analysis. of data recovered by test phases or by the actions of one or more player(s) in real time. However, a soundtrack or an interactive system planned in a sound direction can only be based

on “what is perceived by the developer as effective in a certain game context and/or for a typical player target”: the question of the use of sound in video games as a flow engine, in particular due to its artistic character, therefore remains a subjective element. However, in the same way as over the course of video game experiences over the past 40 years, certain components linked to the appreciation of a game (due to certain mechanics or cognitive phenomena) have been able to be decoded and analyzed, the same idea applies to the sound direction of a video game. We will therefore try, during this dissertation, to understand, list and establish elements of response which would allow us to be guided in the realization of a sound direction, in order to be able to more effectively arouse, and above all maintain the state of flow in the player.

But first of all, it is necessary to establish a differentiation between the concept of “immersion” and “flow”, often misunderstood and wrongly grouped together, even in the professional environment.

2. Distinction between "Immersion" and "Flow"

If we were able, during our first part, to explain with more or less detail the state of flow as well as its context of appearance, that of immersion is much more complex, because it depends on several phenomena and variable factors according to the profiles. The most commonly accepted definition (in particular in the current French dictionary “Le Robert”) would consist of: “the action of immersing, of plunging (in a liquid or medium)”. From there, we can obviously consider the virtual game space as an environment, in which the developer immerses the player through the use of different mechanical, narrative, haptic, visual and sound processes. We can then put forward a definition of immersion specific to video games, which would be as follows: “Immersion in video games is a process that allows, via the sum of different mechanical, narrative and sensory processes, to mentally and physically immerse the player in a digital space, creating an illusion of escaping from real space.” Thus, as soon as this illusion would be broken, the player would be led out of the experience. However, once again, each player profile is singularly different. In this case, a gameplay element (such as a more or less extensive jumping mechanism), a narrative process used (such as a reversal in the story), or even the use of certain sensory feedback (such as haptic feedback from a controller) would be able to break a player's

interest in an experience. If certain basic immersion processes, such as haptic feedback or sound, can sometimes be adjusted via various parameters in the game interface (allowing you to adjust the intensity of vibrations, to adjust the sound mix of the game, to adapt the sound to the listening system used, etc.), others are much more complex to modify according to the particularities of a player, and sometimes frozen during an experiment carried out on a finished product. Through all these issues related to the creation and retention of the state of immersion in the player, the state of flow can then be used as a tool by developers to guide the experience, and remove the friction present on physically, intellectually and/or emotionally.

To make this concept clearer, let's use an example: a player enters an empty room, and the front door closes behind him. A message then appears on the screen: “To exit, you must solve an enigma, and you will remain a prisoner of this room until the resolution of the enigma in question”. This puzzle consists of a Rubik's cube 3 squares by 3 squares on each side. Where this more or less complex enigma could frustrate, or even arouse a desire to abandon the experience in the player according to his abilities, the developer could choose to broadcast soothing sound feedback, such as soft atmospheres of water flow. water or natural space noises, in order to soothe the player and facilitate his concentration, as well as increase his capacity for resilience in relation to the imposed challenge. Also, with each manipulation of the rubik's cube, satisfying and impacting sound feedback, as well as a catchy musical pattern appearing at the completion of each side of the object, would stimulate the player by bringing dopamine peaks to him. Such processes, simple in appearance, would thus make it possible to keep the player in the experience, being able, then, to go so far as to register the repetitiveness of the manipulations carried out as the only pleasure of play, the latter then becoming decorrelated from the success or the puzzle failure.

Via the example used, and the explanations provided in our first part, we can then synthesize a definition of flow as follows: “A mental state characterized by the total psychic investment of a person performing an action, occurring when the subject between boredom and anxiety. Maintained by various stimuli, this state makes it possible to maintain the subject's commitment to the action being performed.

Thus, according to a summary of the words of Christophe Héral during our interview: “If the objective of an experience is a beacon, and the environment to be crossed to reach it is an ocean, the state of flow would then be the rudder that would lead us there, while helping us to navigate between the eddies likely to capsize the boat. However, to stay the course, the sound feedbacks serve as beacons, allowing us to orient ourselves in order to arrive at the right destination, and this without discouraging us. So if immersion and flow are two different states, flow would maintain immersion. But then, breaking the state of flow would be likely to take the player out of the state of immersion, and vice versa: these two states would therefore be interdependent, i.e. the two sides of the same coin without which pleasure and a player's interest in an experience could not exist.

3. Sound, pillar of flow

3-A - The technical and human issues related to the realization of the sound design of a video game

We can simplify the definition of “game design” as follows: “The definition of a set of rules and artistic biases in order to design a game”. However, a multitude of video games have, for technical reasons, budgets and ambition, treated sound as a separate, even secondary element, whether in its function or the resources allocated to it. Without going into a denunciation of these titles, far too numerous to count, the reason was first technical then marketing.

In fact, until the 1990s, the memory contained in game cartridges was not able to contain or restore a substantial quality or quantity of sound material. To overcome this, the sounds were synthesized, and the musical elements were poor in number of polyphonic voices as well as in terms of timbres, in addition to sometimes being very repetitive. From 1990 to the appearance of the CD, certain titles (such as Lucas Art's adventure games, in particular Monkey Island) began to seriously explore the notion of sound interactivity. Indeed, the technology behind the sound chips having evolved, the number of voices that could be played at the same time, as well as the computing power of the machines, made it possible to allocate more robust and complex systems to the sound direction. Also, the instruments being controlled by MIDI, their melodic lines could be controlled and reshaped in real time, according to the needs: for example, in Monkey Island, the audio programmers had created a system making it possible to readjust the midi of the instruments played in such a way that the musical transitions from one table to another take place as smoothly as possible. At that time, the sound was beginning to take on real scope and importance. But, as soon as the first Playstation was released, a drastic change in the design of video game soundtracks took place: the latter were now made up of real audio files played in real time, and no longer MIDI data interpreted by a chip. sound. However, the cost of an interactive approach to audio in the form of samples turned out to be much higher (in terms of calculations and memory required), compared to an approach in the format of synthesis controlled in MIDI, and the consoles then chose to focus on the graphic rendering of the games, in order to offer the so-called “photorealism” advanced in the marketing campaigns of the time. Also, handling audio files in real time was much less flexible. This led audio interactivity in video games to take a place in the passenger

seat for more than a decade, until the appearance and adoption of middleware solutions such as Wwise.

As of the date of writing this memoir, the latest generation consoles as well as the power of PC components allow game creators to allocate much more power to the benefit of sound direction. Nevertheless, the material and human resources allocated to the sound part of the games remains very minimal. Here are some statistics on AAA open-world games released over the past ten years:

- For "Ghost of Tsushima", the audio team was only 4% of the Sucker Punch studio workforce at most (according to Rev. Dr. Bradley Meyer via his interview in The Sound Architect Podcast, Audio Director at Sucker Punch).
- For “Cyberpunk 2077”, the audio team represented 3.5% of CD Projekt Red's workforce (16 people out of approximately 500 employees).
- For the Hitman trilogy (2016, 2018, 2021), the audio team made up 2.5% of IO Interactive's workforce (out of a total of around 200 employees).
- During development of "The Elder Scrolls V: Skyrim", the audio team consisted of a single member, Mark Lempert (later joined by Dave Shreiber), while the total staff working on the game amounted to approximately 100 employees (i.e. approximately 1 to 1.5% of the staff of the team working on the sound part).

We can therefore easily correlate these statistics with a lack of interest, understanding and ambition coming from studio managers and production managers vis-à-vis the importance of sound direction. in a game. Also, since open-world games require the most systems and resources (given their scale and complexity), the statistics provided above are all the more revealing. However, according to Bjorn Jacobsen: “ Audio would not necessarily be pushed into the background by the lack of manpower, but more by the general attitude of the development process towards the sonic dimension of a game, which can then cause the audio team to ignore the decision-making process, regardless of its size. This is usually due to a lack of understanding of how long it takes to craft a great sound design, the amount of code and number of tools required for this task, and the number of programmers are required to achieve these.”

Now that we have seen and understood the origin of the misrepresentation of sound in the video game industry, it is time to explain why the quality of a sound direction constitutes a vital element of a gaming experience, and how the consideration of the sound part as an integral part of the game design process results in a coherent and impacting whole, thus making it possible to reveal and maintain the state of flow in the player.

3-B - Sound direction designed as an integral part of game design

“Eyes lies. But the ears don't.”, says Seth Horowitz, a neuroscientist specializing in hearing. According to him: “If visual information takes about 0.25 seconds to reach the brain, sound information takes 0.05 seconds to be perceived. Also, our brain is so adept at distinguishing between sounds that it can perceive sound changes occurring in less than a millionth of a second. This auditory acuity is due to the evolutionary training of the human brain, once allowing our ancestors to react instantly to the slightest threat via hearing, and those without any visual cue needed.” Through this information, one realizes that sound information constitutes a sensory channel much faster and instantaneous than that of vision. If these two senses each have their advantages and blind spots with respect to the transcription of the environment around us, the immediacy of hearing, its ability to alert us to a threat as well as that its ability to represent and direct us in a space cannot be called into question.

However, if the video game is a broad medium, each experience of which can be approached from a multitude of different angles, most of the titles released to date come together in one point: the notion of challenge, often in correlation with the notion of danger. The video game thus appeals to our prefrontal cortex and the amygdala, namely the parts of the brain that organize our most basic instincts. If some games manage to offer an experience where this notion of danger is practically or even completely absent (for example, the puzzle game The Witness, or even Flower) and where the sound intervenes for other purposes, the game design strives therefore to propose an interesting challenge in a specific context, where danger (failure, death of the character) is often the main driving force. From this, we can deduce that the most effective way to represent this danger in an experiment would be to exploit the many possibilities of suggestion offered by the auditory canal. Unfortunately, unlike cinema where sound is used in a fine and thoughtful way (conveying a range of emotions and complex sensations, script elements, etc.), many current titles are confined to a basic use of sound, without seeking to provide something other than a sensory support to the visual aspect of an experience: the sound then becomes an element which is added in the background to the rest of the game, and whose main systems have therefore been developed without any real consideration of one of our 5 most important senses. So, sound designers have little leeway to use sound appropriately in a video game proposal. It is this lack of consideration or understanding that will naturally cause a regular break in the state of flow in the experience.

But then, what does a production look like that, from the outset, integrates the sound part at the center of the game design concept? Take the case of the game "Hellblade: Senua's Sacrifice", which was released in 2019 and won a nomination for its sound direction at the BAFTA Awards. In this narrative action experience, the player follows the tortured fate of Senua, a young woman prone to psychosis and psychiatric deviance, harassed by many voices and personalities in her head. In order to allow a player totally unknown to these psychological disorders to better understand Senua's

ordeal and thus identify with her, the audio team decided to use binaural recording microphones. Binaural recording is a method of sound recording obtained by positioning two microphones so that a person hears the realistic result (as one would hear a sound in reality, not diffused by loudspeakers). This process was used to capture the voices of different actors and actresses moving around the microphones, in order to give the impression that these voices revolve around the player's head, making him then feel the madness of the character in a way tangible. In Hellblade, the interaction with these auditory hallucinations is permanent, and at the very center of the game design (notably through the various challenges offered). The player is then invested in a personal way, and the hardships endured by Senua are directly perceived and felt by the person taking part in the experience.

By considering sound in the same way as other sensory information channels, the Ninja Theory studio was able to make the most of its subject. Thus, where each game is different in its concept and its needs, a judicious reflection of the sound design from the development of the game design and the mechanics of a game makes it possible to drastically reinforce (see bring) the player's involvement in the game. within an experience, allowing the state of flow to appear.

3-C - The use of sound characteristics as a vector of flow

As we have seen above, the different types of sound, sound manipulation and diffusion used in an experiment have the ability to drastically modify the perception of the subject. Depending on the evocation needs (if the developer wishes to frighten, surprise, galvanize or calm), different sound characteristics can be presented and/or altered in order to bring a whole range of emotions to the player.

3-C-Part 1: Sound amplitude variations

Abrupt or gradual variations in sound amplitude (volume) are often used in order to vary the tension in the player at a time T. The idea of gradual tension conveyed by sound, already present in musical creation since the beginnings of sound appearance, makes it possible to give rhythm to an experience, articulated around “hot spots” (peaks of tension) and “cold spots” (fallout, calm before the storm). In horror games and films, abrupt volume variations (or "screamer") are also used a lot: the player walks along a corridor, outside noises become quieter and quieter, and suddenly , a monster emerges from the set, bringing with it a whole panoply of sounds and musical elements that surprise the subject. Thus, significant variations in the amplitude of a sound are naturally perceived as a danger by the human brain: these variations then

serve to establish a rhythm in the experience, a rhythm that sound designers work to establish and break according to the situations. In DARQ, a horror puzzle game released in 2020 on consoles and PC, the sound designer in charge of sound direction, Bjorn Jacobsen, frequently uses these rhythm break processes specific to horror. The great dynamic range of the title is made up of very weak recurring sounds (like footsteps and those of various mechanisms), contrasted with very loud and occasional sounds (in order to emphasize certain elements of gameplay or scripts). This choice of sound design then makes it possible to keep the interest of the listener vis-à-vis the action in play, while placing him in a low but constant state of alert, without tiring the player unduly.

3-C-Part 2: Dynamic range management

The overall dynamics of a sound design can thus serve as a tool to convey information to the player, and place him in a feeling of calm or constant tension. In most traditional video games, the phases of confrontation/tension are always represented by a very reduced dynamic, due to the ensemble constituted by the music and the numerous sound feedbacks, thus bringing a stressful, almost suffocating side. If these phases of tension become too recurrent, and if their dynamic range is too small, this can also lead to a break in the flow in the experiment, the lack of amplitude variations then becoming very tiring for the human brain. Otherwise, a game with a very high dynamic range, with very little sound information going above -30 to -25 dbfs, can end up creating disinterest in the player over an extended period of time. To compensate for these significant disparities and peaks, sound designers refer, in video games, to a level standard, consisting of maintaining all of the sound information around -23 LUFS. However, this normalization can lead to certain types of sounds (such as natural atmospheres) being represented louder than they are in reality, which can then create cognitive dissonance or more simply a perceptible auditory fatigue over a period relating to each individual. , and can therefore lead to the disappearance of the state of flow. Sound integration and the use of a dynamic mix can avoid these concerns, a subject that we will explore later. Therefore, the use of amplitude variations, as well as the management of dynamic range within a sound design, proves to be a work of detail that can facilitate and maintain the state of flow in the player, as which can have disastrous consequences on the experience in case of bad dosage.

3-C-Part 3: The frequency domain

The use of the frequency domain within sound creation for a utilitarian rather than aesthetic purpose is a dense and complex subject, which we can only skim over here. That being said: the auditory perception of each individual is different and unique to the latter, due to the unique response curve of each pair of human ears. However, we can see that the appreciation of a type of sound, or at least of a particular frequency range, is also cultural: for example, the Japanese language including a lot of high-pitched sounds, Japanese culture is generally more fond of "bright" sounds with a certain amount of high frequency compared to more western cultures, because they are used to this type of sound. This can be noticed by examining the timbre of traditional Japanese instruments, even recent ones (such as pianos and synthesizers), as well as by observing the response curve of audio tools produced on Japanese territory, or by analyzing the distribution of the frequency spectrum in their audio content.

Despite more or less culturally rooted tendencies, we can still identify certain universal facts: the loudness curve (i.e. an average of the hearing curve of an 18-year-old individual) indicating a clear peak of 8 to 8.50khz, we know that this frequency range constitutes the most sensitive point of human hearing. The instincts of reacting to danger therefore manifest themselves more strongly around this frequency, and the cries of human babies also being placed in this range, we can deduce that this characteristic has influenced human evolution (humans being adapted over time to more vividly perceive this frequency range in order to protect their offspring). Therefore, this information can be used to orient the place of sound in the frequency spectrum of sound information for warning (or used to indicate danger), with the aim of providing more effective and coherent feedback in relation to the intentions of game design. Consequently, this allows the player to have a better readability of his environment and the action in progress, giving him the possibility of reacting more instinctively to potential threats: this then helps to keep him in a state of flow.

3-C-Part 4 - Location

A multitude of scientific experiments have proven that the brain relies on the more or less subtle differences in time, intensity and timbre between the left auditory canal and the right auditory canal in order to locate itself in space, as well as to locate all the sound sources around it. The location of the player's avatar in a space, as well as the identification of the position and the distance of a source relative to the position of this avatar, constitute essential elements in order to make a video game experience more fluid. The more accurate the source location information, the faster the various feedbacks and sound information are assimilated. As a result, the mental effort to be provided by the player in order to reinterpret the signals that surround him is less, allowing the latter to react in a more instinctive way, thus anchoring him more deeply

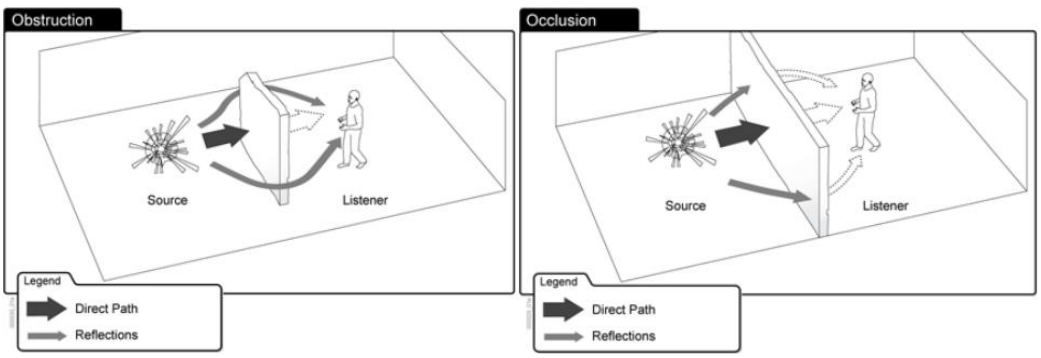
in the experience; the state of flow is then maintained more easily. Thus, the use and precision of simulation of acoustic and physiological phenomena present in real life can facilitate understanding and immersion of the player in a scene. Without going into an exhaustive list, here are some properties, acoustic and physiological phenomena whose understanding and simulation can simplify the localization processes within a video game experience.

3-C-Part 4.1 - Convolution

Convolution is a technique that consists in reproducing algorithmically, from an impulse response (like a pistol shot), the acoustic properties of a space. For a long time, this technology was too expensive in computing power to be used in real time. While the use of convolution reverbs, particularly in first and third person shooters, has become more popular over time, in recent years there has been a growing development of a more dynamic form of convolution in the video game, which would no longer be based solely on the transition from one impulse response to another, but on a reflection of sound signals based on real-time changes in the geometry of a 3D space. At the time of writing, the concept of ray-tracing applied to audio, i.e. the reflection of sound signals based on the rate of reflection and the level of absorption from the surfaces surrounding the source, continues to be explored (notably via NVidia's VRWorks technology). Such an increase in the precision of the representation of the sound space obviously makes it possible to accentuate the tangible relationship that the player weaves between the virtual experience and himself, helping to maintain his state of flow.

3-C-Part 4.2 - Simulation of natural obstruction and occlusion effects

Essentially, the obstruction effect consists of altering the frequency and amplitude composition of a signal, depending on the objects or surfaces that come in the way of a sound source and its listener. However, this effect does not take into account reflections from the sound source, unlike the occlusion effect.



The simulation of these effects makes it possible to increase the realism of a scene, while giving the player the possibility of better representing the space, as well as the various elements which compose it.

3-C-Part 4.3 - The use of binaural and transfer functions

As we have seen above, binaural can be defined as follows: "Any sound recording system capable of restoring a spatial rendering to headphones, with the sensation that the sources come from outside the head" (as with natural perception or when the listener is placed at the center of a multi-channel delivery system). This form of spatialization differs from traditional headphone stereophony (which gives intracranial perception) by reproducing the HRTF transfer functions of our head. The HRTF (Head Related Transfer Function, or "transfer function relating to the head") characterizes, by a mathematical transfer function, the deformations (diffractions, reflections and attenuations) brought to the sound waves by the shoulders, the head and the pavilions of each ear of the listener, which allow us to locate the origin of a sound, both horizontally and vertically. Therefore, these functions depend on the physiognomy of each, and therefore vary widely from one individual to another. While some companies in the video game industry (particularly Sony) are thinking about a way to allow the user to easily create their own HRTF profile, in order to obtain a more realistic localization of sound information in the game space, some audio middleware , like Wwise, can encode a signal in binaural. There are also some software solutions to synthesize a binaural rendering from a traditional audio file, through various filtering techniques unique to each developer.

3-C-Part 4.4 - The simulation of the Doppler effect

This phenomenon is manifested in real life by an accentuation of high or low frequencies coming from a moving source, depending on the proximity between this source and the listener. The closer the moving source is and the closer the perceived waves are (higher), the further the moving source is and the further the waves will be (lower). The simulation of the Doppler effect makes it easier to locate and to locate oneself vis-à-vis the sound sources that make up a scene: in shooting games, races or any other experience requiring the total and instantaneous reactivity of the player to threats can come from all directions, the convincing reproduction of this acoustic phenomenon makes it easier for the player to imagine the action, and therefore to increase control over their movements and decisions, which helps to create or maintain the state of flow.

3-C-Part 4.5: Phenomena of hyper localization, ubiquity and delocalization

The phenomena of hyper-localization, delocalization and ubiquity, relative to the field of psycho-acoustics, can sometimes be ignored when designing a game's sound. However, understanding them can drastically increase the skill of the player to locate the sounds that surround him in space.

Hyper localization is an effect linked to the punctual nature of a sound source, irresistibly focusing the listener's attention on the point of emission. As the source moves, the sound continues to be tracked. It is often the result of structure-borne transmissions, such as a ball rolling on the floor above. If a moving object or an undesirable sound captures the listener's attention, their vision, but especially their hearing, can become disinterested in the action or the scene in progress, thus creating a nuisance that can break the state of flow and immersion in which the player was immersed until then. The game designer must therefore remain vigilant with respect to the sound representation of moving sources, which can facilitate the localization process as well as deconstruct the state of flow in the player.

Ubiquity is a phenomenon that occurs when the sound source cannot be located, or the sound seems to come from all directions without the listener being able to detect its source. If this sound event can be used as a means to build a feeling of tension before the revelation of its location, it is also likely to disturb the listener, and break the state of immersion and flow in which he was immersed. Also, the ubiquity effect can be perceived as a nuisance, coming to distract the player from the action being performed, or away from the emotional intention supposed to be conveyed to the player at a time T. If some sounds can be considered as “global” sound layers, such as natural or urban ambient sounds, their source must be revealed after a certain time, so as not to create an image/sound dissonance that would harm to the experience. For example, if the player is in a city and hears sufficiently distinct exhaust noises, the origin of this source must be able to materialize visually in front of him, after a certain duration. The puzzle game with Escherian architecture “Manifold Garden” uses this phenomenon of ubiquity in particular in its sound design: in game, natural atmospheres of trees agitated by the wind, or natural wild life, are regularly diffused in the player's ears. However, the title continually plunging the player into the heart of themes such as multi-dimensionality and a reversal of the laws of human physics, the use of ubiquity here reinforces the purpose and the atmosphere of the game, rather than creating noise pollution for the player.

Delocalization is a minor form, as well as a component of ubiquity. If there can be delocalization without ubiquity, there cannot be ubiquity without delocalization. This effect implies the recognition of an error in the location of the sound source: we do not

know where the sound comes from, but unlike ubiquity, we know precisely where it seems to come from, while having aware that this localization is factitious and illusory. For more context, in the shooter "Bioshock: Infinite", it is common that during a cutscene where the main character slowly opens a door, while a conversation behind this door can be perceived, no filtering effect come to tell us that we perceive a source through a surface (in this case, a door). The brain then understands that the sound should possess certain characteristics, absent here, in order to sound plausible. This therefore increases the “noise nuisance score” or “NUFS”, a concept theorized by Bjorn Jacobsen and to which we will return in more detail later.

3-C-Part 5 - The Rhythm of Sound Design

“Rhythm” can be defined broadly as follows: “The characteristic of a periodic phenomenon, induced by the perception of a pattern in its repetition.” The question of rhythm within a video game experience is crucial. If part of this rhythm is conditioned, see defined by the game design, it is also arranged, maintained or broken by the sound design, whether at the level of the musical part, or the sound representation of the phenomena appearing in game and actions performed by the player.

An important constituent of rhythm in sound design is the repetition of sounds. “When a recurring pattern emerges, either in a sound that loops or in a multitude of sounds played again one after another, there can be a very fine line between, on the one hand, the acceptance of this pattern as something that we recognize and that makes us feel safe, and on the other hand, something that becomes boring due to its monotonous and repetitive nature.”, says Bjorn Jacobsen in his thesis “The Tuning of the Game”. This makes it possible to identify repetitiveness from the angle of two distinct facets: sometimes that of a tool, sometimes that of an obstacle vis-à-vis the appearance and maintenance of flow. We can then distinguish several primary components of the sound rhythm: the repetition of the same type of sound, and the musical rhythm or rate at which sounds are triggered.

3-C-Part 5.1: Repetition of the same type of sound

To illustrate this aspect, let's take the example of footsteps. The repetition cycle of variant footsteps (induced by the movements of the player), if it can be accepted as a reassuring and comfortable pattern, can just as much come to tire and disgust the latter from the auditory experience, either by frequency components of these sounds, or by their constancy of volume. So these footsteps end up being set aside by the brain, thanks to the famous “cocktail party” effect (allowing the brain to hide information from more or less constant noises in order to focus on one or more specific elements

). However, this mental effort required by the cocktail party effect comes to tire the player, and accelerate, even cause a feeling of weariness of the player vis-à-vis the experience.

Some games develop strategies to avoid this rejection of the repetitive aspect of footsteps: “The Legend of Zelda: Breath of the Wild” has a very subtle dynamic sound mixing system, which consists in particular of gradually reducing the volume of the player's footsteps as the player moves longer and longer on the same surface. However, when the player comes to pass from one surface to another, the volume is raised, making it possible to provide the sound information necessary for the player in order to understand and immerse himself in the world which surrounds him. Thus, the more the brain gets used to the triggering of these sounds and to the sound cycle they constitute, the more the mixing anticipates this habituation, thus preventing the sounds of footsteps from changing into constant noise, or noise pollution.

Another example of a solution would be the one adopted in “Grand Theft Auto 5”, as well as the atmospheric game “Inside”. These two games are singularly different in terms of approach and structure, one being a violent open-world action game with a multitude of gameplays, the other being a platformer with a fixed and very cutscene with relatively static gameplay. However, they share a common point: the decomposition of the sounds of footsteps of the embodied character into a multitude of components, responding to various variables in play, such as the type of clothing or fabric worn, the surface encountered by the feet, the level of water present on the ground, the gait of the character, the breathing of the latter, etc. Naturally, each of these different components themselves have a multitude of variations, making it impossible (or much rarer) to have a negative feeling about a certain sound repetitiveness.

3-C-Part 5.2: Musical Rhythm and Triggering Rate of Sounds

Another major constituent of sound rhythm would be the tempo, ie, on the one hand, the time period separating two occurrences of a sound, and on the other hand, the “musical rhythm”. A scientific study carried out by François Haas at the Langone Medical Center in New York, establishes a correlation between perceived musical rhythm and the behavior (rhythm) of human breathing. The resulting hypothesis of the tests carried out stipulates that an acceleration or a deceleration of the respiratory rate can be provoked according to the variations in tempo presented to a listener. However, this idea is far from new: in music, then in cinema, a more or less high tempo has always been used in order to increase the tension of a moment T. Tempo management, especially within the video game experience is therefore crucial; it can elevate as well as interrupt the state of flow.

- A too fast tempo during a prolonged period tires the player more or less quickly, as can be the case with games like “Thumper” (musical game with very aggressive industrial music and intense rhythm) or with games of the “Shoot Them' Up” type (like ”Ikaruga”) where the many sounds of shots and explosions end up forming a constant noise to the point of detracting from the experience. This then runs counter to the reception of significant sound feedback, which ends up being drowned in a mass of sound.
- A tempo that is too slow, or an absence of tempo, is likely to tire the subject, thus not providing enough sound stimuli to diversify the auditory experience and maintain the player's interest.

If these questions of musical rhythm and rate of sound triggering are often induced by the game design itself, some games set up processes via their sound direction or their implementation in order to circumvent these problems. “Doom Eternal”, the latest shooting game from the famous ID Software license, has frantic, extremely demanding, brutal and dense gameplay. The sound design of the title reflects this well, whether by the riffs of industrial metal and the saturated electronic sound layers of Mick Gordon, or by the sounds of shooting, explosions and cries of monsters coming from everywhere. However, the ID Software audio team, through the implementation of a sidechain system (digital audio practice which consists of reducing the amplitude of a sound signal by triggering another punctual signal, and therefore the reduction is defined by the amplitude of the second signal), manages to bring up the crucial sound feedbacks of the game - when the player shoots, when a dangerous enemy dies, etc. - beyond the "sound mass" present in the game. Also, the space exploration game No Man's Sky manages, despite its high dynamics, to maintain a more or less constant sound presence by procedurally triggering musical layers and ambiences, according to the actions carried out by the player as well as the position of his camera. In this way, the sound space is regularly occupied, without being overloaded or leaving gaps that put an end to a feeling of continuity, and therefore interrupt the state of flow of the player.

3-C-Part 6 - Managing the harmonic ratio of sound elements

Harmonic movement, namely an alternation of moments of dissonance (tension) and moments of consonance (relaxation) is a process used since the beginnings of the history of music, most often constituting the skeleton of a musical work. Within this

movement, we find the “preparation”, or the way of bringing about a dissonance and introducing the tension, and the “resolution”, or the transformation of this dissonance into a new consonance. These concepts have also been used since the beginnings of cinema and commonly adopted in the video game medium, allowing the spectator/player to build anticipation and then release it, forming a more or less predictable cycle and creating interest in the topic.

The problem, as indicated in Bjorn Jacobsen's thesis cited above, arises when the sound elements of a video game experience come into uncontrolled dissonance, either with respect to each other, or with respect to the musical elements which are superimposed on them. . This can lead, in many cases, to auditory and mental fatigue, followed by breakdown of the state of flow. To bring the sound effects into consonance or dissonance with the music in a controlled way, the sound designer then has three options:

- Either adjust the tonality of sound effects and musical elements so as to bring them into consonance with each other, which requires planning for all the eventualities of possible sound combinations, making the process tedious or even impossible depending on the scale of the project.
- Either develop one or more systems during the sound implementation, in order to automatically arrange the harmonic composition of the sound and musical effects between them, to avoid any involuntary dissonance.
- Either break the line separating sound effects from musical elements, in order to place them in a coherent and doubly immersive whole. This process, used in particular in the cinema with “Blade Runner”, is made much more complicated in an interactive medium like that of video games. However, the composer Olivier Derivière manages to include this concept in the sound design of the game “Get Even”, in particular via a soundtrack generated in real time according to the actions of the player: indeed, some of the sound effects (sounds of flickering lamps, door knocking) as well as certain musical instruments (piano, violins) and percussive instruments are played via a midi file, in order to be synchronized in real time with the heartbeat of the main character. Also, some sound effects (such as police sirens) are first played and perceived in a classic, diegetic way, then, as the player moves away from the source of these sounds, transitions to drones of orchestral music being inscribed in the same key. Again, by including sound design in the game design process from the outset, Derivière and the studio “The Farm 51” achieve a homogeneous, coherent and impactful result, which makes it possible to reveal and maintain the state of flow. in the proposed experience.

However, as Joonas Turner points out, dissonance can also be used as a tool, to amplify a warning sound or a sense of danger, making it easier to bring out feedback from the whole sound, and therefore better convey information. It is then up to the

designer to find a happy medium, and to use this kind of process wisely, without endangering the state of flow.

During this part on the relationship between the flow and the sound itself, we could see that the approach of the state of flow by the sound design varied according to each project, whether in the (or the) method(s) used to bring it about and maintain it, or according to the scale of the experience, as well as the different situations that compose it. The sound designer has many ways to act on the sound, including an infinity of solutions, in order to be able to bring the state of flow and remove most of the elements that risk hindering its maintenance during a game. We will now see how understanding the concept of “game feel” makes it possible to link the player even more to the experience through the state of flow.

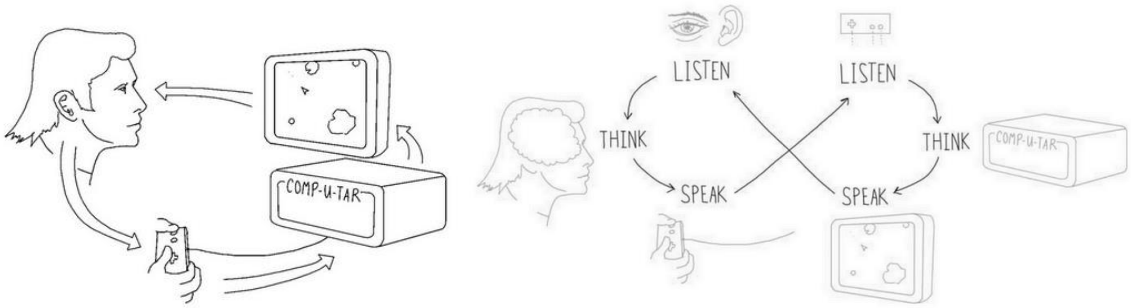
4. Game Feel, Flow and Sound: an intrinsic relationship

4-A - Definition of “Game Feel”

We could define the game feel as “the intangible, tactile sensation felt when interacting with the video game experience”. Resulting from a set of mechanics and systems, this phenomenon can be separated into 3 components: that of real-time control, simulation of the game space, and finally "polish" (i.e. the level of finish experience).

4-A - Part 1 - Real-time control

Real-time control is a specific form of interaction, including two participants - here, the computer and the user - forming a closed loop as illustrated below. Chris Crawford, in his book “Chris Crawford on Game Design”, describes this closed loop as a conversation, a cyclical process in which two active agents take turns listening, thinking and talking.



However, Steve Swink (author of the book “Game Feel”) counters that this metaphor cannot be applied to all situations, and that game feel is closer to driving a car than to a real conversation: “ If the driver wants to turn left, he acts more than he thinks. He turns the steering wheel corresponding to the direction he wants to take, by means of what he sees, hears and feels until the turn is complete. This process is almost instantaneous. However, a conversation takes place over a much longer time frame, below the level of consciousness, in an uninterrupted flow of instructions, while the result of an "input" (control) is perceived at the same time as it is Express. This then constitutes the basis of the game feel: a precise and continuous control exerted by the player on an avatar or a moving object.

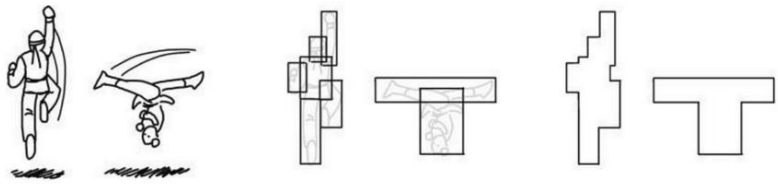
4-A - Part 1 - The simulation of the game space

Game space simulation refers to physical interactions performed in a virtual space, actively perceived by the player. This therefore involves collision detection and response between the avatar and the objects that make up the game world. These interactions give meaning to the avatar's movement, providing it with the ability to move around, between, against objects. of the game, while serving as a reference

point vis-à-vis an impression of speed of movement. This makes it possible to simulate the tactile, physical sensation of interactivity that we experience through the physical spaces that constitute our everyday reality. Using the player's avatar as a channel for expression and perception, the player experiences the virtual space on a physical and tactile level, allowing them to become one with the action and the elements around them by game. Another component necessary for the simulation of the game space is that the latter must be perceived in an active way, unlike the television format which is perceived in a passive way. The game feel therefore revolves around the active and literal perception of space, through the means of interaction that are offered to the player by the game designers. However, many games have detailed collision/response systems. , but the player experiences them only indirectly, since they are decorrelated from the player's action. So, the design of animations, visual effects and sound responses to provide clues relating to their physical composition (their weight, their texture, their elasticity, for example). These “clues” are grouped together under the name of “polish”, ie the process of finishing and refining the video game experience necessary for its appreciation.

4-A - Part 1 - The polish

Polish refers to any effect that artificially improves the interaction, without altering the primary simulation. For example, in "Shadow of the Colossus", as a titan moves, particles of dust rise up relative to the impact of its footsteps, and the noise of the latter is perceived more or less loudly as the player approaches or moves away from the source. Polishing therefore adds interest and emphasizes the physical nature of interactions, helping game designers to “sell the illusion” of these objects to the player, more or less realistically. As can be seen with the example below, a character in a fighting game, without models and animations, is just staring at strange boxes battling each other.



“By assembling these three elements of real-time control, simulation of the game space and polish, we arrive at the following definition: the real-time control of virtual objects in a simulated space, at the through interactions underlined by the level of polish.”, says Steve Swink. But then, how does this notion of game feel relate to the experience of the state of flow during the game?

4-B - The identification of the player with his avatar through sound design

“A sound effect can completely change the player's perception of an object in a video game. [...] For example, one of the animators of God of War had created an animation. Despite the accuracy and quality of the latter, something was missing; the animator was about to start his work over again, when Derek Daniel, one of the game's game designers, offered to add sound to the animation. Obviously, that was the missing piece of the puzzle: Suddenly, Kratos' movement became powerful, satisfying and violent.”, says Steve Swink. As one can imagine, the sound design is the keystone of the game feel: it is what will make it possible to provide an instinctive sensory response to the interactions of the player, vis-à-vis his avatar and objects that make up its environment. “One of the things that seems to be a central aspect of in-game flow retention is when the player receives information about their actions or the control of their avatar through the auditory channel, information that would then be lost or less well perceived if they were only broadcast through the visual channel (animations, interface, etc.). For example, when the player is about to run out of ammo, an audio feedback such as a series of exaggerated metallic clicks that would indicate to the player that his magazine is almost empty, then allows the player to receive the information without the latter needing to take their eyes off the action to examine their ammo counter on the interface.”, says Jonas Turner, sound designer and composer on many games with a very “arcade” look. (“Nuclear Throne”, “Scourge Bringer”, “Downwell”…). Such a process would therefore make it possible to streamline the “man-machine dialogue”, giving a more instinctive aspect to the gaming experience and maintaining the player in a continuous identification relationship with his avatar.

“For the player to have the feeling of actually carrying out an action and controlling his character, the question of the latency between the press of a button and its sound response is crucial. If the slightest delay (real or perceived) comes between these two stages, then the player no longer has the feeling that he is in direct control of the action”, continues Turner, which results in a dissonance on the cognitive level, likely to interrupt the state of flow. However, even if no perceptible delay separates the event from its sound response, it often happens that the sound effect itself includes a certain duration between the beginning of the file and the transient synonymous with the impact of the action. carried out. This relatively common design error gets in the way of the game feel, and therefore of the state of flow. However, the duration separating the input from the transient of an action can be used to

convey certain sensations to the player: for example, if the character is clumsy, or suddenly too heavy due to a full inventory, such a latency makes it possible to convey this feeling of heaviness through sound. The "Dark Souls" series of games uses this process on many occasions: for example, during a roll, the sound will be triggered with more or less latency, in order to suggest a heaviness relative to the physical characteristics of the avatar of the player (weight, number of items in inventory, agility level, etc.). Shortening as much as possible the duration of a sound that is triggered relatively frequently can also make it possible to promote the feeling of response and impact, while reducing the feeling of frustration that can be caused by the repetition of such a sound.

4-B - Interface sounds as a substitute for identification with an avatar

According to Arnaud Roy, composer of the French studio Amplitude, “the challenge of strategy games is to manage to keep the player captive to the experience, by pushing him to play an extra turn. In terms of sound design, this mainly involves the user interface. Thus, any notion of a single avatar being often absent in strategy games, the latter stage the control of a set of characters (troops, armies, civilization, etc.) through the orders given to them. To do this, this command, or decision-making, must provide enough sound feedback to keep the player in a tactile and physical relationship to the experience. “So when the player clicks on an item, they get sound feedback. This makes it possible to immerse the latter in the virtual space and to give him the impression that he is touching something concrete. When I click on a troop, and the general at its head starts shouting "At your orders", the player will have the impression that this troop obeys the player, giving him a feeling of power and additional control. on the action. Also, for all the abstract sounds relating to the various menus of a game, it is the suite of interface sounds that will bring and maintain the player in a state of flow: these sounds each have a texture, relating to the context of history (historical period, realistic or science fiction universe, etc.) and to a set of resulting physical properties (organic, metallic, silky, raw, etc.).

4-C - Sound design at the center of the game feel

Another issue that arises is that of the link that connects the sound elements to the virtual world. For example, if the sound identity of a character's jump sound is perceived as different from that of his footsteps, another form of dissonance can then appear, coming to get in the way of the perception of sound feedback such as a coherent whole in relation to the world in which they occur. Still according to Joonas Turner, an effective way to prevent this dissonance from appearing is to preserve consistency in the analog and digital audio chain of the recording and sound creation process. “At the beginning of the production, I decide which microphone, which preamplifier will be used to record a certain category of sounds, and I maintain this choice for all the sounds constituting this category throughout the sound creation. Then I apply the same processing chain to all the sounds in this category, so as to “glue” them together. Finally, I pass all of the categories of sounds through the same final chain, then coming to unify the whole sound design by attaching it to the same identity. This sound identity allows the player to feel the sounds as belonging to the same sound space, thus creating a consistency and a feeling of coherence which makes all the feedbacks more faithful to the virtual world in which they occur. This process thus makes it possible to reinforce the sensation of game feel in the player's experience”.

Another interesting approach by Turner is to use technical flaws in an implementation, such as distortion caused by heavy overlapping of the same type of sound or multiple sounds, to heighten the tension of a specific situation at hand. portion. “For example, if the player finds himself in a room where knights in armor start attacking the player at the same time, the accumulation of feedback is likely to create a distorted sound mass, due to the overall level of this mass colliding the dynamic processor at the end of the chain used to limit the overshoot of the distortion threshold. But, once the confrontation is over, everything becomes calm and quiet again.” Used judiciously, this kind of approach makes it possible to highlight the moments of tension, while inducing a kind of “logical reaction” of the various elements composing the virtual world, which allows the player to represent it in a more tangible way.

Also, again according to the interviewee, certain implementation techniques, such as ducking (mentioned above with Doom Eternal), can be used as a creative tool to reinforce the game feel, rather than just being used to to circumvent a problem. Turner therefore explains how the use of filters, controlled in sidechain by the triggering of certain sounds, makes it possible to reinforce the feeling of impact of the latter. For example, if an explosion or a high volume sound is heard, a lowpass filter will close and then reopen to emphasize the power of the action caused by the player. This then gives the latter a feeling of power and gaming pleasure, thus helping to maintain the player's commitment during a game.

Finally, Turner says he employs the following technique on each of his projects: First, he cuts an extremely short sound in the middle of the phase of his waveform, in

order to create something similar to a "click" (sound only composed of a transient, without any fallout). Then, it superimposes this element at the beginning of each of the sounds relating to the player's actions. This then creates a sound whose impact is felt instantly on playback, helping to accentuate the feeling of power and response of the sound feedback relating to the actions of the player, while eliminating any possible problem of perceptible latency.

All these techniques and creative processes therefore make it possible to create a dynamic sound ensemble, reinforcing the player's feeling of physical control over his experience, and thus facilitating the appearance and maintenance of the state of flow.

5. The design and implementation of sound interactivity, cement of Flow

Previously, we were able to see that the understanding and relevant use of the sound characteristics making up a soundtrack are generally linked to the creation and implementation of interactive systems when creating the video game experience. Now, we are going to explore, through various sub-sections, how the integration of the different sound elements constitutes a primordial factor in the appearance and preservation of the state of flow.

5-A - The implementation of fluid and natural transitions

Unlike cinema where transitions are fixed and often used as a stylistic effect, transitions in a video game experience are defined by several variables, linked either to a state (victory, defeat, death, reappearance, etc.), or to technical reasons (saving memory, loading visual elements, etc.), or to a narrative evolution (cutscenes). However, these transitions are often subject to a drastic and brutal break in the state of flow, thus breaking the feeling of continuity established until then.

In doing so, some titles try, and those since the 90s, to facilitate this process of transition from one scene to another in order to preserve the state of flow. First, by developing various visual or game design tricks: we can naturally cite the example of “Resident Evil” first of the name which, in order to maintain the tension in play during loading times, concealed the latter by a door opening animation on a black background. It created a sense of anticipation, of mystery, which kept the player on their toes. However, this type of solution is rudimentary: repeated a certain number of times, it still ends up breaking the flow and frustrating the player, sometimes even more than if the game in question had been content with a simple loading time with a still image.

However, 10 years ago, the LucasArt audio team (composed of Michael Land and Peter McConnell) had thought about and managed to skilfully circumvent this problem by using the “iMUSE” music engine (for “Interactive Music Streaming Engine”). , an interactive music system consisting of creating transitions from one musical theme to another in a fluid and natural way, allowing the change from one table to another without interrupting the state of flow and immersion of the player. McConnell described iMUSE as a system acting like a pit orchestra (a type of orchestra accompanying actors in musicals, operas and ballets), triggering longer or shorter musical sequences depending on the speed at which the rendering was setting up the next scene.

However, if this type of system required a significant cost in resources at the time, the tools (Wwise, FMod) now available make it possible to obtain an equally convincing result, by means of a speed of production and much better implementation. Through the sound design of the game “INSIDE”, created by studio Playdead Games and mentioned above, audio director Martin Stig Andersen takes an interesting approach: in order to maintain absolute continuity, the sound engine (here, Wwise) is always active, from the menu screen to the game itself. Thus, each of the sounds are divided into two categories: sounds that are restarted when the player reappears, and those that are never interrupted (like music, ambient sounds, etc.). Such a choice then makes it possible to avoid any abrupt transition, and to keep the player in a homogeneous experience without interrupting the state of flow.

5-B: The use of interactive music as a vector of flow

“The increase in the quality of sound in video games over generations and technological advances has led to a desire to master the sound design of this medium - in particular via interactive music -, so as to be able to more easily introduce this state of flow, or at least the components that constitute it (intense concentration, commitment, temporal alteration, etc.)”, says Masami Komuro. “Already, he continues, through the sound design of the strategy game “Z”, released in 1996 by the studio The Bitmap Brothers, video game developers were beginning to experiment with the notion of interactive music as a means of communication, in order to transmit to the player in a more fluid and instinctive way the information necessary for the comprehension of the action (and thus to make appear and maintain the state of flow), in particular by the sequential reading of MIDI segments following that the player construct new buildings, new troops, etc.”.

Another aspect of interactive music is its ability to bring a form of variety and continuity to a relatively repetitive core gameplay. “The biggest challenge of interactive music is repetition.”, answers Olivier Derivière, following a question asked during a masterclass organized by “Jeux Vidéo Magazine” at the Cité des Sciences in Paris. ”A video game is generally made up of a main gameplay loop, which will repeat itself over the course of the experience. If the aspect of the world in which we find ourselves changes (world of fire, world of ice, etc.), the gameplay loop changes very little, if at all. So, all the video game development professions are trying to break this redundancy. However, music has a great capacity to be able to break the feeling of repetition.”, continues the latter. Thus, the video game composer must, upstream of the creation process, think about the layout of his composition(s), first of all in order to ensure fluid transitions between each of the sound elements, and then, in order to prevent any feeling of frustration related to the excessive repetition of these elements. For this, three approaches are generally employed:

- The horizontal approach, to use the words of an article by Gamekult on adaptive music (written by Valentin Cébo, alias Noddus), “concerns large sections of music, in which the audio engine will be able to draw in order to dress a game sequence based on the “game state” that it identifies. So, keeping the same BPM, sometimes the same melody, the game will transition the music from one state to another, sometimes emphasizing the last 4-8 beats of a track, and thus creating a smooth transition to the following."
- The vertical approach, consisting of stacking different musical layers, in order to activate or deactivate them according to certain variables and changes in the gameplay. For the “Dark Sense” project on which I work, I use this method in particular to “fade in” or “fade out” certain musical and rhythmic tracks, following the evolution of a variable referring to the speed of movement of the player.
- The generative approach, which consists in generating in real time a musical score played by the audio engine, according to the different actions and choices made by the player. It is notably via this method that Richard Vreeland, alias "Disasterpeace", conceives the soundtrack of the game "Mini Metro": in this

contemplative puzzle game consisting in creating and enlarging a metropolitan network, each of the player's actions brings a musical pattern, thus making it possible to create an ensemble that acts as direct feedback to the player's choices.

If these 3 relatively distinct approaches alone represent a very particular philosophy, it is not uncommon for a sound design to use several of these approaches before sequencing its musical elements. However, an important distinction must be made between “adaptive music” and “interactive music”. In my opinion, adaptive music is a music system that adapts to in-game events, correlated or not to the actions of the player, in order to offer transitions and image/sound synchronization similar to that offered by the cinema. However, interactive music is for me a system mainly based on the direct actions and controls of the player, in order to sequence the different musical elements composing the soundtrack. We would therefore have on the one hand, a form of systems coming to adapt, anticipate the events in play, and on the other interactive systems, creating a dialogue fed by the behavior of the player. Some composers, like Olivier Derivière, make it a point of honor to separate these two terminologies, having according to them a definition specific to each. In Derivière's ludography, we find this approach in particular with Get Even, which generates and sequences the soundtrack in real time according to the variables relating to the player (his proximity to certain elements, the heartbeats of the character, etc). However, another game from his ludography explores this concept of “interactive music” even more: “Remember Me”, a sci-fi action/adventure game, released by Dontnod in 2013. During the combat phases that punctuate the progression, the music is sequenced horizontally following the various combos of the player, depending on whether he is moving away from the enemies or getting closer to the fight, who is being attacked... This creates a dynamic result which supports the state of flow of the player at every moment, placing him in a position of conductor. However, according to Derivière, “the question remains how to preserve and maintain consistency behind the systems created. In Remember Me, the different combos all correspond to 8-bar phrases, ending with the same rhythmic pattern. Thus, while breaking the repetition and coming to create a dialogue between the actions of the player and the music in play, the sound design of Remember Me manages to develop effective systems in order to bring and maintain the state of flow of the player in experience. However, such a symbiosis between music and gameplay is only made possible if, according to Olivier Derivière, “the composer understands the language of the developers, and assimilates the gameplay as the central part of the experience”. However, still according to the person concerned, such work is often ignored: “Unfortunately, most video game composers do not play or very little play the game for which they compose.” Finally, he concludes with the following statement: “It is often said that sound design is the poor relation of video games, but in my opinion this is not true: it is those who corroborate this myth who contribute to make it as it is.”

5-C: Interpreting user data as a driver of interactivity

“As a sound designer, I try to bring the various musical and sound design elements in such a way that they seem to react or be the result of the actions carried out by the player. Sometimes it is, and other times it is an illusion. The idea is to ensure that the player feels in symbiosis with the world in which he finds himself immersed. In this way, the latter loses track of time and enters a state of flow.”, says Paul Weir, audio director at Hello Games (“The Last Campfire”, “No Man's Sky”). The music of No Man's Sky is, like the game world, procedurally generated (or rather triggered). By cutting and dividing the entire soundtrack (provided by the instrumental post-rock group “60daysofstatic”) into a multitude of layers, Weir progressively builds a musical ensemble, accompanying the player during its exploration. However, the following problem appears: “In game, one can arrange a sound set which varies constantly, without this variety being perceived by the average player. Thus, despite the complexity and the number of systems making it possible to generate this constantly changing musical ensemble, a form of auditory redundancy is installed in the player's experience. To prevent this, systems analyze the player's actions in real time (movements, camera angles, distance separating the player from certain elements, etc.), and send this information back to the sound engine in the form of variables, which then comes direct the creation and arrangement of musical elements in real time based on these. “In a sense, the inputs to these variables don't really matter, as long as they originate in the behavior of the player,” he continues. Thus, by creating a sound variety that draws on the player's actions rather than randomness, we create a sound design that almost erases any notion of repetitiveness, while establishing a permanent dialogue between the player and the world that is played. he comes to live. This process then makes it possible to never put the sound design in the way of the player's flow experience. However, this is only made possible if, according to Olivier Derivière, “the composer understands the language of the developers, and assimilates the gameplay”. However, still according to the interested party, such work is often ignored: “Unfortunately, most video game composers do not play or very little play the game for which they compose”.

5-D: The need to design a modular and flexible implementation

“In No Man's Sky, the world is procedurally generated. This implies that, unlike a game with a more traditional structure, it is impossible for us to predict the actions of the player in advance, as well as the place and the context in which they perform them. This led me to design a sound implementation arranged around modular systems, rather than an implementation focused on a set of events. “, explains Paul Weir, in a conference given at the GDC on the sound design of No Man's Sky.

- An event-based implementation is designed in direct relation to a specific event: the player performs an action, which triggers one or more sounds. Such a system is relatively fixed: as the size of the project increases, one must continue to bring in a cumulative and basic way a "solution" for each possible situation, making the organization of the implementation complex and narrow, in more than being a significant waste of time and resources.
- On the other hand, an implementation arranged around modular systems, based on the knowledge of the different states of the player and the variables associated with it, is more easily expandable, because it is “indirectly related” to a specific event.

“If, he continues, I use an event-based approach in certain situations, such as for the interface sounds and certain animations of the player, I apply myself to adopting this approach centered around systems for the rest of the sound elements. , especially for ambient sounds: in No Man's Sky, each planet has a defined type (volcanic, oceanic, jungle, desert, glacial, etc.), as well as a set of characteristics (temperature, elemental composition) which are clean. With a systemic approach, if the game designers decide to implement new types of planets, I just have to create a new label in Wwise (called “State”), from which I can allocate new sounds, variables or conditions, all without the need to create additional lines of code in the game engine. Such flexibility in integration allows me, as a sound designer, to achieve the following objectives: in the first place , to create an environment that reacts dynamically and convincingly; second, to include a multitude of layers of detail in order to enrich the player's experience; lastly, to drive and arrange the sound design from game data, thus making it possible to simplify the dialogue between the sound design and the behavior of the player during his game.”

“Such systems then make it possible to bring out key moments charged with emotion in the player's experience, without it being planned, or worse, being imposed on him in advance: a player once contacted me to tell how, after taking shelter in the cockpit

of his ship during a storm, he heard the sound of rain crashing against the wall of his craft, and how it reminded him of certain times of his childhood, waiting in his father's truck for the storm to dissipate. Thus, by basing the majority of the sound implementation of No Man's Sky around modular and complementary systems, Paul Weir manages to maintain the state of flow of the player, while creating moments of play and memories of a rare purity.

5-E - Optimization of the sound implementation and its systems

The optimization of the sound implementation and the systems that constitute it, a subject that appears to be very technical, is also linked to the appearance and maintenance of the state of flow. This process helps prevent and prevent any technical hitches (issues with triggering, sound mixing, etc.) that could get in the way of the player's experience. If, in the previous part, we were able to see how the implementation in the form of systems made it possible to obtain a reactive and versatile sound design, it is now appropriate to review the issues related to the real-time optimization of these systems. .

5-E - Part 1 - The virtualization process

As game environments become more and more gigantic, in particular through the democratization of open worlds over the past fifteen years, game developers can less and less content themselves with loading the same large level and all its components, before letting the player move there (for obvious performance and memory reasons). In doing so, current titles are increasingly focusing on issues related to "asset streaming", i.e. the fact of quickly compressing and decompressing the various elements necessary for the gaming experience (models and graphic objects, sound, visual effects, shading, texture, etc.), depending on the need for their presence at a given time. To give an example, in GTA V, when the player is too far away to see distinctly a visual item (such as a trash can, car, billboard, etc.), that item is momentarily compressed or loaded into RAM, so that it can be quickly uncompressed and loaded if the player gets close to that visual item again element. This concept also applies to sound elements: in order to save memory and computing power, the number of simultaneous voices (i.e. sound channels) is

limited. So, it is necessary to create a clear and organized hierarchy of all the sound elements, relative to their level of importance and occurrence in the scene, in order to decide in real time and according to a multitude of variables, which ones are played, temporarily loaded into RAM so that they can be accessed almost instantly, or disabled. This process is called “virtualization”, and is used in most medium and large video game productions, in order to optimize the performance of the experience. As computer components and console generations have evolved, the number of voices that can be played simultaneously has increased dramatically, making it possible to create sonic environments of ever-expanding quality. That said, the priority of computing power is still given in large majority to the visual part of the experience, thus inducing a certain limitation in the number of sound elements that can be active at the same time; this can then become a problem when one wishes to implement a certain number of sound varieties for the same type of sound. The organization and prioritization of voices then becomes crucial, in order to be able to create a sound environment that is as responsive as it is complete, ready for any type of potential behavior on the part of the player. This prioritization is generally carried out by processing the metadata included in each of the audio and midi files, thus making it possible to provide a modular and extensible organization for all the content relating to sound design.

5-E - Part 2 - Mixing sound sources in real time

Another problem, more or less approached in the previous parts, is that of the dynamic mixing of the different elements between them. In Grand Theft Auto V, the audio team, through Rockstar Studios' in-house game engine titled “RAGE” (Rockstar Advanced Game Engine), creates a set of tools and processes to deliver cinematic impact to the sound direction. , in particular by the balance of the different sources that compose it:

- First of all, the “static” part of the mix is arranged according to a hierarchy grouping together several categories, or groups (ambiences, footsteps, various actions, shooting sounds, car noises, etc.), themselves defined by their level of occurrence, their volume intensity...
- Once this static balance is performed, these groups are then assigned to specific alterations (dynamic, frequency, etc.), most often controlled by the evolution of one or more variables for each of these changes. These alterations can also be applied to all categories of the game, such as when the player's avatar dies.
- Then, different settings of these alterations are assigned to the different types of scenes encountered in game (in the car, in cockpit view, in 3rd person view, on foot on the highway, on foot in nature, etc.), in order to establish the realism of the different sound spaces for the player, as well as to make them

alternate in a convincing way, in order to create a persistent and believable world.

- “The GTA series of games have always used car radios as their main source of music. In GTA V, we have chosen to add an original soundtrack to this to accompany the player even out of context, in order to opt for a more cinematic approach. So we had to find a way to make the transitions between the music coming from the radios and that coming from the soundtrack in a way that was pleasant and natural for the player.” says Alastair Mc Gregor, audio director at Rockstar Games. “Our approach to music is that the player should never perceive the start of a piece, the latter must be part of a continuity with the overall soundtrack in order to avoid any break. ”, he continues. To do this, the audio team separates each of the tracks from the soundtrack into 8 stereo tracks. Predefined mix parameters and precise transitions are then called up by the audio engine, depending on the mood desired to be conveyed at a moment (calm, energetic, tense, etc.). These mood transitions are called by an event system, and certain musical patterns, or “stingers” (ie patterns synchronized to the current tempo of the song being played) are sometimes triggered in order to support a transition or an action.

The goal of these real-time maneuvers is to ensure that they are strong enough to have a positive effect on the experience, but remain implicit and subtle enough that the player does not notice them. in order to install immersion and reinforce the emotional impact of a scene. These dynamic mixing systems then bring a feeling of cohesion and continuity to the sound experience, which naturally contributes to maintaining the state of flow during a game. However, as Paul Weir says: “Sound design dynamic mixing systems are really about game design and the game itself; it is impossible to offer an answer or to develop a structure that can be applied to all the different scenarios, the number of which is almost infinite. It is then up to the sound designer to assess the needs and particularities that make up the gaming experience, in order to be able to develop and put in place the systems necessary to maintain the sound balance during the game.

5-F - The budget, an obstacle to complex sound integration depending on the scale of the productions

“I think it's possible to make pretty interactive music, but it's much longer, more expensive and more complicated. If we want to create musical compositions that manage to go beyond their primary loop structure in order to provide a strong emotion to the player, beyond their utilitarian functions (that is to say, to provide a tension or a feeling of appeasement depending on the context), it is quite difficult, and not all productions can afford it.” Arnaud Roy tells us. According to him, for an equivalent budget, you can have “one hour of music with the standard structure”, and

“twenty minutes of interactive music”, which therefore comes down to costing three times as much. This is why most development studios prefer a quantitative approach to a qualitative approach, in order to break the feeling of repetition with the least possible means. However, this postulate can be disputed: in the same way as an implementation of modular systems, more complex upstream, allows thereafter a growth of the sound variety and the number of more optimal assets, a musical unit less provided in number of assets but then offering a much greater variety in terms of combinations and interactions makes it possible to provide a more reactive, pleasant and impactful experience to the player, by more effectively preventing any feeling of repetition. “That said, the problem with building these kinds of systems is that they are most often based on a large mass of additional lines of code, requiring specialized audio programmers, which is relatively expensive. Beyond a certain scale of production, the composer alone cannot ensure both the creation of the sound design and that of the systems that arrange it, and at the same time take care of the implementation of the latter. , he concludes. Thus, the budgetary question would be directly correlated to the flow potential of a video game experience.

In an article by Gamasutra (using the words of an issue of "Game Developer" magazine published in September 2005), Alexander Brandon, a video game composer who has notably worked on titles such as "Unreal Tournament" or "Deus Ex", tries to give an idea of the fee schedule for sound effects creation and music composition. According to Brandon, the industry standard rate per minute of approved music would be anywhere from $1,000 to $1,200. "When you submit a budget for a composition job, you must estimate the number of minutes required for original music in game. The number of MR (“Minutes Required”) is currently around 2 to 3 minutes of original music per level, but designing an adaptive soundtrack often makes it possible to reduce this number and therefore reduce costs.” Even if the article dates back a good fifteen years, we can assume that this statement still remains true and that in vs In some situations, designing an adaptive soundtrack may be an economical rather than an expensive alternative. Brandon continues, this time describing average prices for sound effects design: “It ranges from about $5 per relatively simple sound (like footsteps), to $50 per more or less complex sound (like a machine gun sound requiring the development of one or more loops read according to a more or less costly system in terms of production time). The formula I use to estimate the cost of a sound effect is: (FX Rate x Production Rate (Simple Effect))+(FX Rate x Production Rate (Complex Effect) = the estimated amount for production and delivery of a sound effect.”.

To end this part, we can once again quote Paul Weir: “The problem with the implementation of increasingly complex and realistic systems is that beyond a certain level of detail, the average player is no longer able to to perceive the difference vis-à-vis a simpler system. It then becomes difficult, for the sound designer as well as the budget manager(s), to justify the implementation of such systems. It is therefore for the sound designer to question each system according to

their cost/increased perceived sound quality ratio, in order to be able to correctly allocate the resources and manpower necessary for the development of the sound design of a game. .”

5-G - The decorrelation between complexity and efficiency of a sound implementation

As we have seen so far, building complex systems for arranging the different elements of a sound design has its advantages in terms of flow potential, as well as its disadvantages in terms of cost. However, depending on the type of game, it is not necessarily necessary to design a multitude of complex systems in order to immerse and maintain the player in a state of flow. Take for example the case of the game “Hotline Miami”, released in 2012 and having marked more than one player by its brutal and galvanizing sound design, embellished with visuals of a rare violence, although visually very minimalist. Here, the music of the game is composed of fixed pieces, never sequenced according to the action, and the various actions of the player and the enemies which surround him (sounds of shooting, of doors which open and close, sounds death of enemies, etc.) are based solely on an event approach (one action/one event = one sound). However, the choice of music and the very design of the sound feedback is such that the player remains glued to his chair from the beginning to the end of a level, and regardless of the number of attempts to achieve it. This is due to two things: firstly, the various music that makes up the soundtrack, carefully chosen by the developers and involving a multitude of electronic artists, all have a very catchy rhythm and structure, although relatively repetitive. The musical part then comes into symbiosis with the rhythm of the game, allowing the player to find himself in a life-size musical clip of which he would be the main subject. Secondly, the impact and controlled sound design of the firearms and the player's actions "cut" through the musical ensemble, sometimes giving the player a feeling of extreme power relative to his actions and sequences. killings, sometimes a permanent feeling of adrenaline, increased tenfold by the "die and retry" side of the game design (that is to say that the player dies at the impact of the slightest blow being dealt to him, forcing him then to start again at the last save point). Thus, as Masami Komuro says, “sound interactivity makes it easier and smoother the dialogue between the player and the consequences of his actions, however, depending on the game design and the game concept, I would not say that it is necessary. Hotline Miami uses other techniques to bring the state of flow, in particular the application of a strong contrast between, on the one hand, the visual area of the screen and the very reduced spaces of the level design, and on the other

hand, the very ample movements and actions of the player served by a very dense soundtrack. However, if the design of complex and interactive systems can be, according to the different scenarios, decorrelated from the efficiency of the implementation, the larger and freer the game space (as in an open world), the more the number of emergent problems come to require a thorough and ambitious reflection of the sound interactivity and the systems which constitute it.

In this part, we were able to see how the design of sound interactivity could constitute the cement of the appearance and the maintenance of the state of flow. We will now explore the dangers of flow breaks in sound design in more depth.

6. The Dangers of Flow Breakage Related to Sound Direction

6-A - Abuse of sound stimuli, or fear of heights

“Recently, I was able to watch several hours of Far Cry 5 gameplay, and I felt like I was watching a hyperactive child gesticulating all over the place. It seemed that the title refused to provide the slightest break to the player, constantly assailing him with activities or altercations with various enemies, coming to get in the way of his movements. It's not the first time I've felt this feeling of suffocation in a video game, but it's the first time it's appeared to me so clearly, when I don't even have the controller. in hand." tells us the YouTuber “Pseudoless”. “Why does calm seem to have become an issue for mainstream AAA gaming? Most high-budget open-world games frequently share this fear of player boredom or disinterest. However, the

surface response chosen to counteract this phenomenon is disproportionate. It's as if the game designers started waving a bunch of keys near the ears of the player, with the aim of constantly stimulating him," he continues. This feeling of frustration linked to the overfilling of space due to a fear of emptiness is unfortunately not new.
However, the sound direction, victim of these unfortunate choices of game design, only accentuates this frustration. "Sometimes preferring a qualitative approach to a quantitative approach helps to enhance the gaming experience." This postulate, naturally applicable to the sound direction of a video game, is notably that adopted by the series of horror thriller games "Silent Hill". "Silence is also a sound.", as the series' flagship composer, Akira Yamaoka, likes to remind us. "In Silent Hill 2, we tried to create a sense of fear sustained by silence. That silence lying within each individual, which is gradually transformed into anxiety. It was a kind of fear that built up little by little. Gradually, it escalated into horror." Through these statements, we understand that the strength of the sound design of Silent Hill 2 does not reside so much in the layout of complex systems (in any case necessarily rudimentary given the period context of 3D games), but more in the player's gradual introduction to dread. Using the sacred rule of "show, don't tell", the sound design knows how to be very light and subtle, as the character of James Sunderland sinks into the ghost town of Silent Hill, a sort of dream as straight out of the tortured imagination of David Lynch. For example, when the player arrives in the city, the player is attacked from all sides by enemies emerging from the suffocating mist, whose presence and the noise of radios coming to alert James of their presence fills the sound space. But suddenly, when the player ventures into the residential building which constitutes the first level of the game, silence comes to engulf the player, leaving audible only the crunch of the footsteps of his avatar on the ground. "I think that selecting moments of silence is another way of producing sound.", says Yamaoka, a concept he applies to the entire sound design of the horror series.
However, the approach to fear through the use of silence and suggestion is twofold here. Indeed, the average player having been unaccustomed to the presence of silence, through the various audiovisual works encountered (games, cinema), suddenly forcing him to confront his internal fear of the void via sound design comes close to the masterstroke. This is what still sets the Silent Hill license apart from other horror games on the market today: this ability to play with the expectations and psychology of the players, in order to increase the horror potential of the work tenfold. It's also this use of sound design that inscribes Silent Hill 2 as a timeless reference to fear, even 15 years later. However, the remake released in 2019 of the famous horror game "Resident Evil 2" adopts an approach that is both similar (especially in its management of the rhythm punctuated by silences), and at the same time opposite: indeed, in the game , our avatar is permanently stalked by a monstrous entity as imposing as it is threatening, named "Mr. X". In order to establish the constant threat that this antagonist is supposed to represent, the sound designers had the idea of continuously hearing the heavy footsteps of our opponent, who is constantly approaching the player wherever he is. . These footsteps are so well localized (in particular via the simulation of obstruction and occlusion effects),

that they allow the player to be immersed in pressure and constant vigilance vis-à-vis danger. Thus, in particular through its judicious sound design, this remake of Resident Evil 2 manages to convey the emotions and moments of tension induced by its game design, without involuntarily boring or frustrating the player.

However, if horror games or rather slow in terms of rhythms can afford to use such a sound design to establish their atmosphere, this is not necessarily the case for other titles, such as action ones. , which, due to their game design, struggle to escape the curse of sound filling. However, some titles, such as Blizzard Entertainment's competitive game "Overwatch", manage to implement through various user data collection systems a dynamic mix via a hierarchy of the importance of the events surrounding the player, which makes it possible to avoid the abuse of sound stimuli, while accentuating the readability of the action.
One of these systems consists, via the “dense clarity, clear density” approach (mantra of Walter Murch, sound editor on Apocalypse Now), in removing elements from the mix in order to leave more room for the main elements of the sound design without that this deletion be noticed by the listener. For this, the Overwatch audio team assigns each of the sound events to a hierarchy of priorities, based on a “threat level” calculated in real time: in summary, the elements with the highest level of threat will be highlighted in the mix, while those with a lower level will be indented or even disabled. Such a system allows the player to identify and focus more easily on the feedback necessary for his survival, which clarifies the action and gives him a feeling of control and additional mastery vis-à-vis his environment. Thus, by orienting the sound design around a clear representation of information, rather than a basic and abusive over-representation of in-game events through sound, the Overwatch audio team manages to encourage the player to refer to his hearing, rather than just the visual feedback provided to him. Such a design then comes to support the state of flow, keeping the player immersed and engaged in a permanent feeling of alertness without tiring him.

6-B - The question of digital latency, separating the input from its sound response

As we have seen in the section on the relationship between sound creation and the appearance of the game feel as well as the flow, it is possible to introduce a perceptible latency relating to the space left between the start of a file sound and the peak of its transient. However, sound creation is not the only element in the chain that can introduce this phenomenon of latency: the code, or even the interaction of the audio engine with the game engine is likely to create digital latency, coming to affect negatively affect the player's experience, and join the list of obstacles to

maintaining the state of flow. Digital latency can be defined as: “The time required to convert an analog signal to its digital representation”. Depending on various situations (implementation, game medium, code, etc.), the game engine may take some time to send information to the sound engine, which in turn may take some time to respond to the game engine. Also, a more common issue may be the time it takes for the sound engine to calibrate the triggering of an event to the image, which is often the case when the sound engine is running asynchronously with the video engine (i.e., regardless of the player's frame rate). This then leads to a desynchronization between the video engine and the sound engine, which each operate on different “threads”. If in some cases, the latency and therefore the desynchronization introduced can be imperceptible, in others, it can harm the immediacy of the feedback, and therefore, inevitably, the feeling of control of the player vis-à-vis his experience.

To circumvent this problem, Jakob Schmid, the designer and audio programmer on the game “140” (but also previously on the famous INSIDE), manages to create a perfect synchronization between the sound design and the animations, by launching the latter directly via the part of the code relating to the sound. Also, all the musical loops and sound elements are triggered at the same time, i.e. when the scene is launched, so as not to have to re-trigger these loops and thus cause a desynchronization between the visual and the sound. The fade-ins and fade-outs are then performed through the code, which makes it possible to conceal any notion of latency relating to the triggering of a sound event, and thus maximize the potential flow of the experience through instant responsiveness.

6-C - Ignorance of psychoacoustic phenomena

As we have seen in our section on localization, psychoacoustic phenomena constitute a usually underestimated, but nevertheless very real, part of the way in which the player perceives and feels the sounds constituting his environment. Ignoring these phenomena during the sound design process can lead to the player's attention being diverted from an event or action being performed, thus interrupting the state of flow. of the player. Let us review some of these phenomena.

6-C - Part 1: The drone phenomenon

This phenomenon is characterized by the presence of a stratum within a sound set, a stratum of stable pitch and without noticeable variation in intensity. Linked to music

in its designation (the drone is a low permanent sound on which certain pieces are based), it is also observed in industrial and urban soundscapes. Many technical systems generate sound constants that come close to this effect (mechanical ventilation, fluorescent tubes, hum of motors, pipes, etc.), even if the frequencies concerned are not limited to the low sounds that originally characterized it. In a video game scene, this effect is most often manifested by the presence of room ambient noises, or “room tones” (generally made up of a set of noises originating from the examples mentioned above) . Representing these types of sounds too intensely in the sound mix, or prolonged exposure to them, can quickly lead to auditory and mental fatigue that can put an end to the state of flow.

6-C - Part 2: The phenomenon of anticipation

While waiting for a situation to come, a person “pre-hears” - that is to say believes that he actually hears - the expected signal, whereas no sound has yet been emitted. This effect is observed both in situations of expectation of unknown sounds, where the slightest rustling becomes a clue, and in familiar circumstances where the listener anticipates a predictable (pre-audible) sound context in his memory. It is with this phenomenon in particular that the sound design of Silent Hill 2 plays, in order to maintain the player in a vigilance and a continuous tension. However, if this phenomenon can participate in maintaining the state of flow in the experience, it can also interrupt it: if a character is moving and the player expects to hear his footsteps, but these don't appear in their sound environment (for example, if the character is too far from the player and their footsteps attenuation settings make the source volume too low), a phenomenon similar to that of “uncanny valley” will occur, which can then come to distract and break the immersion of the player, resulting in this case in a loss of the state of flow.

6-C - Part 3: The phenomenon of irruption

The phenomenon of irruption characterizes an unforeseen sound event, coming to modify the climate of the moment and the behavior of the listener in a marked way. This effect is to time what the intrusion effect is to space. For example, despite the generalization of its use, the ringing of the telephone remains for many people an aggressive sound event, less by its tone which knew how to soften, than by its unforeseen and imperious character: not only a call interrupts the present state, but it dictates a new behavior for a given time. This phenomenon occurs in particular when exploring the game space of certain open worlds, where the player, soothed by the sounds and atmospheres of nature, finds himself uprooted from his state of plenitude by the noise of a patrol. of enemies passing close to him at full speed.

6-C - Part 4: The phenomenon of remanence

This phenomenon is characterized by the sensation of persistence of a sound that is no longer heard. After extinction of the emission and the propagation, the sound gives the impression of still being “in the ear”. This phenomenon occurs in particular when the player passes from a noisy and rich sound environment to a much quieter environment. If this alternation can be used as a stylistic effect, it can also help to amplify the memory of the intensity of the previous environment, which can help to cause a rupture effect, then interrupting the state of flow of the player . The persistence phenomenon occurs in particular during a loading time which follows an action scene, where the difference in dynamics or intensity in volume creates an unpleasant psychological sensation for the player.

6-C - Part 5: The phenomenon of anamnesis

Also called “reminiscence effect”, the phenomenon of anamnesis characterizes a signal or a sound context causing the listener to return to consciousness of a past situation or atmosphere. Effect of meaning, it characterizes the triggering, most often involuntary, of memory by listening, and the evocative power of sounds. This phenomenon is often created voluntarily by sound designers, in particular through the use of certain sounds in horror games such as that of the music box, which evoke the player's childhood, and the vulnerability felt during this period. In this case, a surprise effect generally occurs soon after, in order to rush the player once his guard is down. Like the phenomenon of anticipation, the phenomenon of anamnesis can serve as a tool like appearing as a nuisance: for example, if the player finds himself immersed in a sound set, for example a city, where certain sounds come distract him from the action at hand, then the state of flow is likely to be broken. In order to avoid this phenomenon, it is therefore important to establish a hierarchy in the dynamic mixing of sources (like that of Overwatch), in order to be able to more or less control and keep the player in the flow of the action. Also, if the use of certain sounds can be used to establish a desired reference (as for example via the famous “Wilhelm's cry”), this process tends to harm the experience, rather than to reinforce it.

6-D - The concept of nuisance score and the unit of measurement “NUFS”

According to Bjorn Jacobsen, the “nuisance score” characterizes a measure relating to the level of annoyance produced by one or more sounds. Each of these sounds or set of sounds are then labeled according to their nuisance score and their potential occurrence, in order to define which of these elements must stand out from the others within the mix. Jacobsen describes in particular a case in point, where a sequence of footsteps made up of 5 sound files, each altered in pitch and volume randomly for each repetition and thus providing a theoretical infinity of variations, stood out from the rest of the sound environment in a harmful way. This was caused by one of the sequence's footstep files, which sounded "too different" from the other sounds in the cycle. Over time, the player's attention ended up focusing more and more on this too singular element, resulting in a pattern of repetition that was as predictable as it was unpredictable, and then creating a significant noise nuisance in the player's experience. In doing so, Jacobsen assigned a nuisance score of 8 to footsteps, as well as sounds of doors and explosions. So these sound events came to be considered “immersion-breaking sound bugs”. Such a nuisance score could also be substantial depending on certain playing spaces or levels, if the harmonics of the soundscapes came into dissonance with the harmonic content of the music played at the same time.

The concept of “NUFS” (Nuisance Unit Full Scale), theorized in the work “Tuning of a Sound” by Bjorn Jacobsen, therefore consists in representing the “nuisance score” of a sound design relative to a certain duration. This process, although relatively subjective, amounts to measuring the nuisance score of the different sections of the game and categorizing them by a measurement in NUFS, in order to be able to more easily target the various problems related to sound design, and guide the creation of systems. active.

6-E - Issues of decorrelation between a sound event and its intention

The phenomenon of decorrelation between an action and the sound event associated with it is relatively common. For example, if the sound impact of a sword strike is inversely proportional to that suggested by the animation or the situation in play, this can create a form of cognitive dissonance. This can especially happen

when the sound designer is content to create and deliver a sound file, without having implemented and tested it by himself beforehand. This type of negligence can lead to many inconsistencies in the sound design of a game, becoming detrimental to the appearance and maintenance of the state of flow. As Paul Weir says: “I understand that depending on the scale of the productions and the organization of their hierarchy, it is sometimes complicated to be able to implement and verify your work yourself. However, there is always the possibility of requesting daily checks, in order to be able to prevent any negligence on the level of the design of the sounds or the systems which arrange them”.

It also happens that, depending on the level of repetition or the low level of variety of a sound, the latter comes to lose in importance and impact in the experience, even to be frustrating and then coming against the intention. departure. It is therefore preferable to prioritize a qualitative approach to a quantitative approach: the higher the layout of the systems and the higher the level of perceptible variety, the less it is necessary to add content in order to circumvent the redundancy of the same sound.

7. Without the image, is the sound enough to reach and maintain the state of Flow?

According to Masami Komuro, “it is easier to bring the state of flow into a game that consists only of sound elements and interactions (e.g. responding to the player's voice), rather than bringing the state of flow in a game where any sound aspect is absent.” The increase in the quality of sound in the video game over the generations and technological advances has led to a desire to master the sound design of this medium - in particular through interactive music -, so as to be able to more easily introduce this state of flow, or at least the components that constitute it (intense concentration, commitment, temporal alteration, etc.). Already, through the sound design of the strategy game "Z", released in 1996 by the studio The Bitmap Brothers, the developers began to experiment with the notion of interactive music as a means of communication, in order to convey to the player in a more fluid and instinctive the information needed to understand the action (and thus bring about and maintain the state of flow), in particular by the sequential reading of MIDI segments depending on whether the player builds new buildings, new troops, etc. .”. The following question may then arise: without visual feedback, can sound design be enough to achieve and maintain the state of flow?

If this question may seem secondary in the eyes of an individual in full possession of his visual abilities, it is quite different for visually impaired and blind people. We can only welcome an approach like that of the Overwatch audio team, which encourages the fact of playing by using almost exclusively the sound information conveyed by the game. A video on the YouTube channel “Vice News” tells the gamer journey of Michael Espinoza, blind player since birth, and how he manages to play mainstream titles such as Mortal Kombat, yet not intended or oriented around accessibility for visually impaired and blind people. If the work of learning and memorizing stroke patterns, the delay times between them, as well as the identification of strokes in relation to their sound response is naturally more laborious than that carried out by a player in full possession of his visual abilities, these games are nevertheless made accessible thanks to their sound design. We can therefore say that without the image, the sound can be enough to make a player reach the state of flow, even if the symbiosis of the visual and sound experience is always more effective for a lambda player. An amusing anecdote from Christophe Héral confirms this: “During a conference given at an edition of the Toulouse Game Show, and dealing with video game music, a man complained that in Rayman Legends, the sound does not did not come to inform the fact that the avatar of the player moved against a wall. If this remark had surprised me at first, I realized later that the man in question was blind,” he said. “It's interesting, because thanks to the sound feedback provided by the game, such as the sounds of lava or danger that get louder as you approach it, even a blind person can play Rayman Legends! This is all the more confirmed with the musical levels of the game, where by memorizing patterns and above all concentrating on the music, a player can only use his hearing to go through a level.”. Indeed, the musical levels of Rayman Legends, designed as interactive music videos by Christophe Héral and the sound implementer Adrien Pavageau, consist of a musical platform game test where each jump or action to be performed is announced by the construction of the piece, as well as underlined by sound feedback synchronized to the tempo and the musical structure. This then results in an exhilarating image-sound symbiosis, instantly carrying the state of flow, which nevertheless remains playable by a visually impaired audience.

Finally, if some games allow in a more or less unforeseen way to be experienced and to bring a state of flow only through sound, others, like the independent title “A Blind Legend” released on computers and devices mobiles in June 2015, are directly designed to be experienced by visually impaired people. Smart, the game revolves around simple controls, namely a swipe up, down, right or left of the screen, allowing the player to have a tactile level of interactivity with the action. For the rest, no visuals: everything happens at the level of binaural sound, allowing you to orient yourself and move around in space via sound feedback. While an experience such as this is obviously ideal for a blind person, it is all the more interesting for a player with intact vision, who must then learn a new way of playing, connecting in a deeper way with their hearing. Then, the images become mental, drawing on the player's imagination, thus making the experience unique for each profile. This reconnection to the primary senses of the human being allows the player to immerse himself much more effectively

than if the visual information were offered to him directly, thus maintaining the interest and the commitment of the player for the benefit of maintaining his state of flow.

Conclusion

During this thesis on the relationship between sound and appearance as well as the maintenance of the state of flow in the video game experience, we were able to understand and define the state of flow in the context of a video game. , explore the different ways of creating and arranging sound design to bring about this state, and review a multitude of concepts, methods and implementation systems to maintain it. Through the experience of professionals in the video game industry, as well as deductions based on scientific studies, we have been able to understand that, if there is not an answer applicable to all the different types of video game experiences with regard to the appearance and maintenance of the state of flow, there are on the other hand several tracks or axes of improvement which come together in most sound designs, at least those carried out in a relevant and thoughtful way. We therefore note that the understanding, alteration and management of the different characteristics specific to a sound, as well as the development of a sound design broken down into a set of flexible and interactive systems, make it possible to facilitate the appearance and maintenance flow, and this, even by dismissing the question of the audiovisual report. If the development and use of all these concepts is not easy, and sometimes involves budgetary issues, it is on the other hand crucial: it is this which makes it possible to maximize the potential for pleasure and commitment of the player. vis-à-vis his experience, in order to give game designers the possibility of correctly conveying the desired message(s) (whether emotional, personal or even political).

I take advantage of this last paragraph to thank, without naming them (the list was far too long), all the people who helped me, in one way or another, in the conception and writing of this work, and without which the result would not have been the same.
Finally, I thank in advance the reader(s) or the reader(s), who were able to take an interest in this subject and try to understand it by my side.

Sincerely,

Gabriel Sauvage

Bibliography

"Flow: The Psychology of Optimal Experience" **- Mihaly Csikszentmihalyi**

"Flow in Games" **- Jenova Chen**

"The Tuning of the Game" **- Bjorn Jacobsen**

"Towards an Understanding of Flow in Video Games" **- Ben Cowley**

"Game Feel" **- Steve Swink**

Ludography

“Doom Eternal” - 2020, Id Software

“Manifold Garden” - 2019, William Chyr Studio

“Inside” - 2016, Playdead Games

"Limbo" - 2010, Playdead Games

“140” - 2013, Abstraction Games

“Rayman Legends” - 2013, Ubisoft

“Get Even” - 2017, The Farm 51

“Remember Me” - 2013, Dontnod Entertainment

"Silent Hill 2" - 2001, Konami Computer Entertainment Tokyo

"No Man's Sky" - 2016, Hello Games

“Grand Theft Auto 5” - 2013, Rockstar Games

“Overwatch” - 2016, Blizzard Entertainment

“Z” - 1996, The Bitmap Brothers

“The Legend of Zelda: Breath of the Wild” - 2017, Nintendo

“Thumper” - 2016, Drool

“The Secret of Monkey Island” - 1990, Lucasfilm Games

“A Blind Legend” - 2015, Dowino

“DARQ” - 2019, Unfold Games

“Hellblade” - 2017, Ninja Theory

"Bioshock: Infinite" - 2013, Irrational Games

The relationship between sound and the state of Flow in the gaming experience

Summary

This thesis focuses on the relationship between sound and the state of Flow within a video game experience. I describe, in several parts and subparts, the concept of the “flow” state. Then, I explore the way in which sound designers contribute to recreating this state of flow in the gaming experience, through its sound direction. First of all, I make an introduction to flow and undertake to define this state, distinguishing it from the state of immersion. Following this, I explain how the mastery of sound characteristics constitutes the pillar of the appearance of the state of flow, and how the integration of sound design in the game design process, as well as its implementation via the The development of flexible and interactive systems makes it possible to convey this state to the player.

Keywords: Flow, Video Game, Gaming Experience, Sound Design, Composition, Sound Design, Interactivity, Sound Interactivity, Game Feel, Wwise