The Effects of Combining Video and Music with Conflicting Emotional Content on Self-reported Emotional Experience

Jiang Zhenghua

Graduation Thesis, February 2015
Media Technology MSc program, Leiden University
Supervisors: Peter van der Putten, Maarten H. Lamers
jiangzhenghua1990@yahoo.com

ABSTRACT

In film, video and music are combined to communicate or evoke an emotional response. Often video and music with similar emotional content are used to reinforce each other, but sometimes intentionally video and music with contradicting emotion is used to achieve a specific emotional effect. In this paper we will investigate this by dubbing a film fragment with several types of music with varying emotional content, and letting subjects self-assess the emotional content through the self-assessment manikin and the Geneva emotion wheel, which translate into three dimensions of emotion: pleasure, arousal and dominance. We found that the perceived pleasure dimension of the film fragment is primarily driven by the video image, regardless of the music. Also for film fragments dominance, if the music is not consistent, dominance and arousal are generally lower. Moreover, the perceived emotion varies more across subjects if the emotional content of the video and audio image are in conflict.

I. INTRODUCTION

Movie directors understand the power of music in film. They hire skillful composers to create music which may enhance corresponding movie scenes. Scientific research also showed that when we receive data for more than one sense at one time (e.g. we see two cars crashing and we also hear the "bang" sound), the stimulation will be more enhanced than if those two events happen separately [Meredith & Barry, 1986]. Physiological explanations for this phenomenon have been provided in neuro-scientific research [Stein & Terrence, 2008]. Experiments with video and music also showed that videos combined with music with similar emotional content will enhance people's emotional experience [Thompson et al, 1994].

However, nowadays there is an attempt by some movie directors and composers to try to use music whose emotion is not consistent with the image. One classic example is the "Assassin Fight scene" in "Sherlock Holmes: A Game of Shadows" (2011) directed by Guy Ritchie and composed by Hans Zimmer: the scene in which Holmes is fighting with an assassin is accompanied by cheerful Irish music. One other interesting example can be seen in one episode in English TV Series "Agatha Christie's Poirot" (2008) with music by Stephen McKeon. In episode 4 in series 11 called "Appointment with

Death", when a corpse is discovered, some characters shriek in horror, but the music is very soft and relaxing Arabic music.

Accordingly if the video image and music do not correspond in emotional content, it may lead to a more memorable scene, and a new emotional content may be experienced by the audience. However it doesn't always deliver the desired effect. Therefore, the research question is:

How will the overall perceived emotional content for a film fragment change if the background music is varied across different dimensions of emotional content. Hence the background music variants may be conflicting with the video in terms of specific emotional dimensions.

Unlike previous research that focused primarily on what modality takes priority (mostly visual), we are more interested in how conflicts in emotional content affect the overall experience. It is a novel cinematic technique that is being used more and more often, and it is interesting to analyze and understand its effects better through dimensional models of emotion that are commonplace in psychology. Likewise it may provide a better view on the usefulness of these psychological models for cinematic researchers, film critics and film makers.

The remainder of the paper is structured as follows. In section II we will discuss related work, and next we briefly introduce the experimental approach in section III and the tools of the research in section IV. Results will be provided in section V and this will be followed by a discussion (section VI) and a conclusion (section VII).

II. BACKGROUND AND RELATED WORK

In this section we will highlight relevant background from related work, mostly about cinematic research or psychological research about multisensory perception and integration.

A. Power of Music in the Movie

The introduction has mentioned that videos together with emotional-corresponding music may enhance the effect of impression, and neutral videos with varying music will lead to varying emotions. A considerable amount of research has been done in this area.

Magliano and Dijkstra found that music is one of the six film factors (such as montage) which might contribute to the psychological definition of the situation. In their study, participants were invited to watch a movie clip from James Bond "Moonraker"(1979). While watching the video, participants were asked to predict what would happen next aloud. Finally, the experiment showed that music together with other five film factors might support visual and discourse information which affect participates' prediction of the story of the movie [Magliano et al, 1996].

In this research, music was not tested separately, but together with other film factors. Therefore we cannot say that music is a significant film factor based on this experiment, but it still shows a relationship between music and video.

In another study about music and video clips, Marshall and Cohen studied the effect of two different soundtracks on impressions about three geometric forms: two triangles and one circle [Marshall et al, 1988]. In their experiment, subjects watched the triangles-and-circle animation with one of the two soundtracks or without any soundtrack. The result showed that subjects might connect that music with the "style" of those characters (even if they were just geometric forms). For example, the large triangle

was judged as the most active one among the geometric forms [Cohen & Annabel, 2001].

Besides geometric-form animation, video clips were also used in the experiment. Bullerjahn and Guldenring invited a couple of professional composers to create five soundtracks of different style for the same 10-minute neutral video (video with no obvious emotion content). Both quantitative and qualitative analysis showed that the different soundtracks led to different judgments of the appropriateness of emotional categories, choice of genre, reasons for the actions of the protagonist, and expectations about the completion of the film [Bullerjahn et al, 1994].

More features of music soundtrack were found, in fact. Bolivar noted that music may let us focus attention to a visual object not only from structural congruencies, but also from semantic congruencies [Bolivar et al, 1994][Boltz et al, 1991]. For example, a soundtrack featuring a lullaby may direct attention to a cradle rather than to a fish-tank when both two objects appear on screen; this research showed some advanced relationship between video and music.

This piece of research more or less showed that music may affect the meaning of corresponding image, however in these experiments the video clips were normally neutral (with no emotions), therefore more research should be done in this area.

B. Priority of Different Senses

While we are watching one video and listening to one voice (no matter whether they are in consistent), usually optical sense will take more important part. There is no doubt that optical sense often takes priority over other senses. Scientists also found that compared to auditory sense, optical experiences will be remembered longer, stronger, and more accurately [Bigelow et al, 2014].

Besides, there was also some research about conflict between auditory and optical senses, and the result is called “McGurk Effect” [Munhall et al, 1996]. Research shows that when there is a little conflict between visual and auditory senses (e.g., when a person pronounces “va” while his mouth-lip-motion acts like “ba”), optical part will take priority (the people in the experiment will think they have heard “ba” instead of “va”). Our brains ignore physical sound but completely accept optical image in this case [Smith et al, 2013].

Currently the McGurk Effect relates to oral language. It may be meaningful to be a reference of our current study, due to the fact that they are both related to “receiving multisensory data in conflict”.

In fact data in conflict can happen in one single sense. The Stroop Effect is one of those effects. The main description of this effect is simple as described in Wikipedia: “When the name of a color (e.g., "blue," "green," or "red") is printed in a color not denoted by the name (e.g., the word "red" printed in blue ink instead of red ink), naming the color of the word takes longer and is more prone to errors than when the color of the ink matches the name of the color.” [Wiki: Stroop Effect][Stroop & John, 1935]

Some popular science books describe this effect as “… people will have confusion when asked to name the color of the words”, but in fact there is no “confusion” in this case. People won’t hesitate about what to answer, but they just answer the color of the meaning of the words (maybe not aloud, but silently) without hesitation, then they realize they are wrong. Therefore it shows that the semantic takes priority compared to the optical senses (in this case: color).

Based on the Stroop Effect, the thought with movie and background music with conflict was designed, which may be considered as one other kind of Stroop Effect.

III. METHODS

In our experiments we will show a collection of movie clips with different emotional content, and we dubbed each fragment with various variants of music with different or similar emotional content. Subjects then need to rate their experience. In this section we will describe some of the methods used, such as a standard collection of videos categorizing different emotions, the PAD dimensional model for emotion, and the various measurement tools for measuring emotion. The detailed experimental set up will be described in the next section.

A. Standardized Collection of Movie Clips with a Specific Emotion

The question like "Which video clip means 'happy', and which movie clip will trigger a 'sad' emotion" has been studied by a number of researchers. Robert Levenson and his student took years to create a set of movie clips [Gross et al, 1995], and each movie clip may trigger one specific emotion in people (and only one emotion, not a combination of multiple emotions). Emotions include happiness, anger, fear, shock, satisfactory, sickness, etc.

This collection is now a generally accepted data set for emotion psychology experiments. For example, the movie clip representing "sadness", which is a clip from movie "The Champ" in 1979, has been used by Dutch scientists to test the connection between sadness and binge eating disorders [Dingemans et al, 2009]. The movies clips we used in research also come from this same collection.

B. The PAD Emotional State Model

It is not easy to compare between emotions with one single parameter, hence various dimensional models were created in order to study emotions using multiple dimensions, such as the circumplex model, and the PANA model [Russell, 1980][Watson & Auke, 1985]. Here we decided to use PAD emotional state model, which is a generally accepted model of affective state and used in many other studies. In the PAD model different emotions can be separated into three dimensions: pleasure, arousal, and dominance (PAD in short), which make emotions easy to measure and analyze. In those three dimensions, "Pleasure-displeasure" assesses the affective quality of experience; "arousal-non-arousal" addresses the issue of physical activity and mental alertness, and "dominance-submissiveness" defines the individuals feeling of control, or lack thereof, on the given situation [Caicedo et al, 2006][Mehrabian, 1980]. For example, if someone feels angry, he is considered to be negative in the pleasure dimension, high in the arousal dimension, and independent in the dominance dimension.

In this study, the PAD scale will play an important role, by selecting video clips and measuring subjects' emotions.

C. Measuring Emotions

In order to be more accurate as well as easier reproducible, two methodologies of measuring emotions should be chosen in the experiment, which

1. Can be done accurately.
2. Requires no specific technical skill or tools, i..e. programming or physical measurements.
3. Can be executed in 5 minutes.

The analysis of some scientific methods of measuring emotions can be seen in Table 1.

Method for assessment of emotional state	**Can be done accurately?**	**Need specific technique?**	**Can be executed in 5 minutes?**
Physiological Arousal (e.g. [Caicedo et al, 2006])	No	No	Yes
Motor Expression (e.g. [Friesen et al, 1983])	Kind of	Medical technique needed	Better longer
The PAD Scale (used directly) [Mehrabian, 1996]	Yes but hard	No	Yes
Self-Assessment Manikin [Bradley et al, 1994]	Yes	No	Yes
Geneva Emotion Wheel [Scherer, 2005]	Yes	No	Yes
PrEmo [Desmet, 2005]	Kind of	No	Yes
Affect Button [Broekens & Brinkman, 2013]	Yes	Computer programming needed	Yes

Table 1, analysis of some scientific methods of measuring emotions

We finally decided to use the **self-assessment manikin** [Bradley et al, 1994] and the **Geneva emotion wheel** [Scherer, 2005] to measure emotion, because they are easy to execute, and they both can be transformed into PAD dimensions, since the video and audio clips are chosen based on those dimensions.

C-1. Self-assessment Manikin

The main part of self-assessment manikin (SAM in short) looks like a selection of comics characters, see Fig 1.

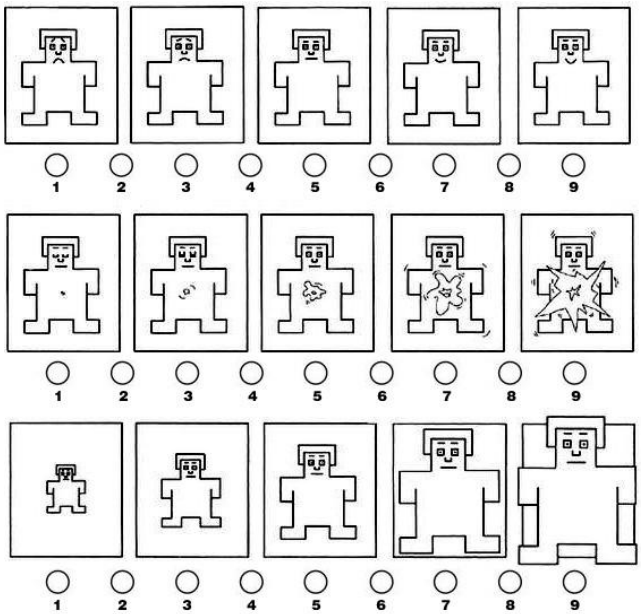


Fig 1, Self-assessment Manikin

When the user has one emotion in mind, he should mark the scales above. The first line corresponds to pleasure, from extremely negative to extremely positive. The second line corresponds to arousal, from low arousal to high arousal. The third line corresponds to dominance, from dependent to independent. The illustrations make users feel easy to comprehend.

In this SAM, only the first line (pleasure) and the third line (dominance) are related to the dimensions of four video clips, but in the research we will analyze arousal as well, so we still keep the second line.

C-2. Geneva Emotion Wheel

In principle SAM alone is accurate enough for measuring emotions, but it may be hard for people to assess these emotions separately along the dimensions, so the Geneva Emotion Wheel is introduced as one other tool of measuring emotion. The spider-web-like the Geneva Emotion Wheel (GEW in short) is another kind of questionnaire. The user should tick the correct circle which describes his emotion. The larger the circle is, the stronger emotion it implies, which is very clear for users. Here we have the GEW in 2.0 version from the creator [Sacharin et al, 2012], see Fig 2.

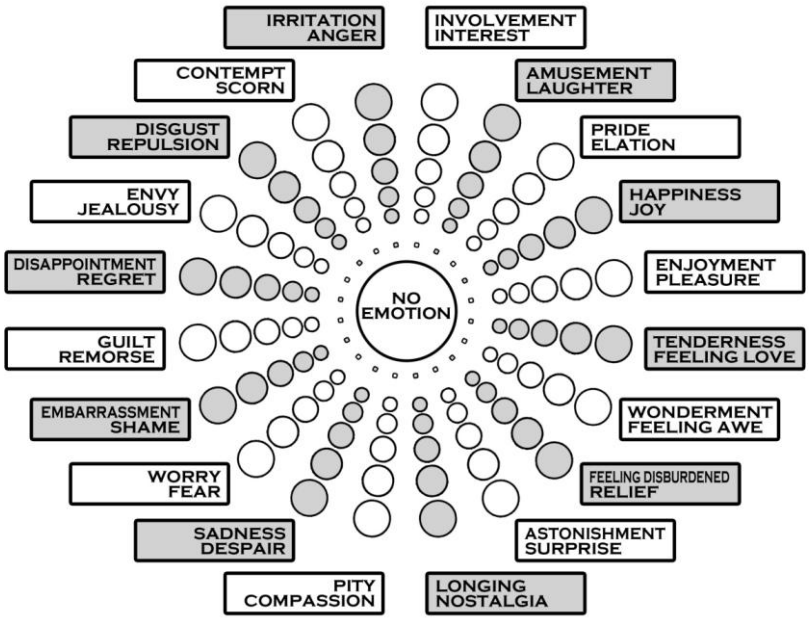


Fig 2, Geneva emotion wheel 2.0

In the original version of GEW, the middle part contains two parts: “No emotion” and “Other emotion (write it down)”. However in this experiment, if too many people tick “other emotion”, there will be problem, so we deleted this part, only to let spectators choose the most similar emotions they have.

In fact, the emotions in GEW are arranged according to pleasure-dominance dimensions, which contain pleasure in horizontal direction and dominance in vertical direction, and the middle part is the origin of pleasure-dominance (P-D in short) coordinates. Therefore we can have detailed study of pleasure and dominance based on GEW.

IV. EXPERIMENTS

During experiment-designing period, different movie-music combinations were made, and a proper questionnaire was designed for subjects to measure emotion. In the following paragraphs, related content will be introduced.

A. Choosing Video and Music

According to the video category in previous paragraphs [Gross et al, 1995], we chose four movie clips, which will trigger people's four emotions, see Table 2.

Movie Source	**Brief Content**	**Year**	**Emotion Involved**	**Time Length**
When Harry Met Sally	Sally fakes an orgasm at a restaurant	1989	Amusement	2m52s
My Bodyguard	"Bully scene"	1980	Anger	5m49s
The Champ	Boy cries at father's death	1987	Sadness	2m43s
/	Beach scenery, with sea wave and coconut tree	/	Relief or relaxation	3m13s

Table 2, four movie clips used in the experiment

Those four emotions are all based on the content of the four movie clips (except the beach scene), therefore subtitles were added for the experiment. Moreover, these four emotions are not chosen in random. When given pleasure-dominance coordinates, those four emotions (together with four video clips) belong to four quadrants. See Fig 3:

First quadrant: amusement (Positive, independent);
Second quadrant: anger (Negative, independent);
Third quadrant: sadness (Negative, dependent);
Fourth quadrant: relief (Positive, dependent).

Fig 3, four movie clips in four quadrants

Music with four corresponding quadrants were also selected and edited, therefore 16 (4 multiplied by 4) movie-music combinations were made, and 16 video variants were produced. The reason why we only used pleasure-dominance dimensions instead of pleasure-arousal-dominance dimensions is because 64 (8 multiplied by 8) videos should be made if so, which needs huge amounts of subjects.

With regards to the music, most of them are chosen from movie soundtracks, and the emotion involved in the music were mostly referred from corresponding movie content and some film reviews about the soundtrack. Based on personal professional movie-making experience, the primary investigator selected the music clips to match the movie clips. Since the rhythms in the four movie clips above are different, we cannot put the same music in two different movie clips, therefore we used 16 sets of music clips altogether.

For the music we selected less known music fragments to minimize the impact of personal emotional experiences. The sources of music clips can be seen in appendix A.

In the following content we will introduce four original movie scenes (without music) as "movie clips", and movie scenes together with music clips as "experimental clips".

B. Experimental Setup

In this experiment we have four movie clips and four corresponding styles of music clips as mentioned before, therefore 16 combinations (which means, 16 experimental clips) will be made. Four of them contain movie and music clips with consistency in terms of emotional content. We will call them "normal" experimental clips, and the others are "conflicting" videos, because their movie-and-music emotional contents are in conflict.

In order to avoid ordering bias, it is better to let the subjects watch the four random experimental clips, no matter whether they are normal or conflicting. However, they are not allowed to watch the same movie clip more than once (for example, watching "the Champ" with music A then watch "the Champ" with music B should not happen in this experiment), therefore each participant will see all four movies, but with a random selection of music style per movie, to avoid bias. Participants should note their emotions (using SAM and GEW) after watching each experimental clip.

Better and clearer results would be evident if more participants participate in this experiment. Here given the limited time available a target of at least 20 participants was set, which means it is better that every experimental clip is watched five times by different subjects even if they are arranged in random. We used a stratified random generation method to control such random situation. In fact, for n people participating in the experiment, n/4 evaluations will be given to each video variant.

This experiment is better to be done face-to-face, however online test is also possible using e-mail or Skype. Due to the fact that some volunteers were not in the Netherlands during the experiment, online test was executed for them.

V. RESULTS

A. Brief Description

In December, 2014, and January, 2015, twenty subjects (11 male and 9 female, 7 online and 13 offline, ages between 19 and 30) altogether participated in the experiment. They watched the four arranged

experimental clips separately and filled in the questionnaire. The subjects were all requested to read the introduction of SAM and GEW first to know how to interpret and use these tools. Then they watched the videos and noted their emotions immediately after watching each experimental clip. The questionnaire can be seen in Appendix D.

They watched four experimental clips with four identical image part (movie part), but different background music. As said in previous paragraphs, each music clip, selected randomly from for quadrants in the pleasure dominance space, was used exactly five times in the experiment.

Finally twenty questionnaires were collected and categorized according to PAD dimensions of movie and audio clips.

B. Analysis

All the results of that experiment were arranged as Excel file and analyzed. The following content is based on the average value and standard deviation based on statistic results, where distinct results can be observed.

In order to be short and easy to comprehend, in the following content, numbers will be given to movie and music clips in each corresponding quadrant in P-D coordinates. See Fig 4.

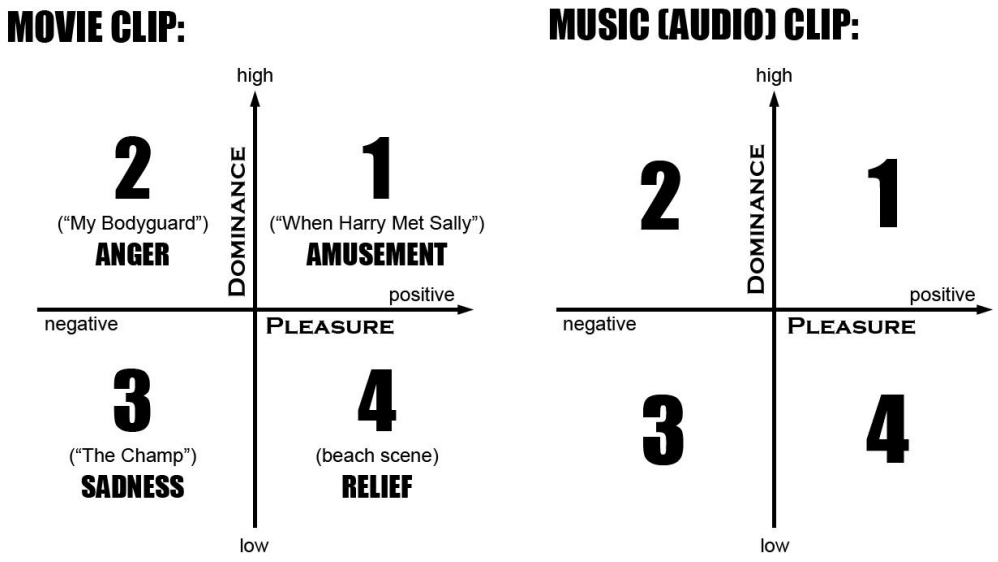
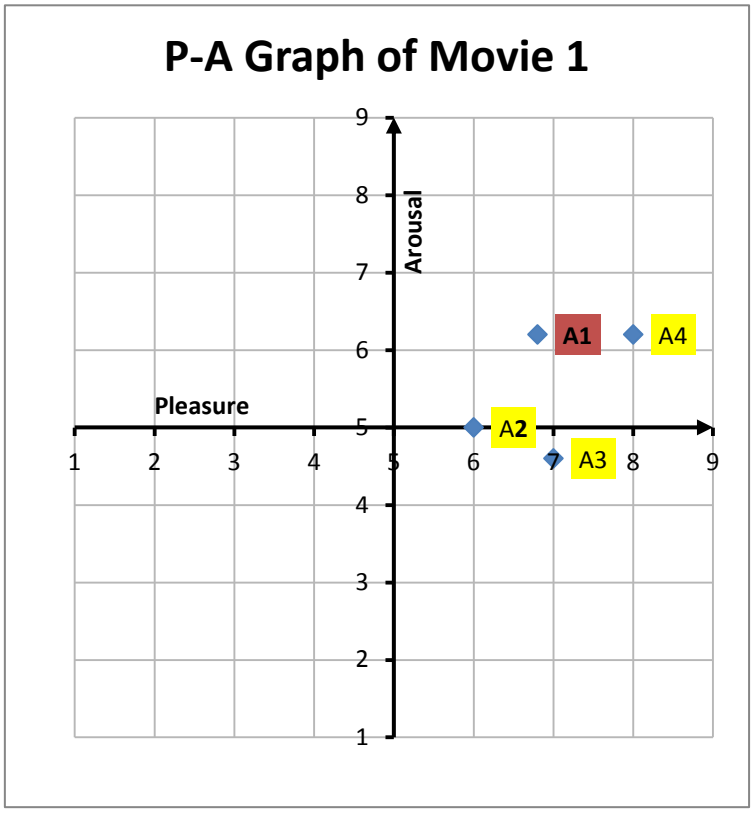


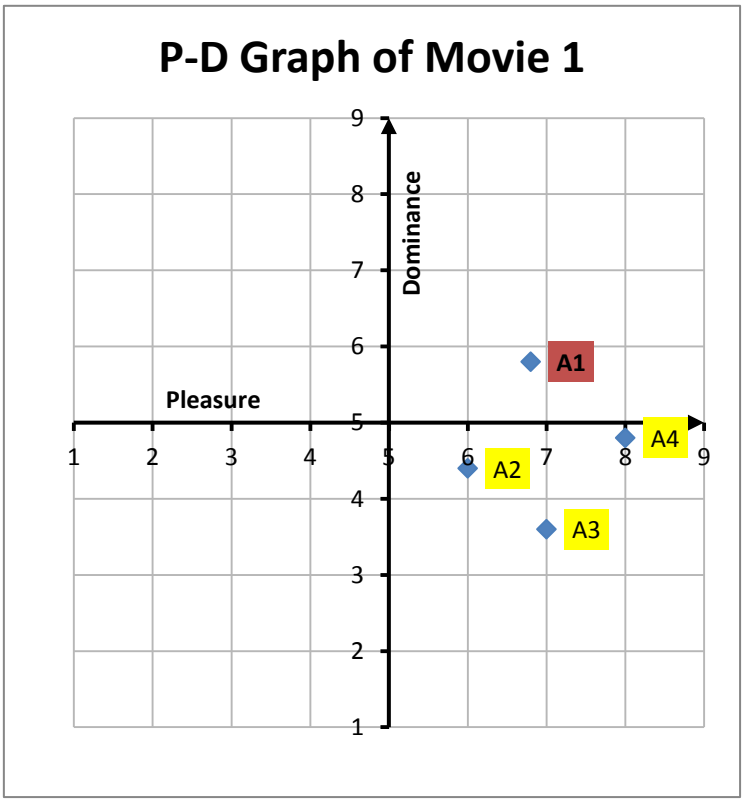
Fig 4, numbers of four quadrants in P-D coordinates. If we put those four movie and music clips into P-A coordinates, they may be in the same quadrants as in P-D coordinates.

B-1. About Average Emotion

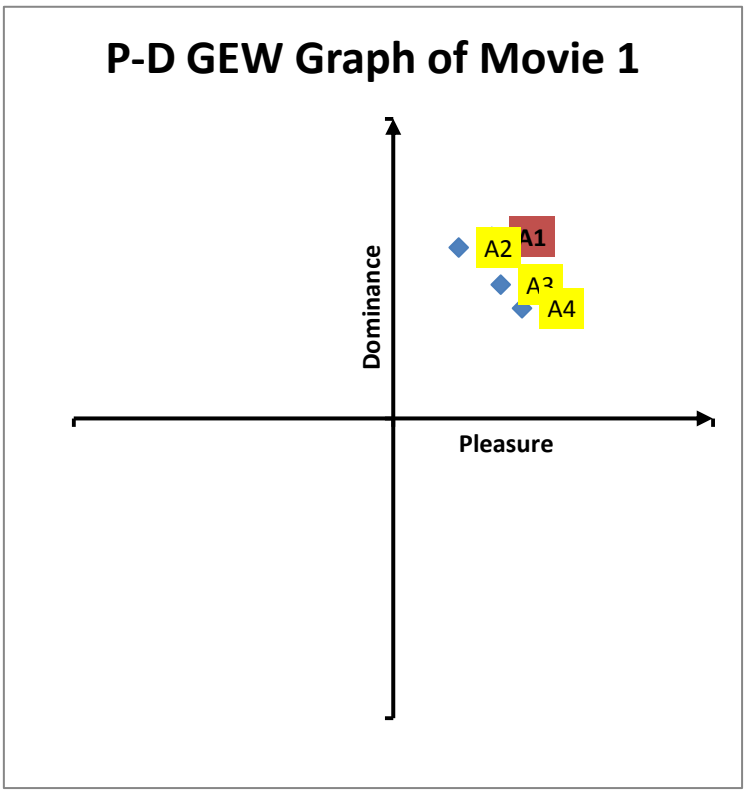
The mean result of the emotion in each movie clip can be shown in three graphs: P-A and P-D graph showing the result based on the self-assessment manikin and P-D GEW graph showing the result based on the Geneva emotion wheel. Graph 1, 2, and 3 shows the three graphs of movie clip 1.



Graph 1, the pleasure-arousal (P-A) graph of movie clip 1 ("When Harry Met Sally")



Graph 2, the pleasure-dominance (P-D) graph of movie clip 1 ("When Harry Met Sally")



Graph 3, the pleasure-dominance in Geneva emotion wheel (P-D GEW) graph of movie clip 1 ("When Harry Met Sally")

In each graph, four dots represent the average PAD value of four certain experimental clips (with different music clips) based on one movie clip, and the number of each dot indicates the quadrant each music clips belong to in P-D coordinates (see Fig 4). The highlighted dots represent the experimental clip with music in consistency, which is the "normal" movie-music combination.

Besides, there is no accurate scale in P-D GEW graph. The ends of two axes are in correspondence to the outer edge of GEW (see Fig. 2), but it is not accurate to put number "5" there. Therefore it is better to see the relative position of the dots in P-D GEW graph, instead of analyzing absolute position.

From these three graphs we can observe and find some features.

1. P-D graph and P-D GEW graph seem somewhat inconsistent. The relative position of A1, A2, and A3 align but A4 doesn't. The reason may be the small sample size or subjects' inability to comprehend "dominance (control)" in the questionnaire.
2. The experimental clip with consistency (the "normal" video) is in the first quadrant in two P-D graphs, which is the quadrant of that movie clip and music clip.
3. Generally the arousal and dominance of conflicting experimental clips are lower than those of normal experimental clips.

However this is just a brief result, we have to observe all of the complete graphs, then. The complete graphs of four movie clips can be seen in appendix B, which can be observed and analyzed together. Graphs of movie clip 1 (Graph 1, 2, and 3) are also in the appendix in order to be easy to be observed as a whole.

From these graphs we can have these observations:

1. People will take "pleasure" dimension (positive or negative) of their emotions based on the movie

clip of the video, instead of the audio part.

2. While seeing videos in quadrant 1, 2 (amusement and angry), the arousal and dominance of the emotion when people watching movie with music with inconsistency will be generally lower than those when watching video with consistent music.

3. On the contrary, when watching videos in quadrant 4 (relief), the arousal and dominance of the emotion when people watching movie with music with inconsistency will be higher than those when watching video with consistent music.

4. While seeing video in quadrant 3 (sadness), that will be different. P-A and P-D graphs show that while watching movie of sadness with inconsistent music, people's arousal and dominance will be generally lower than those while watching that video with consistency. However, strangely, P-D GEW graphs show different result. It shows that while watching movie of sadness with inconsistent music, people's dominance will be higher than those while watching movie with consistent music.

We would like to find out how music of different emotions controls the emotion of experimental clips (e.g. happy music makes video happier, etc.), but we didn't see obvious tendency in these graphs, except in graph of movie 4 (movie of relief). In graph movie 4 we saw strong tendency for high-dominance music pieces: positive music makes relaxing video more positive, and vice versa. This phenomenon can be taken account as reference while editing video, and more experiments can be taken about that.

Here are some more observations of "how music of different emotions controls the emotion of experimental clips" about different movie clips, but they are not significant results, just brief observation. Extra researches can be done about that.

Movie "amusement" (first movie clip): relaxing music (quadrant 4) makes pleasure dimension higher, and music with anger (quadrant 2) makes it lower (compared with the same movie with emotional-corresponding music, and hereinafter). For music from quadrant 3 there is no clear effect of this movie clip.

Movie "anger" (second movie clip): happy music (quadrant 1) and sad music (quadrant 3) may turn pleasure dimension higher, and relaxing music (quadrant 4) may turn down audience's dominance.

Movie "sadness" (third movie clip): music with anger (quadrant 2) may make arousal dimension lower.

Although they are not totally confirmed, they can be used while doing video-editing in order to adjust audience's emotion.

B-2. About Emotional Variance

Apart from the average tendency of emotions based on different music, we expected to observe variation across subjects in the assessment of the various dimensions. Therefore in order to estimate the variance of people's emotion for a single movie-music instance, the standard deviations of the dimension-values in every experimental clip were calculated, and categorized based on movie clips. Afterward, in order to compare the variance between different experimental clips, the standard deviations of corresponding dimension were divided in order to compute ratio. The quotients of standard deviations in movie clip 1 can be seen in Table 3. We will provide full details about the table later.

	Self-Assessment Manikin			Geneva Emotion Wheel			
	Pleasure	**Arousal**	**Dominance**	**Emotions**	**Level**	**Pleasure**	**Dominance**
Music 1							
Music 2	0.59	0.65	0.78	*0.00*	*0.30*	0.18	0.23
Music 3	1.18	0.52	0.92	*0.00*	*0.77*	0.32	0.42
Music 4	1.18	0.44	0.78	*0.00*	*0.33*	0.34	0.15

Table 3, ratios of standard deviation of movie clip 1

The numbers in that table show the quotients of corresponding parameters between music clip 1 and other different music clips of movie clip 1. For example, the first number "0.59" is the quotient that standard deviation of pleasure of video 1 with music 1 divided by that with music 2. "GEW emotions" are the different emotions in Geneva emotions wheel and "GEW level" shows how strong their emotions are in Geneva Emotion Wheel. While calculating, the actual numbers of emotion and level of GEW were noted as in Fig 5, which are based on polar coordinates.

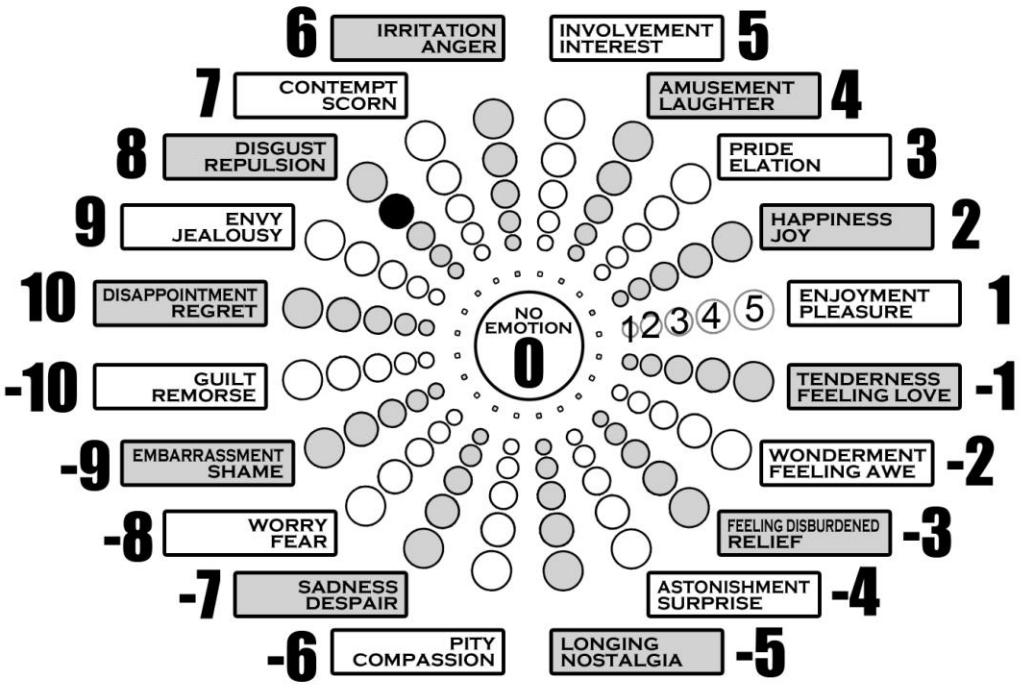


Fig 5, methodology of representing emotions and levels as numbers in GEW. Numbers in thick representing emotions, and thin numbers representing levels. The black dot in this graph represents 8 as emotion, and 4 as level.

In order to compare the GEW results to the Manikin results conceptually, the answers in GEW (which are the filled dots in questionnaires) were transformed into X and Y position in P-D Cartesian coordinates (which means, transformed into pleasure and dominance) based on numbers of emotions (E) and levels (L). Here is the formula of transforming E and L into P and D, based on radian:

$$P = L \cdot \cos[\pi \cdot (|E| - 0.5) / 10]$$
$$D = L \cdot \sin[\pi \cdot (|E| - 0.5) / 10] \cdot SIGN(E)$$

It is kind of a methodology to show the position of those dots in GEW in Cartesian coordinate system,

while pleasure refers to X-axis, and dominance refers to Y-axis in Table 3.

From Table 3 we can observe that the quotients are generally smaller than 1 (H_0: m=1, one-sample two-tailed t-test assuming unequal variances. For PAD, n=9, p=0.042. For P-D GEW, n=6, p=2E-05), thus we can have a basic conclusion that the variance of emotion while watching movie with inconsistent music will be higher than that with consistency.

The complete table of standard deviations can be seen in appendix C, where we can infer the following conclusion:

1. For movie 1 and 3 (representing amusement and sadness), variance of emotion while watching movie clips with inconsistent music will be generally higher than for normal movies (H_0: m=1, one-sample two-tailed t-test assuming unequal variances. For PAD in movie 1, n=9, p=0.042. For P-D GEW in movie 1, n=6, p=2E-05. For PAD in movie 3, n=9, p=0.01. For P-D GEW in movie 3, n=6, p=8E-04).

2. For movie 2 and 4, they still fit previous conclusion, but strangely there are more quotients which are bigger than 1 in those two tables, especially one large quotient in table 2, which is 5.788. We are not sure whether it is just a coincidence or not (H_0: m=1, one-sample two-tailed t-test assuming unequal variances. For PAD in movie 2, n=9, p=0.543. For P-D GEW in movie 2, n=6, p=0.053. For PAD in movie 4, n=9, p=0.963. For P-D GEW in movie 4, n=6, p=0.009).

B-3. Comparing Mean Value of Emotional Variance

In the previous section we estimated and compared the dimensional standard deviation across participants for given movie music combinations; here we would like to compare standard deviation between different dimensions. For that reason the average value about the standard deviations of each dimension across all movie clips (normal as well as conflicting) was calculated. See Table 4.

	Self-Assessment Manikin			**Geneva Emotion Wheel**			
	Pleasure	**Arousal**	**Dominance**	**Emotions**	**Level**	**Pleasure**	**Dominance**
Average	1.22	1.71	1.82	*3.49*	*1.23*	1.18	1.84
Range	1~9	1~9	1~9	*-10~10*	*0~5*	0~5	0~5

Table 4, average value of standard deviations of each dimension, together with the range of the original results.

We should be aware that the value of GEW Emotions and GEW Level are less meaningful in this table. The reason is simple, if one chose 10 as GEW emotion, and one other chose -10, they are next to each other in GEW, but their difference is 20. Therefore we put them here as references only and mainly observe their corresponding pleasure-dominance value.

Compared with three parameters of SAM, we may observe that pleasure is the smallest of the three, then arousal, and dominance is the biggest. Which means the standard deviation of the pleasure dimension is lower than that of arousal and dominance, and the skill of acquainting dominance will vary a lot for different people.

In fact that result even works for GEW, because the average value of standard deviation of X-position, which refers to pleasure, is still lower than that of Y-position, which refers to dominance. In GEW people still make less difference of knowing pleasure, even if they don’t fill in their degree of pleasure and dominance directly.

VI. DISCUSSION

A. Based on the Statistical Result

The result in previous paragraphs may be turned into the following conclusion if described in short:

1. The pleasure dimension is based on the image more than based on the music.
2. Subjects agree more on the pleasure dimension than on arousal or dominance.
3. The arousal and dominance will decrease when people are watching conflicting videos, but if the content of image is already in low dominance, maybe they will increase.
4. Variance of the emotions people have will be higher if people watch conflicting video.

The second and third conclusion may be explained as the fact that people will feel uncertain when facing strange objects, such as a video with music in conflict emotion. When watching a sad movie clip with very funny comedy music, people's mind will always feel unusual and abnormal – it is unusual indeed. Thus the emotion which should appear in people's mind is weakened, and thought of uncertainty like "what is that on earth" appears, therefore the arousal and dominance decrease, and variance increases.

However we are not sure about the change of arousal and dominance in the experimental clips with movie in low dominance. Maybe due to the fact that they are already in low dominance, thus feeling conflicting and "abnormal" will not make it even lower, because it is unusual already – against our usual cognition of video-and-music combination. People will not be immersed in such kind of video, and feel their origin arousal and dominance.

One of the main limitations of this research is the relatively small number of subjects, and the results may be more reliable in a larger scale experiment. When more people participate in the experiment, maybe there will be less difference between P-D and P-D GEW graphs.

While there are enough subjects, perhaps some more principles will be found. There is one assumption that four variants of one movie will be arranged in (similarly) counterclockwise in P-A and P-D graphs, just like four audio variants in four quadrants in P-D coordinates. That assumed principle appears in P-A and P-D graphs of movie 1, but not in that of other movies. We are not sure whether it works for all videos, or just correct in high-pleasure-high-dominance movie clips. More research should be done into this particular topic.

Furthermore, it is also suggested to adopt more than one video in each quadrant in P-D coordinate, otherwise this research may be considered to be related to some certain emotions and movie clips, instead of all the emotions based on the PAD scale.

In the future, in addition to increasing the number of subjects and movie clips, it is recommended to create music consisting of different emotions in the movies, instead of using existing music pieces, because that will make music more suitable for the corresponding movie clip even if their emotion don't match. A number of professional composers should be hired for this research.

B. Based on Effect

The statistic result above may lead to some analysis leading to the effect in cinematic area.

The locations of four different dots in one P-A or P-D graph show that difference may exist when one video is shown with different background music. However generally it may not turn negative movie

scene to positive content, and vice versa. It will not make a tragic killing scene to a comedy unless the image itself is funny. The "pleasure" dimension is not that easy to adjust.

However, it is possible to adjust the arousal and dominance, which means making a sad movie more dynamic or making one lively scene more relaxing. Directors may create some impressive scenes based on this technique. Moreover, judged from the graphs in the appendix, it implies that it is easier to change audiences' emotion with low-dominance movie clips, especially for relaxing movie fragments (one with high pleasure and low dominance), and it will cause greater emotion change than that with high-dominance movie clips. In this case, outstanding effect can be made using the music-and-video-in-emotional-conflict technique.

The arousal dimension is highly variable together with relaxing video. From the graphs we can see that the arousal will be much higher when given high-dominance music than low-dominance music, and for high-dominance music, it can even control the pleasure dimension of video. Based on this point, directors can produce one calm movie fragment with high-dominance music (e.g. exciting, dynamic, etc.), when arousal rises in people's mind, the contrast between video image and people's mind may give them an impressive and unforgettable experience. Since the result is significant, it is highly suggested to do so.

On the other hand, it should be aware of that people tend to have different emotions when given emotional-conflict music, which makes it ambiguous and uncontrollable. Therefore if a director's target is to make audience immersed into one certain emotion, he had better not put music with inconsistency, but put consistent music to enhance the atmosphere.

VII. CONCLUSION

In film industry, video and music with contradicting emotion is sometimes used to achieve some specific emotional effects. We have investigated this by dubbing a film fragment with several types of music with varying emotional content, and letting subjects self-assess the emotional content across key emotional dimensions. Using the self-assessment manikin (SAM) and the Geneva emotion wheel (GEW) tools, we separated people's emotion into three dimensions: pleasure, arousal, and dominance, then we reached following conclusions:

1. For both SAM and GEW measurement, people will take "pleasure" dimension (positive or negative) of their emotions based on the movie clip of the video, regardless of the audio part.
2. The variance of emotions in pleasure dimension is lower than that related to arousal and dominance.
3. While watching the experimental clip with the high dominance, the arousal and dominance of the emotion when people watching the movie clip with inconsistent music will be generally lower than those when watching video with music with consistency. For that situation with movie part in low dominance and positive pleasure, the arousal and dominance of the emotion with inconsistent music will be generally lower than those with music with consistency. The result is still to be determined for movie in low dominance and negative pleasure.
4. Variance of emotion while people watching movie with inconsistent music will be generally higher than that with consistency.

Even with its limited scope and scale, this study serves as an example how dimensional models of emotion might be used in cinematic research or by film producers. Future work may include extending this study with more movies as well as more subjects.

REFERENCES:

[Bigelow et al, 2014] Bigelow, James, and Amy Poremba. "Achilles' Ear? Inferior Human Short-Term and Recognition Memory in the Auditory Modality." *PloS one* 9.2 (2014): e89914.

[Bolivar et al, 1994] Bolivar, Valerie J., Annabel J. Cohen, and John C. Fentress. "Semantic and formal congruency in music and motion pictures: Effects on the interpretation of visual action." *Psychomusicology: A Journal of Research in Music Cognition* 13.1-2 (1994): 28.

[Boltz et al, 1991] Boltz, Marilyn, Matthew Schulkind, and Suzanne Kantra. "Effects of background music on the remembering of filmed events." *Memory & Cognition* 19.6 (1991): 593-606.

[Bradley et al, 1994] Bradley, Margaret M., and Peter J. Lang. "Measuring emotion: the self-assessment manikin and the semantic differential." *Journal of behavior therapy and experimental psychiatry* 25.1 (1994): 49-59.

[Broekens & Brinkman, 2013] Broekens, Joost, and Willem-Paul Brinkman. "AffectButton: A method for reliable and valid affective self-report." *International Journal of Human-Computer Studies* 71.6 (2013): 641-667.

[Bullerjahn et al, 1994] Bullerjahn, Claudia, and Markus Güldenring. "An empirical investigation of effects of film music using qualitative content analysis." *Psychomusicology: A Journal of Research in Music Cognition* 13.1-2 (1994): 99.

[Caicedo et al, 2006] Caicedo, D. Güiza, and M. van Beuzekom. "How do you feel? An assessment of existing tools for the measurement of emotions and their application in consumer product research." *Delft University of Technology, Department of Industrial Design* (2006).

[Cohen & Annabel, 2001] Cohen, Annabel J. "Music as a source of emotion in film." *Music and emotion: Theory and research* (2001): 249-272.

[Desmet, 2005] Desmet, Pieter. "Measuring emotion: Development and application of an instrument to measure emotional responses to products." *Funology*. Springer Netherlands, 2005. 111-123.

[Dingemans et al, 2009] Dingemans, Alexandra E., et al. "The effect of suppressing negative emotions on eating behavior in binge eating disorder." *Appetite* 52.1 (2009): 51-57.

[Friesen et al, 1983] Friesen, Wallace V., and Paul Ekman. "EMFACS-7: Emotional facial action coding system." *Unpublished manuscript, University of California at San Francisco* 2 (1983): 36.

[Gross et al, 1995] Gross, James J., and Robert W. Levenson. "Emotion elicitation using films." Cognition & Emotion 9.1 (1995): 87-108.

[Magliano et al, 1996] Magliano, Joseph P., Katinka Dijkstra, and Rolf A. Zwaan. "Generating predictive inferences while viewing a movie." *Discourse Processes* 22.3 (1996): 199-224.

[Marshall et al, 1988] Marshall, Sandra K., and Annabel J. Cohen. "Effects of musical soundtracks on attitudes toward animated geometric figures." *Music Perception* (1988): 95-112.

[Mehrabian, 1980] Mehrabian, Albert. "Basic dimensions for a general psychological theory: Implications for personality, social, environmental, and developmental studies". Cambridge, MA: Oelgeschlager, Gunn & Hain, 1980.

[Mehrabian, 1996] Mehrabian, Albert. "Pleasure-arousal-dominance: A general framework for

describing and measuring individual differences in temperament." *Current Psychology* 14.4 (1996): 261-292.

[Meredith & Barry, 1986] Meredith, M. Alex, and Barry E. Stein. "Visual, auditory, and somatosensory convergence on cells in superior colliculus results in multisensory integration." *J Neurophysiol* 56.3 (1986): 640-662.

[Munhall et al, 1996] Munhall, Kevin G., et al. "Temporal constraints on the McGurk effect." *Perception & Psychophysics* 58.3 (1996): 351-362.

[Russell, 1980] Russell, James A. "A circumplex model of affect." *Journal of personality and social psychology* 39.6 (1980): 1161.

[Sacharin et al, 2012] Sacharin, Vera, Katja Schlegel, and K. R. Scherer. "Geneva emotion wheel rating study." Center for Person, Kommunikation, Aalborg University, NCCR Affective Sciences. Aalborg University, Aalborg (2012).

[Scherer, 2005] Scherer, Klaus R. "What are emotions? And how can they be measured?." *Social science information* 44.4 (2005): 695-729.

[Smith et al, 2013] Smith, Elliot, et al. "Seeing is believing: neural representations of visual stimuli in human auditory cortex correlate with illusory auditory perceptions." *PloS one* 8.9 (2013): e73148.

[Stein & Terrence, 2008] Stein, Barry E., and Terrence R. Stanford. "Multisensory integration: current issues from the perspective of the single neuron." *Nature Reviews Neuroscience* 9.4 (2008): 255-266.

[Stroop & John, 1935] Stroop, John Ridley (1935). "Studies of interference in serial verbal reactions". Journal of Experimental Psychology 18 (6): 643–662. doi:10.1037/h0054651.

[Thompson et al, 1994] Thompson, William Forde, Frank A. Russo, and Don Sinclair. "Effects of underscoring on the perception of closure in filmed events." *Psychomusicology: A Journal of Research in Music Cognition* 13.1-2 (1994): 9.

[Watson & Auke, 1985] Watson, David, and Auke Tellegen. "Toward a consensual structure of mood." *Psychological bulletin* 98.2 (1985): 219.

[Wiki: Stroop Effect] Wikipedia “Stroop Effect”: http://en.wikipedia.org/wiki/Stroop_effect. Retrieved 2015-01-20.

APPENDIX

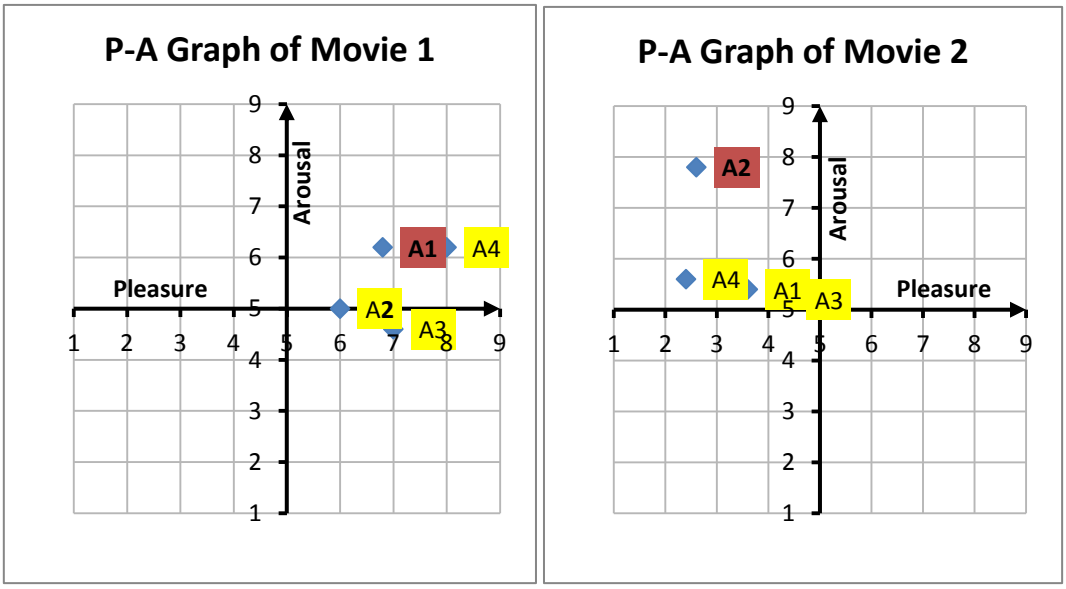
A, Music used in the videos:

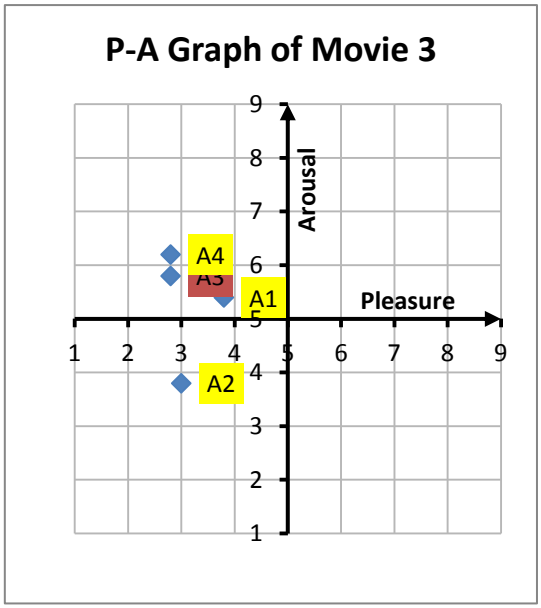
Video Category	Audio Category	Sources of Music Pieces
“When Harry Met Sally” (Amusement)	Amusement	Soundtrack from “Cocoon The Return”
	Angry	Soundtrack from “Mission Impossible 3”
	Sadness	Soundtrack from “Furuhata Ninzaburo”
	Relief	Soundtrack from “Hong he gu”
“My Bodyguard” (Angry)	Amusement	“Would You” by Touch and Go
		Soundtrack from “Furuhata Ninzaburo”

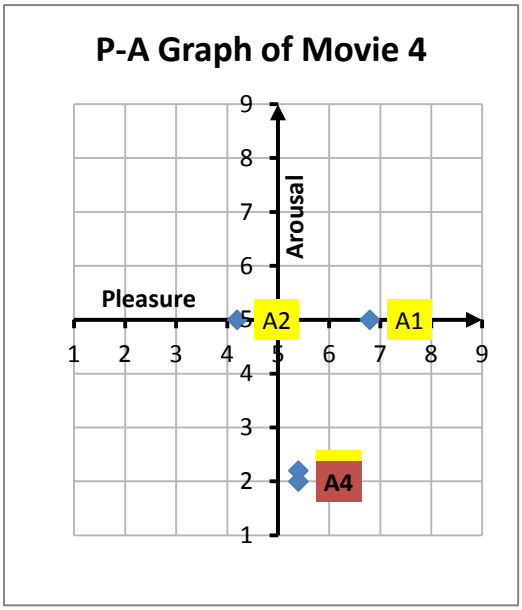
		Unpublished soundtrack from “Frozen”
	Angry	Soundtrack from “Mission Impossible 3”
	Sadness	Soundtrack from “East of Eden”
		“Boadicea” by Enya
	Relief	“Lead a Normal Life” by Peter Gabriel
		Soundtrack from “Prince of Tides”
“The Champ” (Sadness)	Amusement	Soundtrack from “Hustle”
		Soundtrack from “The Adventures of Tintin”
	Angry	Soundtrack from “Mission Impossible 3”
	Sadness	Soundtrack from “Cast Away”
	Relief	Soundtrack from “The Legend of Bagger Vance”
Beach Scene (Relief)	Amusement	Soundtrack from “Furuhata Ninzaburo”
	Angry	Soundtrack from “Mission Impossible 3”
	Sadness	Soundtrack from “X Flies”
	Relief	Soundtrack from “The Legend of Bagger Vance”

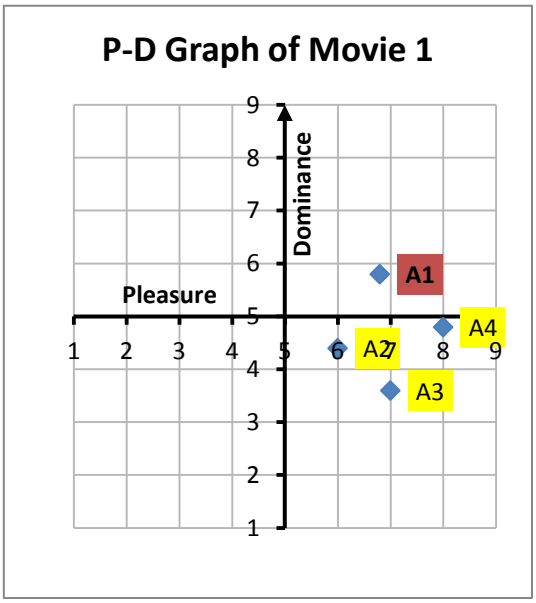
B, Graphs about Average Value of Four Movie Clips

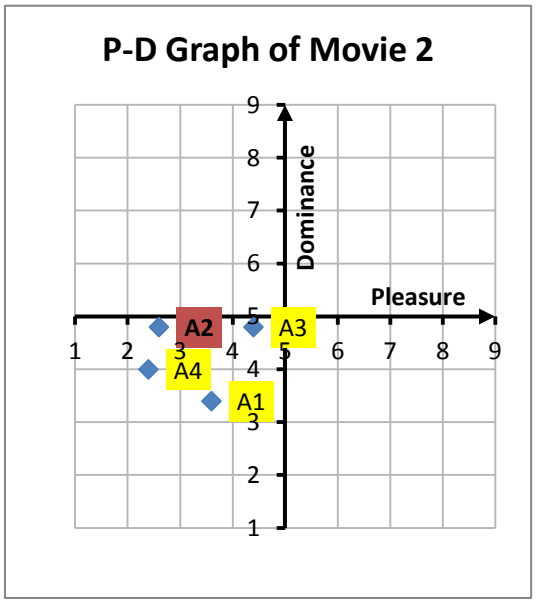
All the graphs are categorized based on different coordinates. All shown mean values are based on n=5.

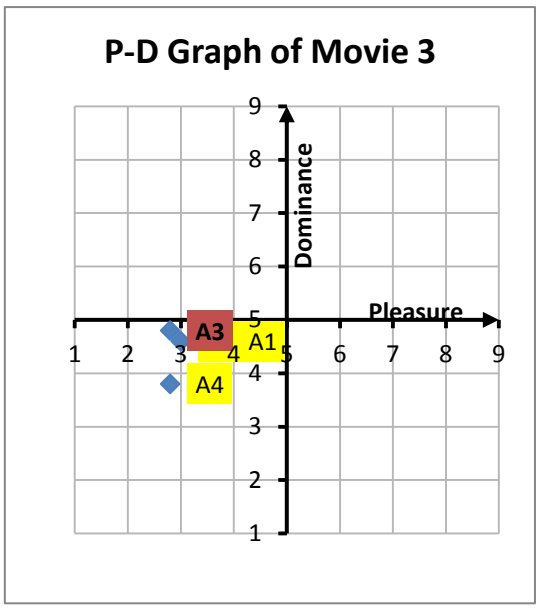


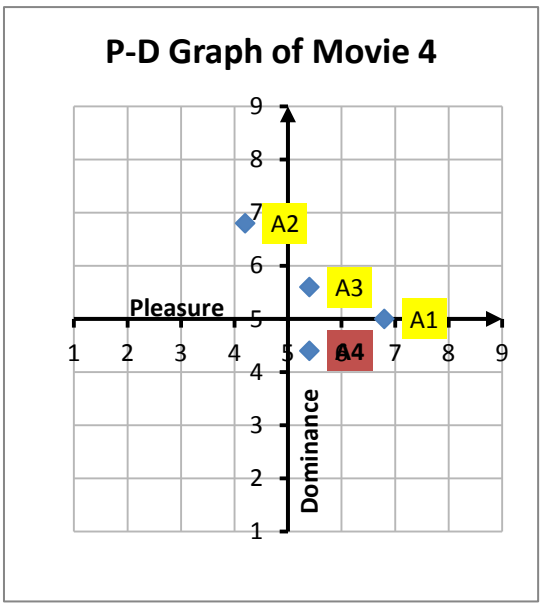

P-A Graph of Movie 3
Arousal
Pleasure
A4
A3
A1
A2

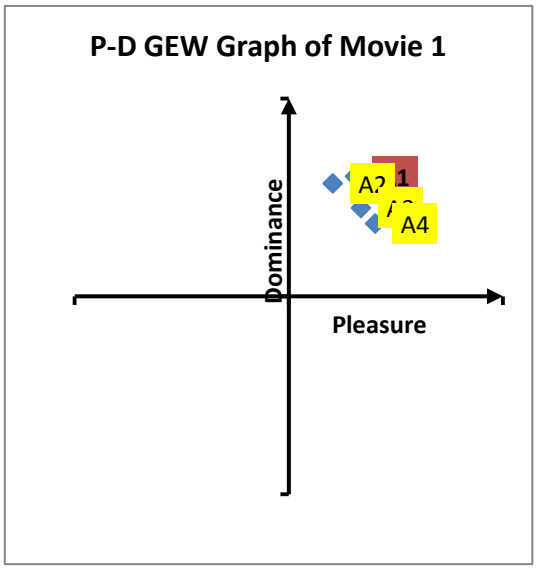

P-A Graph of Movie 4
Arousal
Pleasure
A2
A1
A4

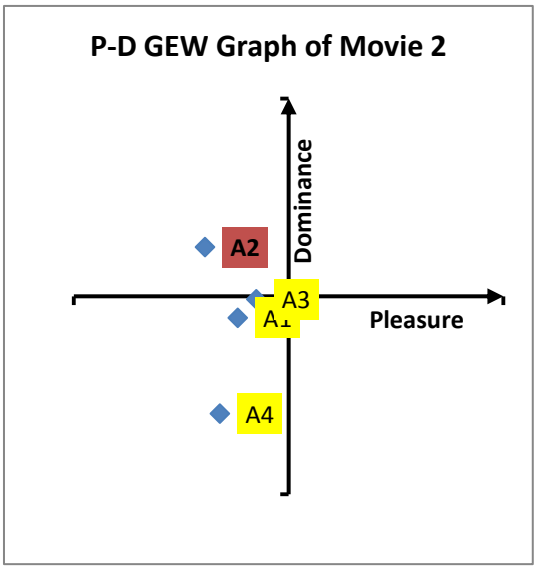

P-D Graph of Movie 1
Dominance
Pleasure
A1
A4
A2
A3

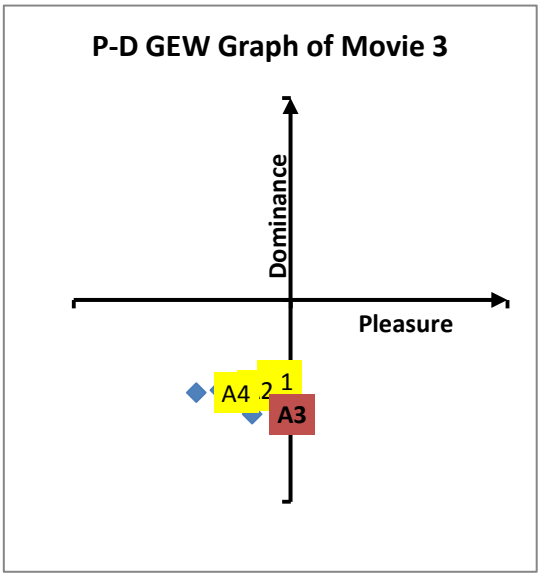

P-D Graph of Movie 2
Dominance
Pleasure
A2
A3
A4
A1

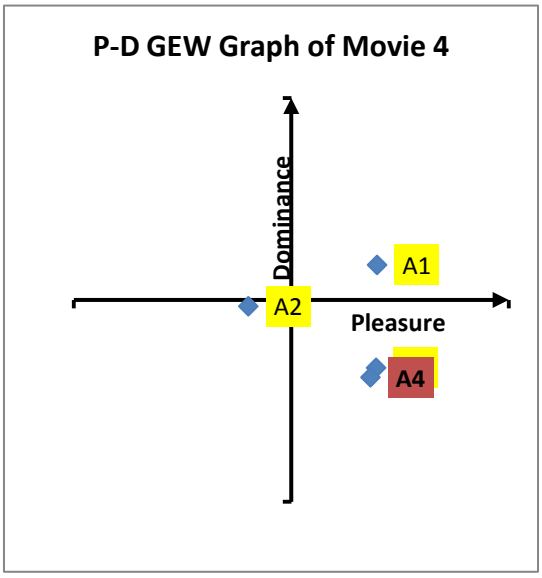

P-D Graph of Movie 3
Dominance
Pleasure
A3
A1
A4


P-D Graph of Movie 4
Pleasure
A2
A3
A1
A4
Dominance


P-D GEW Graph of Movie 1
Dominance
A4
Pleasure


P-D GEW Graph of Movie 2
Dominance
A2
A3
Pleasure
A4


P-D GEW Graph of Movie 3
Dominance
Pleasure
A3


P-D GEW Graph of Movie 4
Dominance
A1
A2
Pleasure
A4

C, Tables about Quotients between Standard Deviations in Four Movies

MOVIE 1	**Self-Assessment Manikin**			**Geneva Emotion Wheel**			
	Pleasure	**Arousal**	**Dominance**	**Emotions**	**Level**	**Pleasure**	**Dominance**
Music 1							
Music 2	0.59	0.65	0.78	*0.00*	*0.30*	0.18	0.23
Music 3	1.18	0.52	0.92	*0.00*	*0.77*	0.32	0.42
Music 4	1.18	0.44	0.78	*0.00*	*0.33*	0.34	0.15

MOVIE 2	**Self-Assessment Manikin**			**Geneva Emotion Wheel**			
	Pleasure	**Arousal**	**Dominance**	**Emotions**	**Level**	**Pleasure**	**Dominance**
Music 1	0.53	0.86	1.12	*1.10*	*0.77*	0.99	1.09
Music 2							
Music 3	1.00	0.57	5.79	*1.36*	*0.53*	1.18	1.39
Music 4	0.59	0.63	1.10	*1.21*	*1.83*	1.17	1.06

MOVIE 3	**Self-Assessment Manikin**			**Geneva Emotion Wheel**			
	Pleasure	**Arousal**	**Dominance**	**Emotions**	**Level**	**Pleasure**	**Dominance**
Music 1	0.41	0.90	0.98	*0.08*	*0.94*	0.67	0.58
Music 2	0.45	0.53	0.78	*0.15*	*0.49*	0.39	0.46
Music 3							
Music 4	0.64	1.00	0.93	*0.42*	*0.91*	0.62	0.74

MOVIE 4	**Self-Assessment Manikin**			**Geneva Emotion Wheel**			
	Pleasure	**Arousal**	**Dominance**	**Emotions**	**Level**	**Pleasure**	**Dominance**
Music 1	1.16	0.58	1.98	*0.61*	*2.10*	0.96	0.57
Music 2	0.74	0.24	1.91	*0.30*	*0.62*	0.48	0.54
Music 3	0.69	0.43	1.36	*0.56*	*0.94*	0.79	0.74
Music 4							

D, Questionnaire Used in the Experiment

(We only show the questionnaire of the first experimental clip here, which contains tutorial of comprehending SAM and GEW. The remaining three are the same)

1-1, Read 1-2-1 and 1-2-2 first to know what you should do during the experiment, then watch the first video.

1-2, Describe your emotion after watching the video using following method:

1-2-1, Rate your emotion using the Self-Assessment Manikin.

You will describe your emotion according to the three variables *pleasure*, *arousal* and *dominance*:

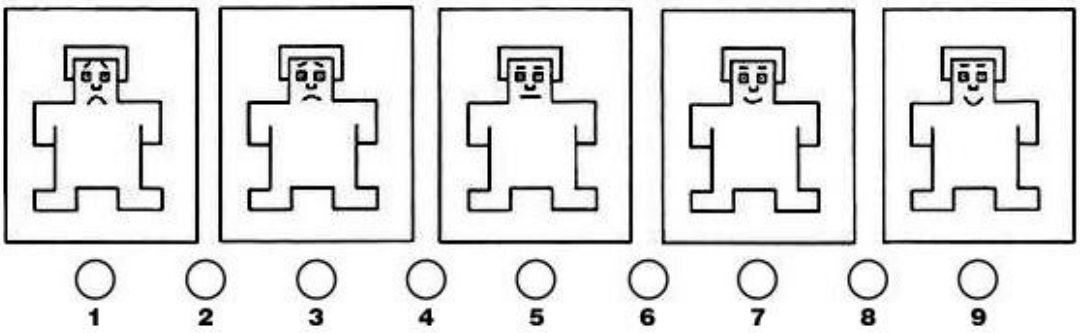
- **Pleasure** refers to the degree of **positiveness or negativeness** of an emotion.
- **Arousal** refers to the **intensity** of an emotion.
- **Dominance** refers to the **level of being in control**.

(1) Select a number underneath the pictures that reflects the level of pleasure in your emotion.

The question is: How much negativeness or positiveness do you experience, during the emotion you have?

Examples, from left to right:

- **Negative pleasure:** unhappy, annoyed, unsatisfied, melancholic, despairing, bored.
- **Positive pleasure:** happy, pleased, satisfied, contented, hopeful, relaxed.

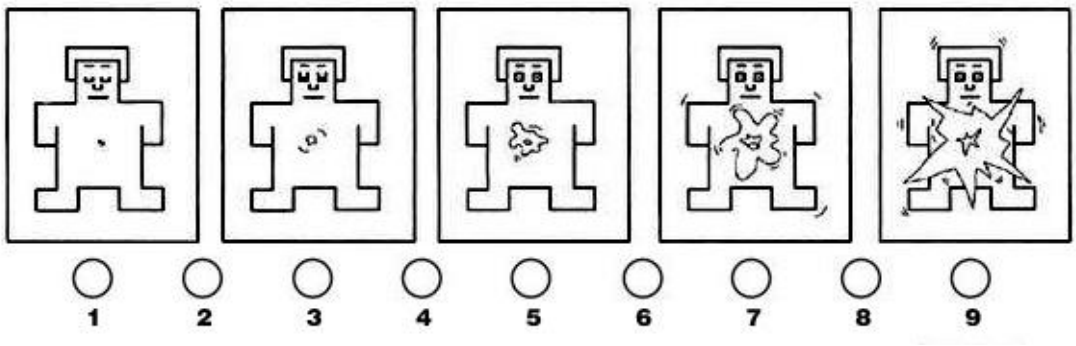


(2) Select a number underneath the pictures that reflects the level of arousal in your emotion.

The question is: What is the intensity level of your emotion?

Examples, from left to right:

- **Low arousal:** relaxed, calm, sluggish, dull, sleepy, unaroused.
- **High arousal:** stimulated, excited, frenzied, jittery, wideawake, aroused.

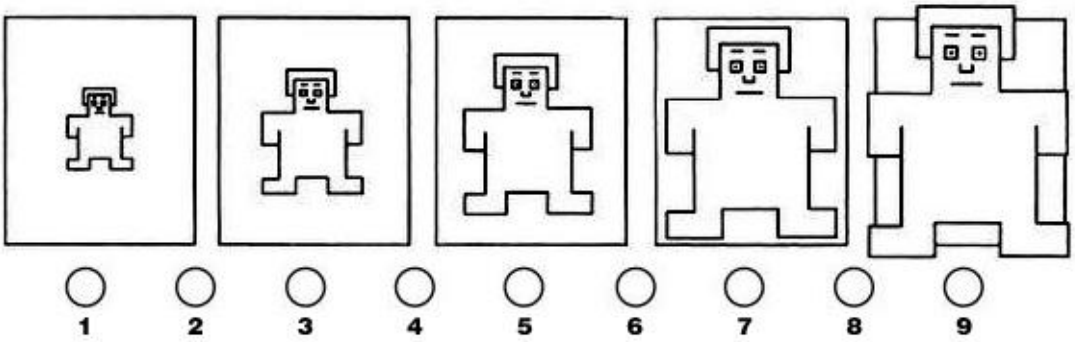


(3) Select a number underneath the pictures that reflects the level of dominance in your emotion.

The question is: How low or high is the feeling of being in control, during the time you had that emotion?

Examples, from left to right:

- **Low dominance:** controlled, influenced, cared for, awed, submissive, guided.
- **High dominance:** controlling, influential, in control, dominant, autonomous.



1-2-2, Rate your emotion using Geneva Emotion Wheel.

Please rate the intensity of the emotion in the wheel which best describes what you felt. If you did not feel any emotion at all, please check the circle in the center of the wheel (labeled "No Emotion").