

# **Music in Video Games**

An Interactive Qualifying Project Report  
Submitted to the Faculty of  
WORCESTER POLYTECHNIC INSTITUTE  
In partial fulfillment of the requirements for the  
Degree of Bachelor of Science  
By

\_\_\_\_\_  
Shiyi Liu

\_\_\_\_\_  
Yizhen Wang

August 25, 2018

Advised by:

Prof. V.J. Manzo, PhD

Worcester Polytechnic Institute

**Contents**

Abstract ..... 2

Introduction..... 3

Background..... 4

    History of video game music ..... 4

    Feature of video game music ..... 6

    Functions of Video Game Music..... 8

    Placement of Music ..... 11

    State the problem..... 12

    Examples of the problem or question ..... 13

    Variations of the problem or question ..... 14

Methodology & Design ..... 16

    Methodology ..... 16

    Design ..... 17

    Data ..... 20

    Sample Base: ..... 23

Analysis:..... 24

    Over View ..... 26

    Individual Analysis ..... 31

    Over-analysis ..... 83

    Retrospection and Improvement ..... 86

Conclusion ..... 89

References..... 90

## 1. Abstract

Through decades of evolution of video games industry, the audio element has become an inseparable part of the video game experience. However, for different game genres, the literature is still sparse regarding how background music (BGM) fits a game well in terms of being compatible and reciprocal. This report explores the function of video game music and what makes the BGM an irreplaceable part of the gameplay experience so that the players do not switch off the background sound. To achieve our goal, an experiment was prepared to evaluate how BGM influences people's perceptions toward the game. By providing video game clips with switched BGM to subjects and surveying their perceptions and feelings of consistency between the clips and BGM, we were able to gain insight into how music may affect people's game play experience.

## 2. Introduction

From 1958, the birth of the first video game Tennis for Two till now, the video game industry grows from zero to an 82 billion dollars industry. As video games become increasingly important in human daily life, the academic study in video games goes deeper. Video game music as an essential part of video game often being ignored in previous studies attract our attention.

Our project could be divided into four main parts, introduction, background, experiment and conclusion. First, the introduction part explains the structure of our project. Second, the background part contains the history, feature and function of video game music help to state our question, what makes video game music effectively affect people's game play experience and figures out the variation of our problem. Third, the experiment part provides methodology of our experiment as well as analyzation of our data. Finally, the Conclusion part states that the consistency between video game music and the game itself could be affected by various factors. It is hard for people to detect which exact part affects the consistency. However, we still hope our study could give the subsequent studies a hint in this field.

## 3. Background

### 3.1. History of video game music

In the past 60 years, a new kind of interactive media called video game were developed rapidly. Considerable attention has been paid to video games on how video games entertain people and how could game developers make better games. However, game audio as an important component of video games are often being ignored. People consider video game music as a kind of useless and substitutable part of a game. For example, Microsoft insistently believed that music in every Xbox 360 game should be replaceable with the user's own music files (Harlin 2018). Thus, to enhance the importance of video game music, it is necessary for us to trace back to the birth of video game music and figure out the role BGM played in video games.

The game industry was born from a war time era with the advanced technology. In 1958, Willy Higginbotham, a young nuclear physicist created *Tennis for Two*, which is believed to be the first video game in the world. By using the computer previously work for the military projects, this interactive entertainment experiences opened the door of the video game industry. The very first arcade video game was called *Computer Space*, from 1971. It only has one sound channel which noise and tones blended in as needed. However, the hit of the video game industry did not exist until, Atari's *Pong* one of the most famous game in the history came out. In *Pong*, beeping sound runs out whenever the ball hits the paddle. Although the existence of the sound in *Pong* was considering to be a small accident, it is still an important factor of making the sound in video games famous (Collins 2008, p8-9).

At the early age of sound in video games, sound had been used as a tool to attract people to the arcade. By 1980, due to the limitation of the technology, although most game

systems had co-processors for sound, continuous music in game had yet to exist. (Collins 2008, p15). The birth of the home console systems pushes the industry into a new direction. In 1985, Nintendo launches its first console with five channels, the *Nintendo Entertainment System*, often known as the NES. Games like *Super Mario Bros* (Nintendo, 1985) and *Legend of Zelda* (Nintendo, 1986) not only helped Nintendo to open the America market but also proved the potential of the home consoles. Nevertheless, music in those games also becomes the memory of the growth of one generation.

Another great leap forward in game audio occurred in 1989 with stereo channels (two fixed voices on one channel) for the *NEC Turbo Grafx 16* (6 channels) and *Sega Genesis* (10 channels). In 1990s, as more and more home console existed in the market, the audio hardware of those home consoles went even further. In 1999 Nintendo 64 has the capability of 100 channels and (48kHz) better-than-cd quality. After entering in the 2000s, 2006 is a landmark year for game audio with Sony's *PlayStation 3* released. With up to 512 channels, 8 outputs (7.1 Dolby Surround at 96kHz) DSP and Super Audio CD capability, *PlayStation 3* gives players a better game audio experience than any other home console at that time. The success of the next generation home console such as *PlayStation 3*, *Xbox One* and *Nintendo Wii* shows the capability of home consoles and provides players awesome auditory effect.

Along with the development of arcades and consoles, the gaming industry also focused on personal computers. In 1977, Home computing gets its first real audio experience with the Apple II Home Computer. Soon after, home computers such as the IBM PC, Atari 400 & 800, Commodore 64 and others followed from 1977 into the late 1980s. At those days, most home computers just have three to four sound channels, and one or more channel capable of both variable tone and noise. For many years audio of personal computer remains almost the same , just like the Atari 2600 sound capabilities. However, breakthrough in music technology appears as Musical Instrument Digital Interface (MIDI) standard exists. (Aska 2017) Due to the fact that a

MIDI recording is made, only the data about that instrument (note on or off, pitch, volume, etc) is recorded and played back, the MIDI file is much smaller than the digitally recorded audio clips. 1983, MIDI became the standard for musical instruments, drum machines, and other audio-related equipment to communicate with one another both live, and on the computer. In the late 1980s, the concept of the Sound Cards (an external hardware dedicated to playing sound so that the CPU could do other tasks) was taking hold. In 1989, Sound Blaster which introduced by the Creative Labs, led sampled audio in games, and a basic synthesizer with the ability of MIDI playback. However, after the existence of CD-ROM technology, game developers abandoned MIDI in games owing to the fact that CD-ROM helped to produce sound more realistic (Collins 2008, p63). Today, sound cards of personal computers have caught up with the current state of console audio which help to provide PC gamers wonderful gaming experiences.

The history of sound in video games were highly effected by technology evolvment. As home console and PC gaming become more and more mature, arcade is on the wane. Other gaming platforms such as mobile phone, handhold console and rhythm-action add-ons also made the game industry more diverse. With less technology constraints, game audio developers are able to create more consistent music and helps to increase the reality of the games. Although the function of video game music is still being evaluated, the importance of game music is above suspicion.

### **3.2. Feature of video game music**

Game audio could be divided into two parts, diegetic sounds and non-diegetic sounds. According to Karen Collins, diegetic sounds are source music, narratives and other sounds for characters that help video games to tell its story. Whereas non-diegetic sounds contain background music and sound effects (Collins 2008, p184).

As for non-diegetic sounds, background music in video games are different from any other audio experience since it is non-linear. It is a kind of dynamic audio with both interactiveness and adaptiveness. Unlike music in films or TV programs, players could engage in the music playing process in video games. Background music may either be interactive (change based on gamer's action) or adaptive (change based on the changing environment) or both at a specific time in the game and the features may change throughout the game play.

### **3.2.1. Adaptive music**

Music in video games could be adaptive when it changes with the gameplay but unaffected by the play's action (Collins 2008, P126). One of the most famous multiplayer online battles arena, Dota2 (Valve, 2011) shows the adaptiveness of music by manipulating it's background music when the "night theme" of the game begins.

### **3.2.2. Interactive music**

Background Music could also be interactive when the music changes due to the player's action. In *Portal2* (Valve Corporation, 2011), the background music changes in mix and volume when the character move in space ("Portal 2'S Dynamic Music - An Interview with Composer Mike Morasky, And Five Tracks To Listen To Now!" 2011). Moreover, when the character jumps on the "repulsion gel", a loop will be added to the music.

### **3.2.3. Interactive and Adaptive music**

When music in a game featured both adaptive and interactive, both play's action and the change in background will affect video game music. The *Devil May Cry* series is a famous action game series of Capcom. In this series, in addition to the gorgeous movements, smooth operation, and awesome sense of percussion, the distinctive rock music which is gradually rung as a player enters the battle has become one of the important factors that make this series succeed.



### **3.2.4. Linear music**

However, not all music in video game are dynamic music. In video game cinematics, music plays the role just like it plays in films. Cinematics are one of the parts in games that could not be interacted, thus music in cinematics are normally linear and non-dynamic. For instance, in *Monster Hunter: World* (Capcom, 2018), every time a player discovers a new Monster, the gameplay will stop and a monster intro cinematic begins. While the cinematic is playing, the player could not interact with the game, thus, the music will not change until the end of the cinematic.

### **3.3. Functions of Video Game Music**

Compared with many other elements in games such as plots, characters and game mechanisms, music is one of the most easily overlooked element in the games. Game scene and gameplay themselves could occupied players' attention easily. However, once players choose to turn off the music in the settings, they might realize that something is missing immediately. Background music usually don't take the lead, but they actually support the games as much as any other elements in games do.

Music served for multiple functions in video games. At the early stage of video game history, arcades played music out loud to attract players to the machines. When time goes by, music now could not only provide entertainment value, but also adds immersion to games, set the mood of the games and provide feedbacks to players. Nowadays, game audio composers have less restrictions owing to the evolvement of technology and nature of game industry. Although the process of creating music in games may vary from company to company, from platform to platform and from game to game, music themselves still generally holds the same types of functions throughout the gameplay.

### **3.3.1. Entertainment Value**

There is no doubt that background music has entertainment value. It could fill in the gaps exist between ambient sounds and sound effects. For example, in *Angry Birds* (Rovio Entertainment,2009), the background music is fun and memoizable. Players enjoy the music in *Angry Birds*, and the music may get stuck in their heads. The casual game soundtrack is relatively fashionable, modern, cheerful and relaxed, allowing the player to experience relaxation and entertainment. Guitars, basses, keyboards and drums are usually used in the selection of musical instruments. Most of the arrangement methods used in casual games are based on relaxed style. The storyline in *Angry Birds* is fairly easy, a group of pigs stooled the bird's egg and made them very angry, thus the wingless birds decided to destroy the pigs ("History Of Angry Birds | Freeangrybirdsgame.Org" 2018). Since sometimes there are few stories or no story in casual games, the entertainment value of music becomes. especially important.

### **3.3.2. Immersion**

Another important function of background music in video games is that it supports players' involvement to games both physically and mentally (Munday, 2007). Background music make games more immersive physically by a well-known effect called cocktail party effect. Colin Cherry's study in 1953 (cited in Sara Rigutti, Carlo Fantoni, and Walter Gerbino, 2015) described this effect as "the ability to tune into a single voice and tune out all others during a crowded party". During the gameplay, music contributes to video game immersion by occupying the area of the brain dedicated to dealing with nonlinguistic sounds (Munday, 2007). Moreover, music also functions as a barrier to prevent noise coming outside the gameplay and reduced the possible distractions from the real world.

Moreover, background music adds immersion mentally by setting the mood and identifying the era of video games. Music is one of the most beautiful artform in the world. It can

be used to convey emotions, highlight the atmosphere and point out the theme of a story in traditional film and television industries. Music in games at the same time maintains most of the functions in those industries. At post-production stage, game audio composers normally work after received the setting of a game. Based on the storyboards, concept arts and other design documents, composers need to decide the style and the mood of background music and create several temp tracks. According to Sadoff, 'temp track' is a temporary mock-up of a film's sound track and is assembled from pre-existing music before the real commissioned score being composed (Sadoff, 2006). This shows the close relationship between the development process of video game music and the development of games. In this way, music strongly tight with video game itself and express same tone of the game. Once a player sees the game scene, they might have certain expectations as the listener for certain musical themes that match game genres. Also, if a game sets in a particular time period or location, it will demand a specific style of music. For example, in massive multiplayer on-line first-person shooter game, *Destiny 2* (Bungie, 2017) , the original sound track in EDZ is a somber orchestral music which revealed the attributes of European Dead Zone, the last city of earth.

### **3.3.3. Provide Feedback**

Furthermore, background music has the ability to provide feedbacks to players. In *Monster Hunter: World* (Capcom, 2018) mentioned above, when the player approaches the desired monster, background music with tension exists. In this way, gamers are able to realize that they have found what they need to hunt.

### **3.3.4. Establish Pace**

Background music may also establish the pace of a game. Tempo can affect the overall pace of gameplay. A peaceful village may have calm and flowing music. However, once the war spread to the village, background music may accentuate the sudden change by speeding up the composition or changing to a more dynamic one. This function appears in the very early stage

of the video game history. In *Space Invader* (Taito, 1978), often known as the first video game which has continuous music, the speed of music becomes faster as the pace of the game becomes faster. This might be the earliest example of how background music set the game pace.

### 3.4. Placement of Music

Unlike music for film or television programs, most video game music are dynamic. The special ability of adjusting music gives game audio developers more freedom to place the music in various places and make changes for specific events. Due to the fact that many games nowadays have similar gameplay process, there will be some commonalities in music in the similar area of games. In our experiment, we divided the in-game areas into 4 parts, start scene (main menu), outdoor fighting scene, boss fight and base (safe zone) to discuss the key feature of those music types.

#### 3.4.1. Start Scene (Main Menu)

Music in start scene for example, should have relatively consistent dynamics. Also, since start scene gives the audience the first impression of a game, the background music should give the players a hint of the game. Music for start scene and main menu may change when players switch between menu screens or various sub-menus.

#### 3.4.2. Outdoor Fighting Scene

According to Alyssa Aska, outdoor fighting scene could also name as overworld scene which often exist in game genres like adventure games or role-playing games in which a large opened world map appears (Aska, 2017). Owing to the fact that players spend most of their time in the out-door fighting scene, and the background music should hold the mood and pace of the game.

### **3.4.3. Boss Fight**

Boss Fight scenes are essential part of gameplays which are usually considered to be tense and drastic. Boss fight music at the same time normally shows exactly the same feature as the boss fight game play. High-energy music with faster pace comes along with player's active interaction with the game. Since most boss fights are short, high-energy music in boss fight are less likely to fatigue the players. Moreover, owing to the fact that bosses are also importance characters in many games, they might have their own background story. Music in boss fight might also reflexes the story of a boss which helped the player to understand the gaming world better.

### **3.4.4. Base (Safe Zone)**

Most of the games contains existed locations that could not enters a battle, and those places are called safe zones (Aska, 2017). In many adventure or role-playing games, safe zone is a place distinct form any other area of the world. In the base of the game, players could do a variety of things, such as trade, get awards and claim tasks etc. However, battles are not allowed in this area. Thus, base is normally the most peace place in a game. Music in base could generally have a low-key, then a little high-energy that makes the music more interesting.

## **3.5. State the problem**

In decades of evolution of video games industry, the audio element has become an inseparable part of the video game experience. However, for different kind of game genre, we lack a definition of how background music fits a game well that they will be compatible and reciprocal. Every year there are a bunch of different games in different game genre get the best audio award. While playing those games, gamers often get an experience of consistency between the background music and game. This feeling separates them from the outside world,

attracts them, manipulates their heart following the pace of the game plot. However, even though some other games have good background music that is pleasant and enjoyable, the interaction between the background music and the game itself is not close that sometimes the game is independent from the background music. Thus, what element defines the consistency between the background music and the game becomes an important question.

The main purpose of this study is to figure out how and why video game music works well with game itself. This may give us the hints for future study in this field. In the discussions above, game music is normally dynamic music which is non-linear, interactive and adaptive. It has an obvious heap of functionality but functioned differently under different game genre and different placement in games. In our experiments, we switched the background music of games and designed questionnaires to collect how our subjective feel towards the edited clips and the music. Although the perception towards music and game scene are very personal, there are still many consensuses we are able to get from their answers. Through this method we could get an insight on how video game music affects the gameplay.

### **3.6. Examples of the problem or question**

In the year 2011, *Baba Yetu*, the main menu music of Civilization IV (Firaxis Games, 2005) won a Grammy for Best Instrumental Arrangement Accompanying Vocalist(s). It is known as the first Grammy award for a video game music. The language of this song is rather strange, the real Swahili, from the language of African tribes and the lyrics are a Swahili translation of the Lord's Prayer (Kuchera 2011). The song can be so touching, mainly because of the combination of the ritual music and the natural way of singing. The whole song is sung by the quaint male voice, and female voice sings the accompaniment. The layering is very strong. From the natural way of singing at the beginning of *Baba Yetu*, the music gradually unfolds the grand atmosphere. The style of the song and the gameplay of the *Civilization* series, from the

establishment of the village to the establishment of the dynasty, fits very well. There is a sense of progress from the original tribe to the future world, as well as praise and awe of the development of human civilization. According to Christopher Tin, the composer of *Baba Yetu*, since the Grammy organization's changed the classification of some of its categories to include music originally composed for video games "it's a great step forward for the video game industry as a whole to get acknowledged along the film and TV as a dominant visual medium" (Narcisse 2011).

For now, *Baba Yetu* is not just a video game theme music or a religious music, it touches so many people other than gamers and religious believers. It is a great example of how video game music works perfectly with the game itself and improves the status of video game music. However, there are still a large amount of people in this industry regard video game music as a less important element in the game and does not treat video game music seriously. Due to the fact that more and more great video game music exists in the industry nowadays, how video game music functioned well with the game becomes a question.

### **3.7. Variations of the problem or question**

According to *Physiological Stress Response to Video-game Playing: The Contribution of Built-in Music*, "auditory input contributes significantly to the stress response found during video game playing" (Hébert et al. 2005). In the experiment done by Sylvie Hébert and her colleagues, the players have been randomly assigned music or silence gameplay experience. The experiment result shows that 15 minutes after the game play, the group which had been assigned music in games shows higher cortisol levels as well as high cortisol level directly related to high stress level after playing video games. This experiment gives people a hint on how video game music functioned neuro-scientifically. Owing to the fact that music is

information in auditory form, what will happen if video game does not have music is a new aspect for people to explore.



## 4. Methodology & Design

### 4.1. Methodology

#### 4.1.1. Overview:

The objective of the experiment is to find out how gamers react to BGM and game clip, and how they feel the consistency\* (in the degree of how the music fits the atmosphere, environment and story of the game) of the video game clips with unoriginal BGM.

#### 4.1.2. The Experiment Design:

During the experiment, we are using video game clips from several different kind of game genre with different scene and switch their BGM to get new clips. After showing the new clips to subjects, we will use questionnaire to collect their thoughts and feeling about the consistency between the game and the BGM. The questionnaire will include several questions like the scale of the consistency, the style of the BGM and game separately, how will they improve the consistency, etc.

#### **1). How does this tie into the background research? Why does this make sense?**

In our experiment, we will select several representative games from each game genre and switch the BGM among them. Because games under the same genre normally have similar mechanics, their BGM are more likely to have things in common. Thus, switching BGM among games may give us an insight on how music fits the game.

The Participants:

#### **2). Who, how many, why these people?**

In our experiment, most of the subject are college students from USA and China, who consist of the most part of the gamer base for all kinds of games (the participants will be referred as *subject* from now on in the following section of the paper). Since most of video games have a particular target audience among the gamer base that are willing to pay for the games, the feedback from our gamer subject is constructive.

### **3). How does this relate to the background research?**

Though all kinds of people could be a part of our experiment, in order to ensure our experiment could be carried out smoothly, we decided to choose college students with various gameplay experience as our target audience.

## **4.2. Design**

### **4.2.1. Proposed Method for Experiment Set Up**

We will extract several clips from different games and combine them with different background music from other games making experiment clips. By showing those clips to the subjects, we will collect their opinion and review of how they feel about the “consistency” of those clips by questionnaire.

Clips in the first part will have game scene from different kinds of games: Role Playing Games, Massively Multiplayer Online Role Playing Games, Racing Games, Strategy Games, Puzzle Games. Background Music (BGM) will all have their sources from games mentioned above. However, the clips will have the scene and BGM mixed up and composed together. For example, a Role-Playing-Games’ battle scene might have a BGM from a Puzzle-Games’ home screen. All of those clips will be watched by subjects.

Questionnaire in this experiment will be used to collect subjects' thought and review about the clips and have same questions all the time. Questions will be mainly consisted by short answer questions asking about feeling about the clips: consistency between the BGM's beats and the game scene, melody and the scene's art style or BGM's intensity and game's battle scene. Other questions will ask for advices of possible improvement of the "consistency" between BGM and game scene.

The question of the clips will focus on the subjects' subjective feeling and thought since everyone might have different perspective on the same clips. Therefore, we will pay attention to the reason why subjects have feeling of consistency or inconsistency of the clips.

Materials will be pressed into a package and delivered to subjects by email, messages or internet transactions. Subjects in the experiment will chose from college students from Chinese and American universities. Subjects will be consisted by students who have played games a lot, who have played games sometimes and who have never played games. All subjective in this game will be officially contacted by email, which will be collected as confirmation to the experiment.

#### **4.2.2. Proposed Method for Analyzing the Data:**

Materials will be composed by two parts: The first part questions will be set up as answering in scale (from 1 to 10). The second part will be set up as quick answering. All materials will be packaged into a zip file and will be delivered to subjects by email, messages or internet transactions. Responding answers from the subjects will be collected by packed questionnaire zip file.

**Step.1. Describe and summarize the data:**

The answers from subjects will be collected and we will analyze the feature and trend in the statistics. Repetitive answers will be extracted from the answers set and be analyzed to find out the most agreed review point.

**Step.2. Identify relationships between variables (Game screen and BGM):**

Although the main argument of this experiment is to find out the effect of miss pairing BGM with the game screen, we will distinguish the degree of unfit the BGM and scene are by comparing the answers from the questionnaire.

**Step.3. Checking the statistics and filtering out useless data:**

Considering extreme case, some responses might be ambiguous, confusing or incomprehensible. Distinguishing those answers is necessary to keep the statistics instructive and representative.

## 4.3. Data

### 4.3.1. Data Component

In our dataset, we collected feedbacks from 19 subjects participating in our experiment. Those feedbacks, or result dataset of our experiment, included 7 kind of answers given by the subjects (those data will be mentioned as Data No. in the following part of this article):

1. the grand matching score between the BGM and the scene;
2. tag (or description word) of each clip's scene;
3. tag of each clip's BGM;
4. interaction between the clip's scene and BGM (if it exists);
5. the details in the clip shows inconsistency between the scene and BGM;
6. the game's classification in their opinion
7. and the frequency of how often they play video games.

Among those feedbacks, Data No.1, No.2 and No.3 are the most important part of our result since they clearly represent subjects' feeling about the scene and BGM. For Data No.4 and No.5, they reflect how different details (tempo, instruments, sound effect) in the BGM affect subjects' feeling. Data No.6 provides a sense of the degree of how subjects' understanding of the game's mechanics and aesthetic style. Data No.7 will represent the frequency of subjects' gaming times.

### 4.3.2. Data Handle, Extraction:

For Data No.1, we analyze the score in multiple different ways: score range, distribution, average and mode. Because Data No.1 straightly represents how subjects score each clip's consistency, we regard it as the most important data across all data

we received from the feedback. We first compiled Data No.1 score into a scatter chart in Figure 1-1, which performs well in representing the score range for each clip.

Scatter chart is just a grand view of the matching score across all clips

For Data No.2 and No.3, we provided several sample tags for subjects to choose to describe the scene and BGM separately. Some of them are “*Exciting, Nervous, Delightful and Scary*”. We recorded how many times each tag appeared in each clip’s feedback and represented the frequency in a Pareto Chart. In this way, we could easily see the most related tag to each clip’s scene and BGM. Since the most related tag is able to represent most people’s first impression about the scene or the BGM, we could use it to classify the BGM and scene by aesthetic and artistic style and analyze whether it affects the consistency of the clip.

For Data No.4 and No.5, this will mainly reflect how subjects regard the scene and BGM as a whole part. Because different kinds of games use different element in BGM for a particular effect. The effect could be setting up atmosphere, aesthetic style or temporary events. For example, Scientific Games often use techno and electronic music to shows up a “futuristic” atmosphere for the whole story background, while some Ancient Time Games use pipe organ in the BGM to fill the game with a Middle-age feeling. That some Role Playing Games use high-speed pace music to notice the player of a drastic battle happening at the time could be an example for the “temporary events setting up” effect.

For Data No.6 and No.7, we mainly focus on how much gaming experience each subject has in their daily lives. This could critically affect how they review the clip: an

experienced gamer will pay attention to the degree of how deep is the interaction between the game and the BGM by examining whether the BGM is changing with the player's actions and progress in the game. However, an casual gamer may give feedbacks only based on whether they feel comfortable to watch a clip with the provided music in the background.

In a word, Data No.1, No.2 and No.3 are the main factors we decided whether subjects found a clip's scene consistent or not. Other data provides additional information like subject's individual gaming experience and focus point on the clip, from which we could find them helpful to decide whether the feedbacks are impersonal.

Although Data No.4, No.5, No.6 and No.7 are supportive data, they provide a side view of our experiment: how does different person react to the same clip with same scene and same BGM. This is a tricky part because subjective view could be totally different for each person. For example, in clip 10, we put a middle-age style BGM with a horror game scene to make an experimental clip which suddenly becomes the most controversial clip even among our group. One thinks the clip is the most un-matching one among all clips while the other one believes it actually looks legible. Even both members in our group, who are all experienced gamers among all kinds of game, could have division on just one clip. We do believe it is necessary to analyze how each person's familiarity with gaming affects their scores. However, this is an extremely tricky part since it involves highly subjective factors like personal favor that it is impossible to come out any clear result from those feedbacks. Therefore, we decided to describe this individual's factor only in a shallow degree and focus on average performance (score) of each clip.

Besides of mentioned statistics we analyzed, we found it necessary to classify clip's scene and BGM by aesthetic style and game genre. This could be explained by the fact that similar game genre will have similar interaction between the game and BGM. We will also talk about this in a shallow degree since this will includes huge amount of time and resources to cover it completely.

#### **4.4. Sample Base**

Throughout our experiment, we passed out 26 requests for experiment participants and got 21 reply for confirmation of becoming a subjects and participating in this experiment. Among all 21 subjects participated in this experiment, we got 19 subjects' questionnaire feedback in time, one late-submitted questionnaire feedback. Totally we got 20 eligible questionnaire feedback from this experiment. However, we received the 20<sup>th</sup> feedback even after we have finished the analysis and the statistics. Therefore, we left the 20<sup>th</sup> feedback un-used. Because of our experiment is focusing on finding out the under what kind of circumstance that subjects would find the BGM and the scene in a gaming clip seem inconsistency, we could ensure that 19 feedbacks each with 13 questionnaire would be large enough for our experiment eligibility.



## 5. Analysis:

The experiment analysis will be broken into four parts.

The first part will be an over-look at the statistics across the experiment. We will analyze the score range, style tag used frequency and subjects' personal characteristics to show a full view of the entire data set. This will support us in understanding the experiment's background and statistics distribution.

The second part "Individual Clip Analysis" will also have two sub-parts under every particular clip. The first sub-part will focus on each clip's score performance and distribution, considering about the game genre, BGM genre, game mechanics, interaction etc. This part helps us to conclude a trend of how people score each clip based on the scene and BGM. The second sub-part will focus on the style tag given to each clip's scene and BGM. Since these tags are the subjects' direct first impression (feeling) about the clip, this will indicate how different game genre and style will influence people's judgement of consistency between the BGM and the scene. These two sub-parts will be combined together to give us a more comprehensive understanding of the factors that influence the matching scale of each clip.

The third part "Over-analysis" will focus on some interesting points found in the experiment. For example, the subjective style tags given to a particular clip or some completely opposite style tags given to the same clip (which might indicate a bi-distinguished feature in the BGM or the scene). We will also try to eliminate the factors that might heavily influence the statistics' balance. Those factors will be mentioned in the following section "Potential Problems".

The fourth part “Retrospection and Improvement” will analyze the advantages and disadvantages of this experiment. We will write about retrospection about the experiment and our opinion of how to improve this experiment.

## 5.1. Over View

In our analysis, we first collect the subjects' matching scale score (Data No.1) of all clips and put all of them into Figure 1-1. We decided to use scatter chart to represent this data which is suitable for showing up score range and score appearance.

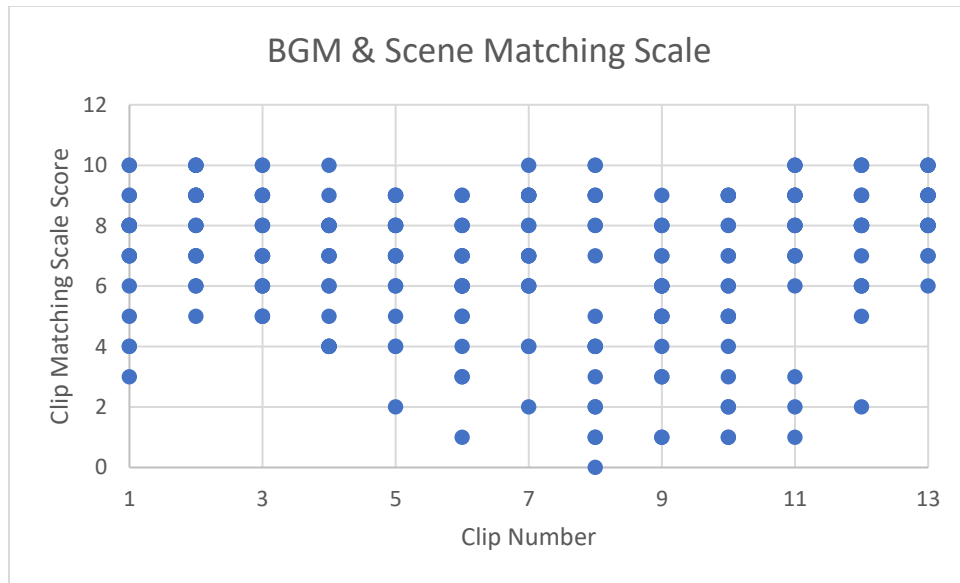


Figure 1-1

In our experiment, Data No.1 is a grand matching scale score given by the subjects to each clip, which ranges from 0 to 10 (10 means perfectly matches, while 0 means not matches at all). This score factors works as the most illustrative criteria of the degree of matching between BGM and scene in each clip. The score range and distribution of Data No.1 will shows up whether the matching scale is public-agreed or controversially polarized.

In Figure 1-1, we have all of our 13 clips' score displayed in the chart. Among all clips, it is not hard for us to find that some clips have a highly centralized score distribution: Clip No.2, No.3, No.4, No.13. Some other clips have a spread score distribution across the whole range from 0 to 10: Clip No.6, No.8, No.9, No.11.

However, the scatter chart is only able to displays the score range but not the density of each score. Therefore, we will analyze each clip one by one separately in their own analysis section.

On the other hand, we calculated and got statistics of the Mean, Mode and Median grand matching scale score of every clip as below: All Score Table *Figure 1-2*, Mean Score *Figure 1-3*, Mode Score *Figure 1-4*, Median Score *Figure 1-5*.

<b>Clip No.</b>	<b>Mean</b>	<b>Mode</b>	<b>Median</b>
1	7.21	8.00	8.00
2	8.21	9.00	9.00
3	7.42	7.00	7.00
4	6.35	8.00	7.00
5	6.89	8.00	7.00
6	6.00	6.00	6.00
7	7.16	9.00	7.00
8	5.26	4.00	4.00
9	5.42	5.00	6.00
10	5.58	9.00	6.00
11	7.37	9.00	8.00
12	7.74	8.00	8.00
13	8.47	9.00	9.00

*Figure 1-2*

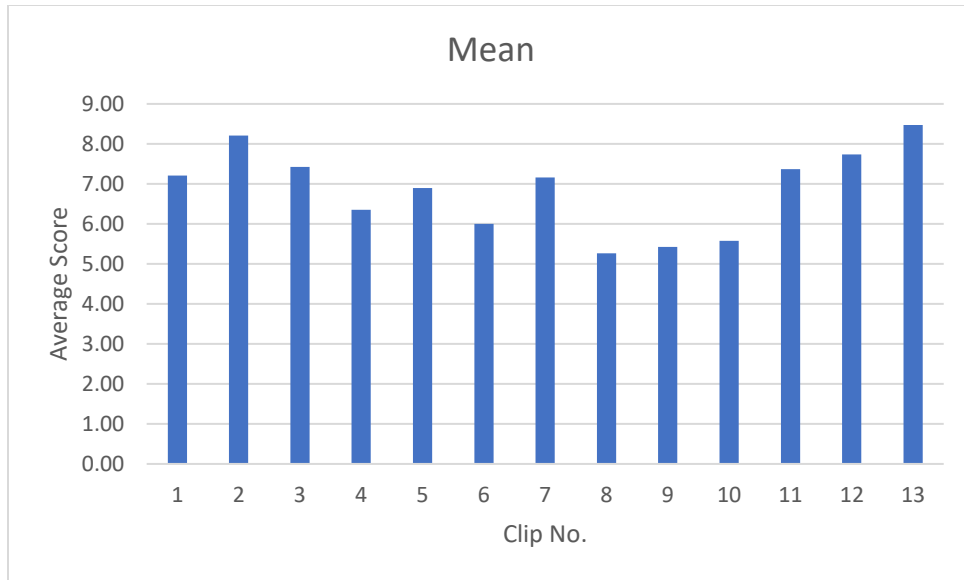


Figure 1-3

Based on the Mean Score Chart of the matching scale score of each clip, we could compare it with the *Figure 1-1* by Clip No. It is not hard to find out some of the most centralized score distribution clip (No.2, No.13) have also the highest average score (8.21, 8.47). This could be explained by the fact they have the most public-agreed degree of matching scale score between BGM and the scene. Other clips with the lowest score of matching scale score (No.8, No.9) are also the clips' with the most spread matching score distribution. This is possibly because they are the least public-agreed degree of matching between the scene and BGM. However, other clips' average matching score is varying randomly which does not correspond with the matching score distribution that we need to talk each clip in details about it.

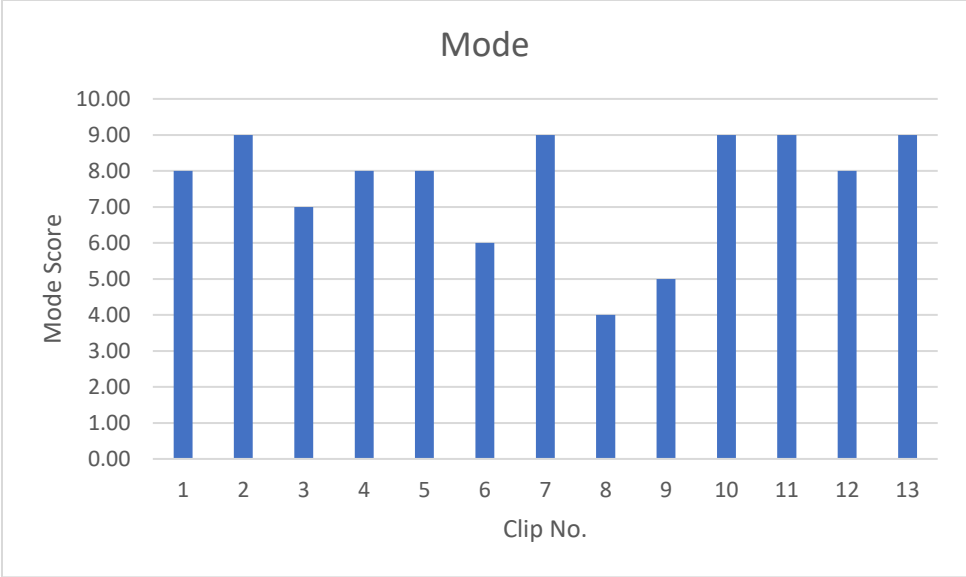


Figure 1-4

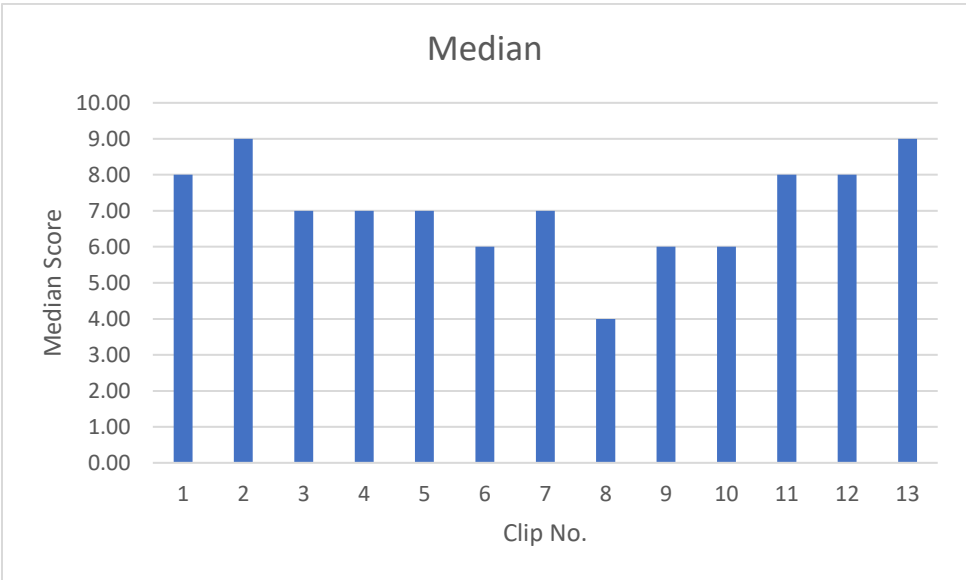


Figure 1-5

Figure 1-4 and Figure 1-5 are graph of each clip’s mode and median matching scale score. We could easily find that the clips with the highest-agreed matching scale score (No.2, No.13) have the highest Mode Score and Median Score; clips with the lowest-agreed matching scale score (No.8, No.9) also have the lowest Mode and Median Score. Most of scores among

Mode and Median score following the trends of score ranking from the Mean Score and the Score distribution, in which we could roughly ensure that the dataset we have does not heavily influenced by extremely high or low value. This prerequisite enables us to continue our following individual clip analysis.



*Experiment Video Start Screen Figure 1-6*

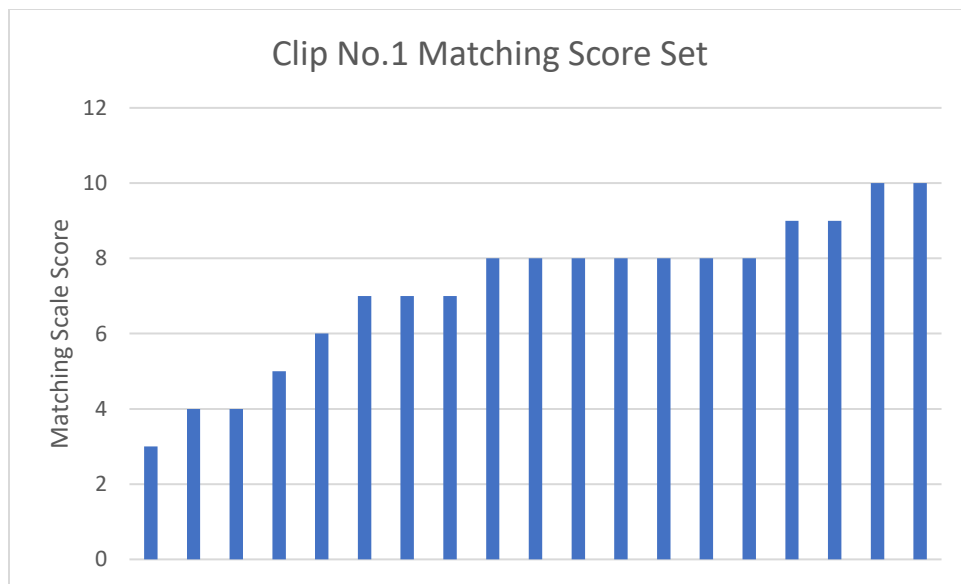
## 5.2. Individual Analysis

**Clip No.1:**

**Video Clip: Bloodborne**

**BGM: Nioh – Disc 1**

**Scene/Stage: Start Screen**



*Figure 2-1-1*

In *Figure 2-1-1*, we have every matching score in an ascending order that given to Clip No.1 as an indication of the grand matching scale between the BGM and the scene. This is score display is not illustrative enough so we compile all score into Score Analysis *Figure 2-1-2* and Score Distribution *Figure 2-1-3*.

<b>Clip No.</b>	<b>Mean</b>	<b>Mode</b>	<b>Median</b>
1	7.21	8.00	8.00

*Figure 2-1-2*



Clip No.1 has the average matching score as 7.21, Mode Score as 8 and Median as 8, which could indicate that Clip No.1 has its score distributed normally. However, it's Mean Score still has a little 0.79 point lower than the other two attribute which might be influenced because of the lowest score (3, 4, 4) in its dataset.

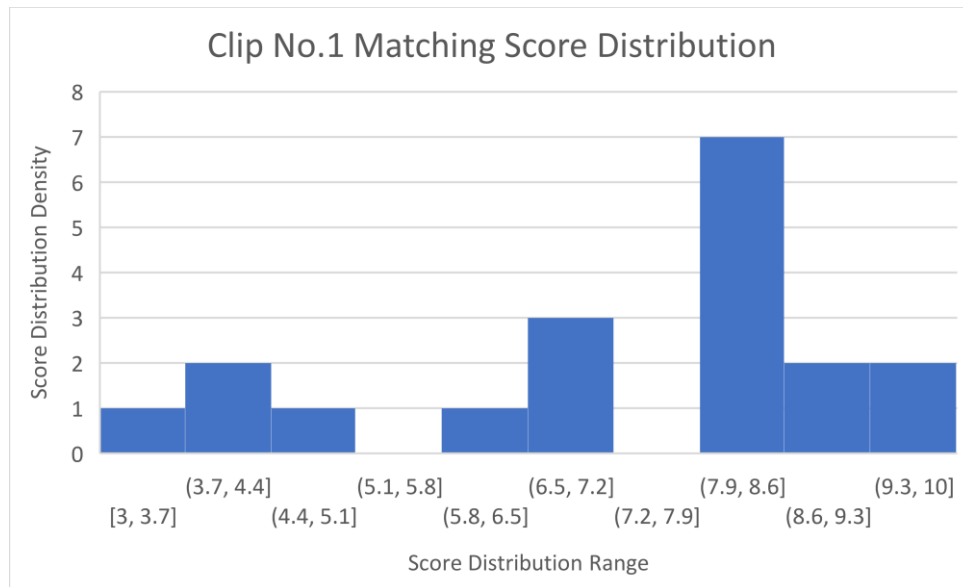


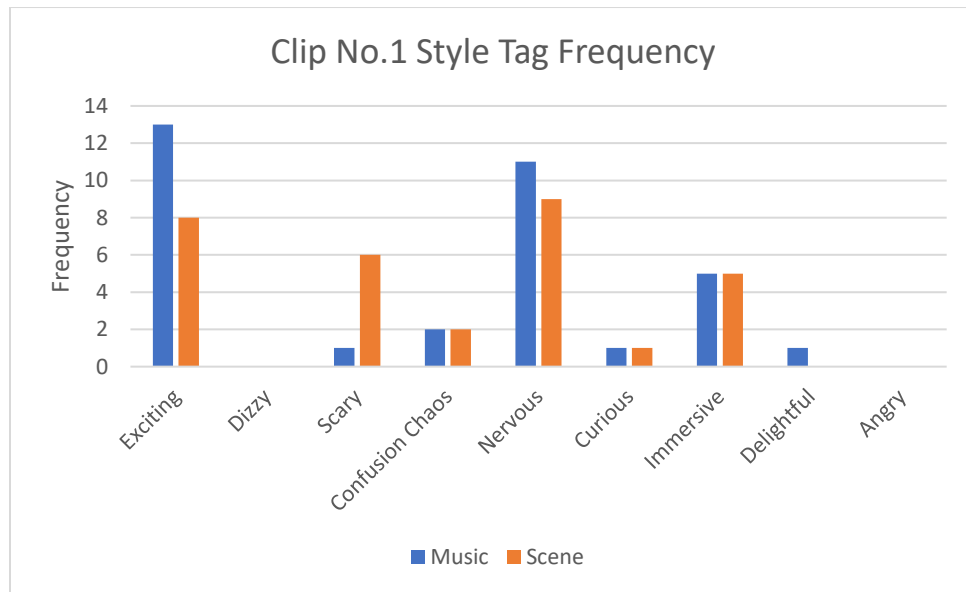
Figure 2-1-3

According to *Figure 2-1-3*, we are also able to see that Clip 1 has a centralized score distribution in the range from 7.9 to 8.6 that 7 out of 19 scores are given in this range. Other scores are divided into three parts. The first part is the score distribution in score range from 8.6 to 10, which holds matching score higher than the 3-M (Mean, Mode and Median) Score. The second part has score ranges from 5.8 to 7.2, in which scores stays in the medium of the whole score range. Score given in this range might indicate that subjects' neutrality attitude toward the matching scale. The third part is lowest score ranges from 3 to 5.1. This part is also responsible for the lower difference between the Mean Score and the other two Score attributes.

Generally speaking, score distribution of Clip No.1 is approximately gathered around score 8, which represent a high agreed matching scale score. Some other subjects may find the

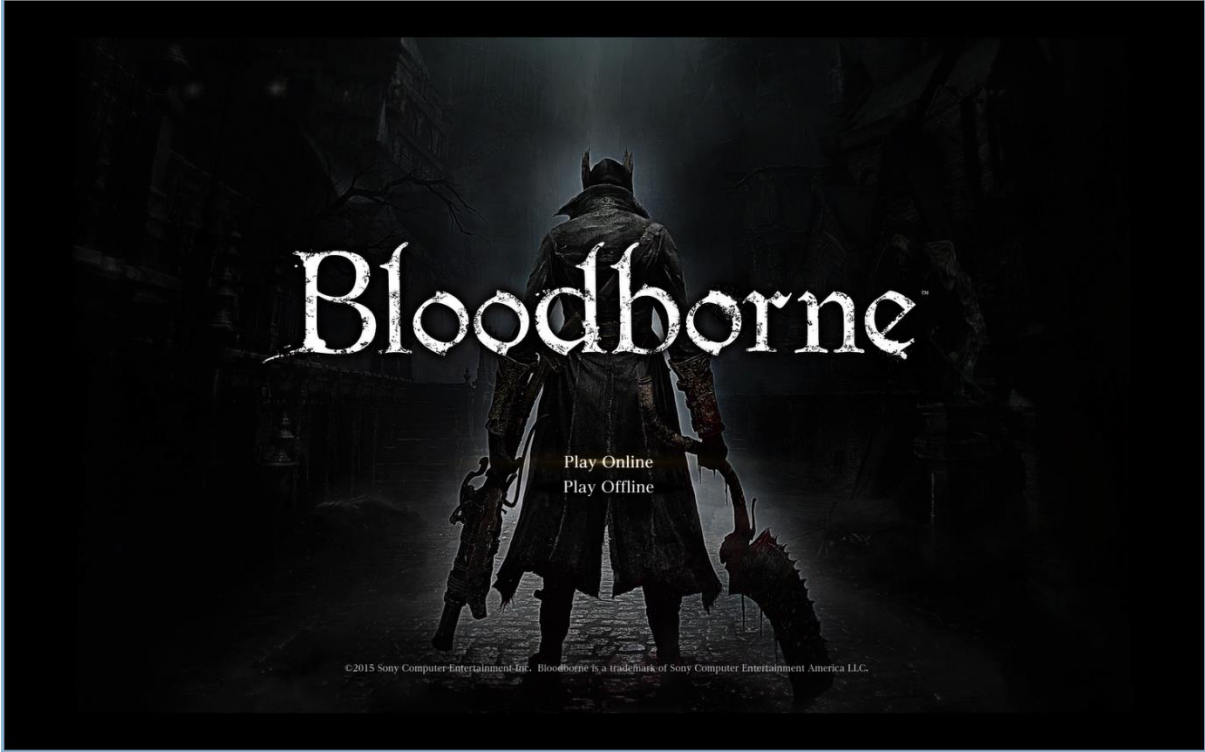
BGM and scene extremely un-matching, while most subjects are believing the consistency in Clip No.1

In *Figure 2-1-4* we have the statistics of style tags about both the scene and the BGM.



*Figure 2-1-4*

The high matching score of Clip No.1 could be seen from the Style Tag Frequency. The two most appeared tags for Clip No.1's BGM are "Exciting" and "Nervous", while the two most appeared tags for Clip No.1's scene are "Exciting" and "Nervous" too. Obviously, the BGM got more Exciting and Nervous than the scene while the scene got Scarier than the music. Clip No.1 also has the "Immersive" tag which have two fields (the BGM and the Scene) with same frequency. However, the difference between the Scary tag could not even compensate the overlapping "Exciting" and "Nervous" tag. Generally, Clip No.1 got a 7.21 Mean Matching Score because the BGM and the Scene both are considered "Exciting" and "Nervous" at the same time.



*Clip 1 Screen shot Figure 2-1-5*

**Clip No.2:**

**Video Clip: Bloodborne**

**BGM: Dark Souls 3 – Firelink Shrine**

**Scene/Stage: Base (Safe Zone)**

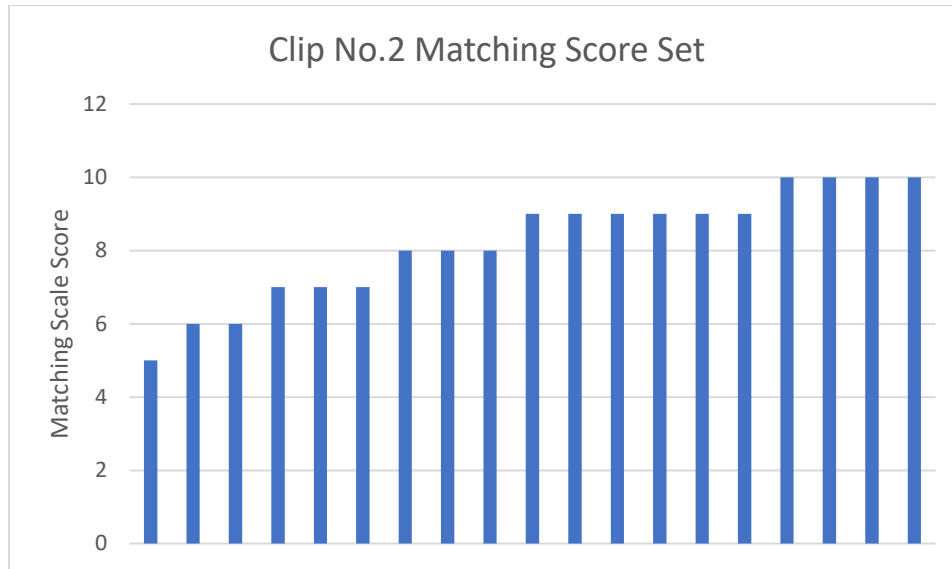


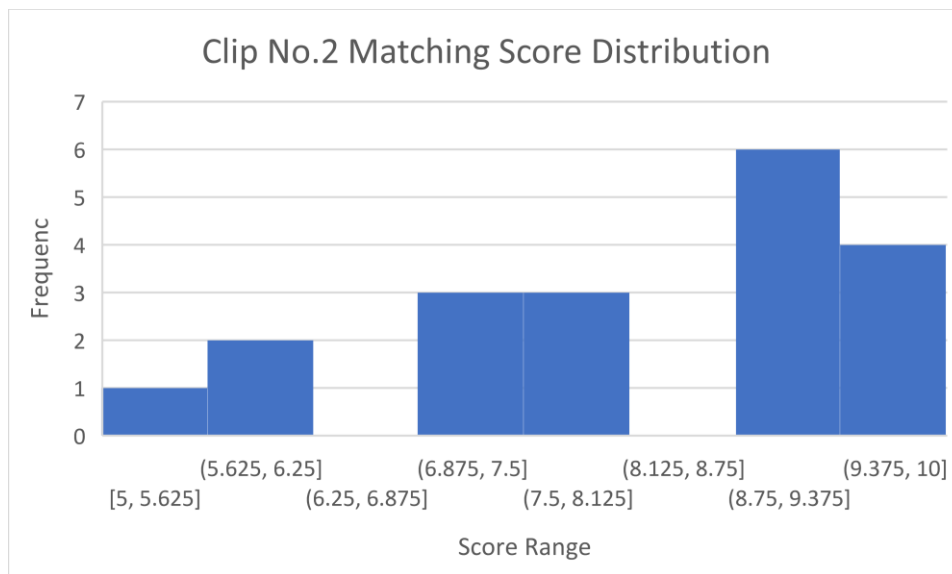
Figure 2-2-1

As shown in *Figure 2-2-1*, we have the data set of matching scale score of Clip No.2. Compared it with the data set of Clip No.1 (*in Figure 2-1-1*), we could see that Clip No.2 has a much higher score on the low score part (which is on the left most part of the chart). This could be seen by the fact that Clip No.2 is one of the clips with the highest matching scale between the BGM and the scene.

<b>Clip No.</b>	<b>Mean</b>	<b>Mode</b>	<b>Median</b>
2	8.21	9.00	9.00

Figure 2-2-2

Based on *Figure 2-2-2*, it is not hard to find that Clip No.2 has an extremely high Mean Score. The Mode Score and Median Score are also high enough that could ensure the data is not influenced by extremely high or low value. The 0.79 difference between Mean Score and Mode, Median Score could be seen from the lower score on the left part in *Figure 2-2-1*. Generally speaking, Clip No.2 still has a high score set in which there are 10 subjects given a 9 or 10 matching score to this clip.



*Figure 2-2-3*

In *Figure 2-2-3*, we have the score distribution of the whole Clip No.2's dataset. As a whole part, we could see that scores given to Clip No.2 are centralized on the right-most part of the chart which represent the highest score in the range. The second largest score centralization locates at the range from 6.5 to 8.1, which is a range a little higher than the mid-point of the whole range, which is score 5. Comparing with Clip No.1, we could see that the lowest score given to Clip No.2 is 5, which is still higher than the lowest score given to Clip No.1 (which is 3). Generally, Clip No.2 has a much higher matching scale score than Clip No.1 in every dimension, which could be seen from the fact that Clip No.2 is one of the clips with the highest matching scale score among all clips.

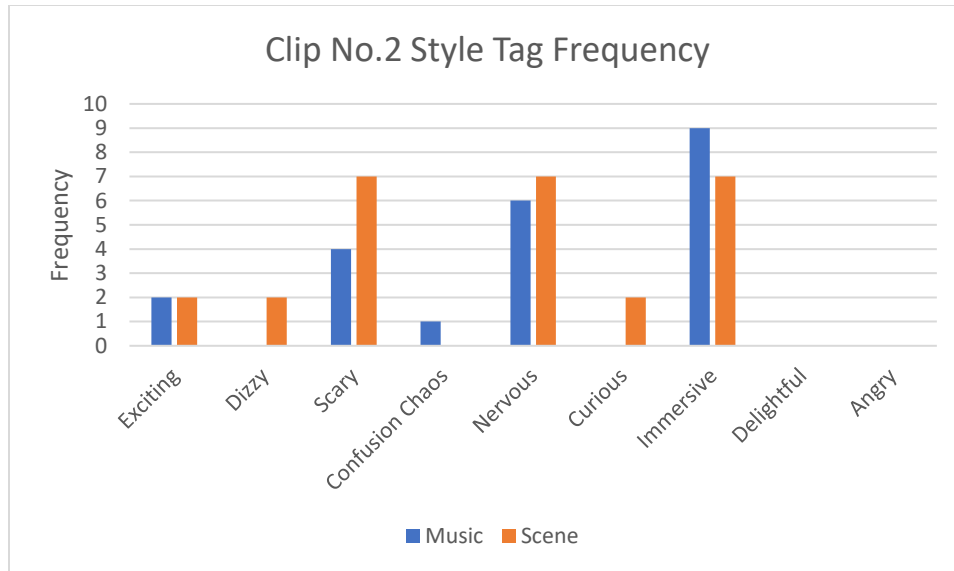


Figure 2-2-4

Based on the *Figure 2-2-4*, conclusion about the style synchronization between BGM and the scene could be made. Clip No.2's scene and BGM are simultaneously getting high frequency in style tag "Nervous" and "Immersive". They also get a large number of "Scary" tag which put them into a perfectly matching states.

Clip No.2 also have the most extra-given style tag among all of the clips. From the data we collected from subjects, Clip No.2's music is tagged with "Sorrowful", "Weird" and "Calm" each for once. It's scene is tagged with "Plain" once. Our highly subjective hypothesis says this might happen because of the consistency and synchronization that leaves places for subjects' imagination. As we see from *Figure 2-2-4*, Clip No.2 is described as "Immersive" the most which is the atmosphere that allows participants to have the maximum possibility for their imagination and get a more comprehensive feeling and understanding.



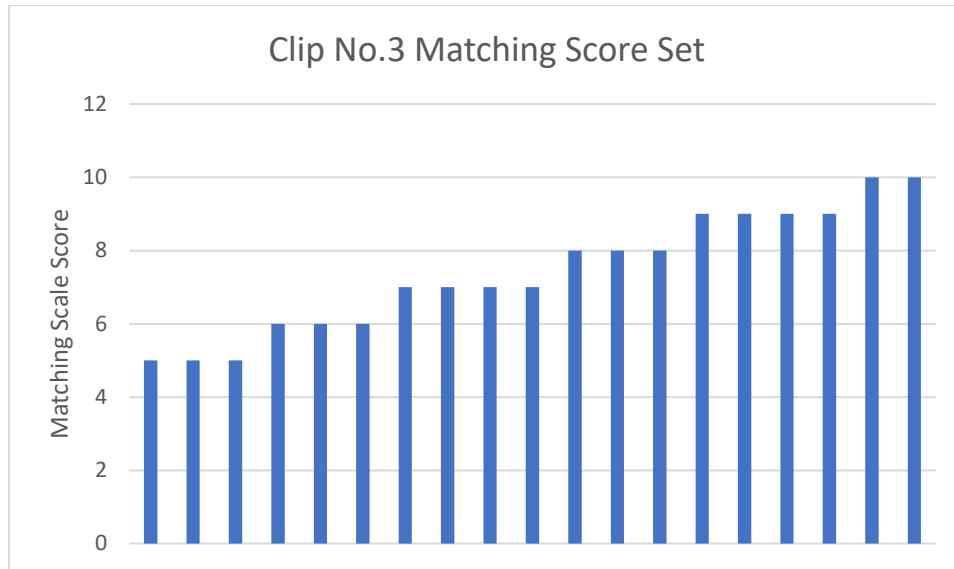
Clip 3 Screen Shot Figure 2-2-5

**Clip No.3:**

**Video Clip: Bloodborne**

**BGM: Dark Souls 3 – Ancient Wyvern**

**Scene/Stage: Boss Fight**



*Figure 2-3-1*

Figure 2-3-1 represent the whole matching score set of the Clip No.3, from which we could see a different trend than that of Clip No.1 and Clip No.2. The lowest score given to No.3 is 5 while the highest score given to it is score 10. This fact indicates that Clip No.3 is not as controversial as Clip No.1 and has a less divided opinion across all subjects participated in the experiment. Comparing it with the chart of Clip No.2, we also found out that Clip No.3 does not have much high score (like 9, 10) as Clip No.2 has. Totally, Clip No.3 is a clip that do not have extremely divided score distribution but not have a highly centralized score point in the range.

Clip No.	Mean	Mode	Median



3	7.42	7.00	7.00
---	------	------	------

Figure 2-3-2

In Figure 2-3-2 we could see the 3-M score of Clip No.3. Obviously, Clip No.3 has its score distributed around score 7 which is a moderately high score in the range 0 to 10. The Mode Score and Median Score are sharing the same score 7 which means Clip No.3's Mean Score is not influenced by extremely high-low value a lot.

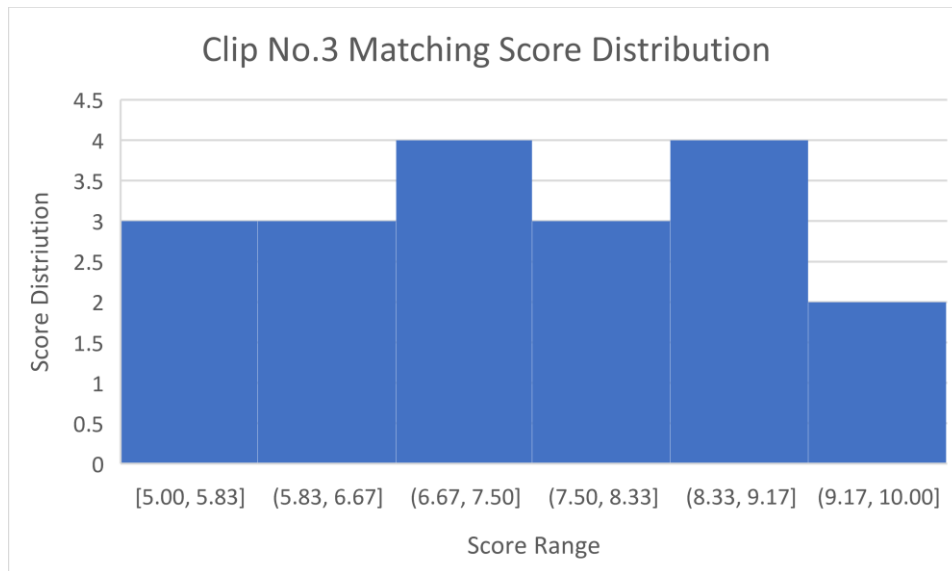


Figure 2-3-3

In the Matching Score Distribution *Figure 2-3-3*, the matching score distribution of Clip No.3 has a rarely even spread on the range from score 5 to 10. The only difference in score frequency we have is less than 2 across all score.

This even distribution trend might indicate the truth that Clip No.3 has a mostly-agreed approximately matching scale which is caused by some un-matching details that makes the clip maintains a moderately score around 7. Being different from Clip No.1, which has a highly divided opinion on the matching scale, and different from Clip No.2, which has a highly concentrated high score distribution, Clip No.3 is a clip with a moderately high score rating, but

varies in a small range. One of the possible explanation is that the main themes of the scene and the BGM maintain the same style and paces, while some details in the BGM (tempo, beats) or in the scene (color choices, characters movement) reflects an inconsistency that lead to score varying.

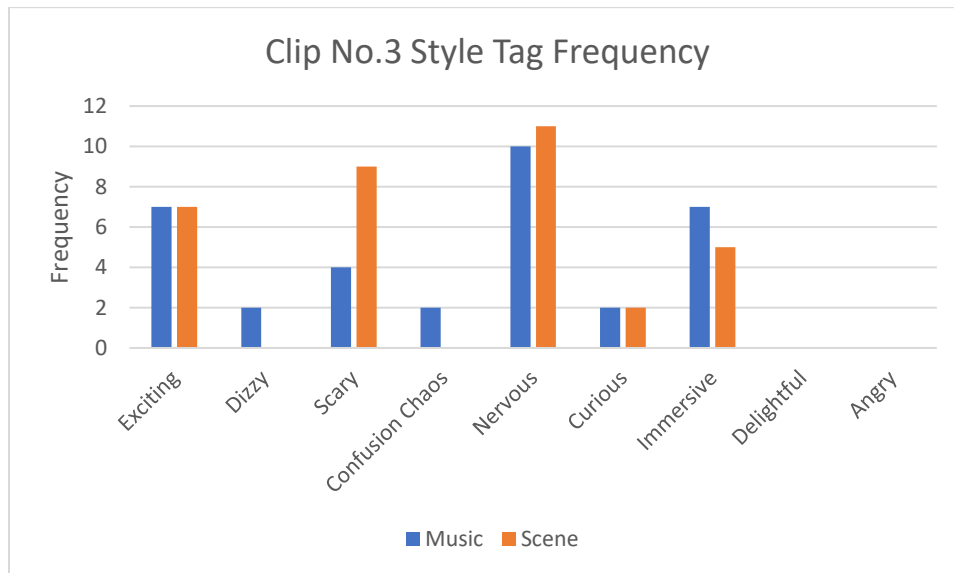


Figure 2-3-4

The Style Tag Frequency Chart (*Figure 2-3-4*) might be able to explain the problem we mentioned above. Clip No.3's scene and the BGM have a highly overlapping area under the style tag "Nervous" which individually have 11 and 10 appearances. The second overlapping tag is the "Exciting" one of which both the scene and BGM have 7 appearances.

However, there are also some highly differenced description between the scene and BGM. First of all, Clip No.3's scene holds 9 "Scary" tag, which are more than twice the number of "Scary" tag the BGM has. Besides, the difference in tag "Immersive" and the appearance of tag "Dizzy" and "Chaos", which appears twice but not only once, might give us a insight that the Music in Clip No.3 performs a more duplicated feeling than the scene in No.3. In other words, the BGM provides more details than the scene and leads to score varying in the matching scale.



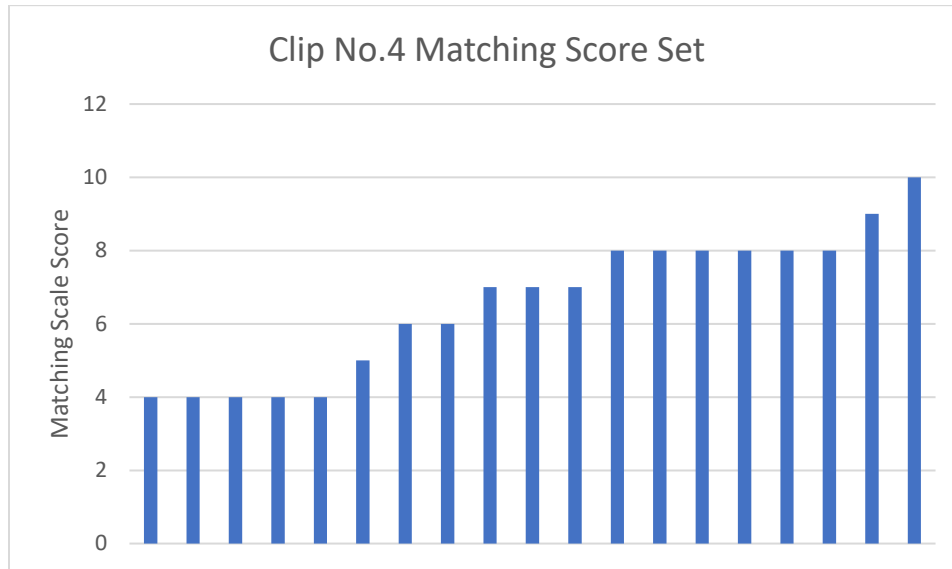
Clip 3 Screen Shot Figure 2-3-5

**Clip No.4:**

**Video Clip: Nioh**

**BGM: Dead Space – Cyanide Systems Offline**

**Scene/Stage: Start Screen**



*Figure 2-4-1*

In *Figure 2-4-1*, we have the whole data set of the matching score set of Clip No.4. It's obviously that Clip No.4 has a lower average matching score across the data set. The highest score it gets is score 10 with only one appearance which indicates it as an extreme edge value. The lowest score it has is score 4 which has 5 times appearance.

<b>Clip No.</b>	<b>Mean</b>	<b>Mode</b>	<b>Median</b>
4	6.35	8.00	7.00

*Figure 2-4-2*

Clip No.4 is the first clip we met that has a relatively large fluctuation between the Mean, Mode and Median score. We could see that Clip No.4 has 1.65 lower difference between the mean score and mode score, which might indicate that the score distribution of Clip No.4 is much more spread out. It also has 0.65 difference between Median Score and Mean Score which could be seen from *Figure 2-4-1* that the Mean Score is dragged down by the lower score part on the left-most part of the chart.

Another feature we found from *Figure 2-4-1* is that Clip No.4's score is distributed across the range in a approximately symmetric pattern: it has several low score at score 4 while having several high score at score 8. This pattern is similar to Clip No.3's pattern but with some differences. Firstly, Clip No.3's score is widely distributed while scores are evenly locating in every minor range inside the whole range. However, Clip No.4's score distribution is spread out based on one mid-point which is around score 6, and has top and bottom score limits on score 8 and score 4. This is an interesting point that might indicate the different opinion of Clip No.4's consistency might only vary in a limited range. In another word, the factors affect the consistency between the BGM and the scene in Clip No.4 are limited, while those factors in Clip No.3 are not.

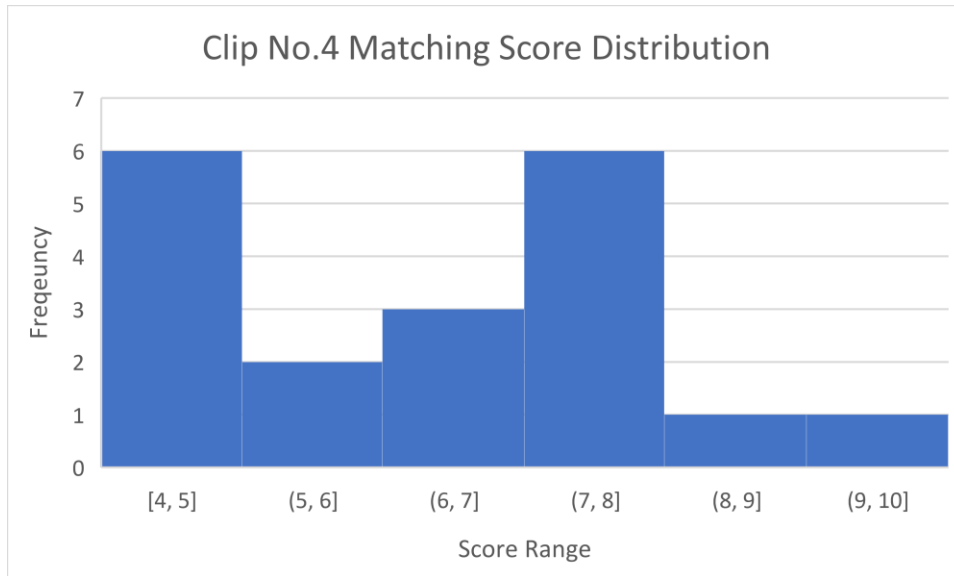


Figure 2-4-3

According to *Figure 2-4-3*, we could prove our hypothesis mentioned above. The matching score are centralized on score 4 and score 8, which behave as limits for other score distribution. There are only two scores located outside this range. If we want to describe this distribution, we will say: “the score distribution is highly divided in the limited range from score 4 to score 8.

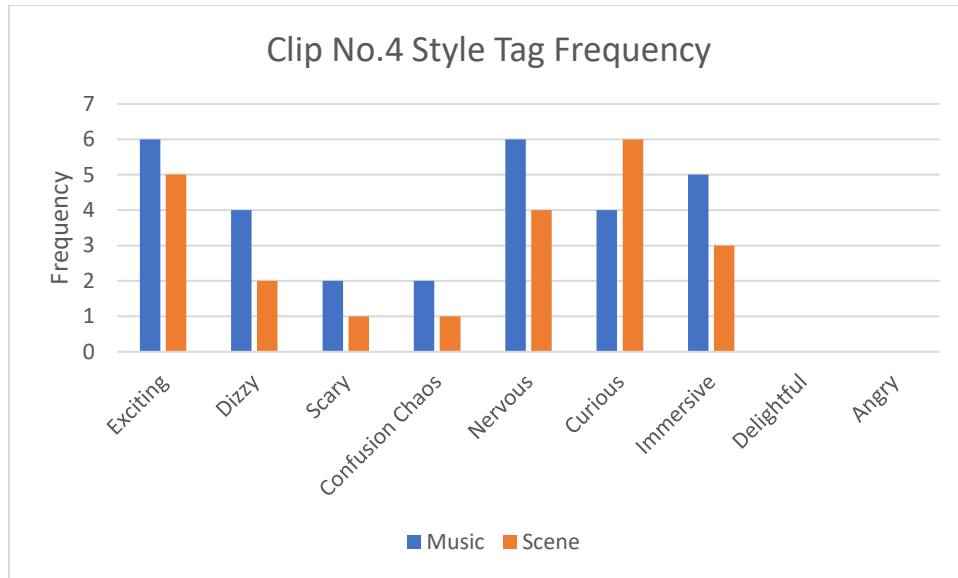
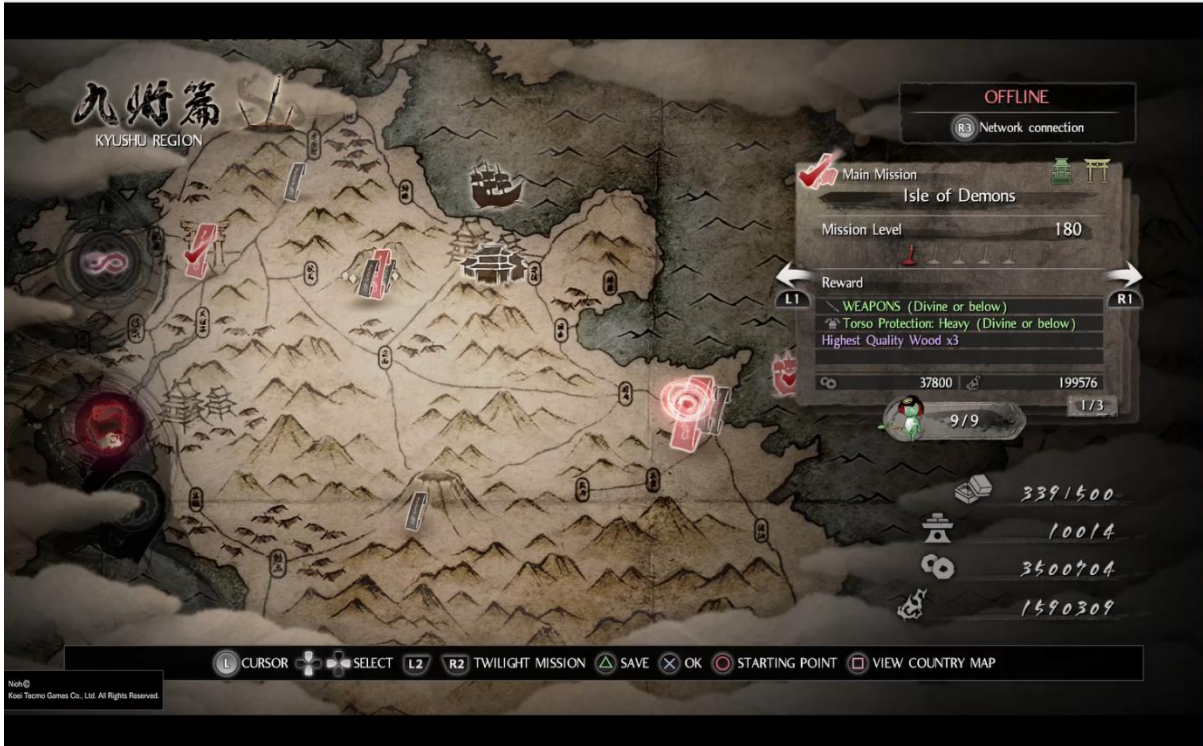


Figure 2-4-4

From *Figure 2-4-4* we could find out that Clip No.4 has a unique style tag frequency distribution. Unlike Clip No.2 or No.3 which have two or three style tags that appear much more times than all of other tags, Clip No.4 has all kinds of tags appearing varyingly. Tags like “Exciting”, “Nervous”, “Curious” and “Immersive” all appear several times in both the Music and the scene rating. This might indicate that Clip No.4 has multiple factors that subjects found the BGM and the scene are matching with each other. However, all of the tags given to Clip No.4 do not have the same appearance in the BGM and the scene at the same time, which means the BGM and the scene are not completely matching in every individual factor that might affect the matching scale between the BGM and the scene.

In a conclusion, the grand matching degree among all tags sets up limits for the matching scale score because Clip No.4’s BGM and scene are approximately matching in multiple ways. Nevertheless, the BGM and the scene are still different from each other that they vary a lot in each individual style field that causes opinion divided heavily in the limited range.



Clip 4 Screen Shot Figure 2-4-5

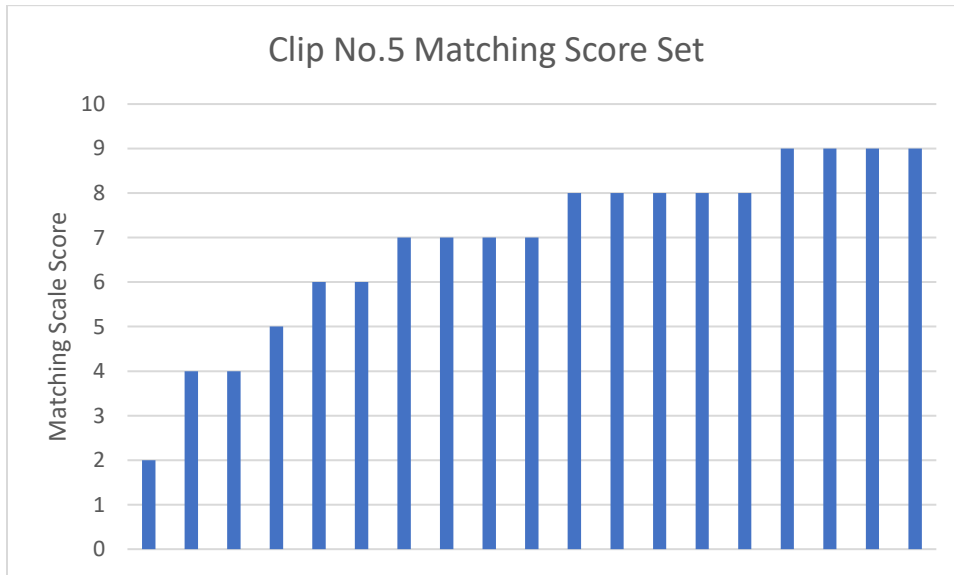


**Clip No.5:**

**Video Clip: Nioh**

**BGM: Destiny 2 – Spark**

**Scene/Stage: Outdoor Fighting Scene**



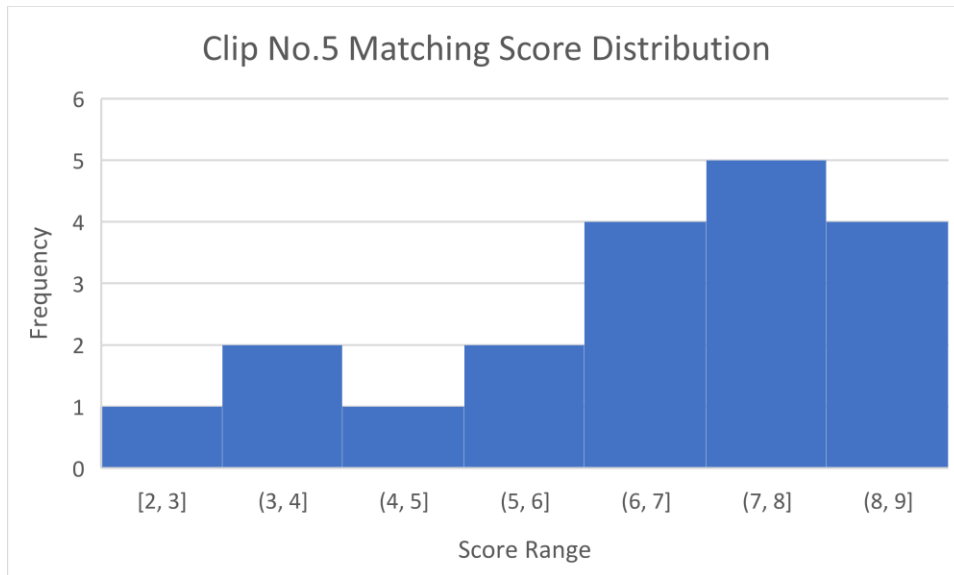
*Figure 2-5-1*

In *Figure 2-5-1* we have the whole matching score set of Clip No.5. Clip No.5 has another unique score distribution, in which there are multiple high score (score 8, score 9) appears several time while having one of the lowest score 2 at the same time. This probably means the opinion of the matching scale is heavily divided on Clip No.5.

<b>Clip No.</b>	<b>Mean</b>	<b>Mode</b>	<b>Median</b>
5	6.89	8.00	7.00

*Figure 2-5-2*

From *Figure 2-5-2* we could see that Clip No.5 has a similar score performance in Mean, Mode and Median Score as that of Clip No.4. The Mode Score is 8 which means the mostly agreed matching scale score is 8 that the median is the score indicates that there are a lots of score given to Clip No.5 is lower than 7 because there are a lot of high score in score 8 and score 9. The low Mean score 6.89 indicates that Clip No.5 has some extreme low value that influence the score performance in this place.



*Figure 2-5-3*

From *Figure 2-5-3* we could see a different fact than our hypothesis based on *Figure 2-5-1*. Though the score distribution of Clip No.5 is divided across a wild range from score 2 to score 9, most score given to Clip No.5 still centralized on range from score 6 to score 9. In other words, Clip No.5 still got a high matching scale score among most subjects. The reason why the Mean Score is very low is that the Mean Score is affected by the extreme low score located at the left half part on *Figure 2-5-3*. The opinion is spread out to wilder range while it still gets a high score among most participants.

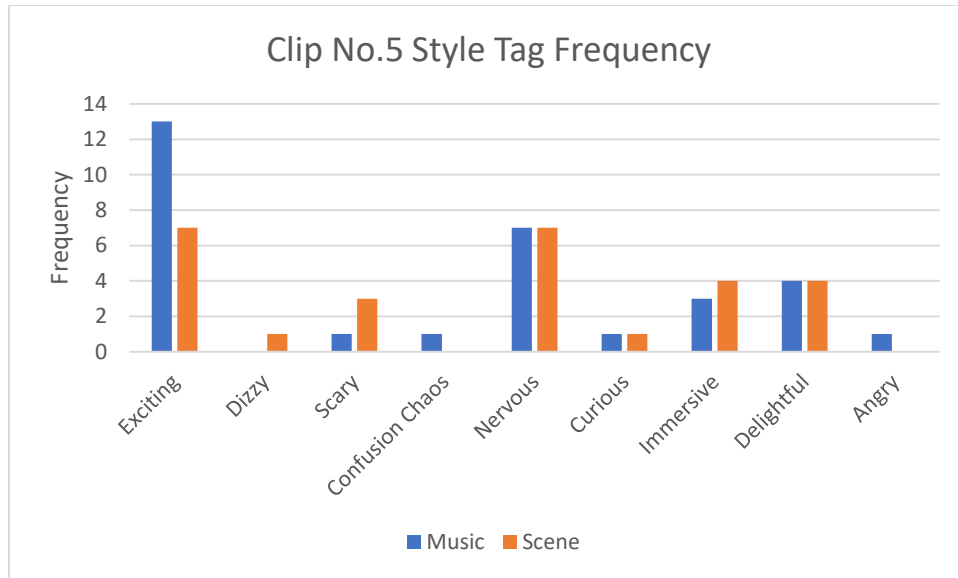


Figure 2-5-4

From *Figure 2-5-4* we might get an insight of why Clip No.5 has such score distribution. The biggest difference we sensed is that tag “exciting” appears more than twice of the BGM than that of the scene. This highly frequency difference represent that the BGM is much more exciting and drastic than the scene in some ways. For other tags, most of them matches at the same rate: “Nervous”, “Immersive” and “Delightful”. The difference could explained the reason why there are some low score rating about Clip No.5, while the matching tags are able to explain why Clip No.5 has highly matching scale score.



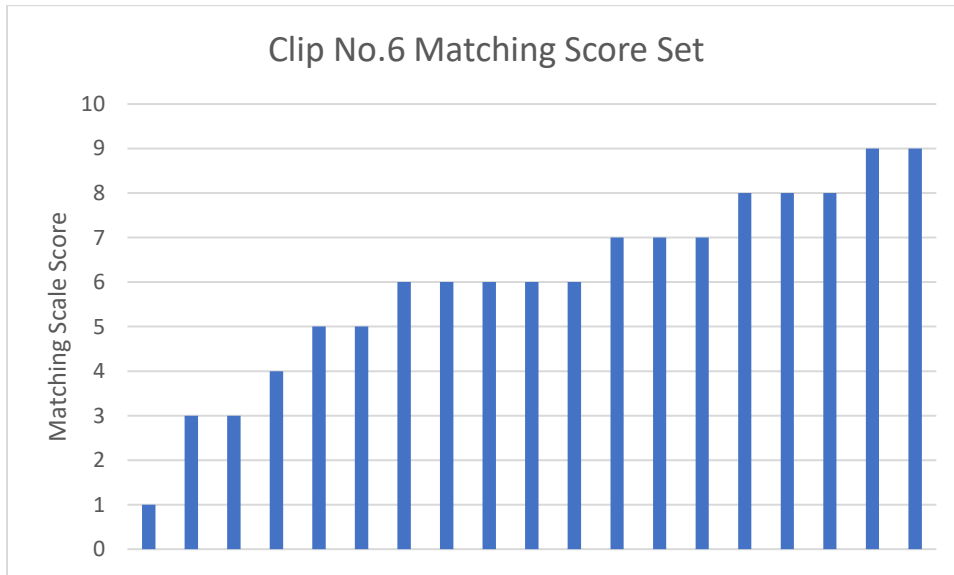
Clip 5 Screen Shot Figure 2-5-5

**Clip No.6:**

**Video Clip: Nioh**

**BGM: Tekken 7 – Heat Haze Shadow**

**Scene/Stage: Boss Fight**



*Figure 2-6-1*

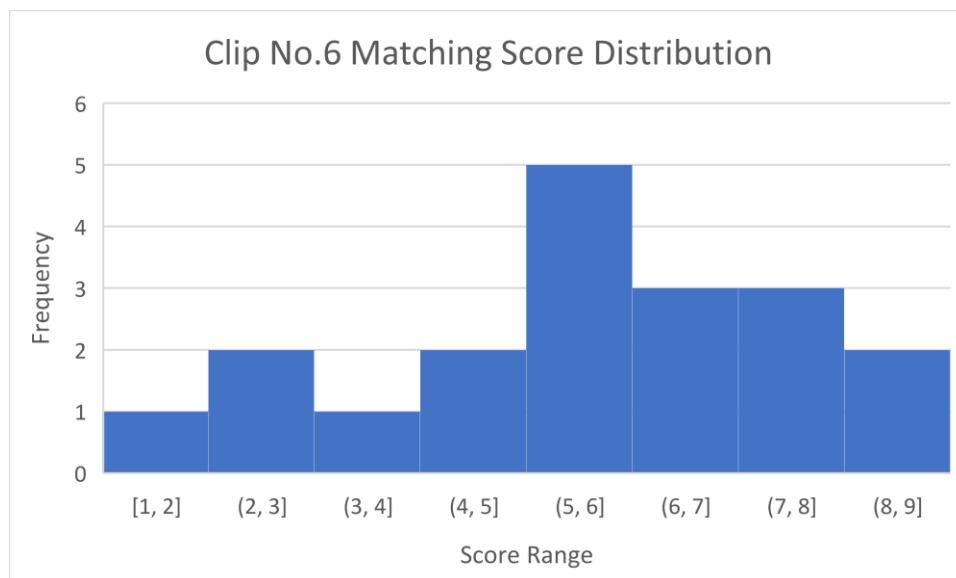
In *Figure 2-6-1*, we could see a similar score distribution trend as we found in that of Clip No.5. However, Clip No.6 does not have that much high score rating but also has a lot of medium score at the same time. This might roughly indicate that Clip No.6 does have a low matching score.

<b>Clip No.</b>	<b>Mean</b>	<b>Mode</b>	<b>Median</b>
6	6.00	6.00	6.00

*Figure 2-6-2*

In *Figure 2-6-2*, we firstly found out the lowest 3-M score we have now. Clip No.6 does not perform well and have an all score 6 performance in Mean, Mode and Median Score. The score clearly tells that fact that Clip No.6 is a clip with a low matching scale score.

However, after retrospection of *Figure 2-6-2*, we found out that the Mean Score of Clip No.6 is still affected by some extreme low score like score 1, and some extreme high score like score 6. This means that opinions are moderately divided upon Clip No.6 but most opinions are agreed that Clip No.6 does not have its BGM and scene matching well.



*Figure 2-6-3*

From *Figure 2-6-3* we could see that most Clip No.6's score are centralized at score 6 while some score are spread to score 1 and score 9.

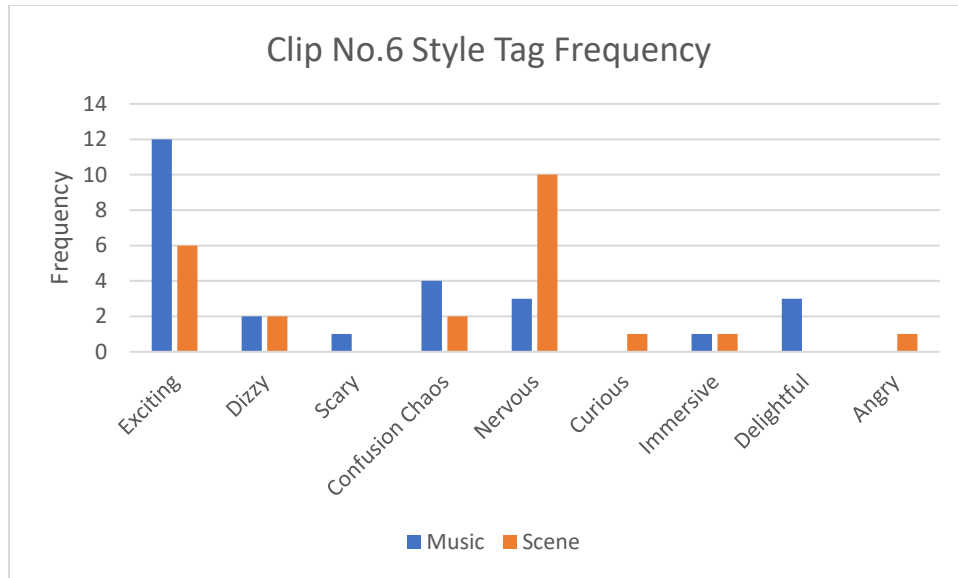


Figure 2-6-4

Considering *Figure 2-6-4* simultaneously, we found some interesting point. First of all, Clip No.6 has some really controversial difference between the style tag distribution of the BGM and that of the scene. The BGM has twice “Exciting” tag than the scene while the scene has third times “Nervous” tag than the BGM. Although in our own opinion that “Exciting” and “Nervous” are similar in some ways, they still present totally different meanings that we have to separate and treat them individually. Some other differences in Clip No.6’s style tag appearance could also be seen from different number of tags of “Scary”, “Curious”, “Delightful” and “Angry”. There is no doubt that the number of style differences in the BGM and the scene cause a low score for the Clip No.6 matching scale.



Clip 6 Screen Shot Figure 2-6-5

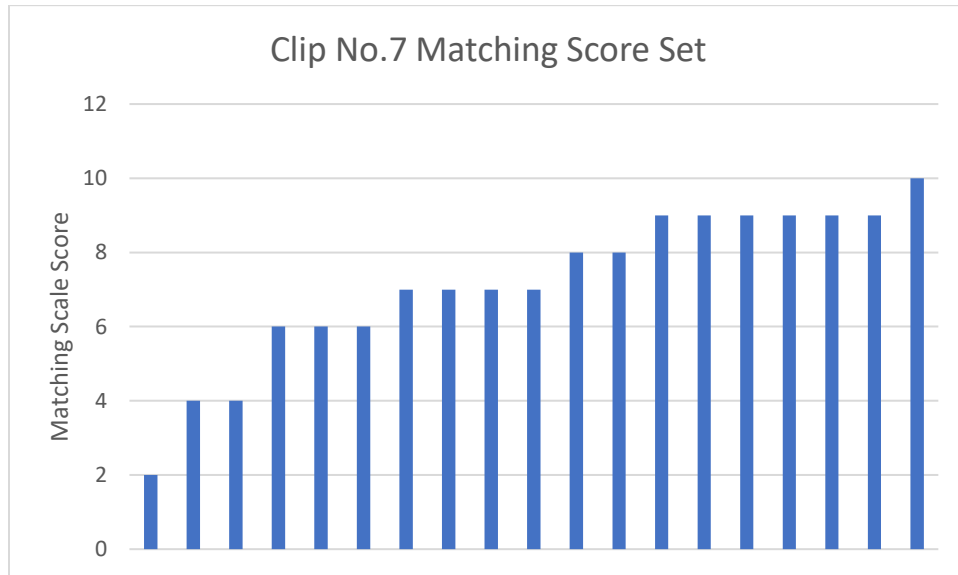


**Clip No.7:**

**Video Clip: Tekken 7**

**BGM: Dead Space – Dead Space Theme**

**Scene/Stage: Start Screen**



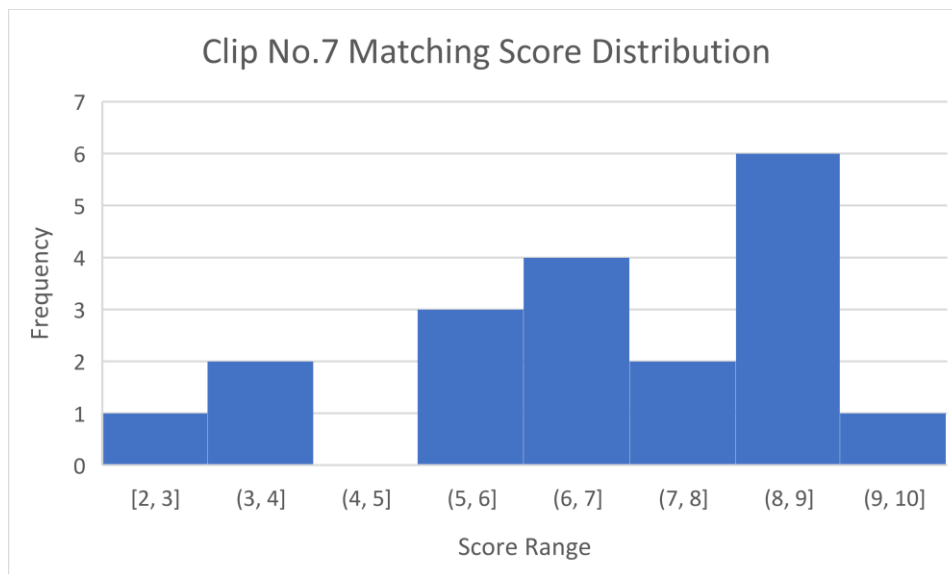
*Figure 2-7-1*

Comparing Clip No.7's Matching Score set *Figure 2-7-1* with Clip No.6's Matching Score Set *Figure 2-6-1*, we found out both graph has similar trends. They both have some high score, several number of medium score and a little number of extremely low score. However, Clip No.7 still holds more high score than high score around score 9 than Clip No.6 which represent a higher agreed matching scale.

<b>Clip No.</b>	<b>Mean</b>	<b>Mode</b>	<b>Median</b>
7	7.16	9.00	7.00

*Figure 2-7-2*

Checking *Figure 2-7-2* proves our hypothesis. We could see that Clip No.7 has a moderately high Mean, Mode and Median Score while its Mode Score reaches even score 9. The Mean Score has 1.84 lower than the Mode Score and 0.16 higher than the Median score, which is caused by a large number of high score like score 9 and some extremely low score like score 2. In other words, Clip No.7 has its score spread in wild range but have more centralized score on high level score than that of Clip No.6



*Figure 2-7-3*

From *Figure 2-7-3*, the trends of highly centralized score on score 9 is obvious. The frequency of low matching scale score is decreasing with the decrease of score itself. This phenomenon is normal since there should be less people agree with extreme score rating. Therefore, Clip No.7 still has a highly rated matching scale score. However, it is still noticeable that Clip No.7 has a large number of moderate score that does not highly agree with Clip No.7's matching scale. This scores are distributed in range from score 2 to score 7 which makes them seemingly less than the high score in score 9.

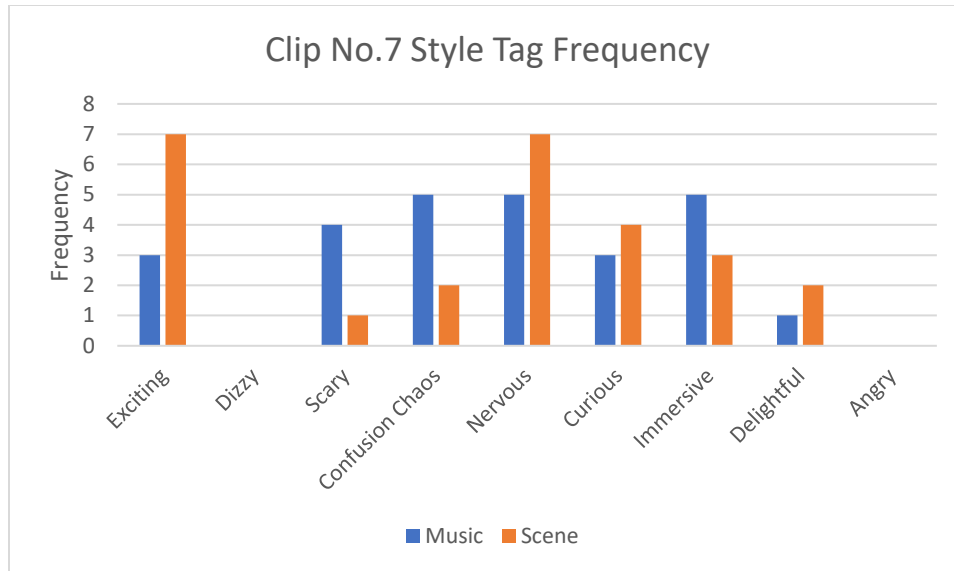
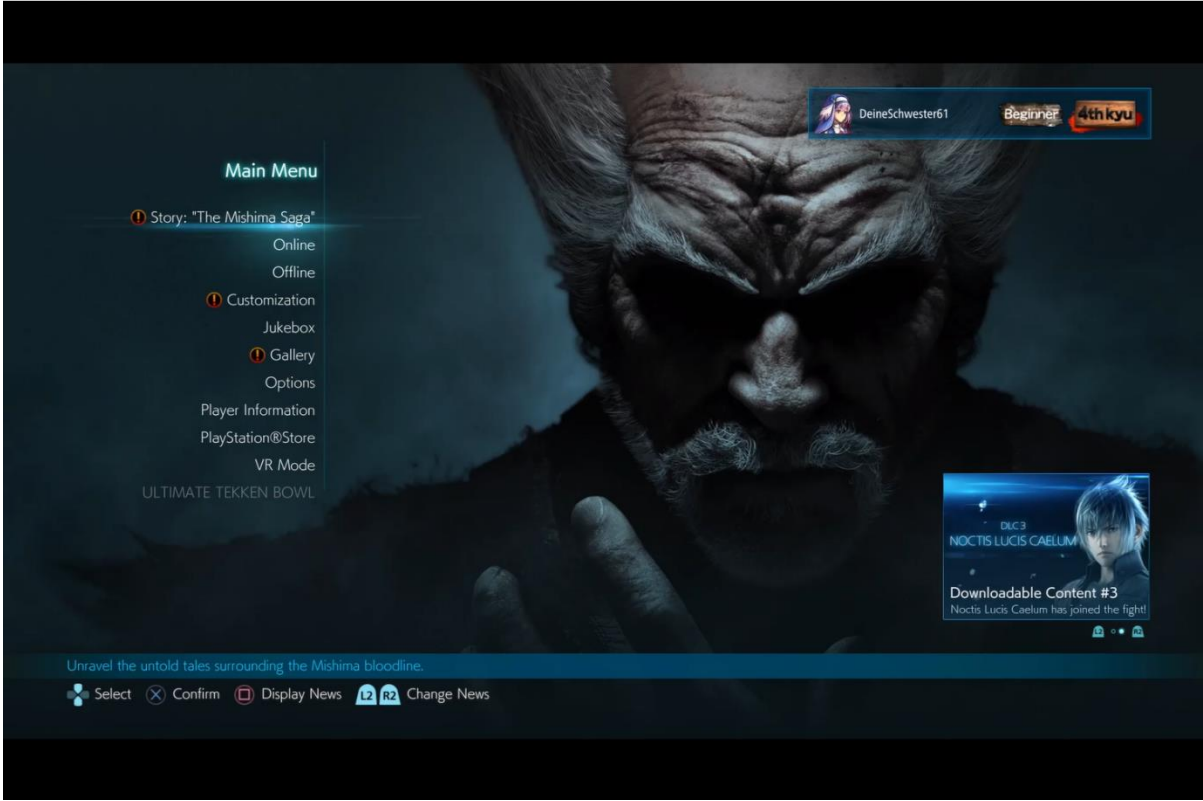


Figure 2-7-4

Comparing *Figure 2-7-3* and *Figure 2-7-4* enable us to find the interesting point of this clip. Clip No.7 have too much high appearance frequency tag like “Exciting”, “Scary”, “Chaos”, “Nervous”, “Curious” and “Immersive”. However, none of them have their number of appearance the same. Some of them even have a large different number between their individual appearance in the BGM and the scene. In other words, Clip No.7’s BGM and the scene both provide a multi-factors feeling to subjects enable Clip No.7 to get a high score from some subjects because its multi-layer feeling from factors mentioned above. It also gets low matching scale score from different degree of each factor that makes Clip No.7 an opinion divided Clip.



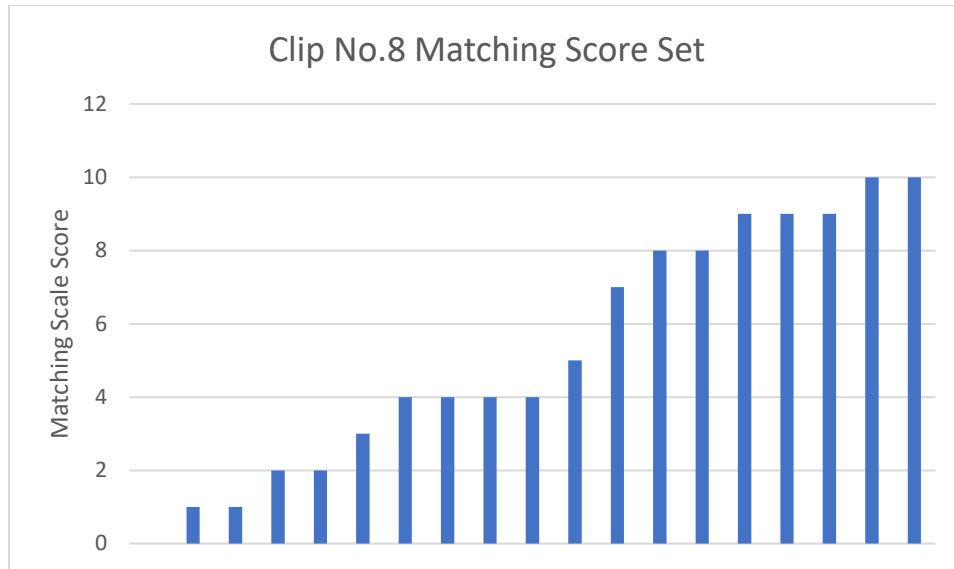
Clip 7 Screen Shot Figure 2-7-5

**Clip No.8:**

**Video Clip: Tekken 7**

**BGM: Bloodborne – Ludwig, the Holy Blade**

**Scene/Stage: Fight Game (Outdoor Fighting Scene)**



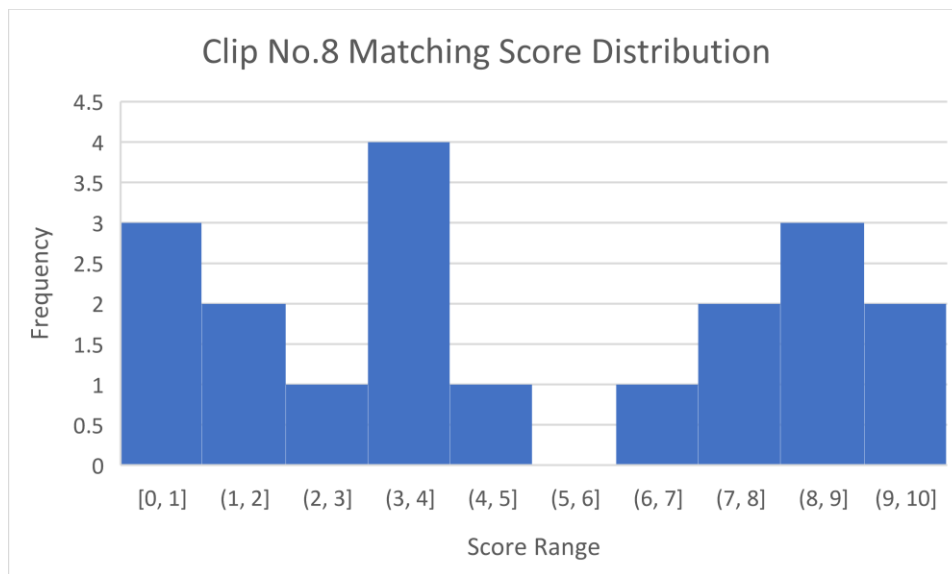
*Figure 2-8-1*

In *Figure 2-8-1*, the trends of low score across the score range is clear for us. We could easily find out there is a drastic decreasing at score 5 that the score performance does not last for long at this point. Besides, Clip No.8 is the only clip in our experiment that received a score rating of score 0. Clip No.8's matching scale score has an interesting score performance. It has lowest score at score 0 and the highest score at score 10. Across the whole score range We could see a totally low score performance.

<b>Clip No.</b>	<b>Mean</b>	<b>Mode</b>	<b>Median</b>
8	5.26	4.00	4.00

*Figure 2-8-2*

From the statistics of 3-M score of Clip No.8 from *Figure 2-8-2*, we could find out that Clip No.8 is indeed the lowest score performance Clip we have among all clips in the experiment. The average score is only at score 5.26. The mode Score and Median score both located at score 4. Mean Score has 1.26 more than the Mode and Median score because of extreme high score value at score 10 and score 9. This represents the truth that Clip No.8 is the most controversial clip with the lowest matching scale score.



*Figure 2-8-3*

From the distribution graph of Clip No.8 *Figure 2-8-3*, we could see that the high score and low score of Clip No.8 are not even negotiable. The two extremely opinion are heavily divided on score 1 and score 10. The most appearance frequency score we have is score 4, which is also a low score. According to our hypothesis. Some subjects might find the clip's BGM and the scene is matching well while some them find the clip is undoubtedly un-matching.

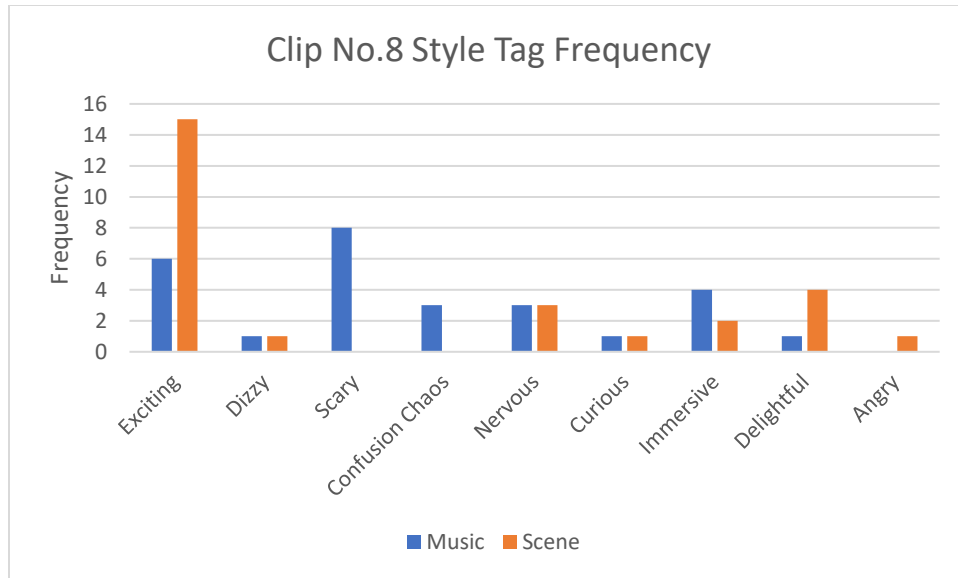


Figure 2-8-4

However, the only obvious data we could get is the style tag “Exciting” and “Scary”. Clip No.8’s scene has twice tag “Exciting” than that of the BGM. The BGM also feels a little bit “Scary” while the scene is not “Scary” at all. We guess is it because of the “Exciting” tag gives a highly divided opinion among subjects. First half of subjects found both the BGM and the scene “Exciting” and gives a high score to the Clip’s matching scale. The other half of subjects found out the scene is much more “Exciting” than that of the music and gives a low score to the matching scale. This might actually be the same reason that Clip No.6 and Clip No.7 looks like having a heavily divided score spread distribution (but they only have a minor-divided score spread distribution).



Clip 8 Screen Shot Figure 2-8-5

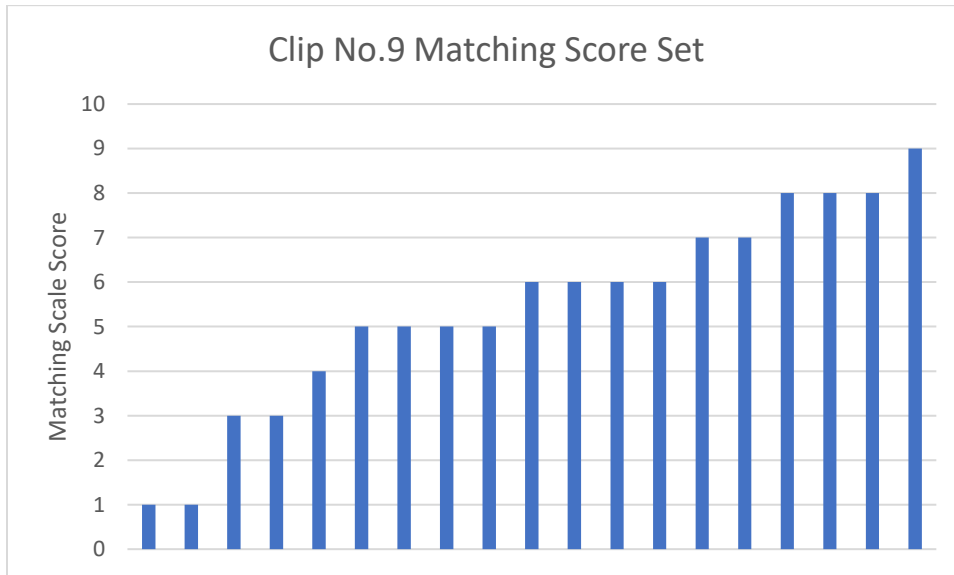


**Clip No.9:**

**Video Clip: Destiny 2**

**BGM: Tekken 7 – Arctic Snowfall 2<sup>nd</sup>**

**Scene/Stage: Base (Safe Zone)**



*Figure 2-9-1*

<b>Clip No.</b>	<b>Mean</b>	<b>Mode</b>	<b>Median</b>
9	5.42	5.00	6.00

*Figure 2-9-2*

From *Figure 2-9-1* and *Figure 2-9-2*, we could see that Clip No.9 has a similar score performance as that of Clip No.8. Low Mean, Mode and Median Score, high score limit at score 9 with low score limit at score 1.

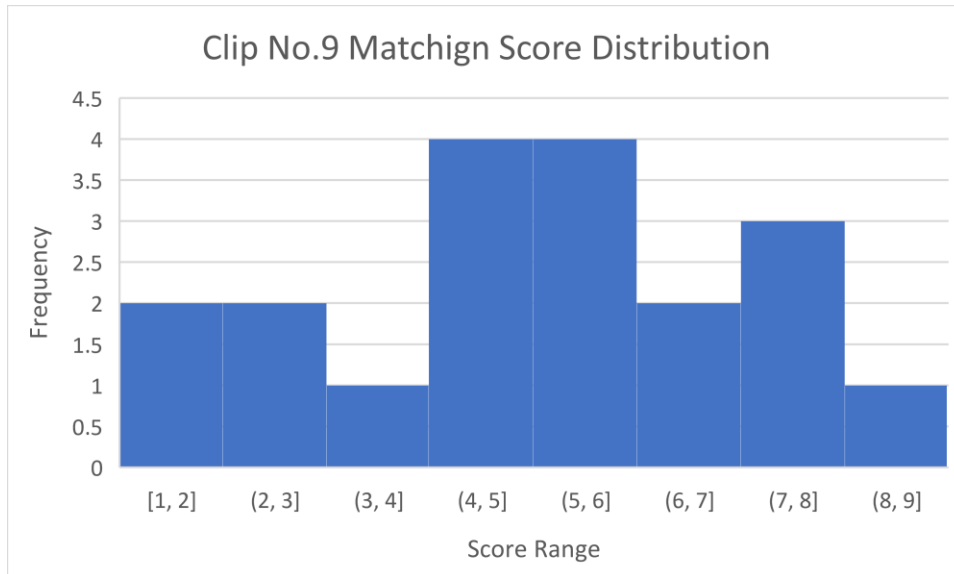


Figure 2-9-3

The difference between Clip No.8 and Clip No.9's score distribution performance is that Clip No.9 has a more centralized score distribution around range from score 4 to score 6. We could see from *Figure 2-8-3* that Clip No.8's score is incomparably divided and gathered at both end of the score range separately. However, we could see from *Figure 2-9-3* that Clip No.9's score is centralized and sharing some same opinion on the scale.

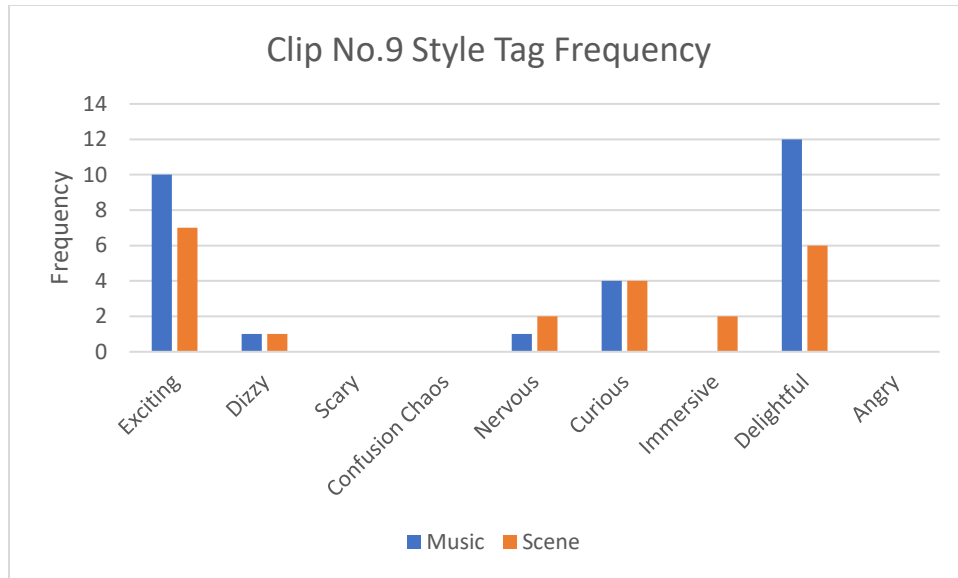
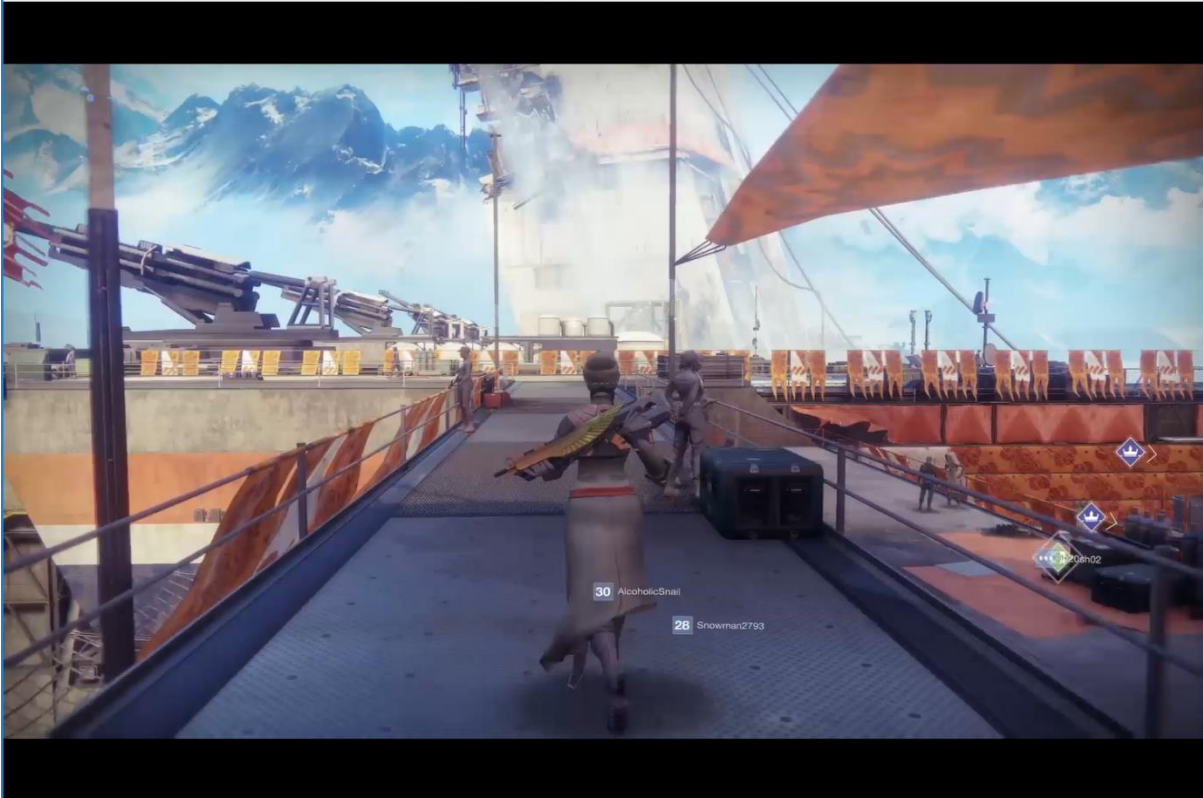


Figure 2-9-4

Considering *Figure 2-8-4* and *Figure 2-9-4* at the same time, we could see that Clip No.9 has more sharable style tag between its' BGM and scene while Clip No.8 only has the style tag "Exciting" massively sharing between its' BGM and scene. This could be a possible explanation for the centralized score phenomenon we found from *Figure 2-9-3*. Since there three matching style tag for Clip No.9: "Exciting", "Curious" and "Delightful", we could conclude that: although Clip No.9 has a great difference of tag appearance of tag "Delightful" between that of the BGM and the scene, the multi-matching tag enable it to have a centralized opinion (matching scale score) in its range and differentiate it from Clip No.8's matching scale core performance.



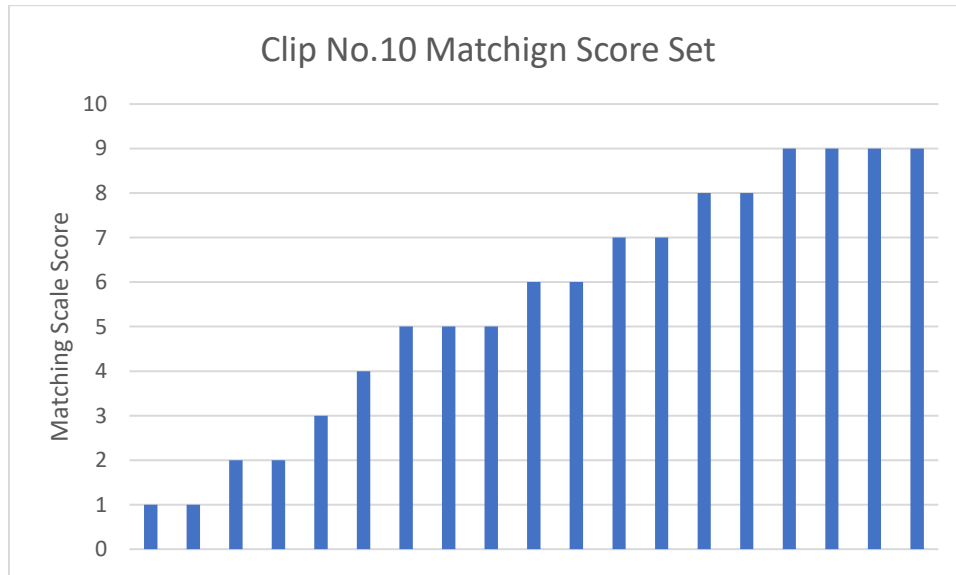
Clip 9 Screen Shot Figure 2-9-5

**Clip No.10:**

**Video Clip: Destiny 2**

**BGM: Bloodborne – Bloodborne**

**Scene/Stage: Outdoor Fighting Scene**



*Figure 2-10-1*

<b>Clip No.</b>	<b>Mean</b>	<b>Mode</b>	<b>Median</b>
10	5.58	9.00	6.00

*Figure 2-10-2*

Observing just *Figure 2-10-1* will give us a feeling that Clip No.10 has a similar score performance as we find in Clip No.8 and Clip No.9. However, we just find out the weirdest score performance in *Figure 2-9-2* that the Mean Score has a 3.42 difference with the Mode Score. Although Mode Score only represents the score that appears the most of times among all score, it is still a data that could only be used as a reference. The difference between the Mean Score and Median Score is small enough to be accepted (0.42). This could be explained by the fact

that Clip No.10 has a separately score distribution that locate at every score but score 9 has most appearance. Therefore, the large difference between the Mean Score and Mode Score could be understood.

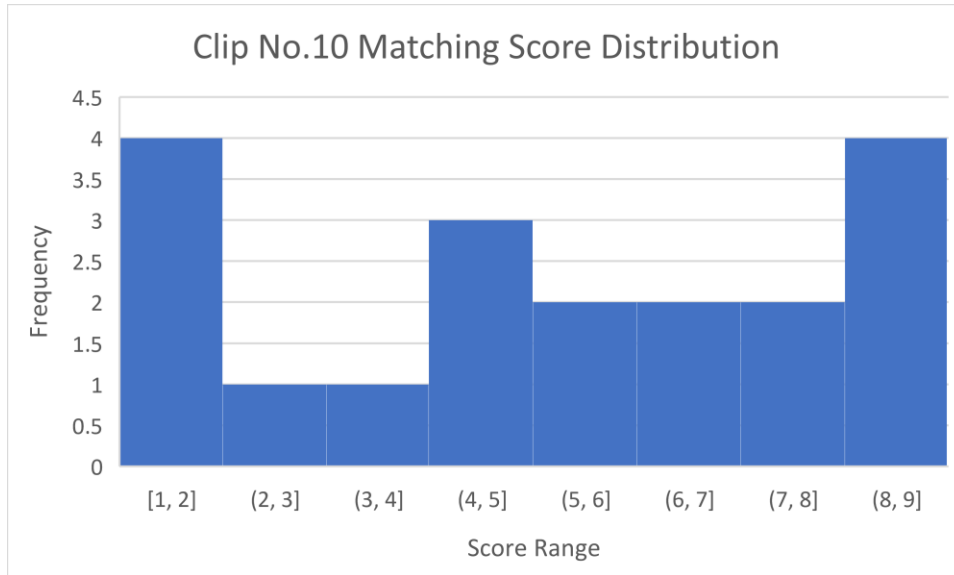


Figure 2-10-3

Evenly separately score distribution could be seen from this *Figure 2-10-3*. Clip No.10 have score 1 and score 9 as the most appeared score across the range while other score from score 2 to score 8 are roughly evenly distributed.

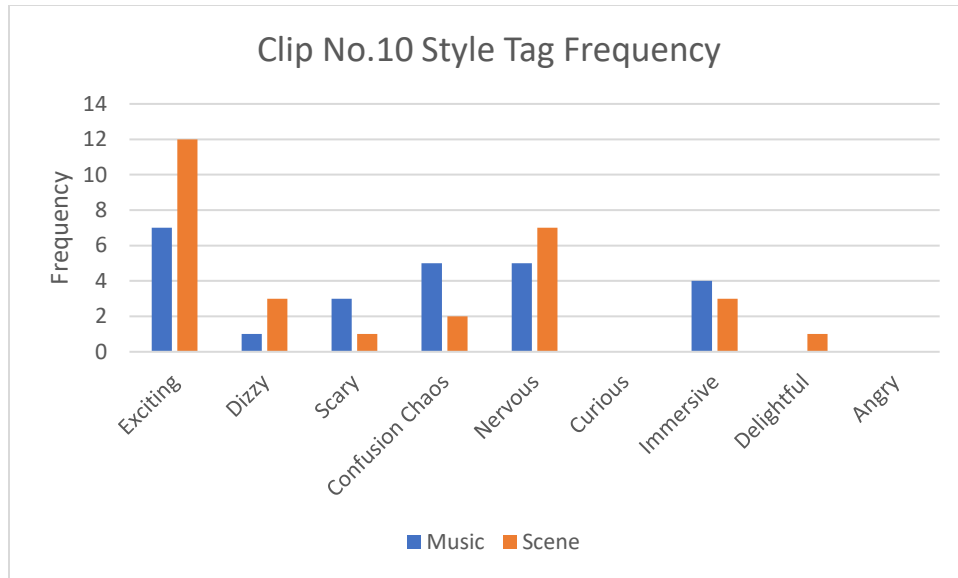


Figure 2-10-4

From *Figure 2-10-4* we could see that both Clip No.10's BGM and scene gives multi-factors in feeling and mood to subjects. However, all factors in their field are different from that of each other. The difference could be the main factor that leads to a divided matching scale score across all score range, but also a reason for some high matching score in the data set.



Clip 10 Screen Shot Figure 2-10-5

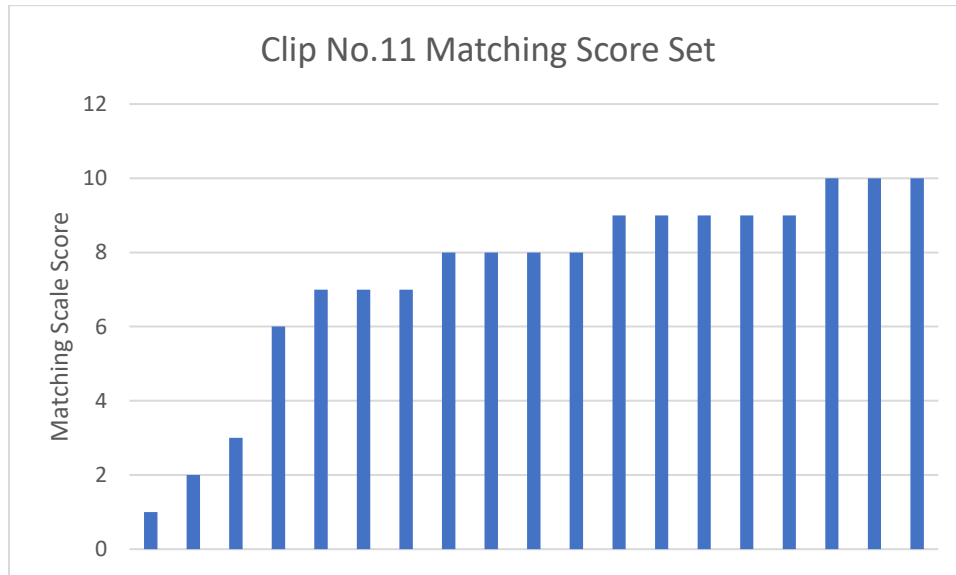


**Clip No.11:**

**Video Clip: Dead Space**

**BGM: Destiny – EDZ**

**Scene/Stage: Outdoor Fighting Scene**



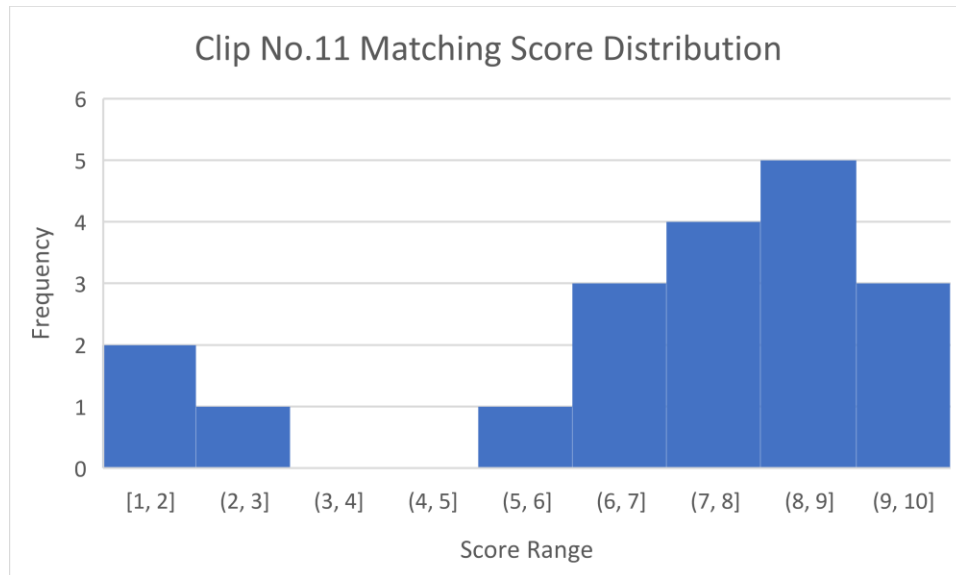
*Figure 2-11-1*

Clip No.11's Matching Score Set in *Figure 2-11-1*, represents Clip No.11's high score across the range. There are a large number of score 8, score 9 and score 10. There are also some low score at the left-most of the graph, but the number of them is really small that does not really affect the score performance of Clip No.11.

<b>Clip No.</b>	<b>Mean</b>	<b>Mode</b>	<b>Median</b>
11	7.37	9.00	8.00

*Figure 2-11-2*

Difference of score 1.63 between Mean Score and Mode Score could be seen from statistics in *Figure 2-11-2*. From both Mode Score and Median Score, we could ensure that Clip No.11 is a clip with high matching scale score between its' BGM and scene. The more than 1 score difference between the Mean Score and the Mode Score could be explained by the extreme low value we have at score 1 and score 2. These low scores drag down the Mean Score value.



*Figure 2-11-3*

In *Figure 2-11-3*, the highly given matching scale score could be seen clearly. We could find some extreme low value at the left part of the graph but their number are pretty small comparing with the number of high score at the right part of the graph.

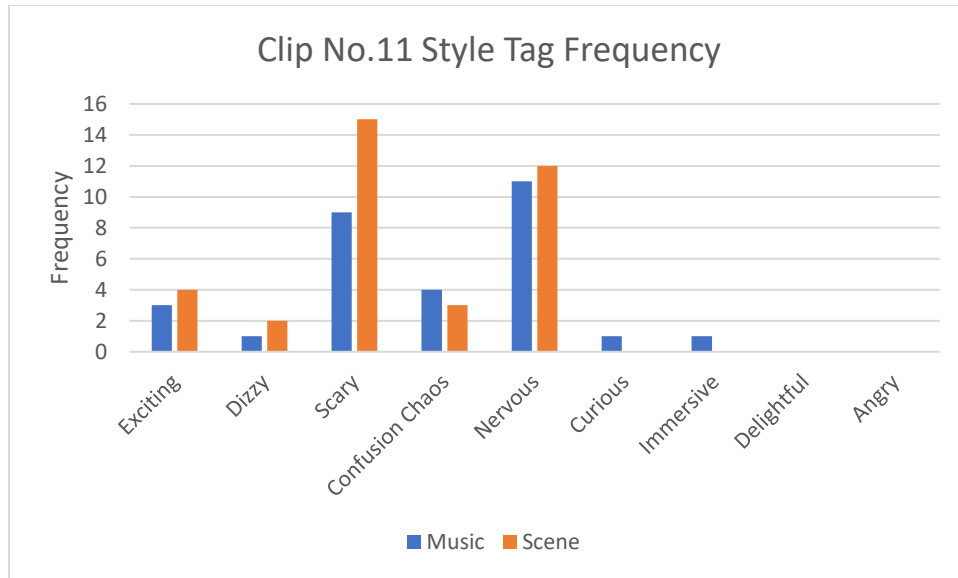


Figure 2-11-4

Figure 2-11-4 provides us a deeper insight about the style tag of Clip No.11. The highly overlapping number of tag “Nervous” should be the main factor leads to the high matching scale score of Clip No.11. The other moderately matching tags like “Exciting”, “Scary” and “Chaos” also contribute to this score. However, the difference between the tag “Scary” in the BGM and the scene could become the reason of those low score at the left part of the graph in Figure 2-11-3.



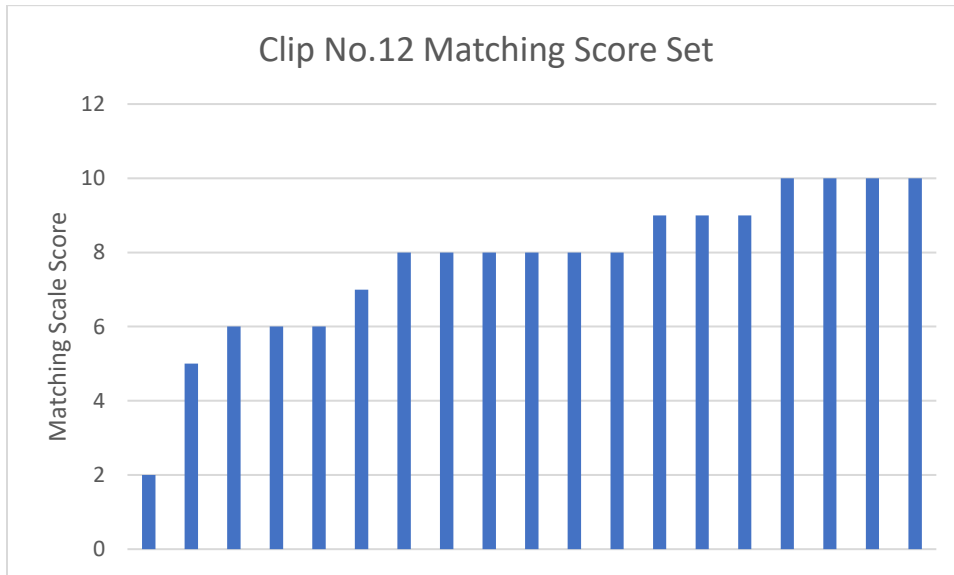
Clip 11 Screen Shot Figure 2-11-5

**Clip No.12:**

**Video Clip: Dark Souls 3**

**BGM: Destiny 2 – The Last City**

**Scene/Stage: Base (Safe Zone)**



*Figure 2-12-1*

<b>Clip No.</b>	<b>Mean</b>	<b>Mode</b>	<b>Median</b>
12	7.74	8.00	8.00

*Figure 2-12-2*

From *Figure 2-12-1* and *Figure 2-12-2*, we could see that Clip No.12 has an extremely similar score distribution as Clip No.11. The Clip No.12's Mean, Mode and Median Score are also really similar to Clip No.11's score performance.

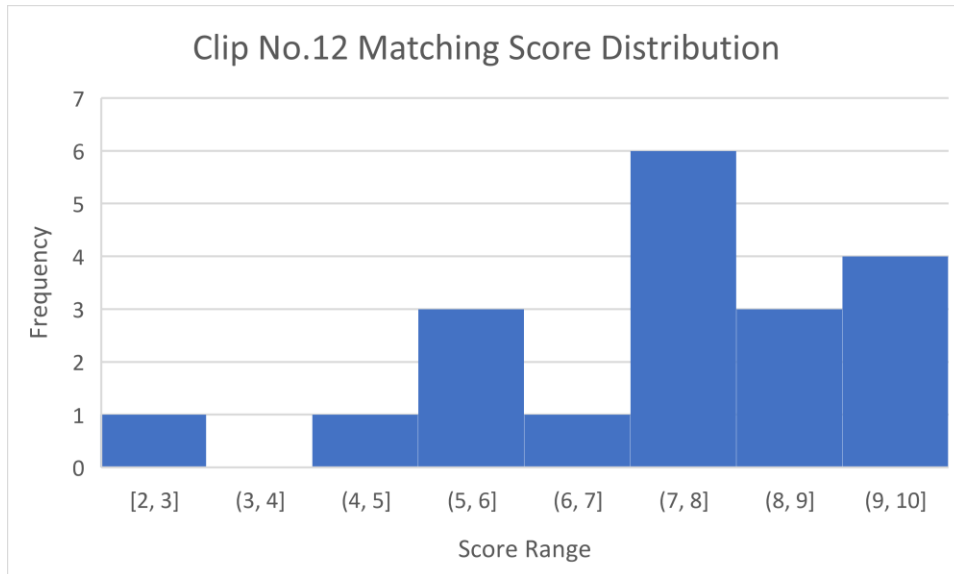


Figure 2-12-3

From Figure 2-12-3, we could easily found out that Clip No.12 has a similar score distribution across the range as we found from Clip No.11. There is a high score centralization at score 8 and there are also a small number of low score at the same time.

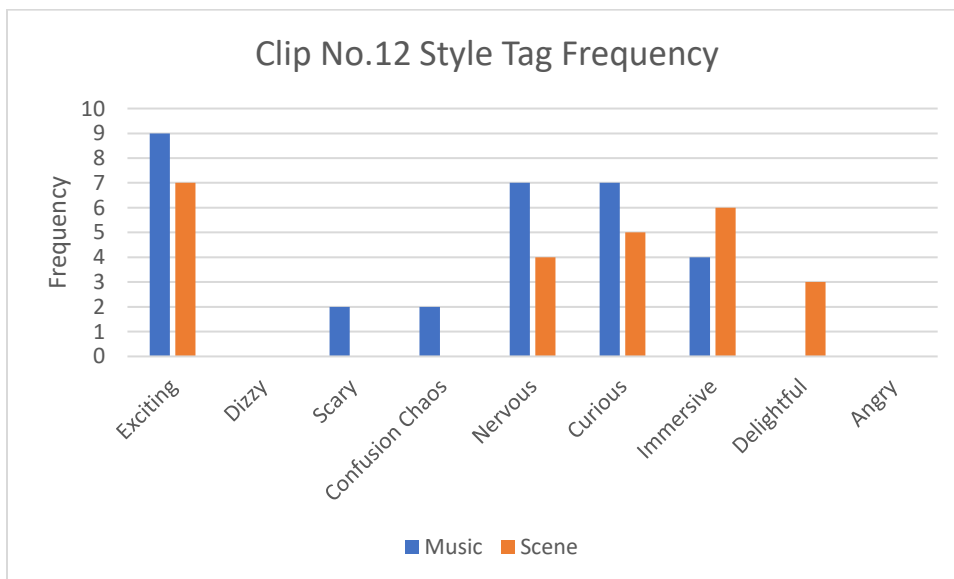


Figure 2-12-4

Despite the different appearance of style tag of Clip No.12, it still has a similar trend in each style tag compared it with Clip No.11. There are multiple style tag has highly overlapping number in multiple different tags at the same time while none of them holding the same number of appearance frequency. In other words, Clip No.12 has a nearly same performance as Clip No.11.



Clip 12 Screen Shot Figure 2-12-5

**Clip No.13:**

**Video Clip: Dark Souls 3**

**BGM: Dark Souls 3 – Iudex Gundryr**

**Scene/Stage: Boss Fight**

Clip No.13 is actually the only clip in this experiment that pair the original sound track with the game scene into a new clip. Therefore, we concentrate a lot on this one clip.

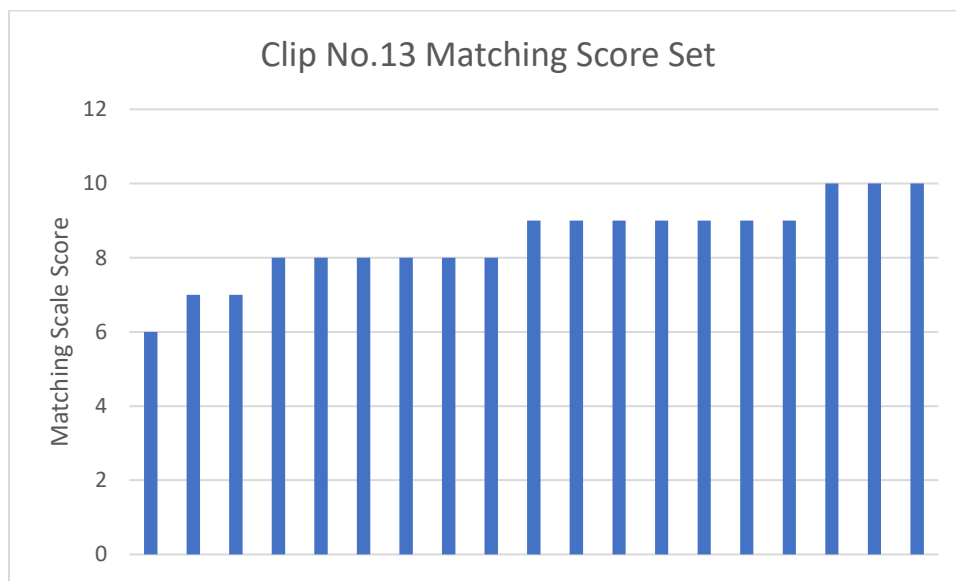


Figure 2-13-1

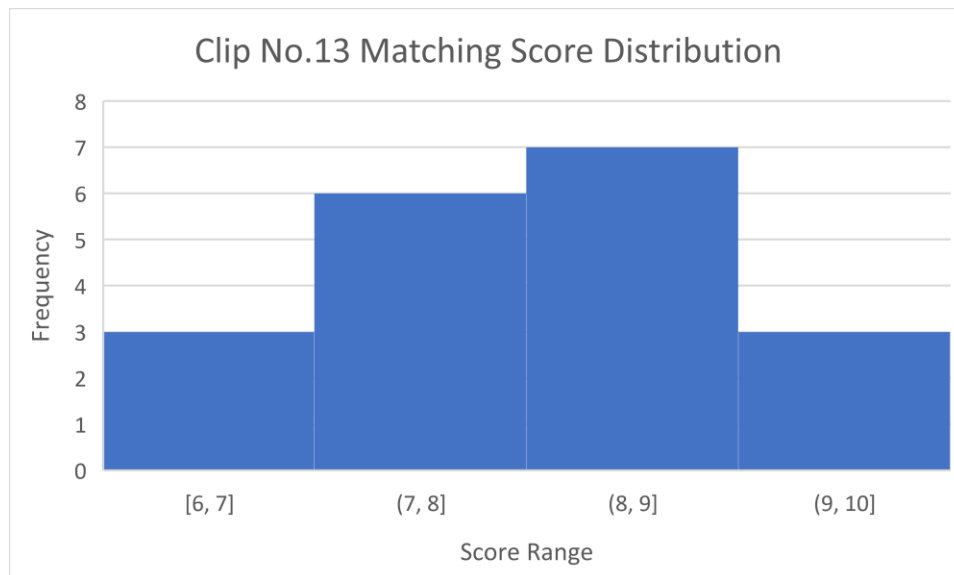
Even just from *Figure 2-13-1* we are able to find out that the score distribution range is smaller than all of other 12 clips. The lowest score Clip No.13 gets is score 6 (which is already a moderately high score) and the highest score it gets is undoubtedly score 10.

Clip No.	Mean	Mode	Median
13	8.47	9.00	9.00

Figure 2-13-2



In *Figure 2-13-2*, the Mean, Mode and Median Score all provides the highest score we have in the whole experiment. The Mode and Median Score both reaches score 9 at the same time while the Mean Score hold only 0.53 lower than the other two scores. We could see that some of the most given score are score 9 (appears 7 times) and the score 8 (appears 6 times), but not score 10. This could be explained that even pairing a scene with the original sound track, subjects might still feel some details that keep them from giving a score 10. However, this factor does not really affect Clip No.10's score performance at all. Since it has the highest score in any ways, we could easily believe that pairing a scene with the original sound track does provide a nearly perfectly BGM-scene-matching clip.



*Figure 2-13-3*

Even though it seems like that the matching scale score of Clip No.13 is centralized at the middle point of the range, it is necessary to notice that this is a range that all locates beyond score 6, which is already a moderately high score point. Clip No.13 is undoubtedly the clip with the highest matching scale score in this experiment.

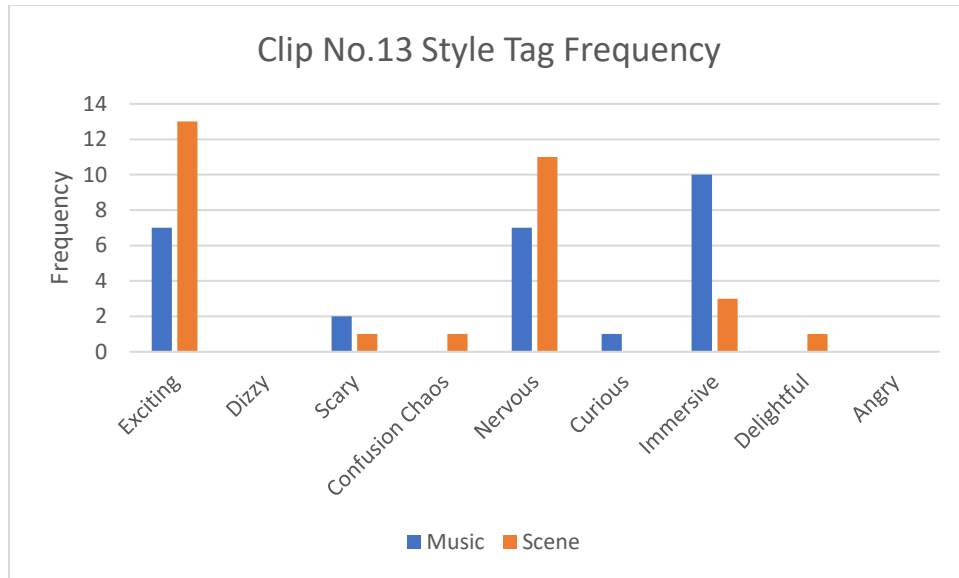


Figure 2-13-4

The only interesting point we could not find out a convincing explanation is the uneven appearance frequency of style tag of Clip No.13. Clip No.13 has several tags overlapping with each other like “Exciting”, “Nervous” and “Immersive”. However, all of these tags have a noticeable difference between their number of appearance in Clip No.13’s BGM and the scene. One of our hypothesis is that the style tag does not really represent a 100 percent persuasive and effective factor that influence the matching scale of the Clip.



Clip 13 Screen Shot Figure 2-13-5

### 5.3. Over-analysis

In our retrospection of the score performance of every clip in this experiment, we have found out some interesting and meaningful facts that might be helpful for us to get deeper insight of the matching scale between the BGM and the scene.

First of all, we found out some rules for high matching scale score. In the two clips, Clip No.2 and No.13, with the highest matching scale score, the BGM and the scene are heavily overlapping in style and atmosphere. Clip No.2 used a *Bloodborne*'s home base scene pairing with a *Dark Souls 3*'s home base BGM. The point is that *Bloodborne* shares a nearly similar theme with *Dark Souls 3* in multiple dimension. For Clip No.13, we even used the same BGM on the same scene, and undoubtedly get the highest score in the experiment. Other Clips like No.1, No.3, No.7, No.11 and No.12 (with a mean matching scale score above 7) have their scene and BGM sharing the similar style too. Some of the weird pairing like Clip No.7, which is using a fight game (*Tekken 7*) scene pairing with a horror game (*Dead Space*) BGM also works well. The possible reason for this phenomenon is that the BGM provides deep, low and scary atmosphere which resonates with the scene well. Another example is Clip No.11, which is using a horror game (*Dead Space*) scene pairing with a MMORPG (Massively Multiple Online Multi-Player Game), *Destiny 2* BGM. However, the BGM we used is portraying a dead zone which somehow provides a sense of danger and frightening and perfectly interacts with the horror game's scene. The opposite example could be Clip No.8. In this clip, a scene from a fighting game (*Tekken 7*) is paired with a BGM from a RPG (*Bloodborne*). The BGM used is originally a piece of BGM from a boss fight from *Bloodborne*. However, this BGM mainly focus on majestic and glorious theme, while the fighting game mainly requires drastic and furious beats. The difference between these two elements brings the lowest score in the experiment. Therefore, it is not difficult to conclude the grand rule for the matching scale between the BGM and the scene that: the style and atmosphere is the main factor of matching scale.

Secondly, the score range of each clip corresponds with the mean score but not acts as expectation. Clip with Highest score, like Clip No.2 and Clip No.13, have their score range centralized at a high place (*Figure 1-1*). On the other hand, clip with lowest score, like Clip No.8 and Clip No.9, have their score distribute across the whole range (*Figure 1-1*). This phenomenon could also be seen from other high score clips, like Clip No.3 and Clip No.7 (with a few exception), and low score clips, like Clip No.6 and Clip No.11. Our discussion provides a legible explanation that the agreement of a highly matching scale in the clip is public-agreed and shared among all subjects, while the disagreement of the matching scale in one clip is subjectively dependent since every participant has their distinguished aesthetic value and feeling. This could also be seen from some high score clip like Clip No.9, which is given several low matching scale score that drag down the mean score.

Thirdly, clips with a relatively high matching scale score often have their BGM and scene share multiple style and mood. Good examples include Clip No.1, No.3, No.7 and No.12. Although they may not match perfectly in each style or mood, they both provides a multi-layer feeling to participants that enable them a high matching score. However, one of the exceptions is that Clip No.13, which is the clip with the highest matching scale score, only has the style “Exciting” shared between the BGM and the scene and even did not have the same matching degree. We think about this one and give some possible explanation. Firstly, Clip No.13 has its BGM and scene highly overlapping in the “Exciting” mood, that participants might get different feel but still agree with the consistency. Secondly, there are some other mood or factors that resonates between the BGM and the scene, but participants did not reflect them on their questionnaire. In other word, we still believe that the BGM and the scene of Clip No.13 heavily share some aesthetic style and emotional mood which are not displayed on subjects’ questionnaire.

In conclusion, throughout the experiment we conducted, we could see through the dataset that the main factor influences the matching scale score between one clip's BGM and the scene is the matching scale between their style and mood. If the BGM and the scene share a recognizable mood that allow player resonate with the atmosphere and immerse into the story, then the BGM performs its character well and will undoubtedly get a high matching scale score. Despite the different function of music around different game, a good BGM should evoke some feeling inside of players' hearts. The feeling could be emotional, angry, sorrowful or even scary. Only by igniting players' heart could a BGM drag players into the game's world and provide an enjoyable atmosphere.

## 5.4. Retrospection and Improvement

### 5.4.1. Retrospection:

We finished our experiment in a few weeks that we spent one week collecting resources like clips and music sound track, one week designing and producing the experiment video and questionnaire, one week for subjects' participation and one week for data collecting and analyzing. In our experiment, we did meet some problems that wasted our time or slowed down our progress. In our retrospection, we write them to avoid similar problems in the future.

1. Choosing multiple games that are under the same game genre. *Bloodborne*, *Dark Souls 3*, *Nioh* are all from the RPG game category. Those games occupy a lot parts in the experiment that they are often feels exciting and nervous, which leads to our overwhelming number of "Exciting" and "Nervous" style tag across the experiment. The style tag "Exciting", "Nervous" are overwhelming across all clips. One of the possible factors causing this is that the BGM and game scene we used in the experiment is too concentrated that we don't have enough game diversity.
2. Did not provide enough style tag choice for participants. The questionnaire provided 9 choices in the BGM and the scene style tag and a blank place for any extra mood or feeling possible. However, participants mainly focus on picking up tag from the list rather than given other mood in their mind. Although provided tags makes this question easier to analyze statistically, it also decreases the diversity and subjective feeling of the mood from each clip. Therefore, the statistics shows up a number of "Exciting" and "Nervous" in the result.
3. The questionnaire we used is a word file and requires the participants to fill it by themselves, which slows down our analyzing process. Because there is no way to extract data from the word file quickly, we manually checked each questionnaire of each participants and store it in the excel file to have the final statistics of the

experiment. This actually took about more than 20 hours to finished because of the details in word using and graph making.

4. Questionnaire Design is not perfect. Some of the question like question 4 and question 5 does not gives a clear instruction that some participants just skip those part which leaves us a great blank in the statistics. Those data will be extremely helpful if answered and collected correctly.
5. We did not have analysis of each person's gaming experience and personality. Throughout the experiment we did notice that some people with little gaming experience give mediocre scores to most of the clip and does not varying much. Some other participants who have some gaming experience will write details they hear from the BGM and tell us about why they feel inconsistency from the clip.

#### **5.4.2. Improvement:**

According to our retrospection, we have had several disadvantages and weakness in the experiment that could be corrected and improved. Therefore, the following several tips are our improvement corresponding each weakness we mentioned in the previous part.

1. Using a more board game base and BGM base. Choosing games clip and BGM from all kinds of video games: RPG, MMORPG, Action Game, Puzzle Game, Adventure Game, Sports Game, Casual Game and Horror Game. The more board game choices we have in our experiment, the more accurate answers we will receive from the data extraction.
2. Asking participants answering questions using their own words or providing a much more versatile tag base in the questionnaire. The first choice will leave the research a team a large number of vocabulary to analyze while the second choice requires a much longer questionnaire to provide enough place for a large number of words.



Either of them will become an inevitable problem since this the experiment do requires different and independent feeling of every clip.

3. Asking participants answer the questionnaire in table using the excel in a regulated format. We could save a lot of time of analysis by extracting those data by using the embedded excel functions. Although analyze them manually will be easier since all the answers are given in format that we could easily compile them.
4. Redesigning question 4 and question 5 into more instructive questions. It is helpful to describe question 4 and question 5 more in details and asking about subjects' opinion about each experiment clip because every clip has different inconsistency between the BGM and the scene.
5. Analyzing each subject's personality and experience to get a deeper insight on the score they given to those experiment clips. However, this will exponentially enlarge the number of data we have to extract, combine and analyze. If we have much more time and resources, we will be able to process this experiment individually and gets a much more comprehensive result report.

## 6. Conclusion

Game industry has become a mega industry these days. Each year, large amount of people become either casual or hard-core gamers. However, no matter what kind of games they would like to play, gamers always enjoy a thing that come along with the gameplay experience, which is the video game music. Video game music is different from other traditional media music since it is mostly non-linear dynamic music. Although in 2000, the National Academy of Recording Arts and Sciences (NARAS) announced that interactive games are able to compete in the annual Grammy awards (McDonald 2005), video game music still does not have the same status as film music does.

Video game music functioned differently under various game types and different area in games. However, we believe there must be some common characteristics between video game music which functioned similarly, and this may explain the question why some video game music matches the game better than others. In our experiment, by switching the background music of gameplay clips, we could get a clue on what effect the consistency between video game music and the game itself. According to our experiment, music that shares the same style and mood with the gameplay often matches better. In 2017, NieR: Automata (Platinum Games, 2017) won Best Music in The Game Awards. Since many people considered NieR's music very immersive, this supports our argument that well composed game music needs to match the game itself. Despite the fact that it is almost impossible for us to evaluate how video game music functioned cause human perception is very personal, we hope our study could give the subsequent studies an insight in this field

## 7. References

Harlin, Jesse. 2018. "DOES IT SOUND NEXT-GEN? – Jesse Harlin (February 2007)". Audiogang.Org. <http://www.audiogang.org/does-it-sound-next-gen-jesse-harlin-february-2007/>.

Collins, Karen. 2008. Game Sound. Cambridge: The MIT Press.

Aska, Alyssa. 2017. Introduction to The Study Of Video Game Music.

"Portal 2'S Dynamic Music - An Interview With Composer Mike Morasky, And Five Tracks To Listen To Now!". 2011. Gamesradar. <https://www.gamesradar.com/portal-2s-dynamic-music-an-interview-with-composer-mike-morasky-and-five-tracks-to-listen-to-now/>.

"History Of Angry Birds | Freeangrybirdsgame.Org". 2018. Freeangrybirdsgame.Org. <http://freeangrybirdsgame.org/blog/2-history-of-angry-birds>.

Munday, Rod. "MUSIC IN VIDEO GAMES." In Music, Sound and Multimedia: From the Live to the Virtual, edited by Sexton Jamie, 51-67. Edinburgh: Edinburgh University Press, 2007. <http://www.jstor.org.ezproxy.wpi.edu/stable/10.3366/j.ctt1q0b5k8.9>.

Sara Rigutti, Carlo Fantoni, and Walter Gerbino. "Web party effect: a cocktail party effect in the web environment." PeerJ 3, no. 3 (n.d.): e828. <https://doaj.org/article/7a68b9ed7cd24918b55b3fd43e6e009d>.

SADOFF, RONALD H. "The role of the music editor and the temp track as blueprint for the score, source music, and scource music of films." Popular Music 25, no. 2 (n.d.): 165–183.

Kuchera, Ben. 2011. "How A Song From The 2005 Civilization IV Won A 2011 Grammy". Ars Technica. <https://arstechnica.com/gaming/2011/02/how-a-song-from-the-2005-civilization-iv-won-a-2011-grammy/>.

Narcisse, Evan. 2011. "Video Games' First-Ever Grammy Winner Talks About 'Civilization' And Music | TIME.Com". TIME.Com. <http://techland.time.com/2011/06/28/video-games-first-ever-grammy-winner-talks-about-civilization-and-music/>.

Hébert, Sylvie, Renée Béland, Odrée Dionne-Fournelle, Martine Crête, and Sonia J. Lupien. 2005. "Physiological Stress Response To Video-Game Playing: The Contribution Of Built-In Music". *Life Sciences* 76 (20): 2371-2380. doi:10.1016/j.lfs.2004.11.011.

McDonald, Glenn. 2005. "A History Of Video Game Music". *Gamespot*.  
<https://www.gamespot.com/articles/a-history-of-video-game-music/1100-6092391/>.