Title	A movie trailer testing method combining traditional questionnaire and psychophysiological parameters
Sub Title	
Author	Shi, Shuoping Kunze, Kai
Publisher	慶應義塾大学大学院メディアデザイン研究科
Publication year	2022
Jtitle	
JaLC DOI	
Abstract	
Notes	修士学位論文. 2022年度メディアデザイン学 第935号
Genre	Thesis or Dissertation
URL	https://koara.lib.keio.ac.jp/xoonips/modules/xoonips/detail.php?koara_id=KO40001001-00002022-0935

Master's Thesis
Academic Year 2022

A Movie Trailer Testing Method Combining Traditional Questionnaire and Psychophysiological Parameters

Keio University
Graduate School of Media Design

Shuoping Shi

A Master's Thesis
submitted to Keio University Graduate School of Media Design
in partial fulfillment of the requirements for the degree of
Master of Media Design

Shuoping Shi

Master's Thesis Advisory Committee:

Professor Kai Kunze	(Main Research Supervisor)
Professor Akira Kato	(Sub Research Supervisor)

Master's Thesis Review Committee:

Professor Kai Kunze	(Chair)
Professor Akira Kato	(Co-Reviewer)
Project Associate Professor Tatsuya Saito	(Co-Reviewer)

Abstract of Master's Thesis of Academic Year 2022

A Movie Trailer Testing Method Combining Traditional Questionnaire and Psychophysiological Parameters

Category: Science / Engineering

Summary

This thesis attempts to propose a new method of surveying audience preferences and behavior in the film industry, combining conventional market research methods such as questionnaire or interviews and the measurement of physiological parameters of the audience. To make up for the deficiencies brought about in traditional film market research and to reveal more about the behavioral and emotional patterns of the audience.

After performing the experiment with 4 pretests and 25 subjects and analyzing their data it could be observed that this new method enables the identification of the type of task from the psychophysiological data of the participants. Also there was a correlation between the physiological data of the experimental participants and their preference for the movie trailers they watched. The audience is more likely to give high ratings when they have a lesser number of EDA peaks and when their HRV shows an increasing trend. This findings provide us with richer information than traditional testing method and make it possible to improve film production or marketing strategies at an effective yet lower cost way.

Keywords:

wearable sensing, physiological signal, movie trailer testing

Keio University Graduate School of Media Design

Shuoping Shi

Contents

List of Figures

List of Tables

Acknowledgements

First of all, I would like to express my gratitude to Prof. Kai for the guidance over the years, although I always wavered in my choice of topic, I was able to receive a lot of valuable suggestions each time. Thanks to Kato and Saito sensei for their advice on thesis writing.

I am also especially grateful to Dingding senpai for her great support not only in academics but also in life. Although she was very busy, she helped me many times when I was in emotional problems. I have learned a lot from working together. Thanks to her help I was able to get my life back on track.

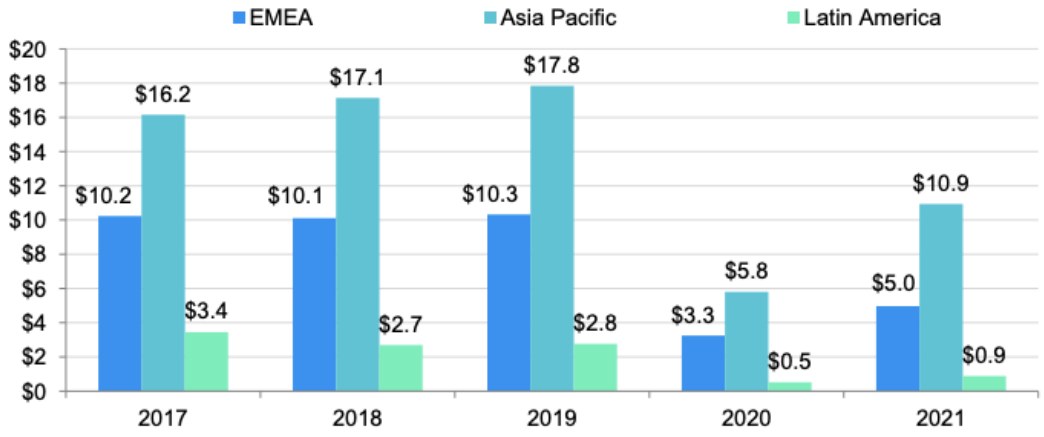
Then I would like to thank Xiongqi for helping me conduct the experiment together and recruiting many participants and George senpai for his long-term technical support. Thanks to all the people from KMD or Keio who participated in the experiment, without their help this study would not have been completed.

Finally, I would like to thank my parents for their continued understanding and encouragement. Although I haven't been home for a long time because of COVID-19, I still received concerns from far away.

Chapter 1
Introduction

1.1. Background and Motivation

The global movie market suffered a huge impact under the impact of COVID19 and still has not recovered to its previous performance level. The film is still one of the main forms of mass entertainment consumption, but the development of mobile Internet and the emergence of new forms of home entertainment in recent years have slowed down the growth of its market size.



(Source: Comscore, Omdia, MPA sources)

Figure 1.1 International Box Office Market by Region – All Films (USD Billions)

With the local market becoming saturated, how to expand externally and export influence has become a common concern in many regions, but it is not easy. Take the Japanese film market as an example. Imported films are in a poor situation in the Japanese market. According to a report[1] from the Motion Picture Producers

1 Statistics of Film Industry in Japan (2021) `http://www.eiren.org/toukei/index.html`

Association of Japan, 959 films were released in Japan in 2021, 51 per cent of which were Japanese films. However, imported films accounted for only 20.7% of the market share. This trend is not only observed in Japan, but in the last two years, China's domestic films have accounted for more than 80 per cent of all box office revenues. China's film market also faces the problem of imported film quotas.

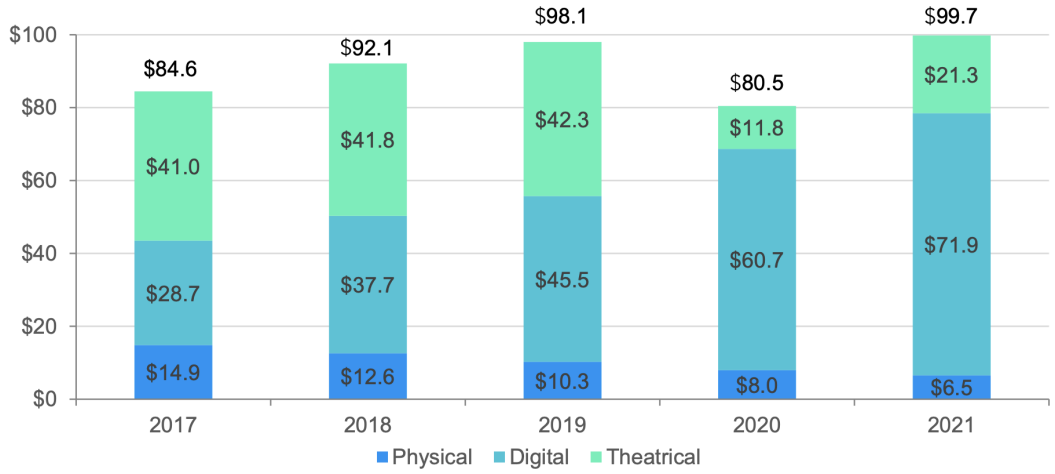
Not only imported films, regardless of the impact on the film industry in the last two years due to the COVID, the Japanese film industry has not made great significant progress in terms of revenue and movie-going since 2000 with the background that the total number of releases has grown and the average ticket price has become higher. China overtaken Japan as the second largest film market in the world in 2012.

Moreover, the Japanese do not seem to be enthusiastic about new film technologies such as 3D. NTT Communications' Survey released in June 2019 shows that 87.6% of respondents who have had cinema viewing in the last year have had a 2D viewing experience, while only 13% have seen 3D movies and 4.3% have seen the IMAX 3D version. Meanwhile, the share of the 3D movie box office in the total box office reached a peak in 2009, after which the total share of the 3D movie box office in the U.S. market began to trend downward, even as the number of 3D movies as a whole was on the rise. It seems that movie technology like 3D has entered a difficult period, as audiences have lost their enthusiasm for it a decade after the release of Avatar. 3D movies are technically stagnant; not only have they not yet achieved high-definition picture quality, but they are still seen as an inferior "money grabbing" deal. The Japanese industry takes into account the input-output ratio, which in turn limits the diversification of film themes.

In the meantime, however, the development of the mobile internet has created a lot of new types of entertainment over the years. Online streaming platforms such as Netflix and Disney plus have also led to a dramatic change in the audience's movie-watching needs and habits before the COVID19 (Figure1.2) [2].

As a movie lover, I still believe there is no substitute for the experience in the theatre. I also hope to see more great productions from different cultures, which

2 2021 THEME Report: `https://www.mpa-apac.org/wp-content/uploads/2022/03/MPA-2021-THEME-Report-FINAL.pdf`



(Source: Omdia (International), Comscore – Box Office Essentials (Theatrical), Digital Entertainment Group)

Figure 1.2 Global Theatrical, Home/Mobile Entertainment and Pay TV Market (USD Billions)

are windows to the world and can build bridges of communication. And given the general environment of the film industry as described above, what can be done to improve the movie-going experience, bring good quality but small budget films to audiences and re-energize the film market is what I would like to learn.

Also, as a person who comes from a sociological background with a basic understanding of market research methods, I increasingly feel that the information obtained traditionally is limited in understanding people's choices and behaviours. Many choices are made unconsciously or subconsciously, and viewers may not be able to accurately communicate the deeper reasons for what they are feeling. Thus, we need to leverage new technologies to gain a greater insight into people.

1.2. Objectives

The main purpose of this research is to:

(1) To explore a new method that combines audience physiological data with traditional questionnaires for movie trailer testing

(2) To evaluate whether this new approach can acquire valid data to analyze audience behavior patterns

(3) To uncover whether there are any correlations between audience expectations or preferences and the physiological signals

1.3. Thesis Overview

- Chapter 1 introduces the current situation of the global film industry, especially in Japan. As well as the motivation and the general orientation of this study.
- Chapter 2 provides a summary of traditional film market research and its limitations. The literature review of methods for emotion analysis and existing techniques of physiological data collection is presented.
- Chapter 3 describes in detail the experimental materials and process design of this study.
- Chapter 4 analyzes the data collected from the experiment, including the questionnaire results and physiological parameters.
- Chapter 5 summarizes the study, discusses the limitations as well as the work that will proceed in the future.

.

Chapter 2
Literature Reviews

In this chapter, the market research methods typically used in the film industry will be reviewed, as well as the methods used for emotion recognition in the field of human-computer interaction.

2.1. Traditional Methods of Experience Testing in Film Industry

The production of a film requires huge time and labor costs in all stages, and outside of the production of the film content itself, the marketing campaign is crucial to whether it will pay for itself and how much revenue it will generate. The traditional methods of market research in the film industry are these, audience test screenings, advertising testing and audience tracking and exit polling.

Audience Test Screenings

Audience test screenings refer to previews that take place before a general release in order to gauge audience reaction. The audience is usually asked to fill out a questionnaire or feedback after the screening. Based on the feedback, the marketing concept is usually fine-tuned, or changes are made to the editing, special effects ending.

Movie trailer testing

Movie trailer testing measures viewers' reactions to each by providing them with multiple versions of the trailer for the purpose of determining better audience engagement. It is like the qualitative studies in other field, is uncertain. It usually

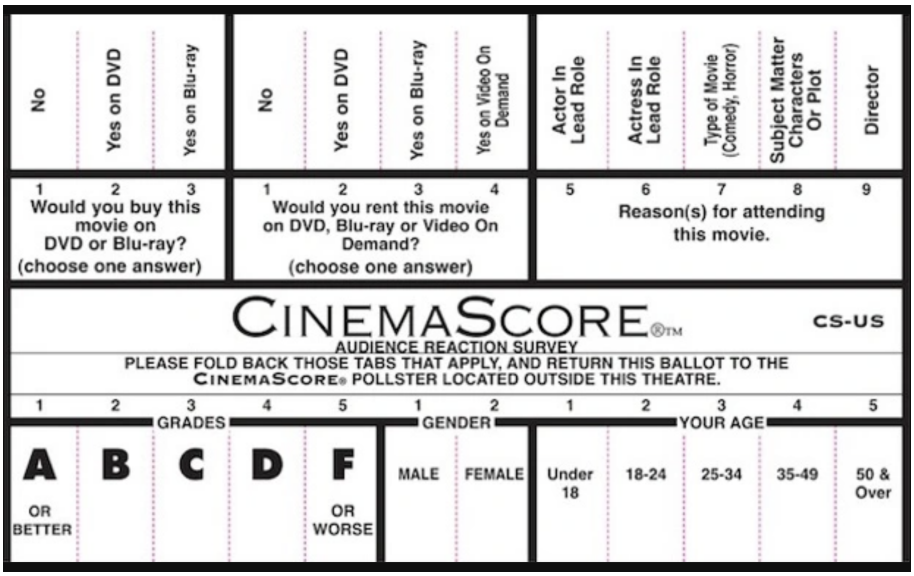


Figure 2.1 CINEMASCORE: Audience Reaction Survey

takes a small sample size to predict the overall audience preference tendency. Thus the accuracy of its results relies on the size of the sample. But again, a larger sample means higher costs.

Film is a complex work that contains many elements such as plot, soundtrack, characters, and special effects. And the combination of these elements will produce different outcomes. Traditional film market research methods get limited information, and audiences could only give a general impression and score as feedback. Of course, many people would make valid suggestions, but sometimes it does not work. An example is the 1960s Batman television series, where Adam West mentioned in his book that a hundred audiences were recruited to watch the pilot episode. But the episode received "the worst score in the history of pilot testing". After that they made adjustments on a laugh track and narration, and redid the test screening. The result was still the same. Eventually it was decided to add "huge new special effects gags that would look great in promos. [1]"

2.2. Emotion Analysis Based on Physiological Signals

Emotions are transmitted in all modes of human communication, and in most cases all these modes can be measured with different devices and sensors. Past studies have explored methods to conduct emotion analysis without direct contact, based on parameters such as facial expression [2], body gesture or speech features. This allows common devices with cameras such as cell phones or computers, as well as with microphones, to be used for the collection of relevant data. These physical parameters can be easily accessed, but at the same time they can also be easily misinterpreted. In different social or cultural contexts, people sometimes hide their emotions or pretend to be in other states, which makes relying only on these visible parameters may cause failure in the results of emotion classification.

Those emotional states that cannot be observed by the naked eye can be classified by measuring internal parameters. This is because human emotions affect the Autonomous Nervous System (ANS), which regulates multiple body parameters [3]. There are already many wearable devices that can be used to capture physiological data in an unobtrusive way, such as eda, hrv, temperature or respiration patterns. Some studies have also used multi-modal approaches that combine physical and physiological signals to recognize emotions. Multimodal measurements usually offer a higher degree of accuracy. Because emotions originate in the brain, many studies have also been conducted using Electroencephalography (EEG) method. The specific method chosen depends on the field of application and the characteristics of the task, and different methods have their own usability and limitation.

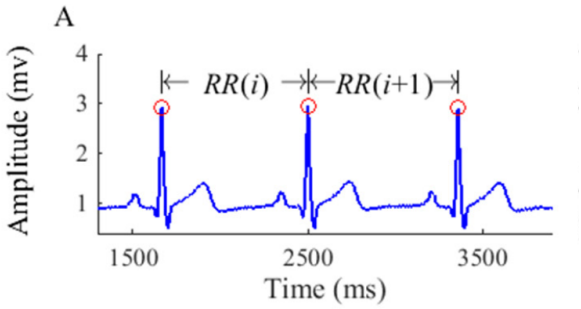
2.2.1 ElectroDermal Activity (EDA)

ElectroDermal Activity (EDA) measures the electrical conductivity of the skin by applying two electrodes to the finger to measure the skin's resistance. This data is obtained by placing a low, undetectable, and constant voltage on the skin and then measuring the change in skin conductance [4]. The acquisition of this data is performed using sensors covering the fingertips, and these sensors are usually fabricated into gloves or finger sleeves for applications to dynamic tasks.

EDA is associated with the regulation of our body temperature, but many studies have also shown that it can be used to measure emotional arousal [5]. The skin electrical conductivity decreases in the relaxed state and increases in the tense state [6]. Features include range, amplitude, rise duration, initial value, time and frequency [7]. Positive or negative stimuli can lead to an increase in arousal and result in an increase in skin conductance. EDA represents not the genre of emotion, but the intensity of the emotion.

The initial studies using skin conductivity on audience watching video were done by Kaiser and Roessler [8] while the survey by Lisetti et. al. [9] contains a detailed review of previous studies in this area. A more frequent method of research using EDA signals is the classification and evaluation of movies through the physiological data of the audience while watching the video [10].

2.2.2 Heart Rate Variability (HRV)



Heart Rate Variability describes the changes in time intervals between each consecutive pair of heartbeats [11]. It is usually represented by the variation in RR intervals (the intervals composed of two adjacent R wave peaks of the cardiac cycle) collected from electrocardiogram (ECG) data [12]. Heart rate variability reflects the transient output of the central autonomic network and the ability of individuals to regulate emotional expression through sympathetic and parasympathetic activity [13]. Therefore, many studies have suggested that HRV can serve as a useful tool for evaluating emotional reactions.

2.3. Summary

As mentioned in 2.1, traditional film market research methods have various problems that are common to both qualitative and quantitative analysis. Specific modifications or marketing strategies still need to rely on practitioners to repeatedly speculate and polish. There are also tests that use focus group screenings to gather more detailed audience responses, including interview and involuntary physical reactions, but these are either very expensive or strictly limited in location and therefore less commonly conducted.

There are some studies to predict the box office success of movies based on different characteristics of the movie, such as budget, movie crew, critics and audience ratings. And techniques such as deep neural networks are used to predict the likelihood of a movie's success [14]. This type of prediction based on film features is broad, and it can somewhat help practitioners understand the placement or status of a certain type of film in the market, but has little effect on a specific film. It does not provide more detailed information to help improve film production or promotional releases. Therefore a demand for a new way of film market research is necessary.

As for how to get the data needed for the study, there has been a lot of research in the field of human-computer interaction for emotion recognition. Observations, even from electronic devices, can have a psychological impact on the subjects, and this study aims to propose a new method that can be applied in a realistic movie viewing scenario, which does not disturb the audience too much and records the data in an unobtrusive way.

In the theater setting because of the darker environment, there may be face coverings such as masks. This is coupled with the fact that different cultures have different interpretations of physical expression. And it cannot be presented in real time, for example, there is no way to know how the audience feels at a specific point in time. This all causes facial or pose recognition to not be an ideal option.

Using EEG enables the acquisition of more information, but its setup process is time-consuming and its noise-sensitive nature makes it difficult to leave the laboratory environment for application in real viewing activities. This is why this study ultimately chose EDA and HRV as the primary psychophysiological data for analyzing audience preferences and behaviors. These data are available in

everyday life and are not restricted to laboratory settings. And there exist many devices for data collection that can be used directly.

Chapter 3
Experiment Design

This chapter will describe in detail the procedure of the trailer testing method proposed in this study and the experimental design of how to obtain the required questionnaire and physiological data.

A brief description of the experimental design will be presented first.

3.1. Ideation

This study is exploratory research and therefore does not initially set very specific research hypotheses. This study believes that many more fundamental issues need to be clarified before further research can be conducted. For movies (or movie trailers), different people may evaluate them in various ways depending on their interests and thus may present different physiological reactions. This research hopes to first discover whether there is a correlation between audience preferences and their physiological data in general without focusing on a certain genre of film in the first place.

The experiment's main purpose was to have participants rate the trailers after they watched them and to collect physiological data while watching the trailers. But what kind of trailer should be played to them? This was the first issue that needed to be settled.

Because different people have different interests, some like science fiction movies and others prefer drama movies, a database covering enough movies of various genres is needed first. It should include not only movies with good reviews, but also movies with average production or average reviews, so as to obtain more comprehensive data.

Once the database is available, how do we select the list of movie trailers for the

participants to watch? In order to more closely resemble the reality of a new movie release, we wanted the trailers played to be ones that the participants had not seen. This is because if they had already seen the movie, their anticipation, viewing experience and evaluation would be influenced by their knowing the content, as well as their physiological reaction. It is for this reason that this experiment did not choose to show the same movie trailer to all subjects.

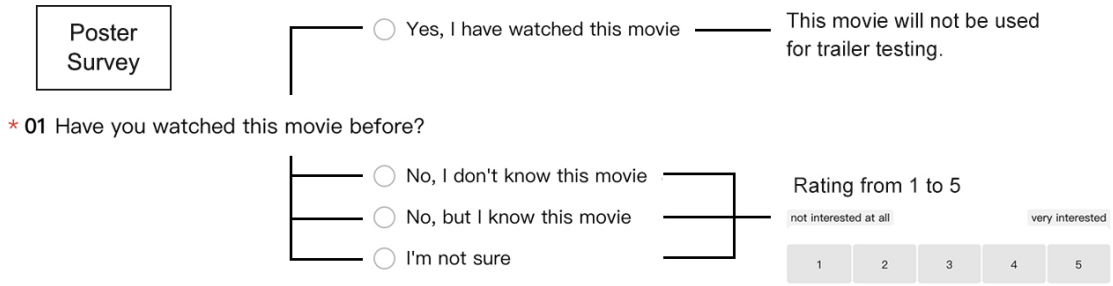


Figure 3.1 Logic Diagram of the Poster Survey

Before showing the trailer to the participants they will be asked to fill out the first survey, asking them to tell if they have seen the movie based on the title and poster. The reason for this approach is that, as in real life, we are usually initially attracted by the publicity of words and images. Through the posters and titles we can identify the general genre of the movie, its director, screenwriter and starring actors, and based on this information we can then make a decision whether to watch or learn more about it.

If they had not watched it, they were then asked if they were interested in watching it. The 5-point Likert scale [15] (1. not interested at all, 2. not interested, 3. neutral, 4. interested, 5. very interested) was used in this survey. All movies will be divided into five groups based on their ratings. This experiment will define the group that gives five points as the Like Group, the group of movies that give three points as Neutral Group, and the group that gives one point as Dislike Group (See the left half of Figure 3.2). This grouping will be used for subsequent playlist generation and data analysis.

To avoid the effect of time on the results of the experiment (e.g., subjects may

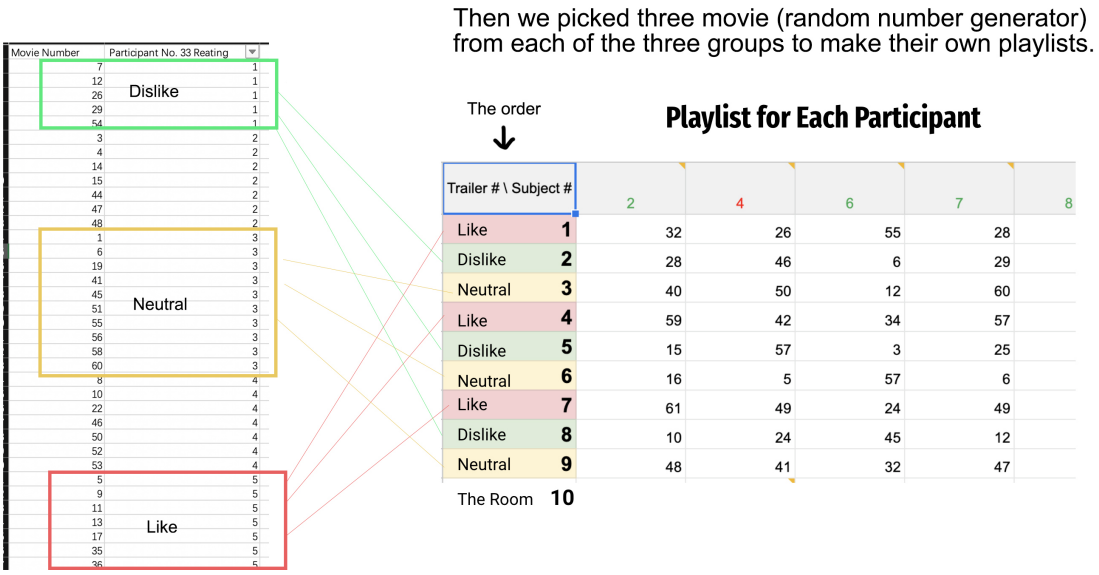


Figure 3.2 Logic of Trailer Testing Playlist Generation

feel impatient or increasingly calm in the latter part of the experiment), different types (like, dislike, neutral) of trailers will be played in a crossover. There are 3 sets of trailer testing. A total of 9 movie trailers were played. The order of each set is the same, all are like - dislike - neutral (See the right half of Figure 3.2). To avoid the bias of subjective selection by the researcher, the selection of the trailer will be done by a random number generator.

After completing the trailer selection, then the trailer testing is performed. After participants watched each trailer, they were asked to rate again based on whether they had the intention to watch the full film on a scale from 1 to 7. So in this experiment there are two ratings in total, one for the poster (1 to 5) and one for the trailer (1 to 7). To avoid misunderstanding, I will refer to them as the Poster Rating and Trailer Rating below. At the same time, their physiological data during the watching process would be recorded in real time.

The above is the main part of the experimental design, more details will be described in the subsection of the experimental procedure.

In order to improve the experimental procedure and control the duration of the experiment as well as the reliability of data collection, 4 pretests were performed before the formal recruitment for the experiment.

The following introduces the materials involved in this experiment.

3.2. Materials

- Fu wristband (Figure 3.3): BVP sensor, EDA electrodes, and hand motion sensor in the main board for obtaining physiological data

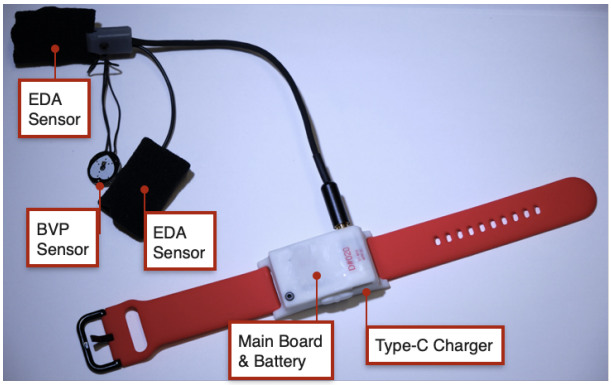


Figure 3.3 Prototype of Fu Wristband

The BVP sensor and EDA electrodes are worn on two fingers through a soft elastic sleeve, and the accelerometer that records movement is on the main board, worn on the subject's wrist like a watch. Apart from the necessary wires for connection, few other factors would interfere with the watching.

EDA is monitored using a Wheatstone bridge with 2 electrodes to measure the skin conductance. BVP is sampled at 50 Hz from a photoplethysmograph. And the accelerometer is located on the main board. Finger sensors are connected to the main board with a cable and a 3.5mm 4-pin audio connector. The main unit is based on the ESP32 module and streams the data through WiFi to the server [16]. This equipment has an accompanying software that allows the raw data to be recorded directly and simultaneously.

- Tablet (iPad Pro 10.5-inch) for the subjects to answer the surveys
- 27-inch Dell monitor screen for movie trailer playback

- Webcam for recording real-time facial responses (such as laughing, intense or other facial expressions)
- Online random pick generator: In order to eliminate the bias brought about by the subjective choices of the researchers, Random Picker[1] was used for generating the playlist and sequence
- Questionnaires: 4 online surveys
  - SD Movie Poster Rating Survey (Appendix B)
  - Movie Trailer Rating Survey (Appendix D)
  - General Questions (Appendix D): Basic demographic information and movie watching habits
  - Follow-up Questions (Appendix E): Self-evaluation of English proficiency
- Movie list: 61 movies (Figure 3.4)

  The 61 movies were selected through a pilot study from a total list of 3,878 films, mainly after 2010 onwards when trailers in HD or above 720P resolution were released online to begin with.

  First of all, the movie was selected based on on the ratings they are received on TOMATOMETER[2] with both critics and audience reviews, IMDb[3], and Douban[4](Chinese IMDb but wuth more users, one of the most used movie review sites in China) Other indicators were also taken into account, such as the budget of indie films versus big label movie companies, major distribution regions (European only, Pacific area, India, US), etc.

  To try to simulate a real movie market research scenario, we tried to detect the subjects' first reaction when they saw the movie clips. Therefore, the

1 https://www.randomlists.com/random-picker

2 https://www.rottentomatoes.com

3 https://www.imdb.com

4 https://www.douban.com

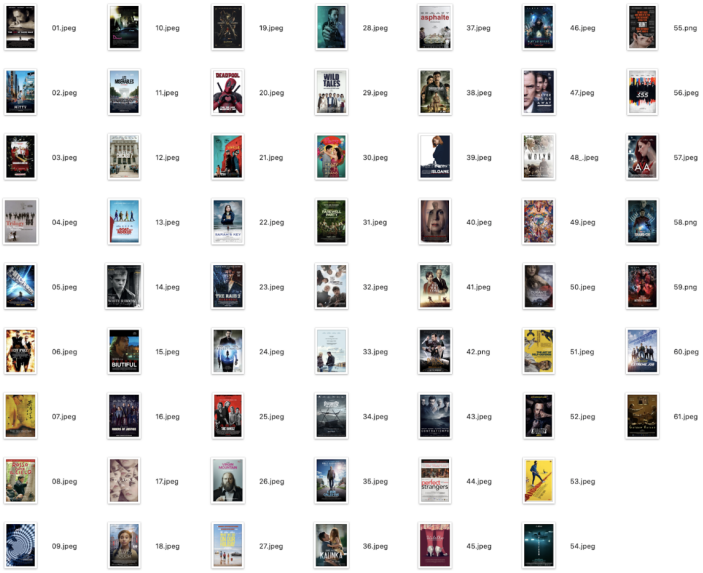


Figure 3.4 The Final 61 Movies Selected for the Experiment

films chosen should not be too popular, less likely to be exposed to the public or still not released in the Japanese market. We then asked people close to the film industry, frequent moviegoers, or general audiences to evaluate their familiarity with the given films. In addition, it is necessary to take into account that this list should contain a relatively rich variety of film genres, reviews, languages. Based on these above concerns and several pretests of the movie posters survey, we settled on a final list of movies.

- Video clips: 1 example movie clip, 1 fan-made movie trailer and 55 movie trailers with subtitles
  - The first example movie clip was used as an example to help the subjects become familiar with the experimental procedure. "The Arrival Of A Train At La Ciotat" is a 50 seconds black and white movie directed by The Lumière Brothers in 1895 which was the first movie in human history [17].
  - 55 movie trailers: From the 61 movies mentioned above excluding those without HD trailers and subtitles, the last 55 are left for movie trailer

testing.. Based on the ratings given by the subjects in SD Movie Poster Rating Survey, nine of them will be selected. The trailers are generally between one minute and thirty seconds to two minutes in length.

– The last video clip is a 4-minute 9-second cut of movie clips fan-made movie trailer for "The Room" from YouTube. A movie is considered one of the worst movies that became a classical cult movie trendy among certain groups because of being too poorly made. And still on cinema at mid-night show of The Landmark Westwood in Los Angeles.

3.3. Participants

A total of 30 subjects were recruited from Keio University Hiyoshi Campus for this study. Among them, 29 people participated in the experiment after signing the consent form (Appendix A). After excluding cases where the experimental procedure went wrong (e.g., the experiment was interrupted in the middle affecting data collection, or the researcher played the video in the wrong order), complete data from a total of 25 participants (11 males, 12 females, and 2 prefer not to say) were used for data analysis.

Regarding their movie-watching habits, 44 per cent of them (11 people) go to the cinema one to three times per month before COVID-19. Forty-four per cent of people go to the theatre less than once a month. Another 12% go once a week or more often. Figure 3.5. shows the distribution of participants' favorite movie genres.

For various factors influencing movie viewing choices, such as the trailer, director, actor or actress, etc., 40% of the subjects indicated that the trailer would influence their decision.

As for the viewing habits of the movie trailer, 28% of the subjects reported that they sometimes watch it. 20% of people watch it often, and 40% watch it occasionally. One in 25 participants would always watch the trailer in advance. As for the reasons for watching, most people reported that they wanted to judge by the trailer whether it was worth spending time or money to see the full-length film online or in the cinema. Finally, two participants expressed that they would never watch the trailer before watching the full movie. The reason is that the

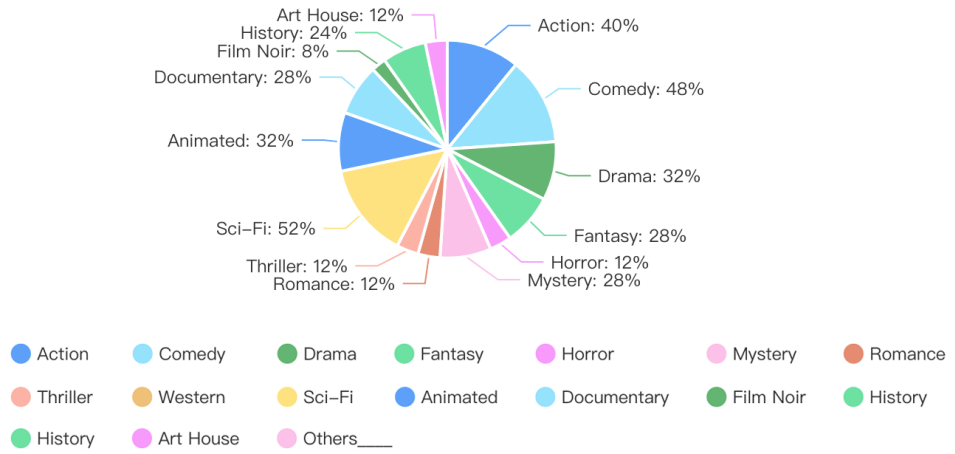


Figure 3.5 Distribution of favorite movie genres

trailer is considered to have leaked too much important content.

The participants came from different cultural backgrounds, with their first languages being Chinese, English, Japanese, Korean, Italian, Turkish, Russian, German, Macedonian, and Indonesian. Because all videos used in this experiment were subtitled in English, they all needed to have some command of English.

And after preliminary data analysis, it was found that mastery of English affected participants' ratings of the trailers they watched. Therefore, a followup questionnaire (Appendix E) on English mastery was added to this study. There were 18 participants with a good command of English and 7 participants with an average command of English. This grouping will be applied in the subsequent analysis.

It is worth mentioning that 2 of the 25 participants did not respond to the EDA sensor, which means we could not obtain their valid EDA data for analysis. Their EDA data, shown in Figure 3.6, showed almost no change on the timeline. This is consistent with the reality that 10% of the population does not respond to EDA sensors. In the part of the analysis performed for the EDA data, these two will be excluded (the questionnaire and HRV data are still valid).

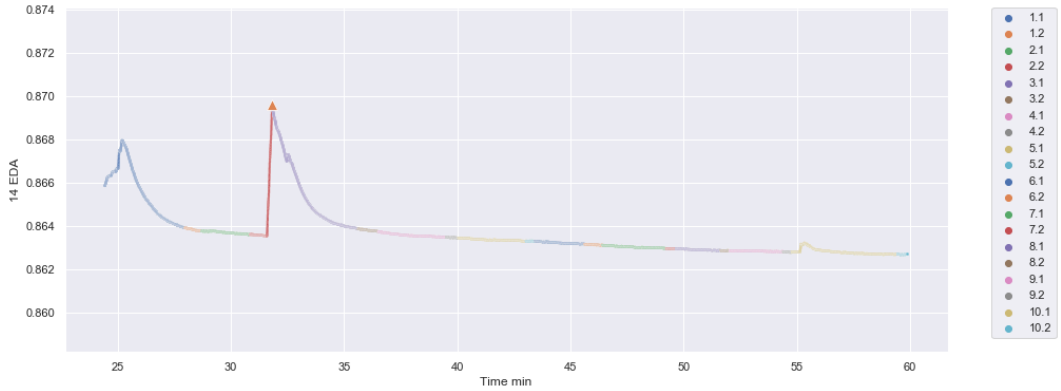


Figure 3.6 Subject does not respond to EDA sensor

3.4. Procedure

The experiments were conducted at Hiyoshi Campus of Keio University. After the arrival of the experiment participants, the researcher briefly explained the procedure and inform the data that will be collected for the experiment. The experiment would take approximately 60 minutes. Participants were required to wear the Fu wristband which would record the physiological data of hand movements, HRV and EDA. In order to avoid participants having prior knowledge of the purpose of the experiment and therefore adjusting their responses, the researcher did not need to mention the relationship between the physiological data and emotions or personal preferences that this study sought to discover (the purpose of the experiment could be explained to them at the end of the experiment).

The experimental protocol is illustrated in Figure 3.7.

After the participant has read and signed the consent form, one researcher helped them put on the device. It will be wore on their non-advantageous hand. And asking them to not move much of that hand if possible during the data recording. If they feel uncomfortable when wearing it, the researcher needs to help them adjust or replace the device with one that fits their finger better. All participants wore the device throughout the experiment, and most reported that they did not feel uncomfortable caused by the sensors.

When the device was worn, the other researcher checked that the relevant data

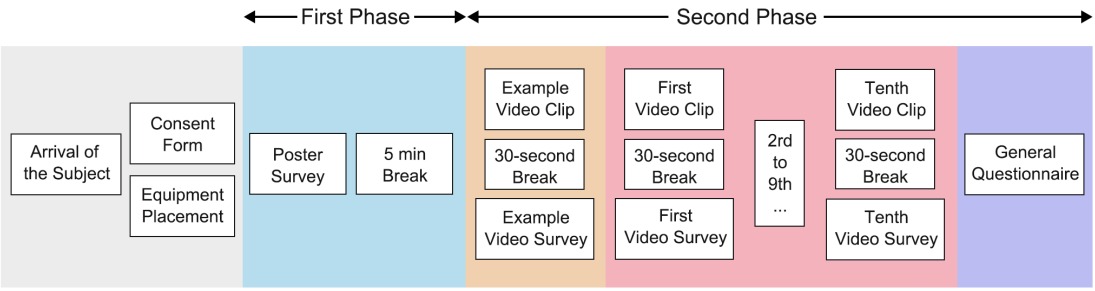


Figure 3.7 Experimental Protocol

was properly transferred to the computer where the data was saved.

Physiological signals were recorded before the first survey started and their reactions during the experiment would also be recorded. The researcher needed to remind participants that tell them not to use their cell phones during any phase of the experiment. Then the researcher inquired if the participant has any additional questions about the data collection and answer them.

The experiment was divided into two main phases, with a 5 minutes break between the two phases:

In the first phase, participants were required to fill out a survey of movie posters. They were presented with the movie posters and questioned whether they had seen the movie. If not, they were asked to rate their interest in the movie.

A 5-minute break began after completing the survey about the movie poster. Participants remained relaxed and calm during this period. This was to record their baseline data, that is, their emotional state of that day.

In the second phase, participants sat in front of a display, watched 11 video clips (including an example clip) followed by a 30s of break and answered the relevant questions based on what they have watched. There was a short demographic survey after all video clips had been played and the survey had been filled out.

3.4.1 Phase 1: Poster Survey and Break Time

Before the experiment started, the researcher ran the system QT on the desktop of the computer and selected the boilingmind server3. Then changing the directory where the experimental data is saved to "trailer-test" folder, putting subject number in the file name slot.

Once the participant had signed the consent form and was wearing Fu wristband, the researcher turned on the computer's hotspot and connected it to the device.

Figure 3.8 A subject wearing the Fu wristband and answering the first survey about the posters

The researcher turned on the device in the system, checked "Plot" and confirmed if the participant's physiological data was being plotted in real time. When the data was plotted properly, The researcher clicked "Start" button to save the data in local file.

Then the other researcher guided the participant to begin answering the first survey about the movie posters. The 5-point Likert scale (1. not interested at all, 2. not interested, 3. neutral, 4. interested, 5. very interested) was explained to

them and the first poster was taken as an example, after which they they were asked to complete it independently.

When participants completed the first survey about the movie posters, they were told to now start a five-minute break which is for base line physiological recording. During this period they do not have to do anything but relax and do not move their hand which was wearing the device substantially. The researcher would not have a conversation with the participants.

The other researcher responsible for labeling needs to input "0" in Fu band software system at the beginning of 5 minuets to indicate the start of the break time.

During their break time, the researcher needed to filter out the video clips and created a playlist that were played in the second phase based on the survey that the participants had just completed by viewing the movie posters. The playlist consisted of 11 video clips , including one example, one video that is the same for all, and 9 movie trailers selected according to the interests of each person. 3 movies they were very interested (rating 5 out of 5 in the poster survey), 3 movies they were not interested at all (rating 1 out of 5), and 3 that were neutral (rating 3 out of 5). In order to avoid the influence of artificially preferred choices and numerical sequencing, the selection and playback order of the trailers are selected by a random generator Random Picker. Recording the selected movie trailer numbers on Google Sheets subject trailer order.

The researcher found these movie trailers and created a playlist using VLC media player in the order: example / 513 / 513 / 513 / the 10th video clip . Each video clip was followed by a 30 second video clip with no sound or picture, but with the guiding text "Please take a break for 30 seconds" and "Now please answer the questionnaire" on a black background 3.9. After finishing the above steps, the researcher confirmed in advance that the headphones and the player were functioning normally.

A 30-second break was inserted after each video firstly because HRV data needed at least 2 minutes to be recorded, and secondly to avoid participants' emotions from watching the previous trailer affecting the data afterwards.

Once the break time is over, one researcher cleared up the label "0", with empty label and click the enter key. The other researcher guided participants to

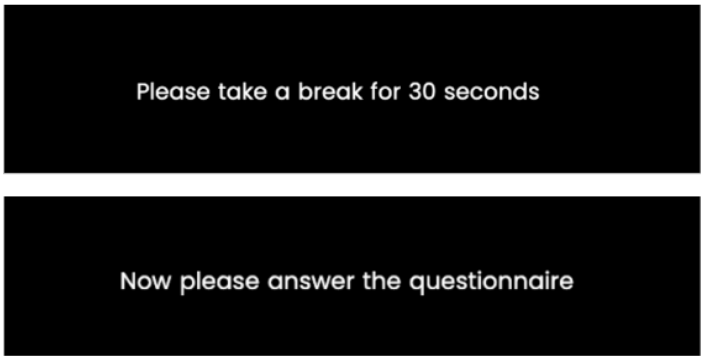


Figure 3.9 The screen for 30-second break after the trailer is played

the second phrase of the experiment to take their seats.

3.4.2 Phase 2: Movie Trailers

When the subject was seated, the researcher explained to the subject that they would start watching an example clip and was asked to fill a survey on a tablet once finished watching. They could ask questions while or after the first example clip, including the period of filling out the example survey.

Before starting to play the first example clip, one researcher turned on the web camera video recording. The other one labeled "0.1" on Fu band system while beginning to play the example video followed by a 30 seconds break time with dark screen, and labeled "0.2" when the break time was over.

The researcher then instructed the subjects to fill out the feedback survey on the tablet. The participant was presented with all possible questions and options. A 7 point Likert scale was used in this phase of the survey. It is considered more accurate and gives a better reflection of a respondent's true evaluation than other Likert scales [18].

Once the real movie trailer began, in order to avoid emotional impact on the participants due to factors other than the video clip, the researchers could not engage in conversation with the participant unless a necessity or emergency occurred. And the participant was required to wear headphones the entire time.

After the subjects indicated that they had no further questions and were pre-

Figure 3.10 A subject sitting in front of a monitor with a headphone on

pared, they moved to the real movie trailers session.

From this stage, the researchers paid attention to and noted the participant's reaction when watching the trailers, such as laughing, being surprised, exclaiming, etc.

The researcher labeled "1.1" on the Fu band system while starting playing the first trailer video and labeled "1.2" once the 30 seconds break was over and the screen showed "Now please answer the questionnaire". The researcher then helped the participant turn on the tablet and the participant began to answer the questions. This process was repeated for the next 9 video clips.

Throughout the trailer testing period, the researchers kept an eye on the Fu band data recording there were unexpected conditions such as the system crushed, the data stopped logging, the device lost connection or lack of power. When the participant proceeded to fill out the feedback survey for the last video clip, the researcher labeled "10.2" in the system.

Once the feedback survey was completed, the researcher cleared the label "10.2" and labeled "-1". Participants were then cued to remove their headphones and filled out the final questionnaire. The research could talk to them if they asked

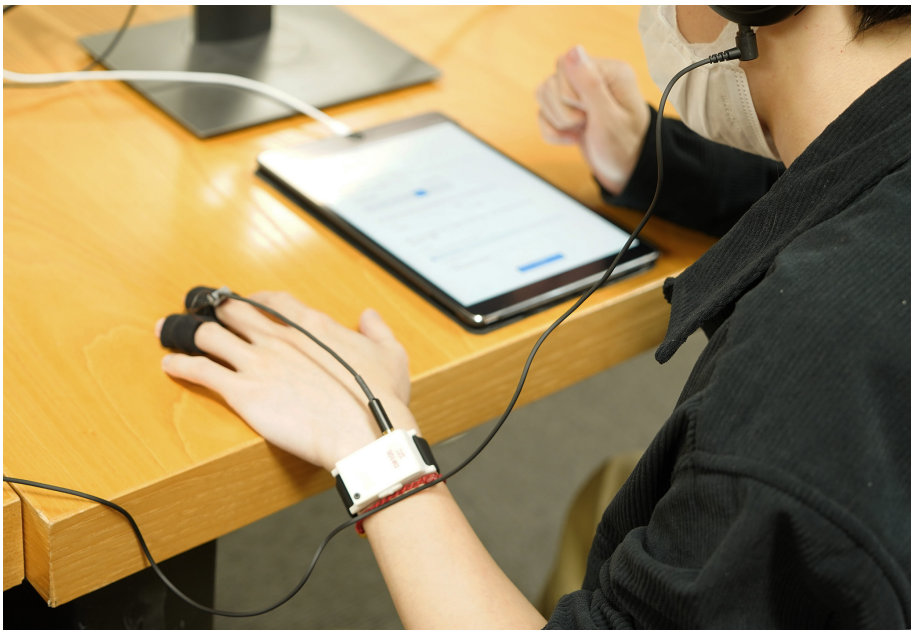


Figure 3.11 A subject answeing the movie trailer feedback questionnaire

questions for the demographic survey during this part.

When the participant completed the last questionnaire, one researcher was responsible for ending the Fu band recording. To uncheck the “Plot” box to cease the real-time plotting of physiological data, uncheck the ”ON” to end the data transfer between the device and the system, and click the “Stop” button to quit the reading and writing of data files. Finally, to close the video recording.

The other researcher took off the Fu band for the participant gently. First, turn off the sliding power button on the side of the device. Then unplug the cable connecting the finger part and the wrist part, remove the finger part, and finally remove the wrist part.

After the experiment, the researchers thanked the subjects for their participation, handed them an Amazon gift card and asked them to fill out the receipt.

The purpose of the experiment can be explained in detail to participants if they ask or want to understand it. If participants had any suggestions or comments about the experiment, the researchers were required to note them down to improve the experimental process.

Before starting the next subject, taking a look of csv file see if the data from

the experiment just performed is safely stored and charging the device just used.

Chapter 4
Discussion

4.1. Questionnaire Results

First of all, let us look at the results of the questionnaire data. As mentioned in the previous chapter, participants were asked to fill out a poster survey at the beginning of the experiment, which was used to screen the movies they had not seen and to rate the movies they had not seen as interesting or not. Participants took an average of 11.9 minutes to answer the first survey about the movie posters, spending an average of 11 seconds per poster. The study did not aim to limit the time participants took to complete the survey, but they were expected to rate the movie in interest through the posters by first response if possible. The time spent is within expectations.

For those movies that participants gave a score of 5 in Poster Survey, they would be defined as the Like Group. Those movies that got 1 point were the Dislike Group, and those got 3 point were the Neutral Group. After the initial sorting, we will select three movie trailers from each group, for a total of 9 trailers for trailer testing. After watching each trailer participants were asked to rate it again on a scale from 1 to 7. This is the Trailer Rating that was stated in the previous chapter.

Table 4.1 The mean and standard deviation based on the Trailer Rating

N = 75 (25x3)	ave	Round 1	Round 2	Round 3	SD
Like Group	5.053	5.5	4.5	5.2	1.7850
Dislike Group	3.800	4.5	3.5	3.4	1.7476
Neutral Group	4.520	4.9	4.3	4.4	1.7271

From the results of the survey (Table 4.1), we can see that there are obvious

differences in the ratings given by the subjects to different groups of movies. Overall, movies categorized in the Like Group in the Poster Survey received the highest average Trailer Ratings (which means they expressed interest in seeing the full film after seeing the poster and remained interested after watching the trailer), followed by the Neutral Group, and finally the Dislike Group. Trailer Rating and Poster Ratings show an overall consistency. This suggests that through the posters, people can roughly discern which movies they are more interested in.

Table 4.2 Comparison of different groups' rating of all subjects

	f-ratio value	p-value	Statistically Significance
Like vs. Dislike	18.87925	.000026	The result is significant at p <.05
Neutral vs. Dislike	6.44054	.012189	The result is significant at p <.05
Neutral vs. Like	3.45815	.064926	The result is not significant at p <.05

Then by performing the one-way analysis of variance on the different scores given by the participants after seeing the posters and trailers, the results of Like and Dislike Group, Neutral and Dislike Group were statistically significant at p <.05. However, the p-values for Neutral and Like Group, although less than 0.05, were not statistically significant. This means that in a statistical sense, it is possible to distinguish those movies that people do not like just by the posters, but more information is needed to judge the movies that they like or are neutral.

Table 4.3 Comparison of different groups' rating of English subjects

Non-English	N = 54 (18x3)	ave	SD
1	Like Group	5.222	1.7339
2	Dislike Group	3.667	1.6823
3	Neutral Group	4.704	1.5496
	Q	p-value	Statistically Significance
Like vs. Dislike	6.90	.00001	The result is significant at p <05
Like vs. Neutral	2.30	.23774	The result is not significant at p <05
Dislike vs. Neutral	4.60	.00398	The result is significant at p <05

Furthermore, it is noteworthy that by grouping the subjects' English mastery, the ratings they gave for the trailers varied considerably. The difference in mean

scores given by participants with better English proficiency in the self-evaluation was more notable when rating the movie trailers (Table 4.3). For their liked trailers, the average score was 5.222, those they disliked were 3.667, neutral were 4.704.

Table 4.4 Comparison of different groups' rating of non-English subjects

Non-English	N = 21 (7x3)	ave	SD
1	Like Group	4.619	1.8835
2	Dislike Group	4.143	1.9049
3	Neutral Group	4.048	2.0851
	Q	p-value	Statistically Significance
Like vs. Dislike	1.11	.71232	The result is not significant at p <05
Like vs. Neutral	1.34	.61438	The result is not significant at p <05
Dislike vs. Neutral	0.22	.98641	The result is not significant at p <05

However, participants with average English proficiency gave closer ratings to all three types of films after watching the trailers (Table 4.4). Like were 4.619, Dislike were 4.143, Neutral were 4.048). Given that we were playing videos in English original with English subtitles or other languages with English subtitles, it seems that the level of language mastery influenced their decision preferences for the movies. We think this is an interesting and worthwhile comparison that may help us understand the differences in audience behavior in movie preferences and the modes in which such differences are presented.

Why did participants with average English proficiency give close scores in the experiment? Was it because of the language barrier that prevented them from understanding the content of the trailer. Or was it because the current English narrative-based promotional materials were at odds with their preferences. Or maybe it is some other reason, which requires further research.

But the results of this experiment also show that the uniform questionnaire method used in traditional trailer testing has limitations in dealing with audiences with different cultural backgrounds.

The following section will continue to explore in depth how this difference will present itself in terms of physiological parameters.

4.2. Feature Extraction from Physiological Signals

Next, we will analyze the physiological data obtained in the experiment. First, the EDA data, followed by the HRV data. We will then examine whether there is some correlation between participants' physiological data and ratings while watching the same movie trailer.

4.2.1 ElectroDermal Activity (EDA)

First of all, focusing on the EDA data, Figure 4.1 shows the complete EDA data of one subject in the experiment, with the horizontal coordinate being the time and the vertical coordinate being the skin conductivity, which is the EDA data mentioned throughout this paper. The different colors on the curves mean that the subject was in a different task. The small triangles appearing on the graph are the times when the EDA peak appeared.

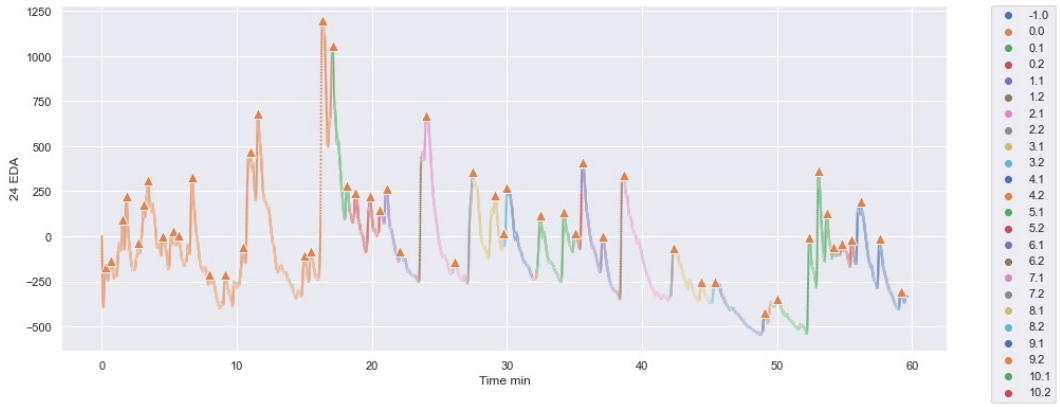


Figure 4.1 Example of EDA Data by Time Plotted

EDA is commonly used to measure the level of emotional arousal. In this experiment, we focused more on the timing and number of times EDA peaks appeared when participants watched the trailer rather than the overall EDA change. Next, EDA Peaks Detection was conducted on participants at different stages.

Figure 4.2 shows that EDA peaks in total are different when watching the video clips and when filling out the survey. Where the phases labeled 1.2, 2.2, 3.2, 4.2,

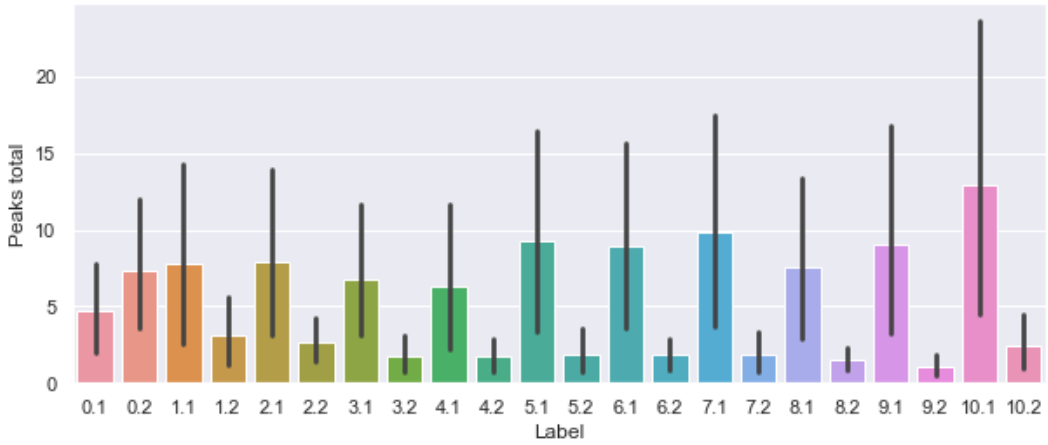


Figure 4.2 EDA Peaks in total of all subjects in Trailer Testing Phase

5.2, 6.2, 7.2, 8.2, 9.2, 10.2 are all the times that participants were filling out the survey. We can clearly see that there are fewer EDA peaks in these times than in other periods. This suggests that this method can be used to identify different dynamic tasks. The arousal level of the participants was higher when watching the video.

Let us then look at how the EDA data presented differently when the experimental participants watched the trailer. Because the length of the trailers varied, the EDA peaks per minute data were used here. From Figure 4.3 we can see that the EDA peaks per minute are less at labels 1.1, 4.1, and 7.1, that is, when the first, fourth, and seventh trailers in the list are played, overall the participants have fewer EDA peaks per minute compared to the other trailers. And the first, fourth, and seventh trailers in the playlist were exactly the movies that the participants thought they were interested in. Also, the group of movies that were considered disliked by the participants, the trailers labeled 2.1, 5.1, and 8.1 appeared to have relatively more EDA peaks. This implies that there is a correlation between audience preferences and physiological data.

Next, an ANOVA test was performed on the data from the Like Group (1.1, 4.1, 7.1) and the Dislike Group (2.1, 5.1, 8.1). The mean value of EDA peaks per minute was 0.648 for the Like group and 0.8324 for the Dislike group. The f-ratio value is 3.52989. The p-value is .062511. The result is not significant at p

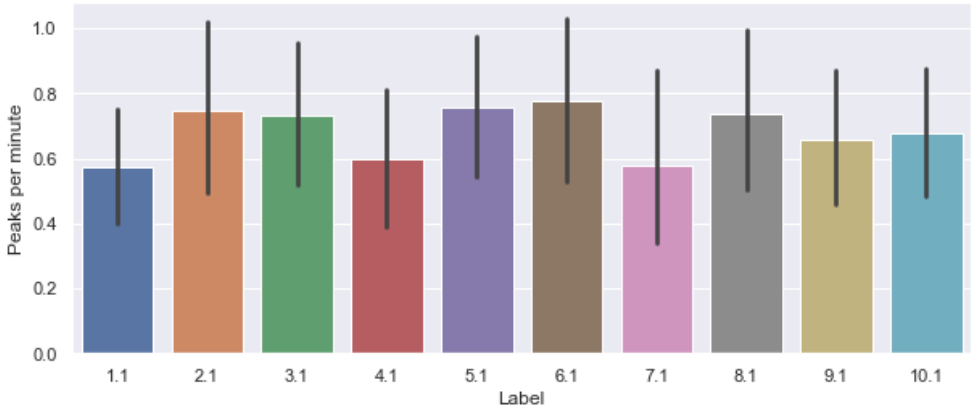


Figure 4.3 EDA Peaks per minute when Watching Trailer

<.05. Although we can observe differences in the graphs, they are not statistically significant.

As already mentioned in the analysis of the questionnaire data, the level of mastery of English influenced the ratings participants gave to the trailers they watched. Here again, group analysis was conducted according to the level of English mastery.

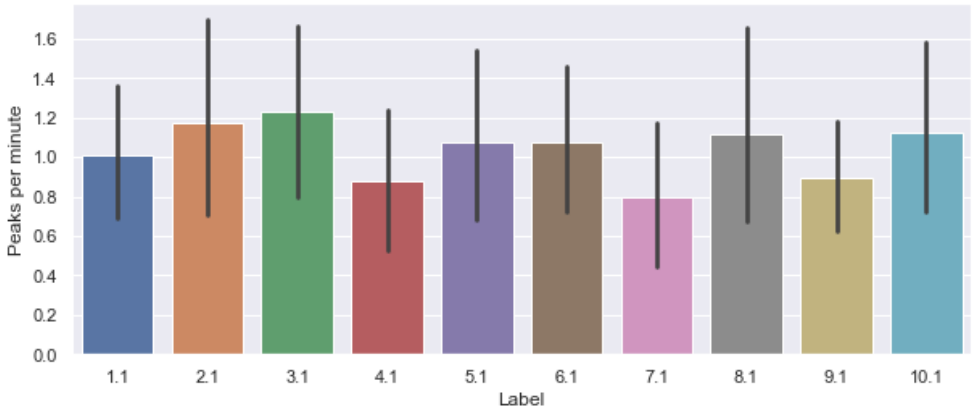


Figure 4.4 EDA Peaks per minute - English Group

Figure 4.3 illustrates the data of EDA peaks per minute for the subjects with better English proficiency (16 participants). From the chart, it shows similar characteristics to the data of all participants. While watching the trailer for the

Like Group (1.1, 4.1, 7.1), participants had relatively few EDA peaks per minute. Here again, ANOVA tests were conducted for Like Group (1.1, 4.1, 7.1) and Dislike Group (2.1, 5.1, 8.1). The mean value of EDA peaks per minute was 0.61 for the Like group and 0.85 for the Dislike group. The f-ratio value is 4.91548. The p-value is .028882. The result is significant at p <.05.

The outcome was statistically significant. This means that we can predict whether viewers will like the movie by the data of EDA peaks per minute. Because EDA measures the level of arousal, the data results suggest that viewers will prefer movie trailers that give them less emotional stimulation.

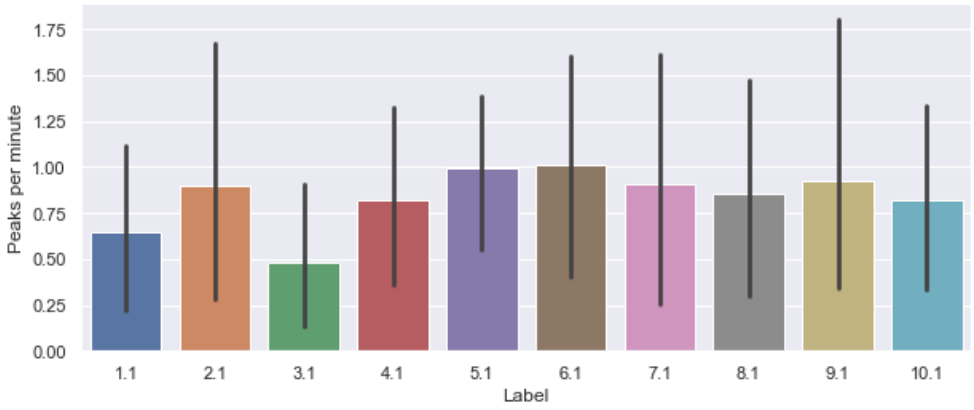


Figure 4.5 EDA Peaks per minute - Non-English Group

Then we turn to the group that self-rated their English as not so good (7 participants). There was no visible correlation in their EDA data and preferences. Their mean values of EDA peaks per minute were also less than those of the English group. It seems that the understanding of the content of the movie trailers was influenced by the language, which in turn affected the presentation of their physiological data.

4.2.2 Heart Rate Variability (HRV)

The HRV data used in this study was calculated as SDNN, which is the standard deviation of the NN (R-R) intervals. Figure 4.6 presents the HRV data change process of one participant throughout the experimental phase. The horizontal

coordinate is the time and the vertical coordinate is the SDNN data. Different colored curves mean he or she is in a different stages of the experiment.

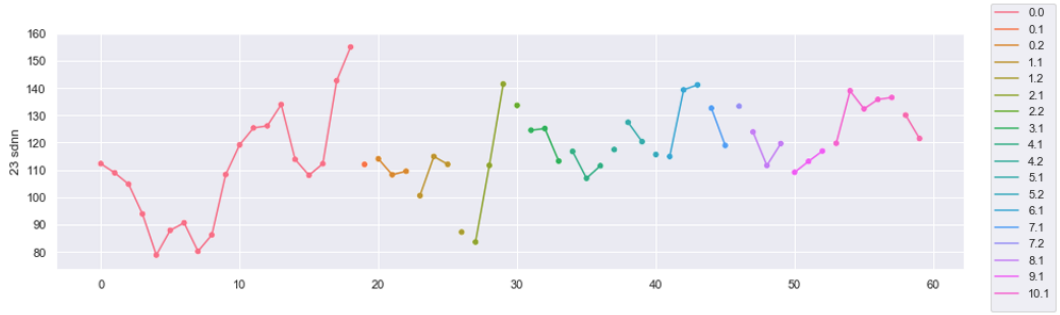


Figure 4.6 Example of HRV Data throughout the Experimental Period

To make it more obvious, the labels are marked next to the HRV data curves of the corresponding time periods (Figure 4.7). The first continuous red curve with a label of 0 is the HRV data of the participants when filling out the Poster Survey and the five-minute break time. Label 0.1 is the example video and 0.2 is the example of the trailer testing survey. The other labels have been introduced before for the 10 trailers watched by the subjects.

Figure 4.7 Example of HRV Data with Label throughout the Experimental Period

For the HRV data, the main focus is on the trend of its changes. The increase of HRV value means that the audience becomes more relaxed while watching the video, whereas the decrease of HRV value means that the audience goes into a tense state. We can observe from the graph that in the latter period of the first

phase of the experiment, that is, the five-minute break, the subject gradually entered a more relaxed state and his HRV value reached the highest value of the whole experiment.

During the process of watching the trailer, this participant's HRV sometimes tended to increase and sometimes to decrease. Here we will compare their HRV data with the ratings given after watching the trailer.

Table 4.5 Ratings given by subject 23 after watching each trailer

Lable	1.1	2.1	3.1	4.1	5.1	6.1	7.1	8.1	9.1	10.1
Rating	6	6	4	5	4	7	6	3	5	7

This participant showed an upward change in his or her HRV data when watching the first trailer, and he or she gave the trailer a rating of 6, which is a relatively high score. This pattern we can also detect in the data of the 2nd, 6th and 10th trailers. And when its HRV value shows a downward trend, such as in watching the 3rd, 5th, and 8th trailers, he or she gives the trailer a relatively low score of 4, 4, and 3. We can also find a certain correlation between HRV data and audience preferences. Although there are exceptions to this, such as the 7th trailer where the HRV data trended downward, but the viewer gave a high score. But overall, the higher HRV data of viewers means they are more relaxed in watching the trailers and more likely to give high ratings. The opposite is also true.

In addition, the second trailer was initially a movie that this participant had indicated in the Poster Survey that s/he had no interest in and did not want to see the full movie. But after watching the trailer, s/he gave a high score of 6. By looking at the HRV data, we can notice that the upward trend is very steep compared to the other trailers that also received high scores. This pattern was also found in the HRV of other participants. This might mean that the physiological data not only reflects whether the audience liked the trailer they watched or not, but also the shift in their preferences can be shown.

Since the HRV data required at least 2 minutes to record and the trailer was not long, not enough data points were recorded in the experiment for analysis. Only preliminary findings are presented here, and future studies are needed if the results are to be validated.

4.2.3 Physiological Data when Watching the Same Trailer

So far, some first conclusions have been drawn from this research. Next let's examine the physiological data of all experimental participants watching the same trailer.

Of the ten movie trailers viewed by the participants, the last one, the tenth one, was the same. The tenth trailer was a fan-made trailer for the movie The Room, which was added to the experiment because of the researcher's personal interest. It was subtitled in both English and Chinese.

EDA

Firstly, the participants were divided into High Rating Group and Low Rating Group based on the score they gave after watching the trailer. Participants in High Rating Group gave the trailer a score of 6 or 7. Participants in Low Rating Group gave a score of 1 or 2.

Then we look at the EDA data. In terms of the results from the analysis of overall EDA Peaks in the previous section, we found that participants had fewer EDA PPeaks per minute when they were watching movie trailers that they liked or were interested in. Assuming this finding is correct, this pattern should also be found when watching the same trailer, as participants who gave a high score had fewer EDA Peaks per minute during watching the trailer. Those who gave low scores had more EDA Peaks per minute.

Table 4.6 EDA Peak Detection - High Rating Group (Score 6 or 7)

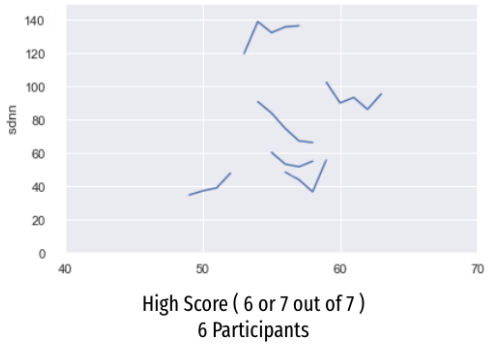
Subject No.	Start	End	Peaks	Duration	Peaks Per Min
4	58.76662	63.45847	4	4.69185	0.853
9	54.72042	59.3774	2	4.65698	0.429
10	133.10195	137.77973	1	4.67778	0.214
11	68.14948	72.81723	1	4.66775	0.214
23	82.94837	87.6633	1	4.71493	0.212
30	49.01958	53.68915	6	4.66957	1.285
mean			2.5		0.535

Table 4.7 EDA Peak Detection - Low Rating Group (Score 1 or 2)

Subject No.	Start	End	Peaks	Duration	Peaks Per Min
2	51.89268	56.61815	5	4.72547	1.058
7	50.93880	53.08618	3	2.14738	1.397
8	46.62457	51.55220	6	4.92763	1.218
26	47.39458	52.04827	2	4.65368	0.430
27	60.31972	64.99770	2	4.67798	0.428
31	96.92792	101.59452	3	4.66660	0.643
mean			3.5		0.862

An examination of Table 4.6 and Table 4.7 reveals that this is indeed the case. Participants who gave the trailer a high score had an average peaks of 2.5 and peaks per minute of 0.535, while those who gave the trailer a low score had an average peaks of 3.5 and peaks per minute of 0.862. This is consistent with the previous results.

HRV



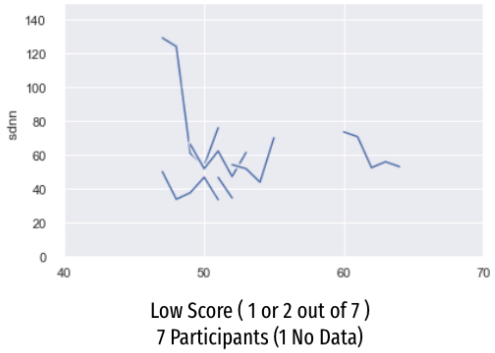
Subject No.	Label	min	max	change
4	10.1	86.27	102.54	-7.04
9	10.1	66.32	90.90	-24.58
10	10.1	36.75	56.17	7.49
11	10.1	51.79	60.44	-5.28
23	10.1	119.77	138.98	16.76
30	10.1	34.87	47.91	13.04
mean				0.063

Figure 4.8 HRV Data - High Rating Group (Score 6 or 7)

Then to check if the HRV data are also consistent. First of all, let's observe the trend of HRV change for the high grouping. Among the 6 participants who gave this trailer a high score, 4 of them showed an upward trend in HRV data, which

means they became more relaxed and enjoyed the process more while watching it. The mean value of the change in HRV data for the High Rating Group was 0.063.

Then the trend for the 6 participants in the Low Rating Group, 5 out of 6 showed a decreasing trend or a decrease in HRV data in comparison to the value at the beginning of the watch. The mean value of the change in HRV data for the Low Rating Group was -15.23.



Subject No.	Label	min	max	change
2	10.1	44.01	70.23	15.89
6	10.1	47.32	66.42	-4.90
7	10.1	34.67	46.83	-12.16
8	10.1	53.85	129.26	-53.10
26	10.1	33.67	50.20	-16.53
27	10.1	52.61	73.73	-20.56
mean				-15.23

Figure 4.9 HRV Data - Low Rating Group (Score 1 or 2)

From the physiological data, we can see that the viewers who gave the trailer a low score were less relaxed than those who gave it a high score. They may not have enjoyed the viewing process and therefore gave it a low score. This result is also consistent with our previous finding.

Chapter 5
Conclusion

5.1. Summary

In this study, it begins with an introduction to the stagnant growth problems currently encountered in the global film market and an understanding of the current state of the film industry's struggles with costs, technological innovation, and export to foreign markets. We also look at the limitations of traditional film market research methods and their impact on marketing strategies.

This exploratory study proposes a new method for movie trailer testing which combines quantitative data from questionnaires and physiological signal parameters to discover how audiences response when watching trailers or video clips. An experimental design was determined after first conducting four pretests, and then 30 participants were recruited for the experiment. Data from 25 participants were used for data analysis.

From the data analysis, a correlation was found between people's physiological data and their preferences. When people watch a movie trailer, the lower their EDA peaks are, the more likely they are to give this trailer a high score and be more willing to watch the full movie. This is a new finding because in other studies that have also analyzed for EDA peaks, such as participation in conferences. People would give higher ratings when they had more EDA peaks. This may be attributed to the difference in the type of task, where the audience is in a passive state of receiving information while watching the trailer, and they are engaged in entertainment and recreational activities. Rather than engaging in activities like meetings that require active thinking and communication.

And they were more likely to give high scores if their HRV values tended to increase while watching the trailer, which means they were more relaxed and engaged in the watching process.

The results of the study suggest it as a possible new method for movie trailer testing to compensate for the shortcomings brought by traditional methods. For example, the use of a combination of physiological data and traditional questionnaires allows for the observation of specific time points in a way that is not possible with traditional methods. In reality it can be used to test the effect of different versions of the trailer and also to assist the editor in making more accurate judgments.

5.2. Limitations

This study, while hoping to emulate a realistic viewing experience, is still being conducted in a laboratory environment isolated as much as possible from external interference. If it is expected to be applied in a real-life scenario it needs to be validated by organizing activities with many people involved and with dynamic tasks.

In addition, since the experiment is an exploratory research, the questionnaire was set up to cover as many questions as possible about movie viewing habits in order to verify the usability of this new method. It is slightly lacking in specialization, such as not testing different genres of films or subjects from various cultural backgrounds separately, which makes it more difficult to carry out in-depth comparisons.

5.3. Future Works

To optimize the experiment process based on logs to enhance data usability. For example, try to automate operations such as movie selection and playback, labeling the time points in the system to reduce possible mistakes in data collection caused by human factors.

I would like to take this research as a starting point to also engage in related investigations after graduation and explore its potential applications. For example, what is the cause of the lower market share of imported films in the Japanese market. What different physiological data characteristics differences are induced by different types of films in English speakers and Japanese speakers. Based on

this difference, is it possible to create more targeted promotional programs based on different regions. Or to select a real unexposed film for an actual screening test using the method presented in this thesis, and based on the feedback results make modifications in the test and evaluate its validity.

References

[1] Adam West and Jeff Rovin. *Back to the batcave*. Berkley Publishing Corporation, New York, NY, September 1994.

[2] Paul Eckman and Wallace Friesen. *Unmasking the Face*. Prentice Hall, Old Tappan, NJ, February 1975.

[3] Maria Egger, Matthias Ley, and Sten Hanke. Emotion recognition from physiological signal analysis: A review. *Electronic Notes in Theoretical Computer Science*, 343:35–55, 2019. The proceedings of AmI, the 2018 European Conference on Ambient Intelligence. URL: https://www.sciencedirect.com/science/article/pii/S157106611930009X, doi:https://doi.org/10.1016/j.entcs.2019.04.009.

[4] Mathias Benedek and Christian Kaernbach. A continuous measure of phasic electrodermal activity. *Journal of Neuroscience Methods*, 190(1):80–91, 2010. doi:https://doi.org/10.1016/j.jneumeth.2010.04.028.

[5] Hugo D. Critchley. Review: Electrodermal responses: What happens in the brain. *The Neuroscientist*, 8(2):132–142, 2002. PMID: 11954558. URL: https://doi.org/10.1177/107385840200800209, arXiv:https://doi.org/10.1177/107385840200800209, doi:10.1177/107385840200800209.

[6] María Viqueira Villarejo, Begoña García Zapirain, and Amaia Méndez Zorrilla. A stress sensor based on galvanic skin response (gsr) controlled by zigbee. *Sensors*, 12(5):6075–6101, 2012. URL: https://www.mdpi.com/1424-8220/12/5/6075, doi:10.3390/s120506075.

[7] Guanghua Wu, Guangyuan Liu, and Min Hao. The analysis of emotion recognition from gsr based on pso. In *2010 International Symposium on*

Intelligence Information Processing and Trusted Computing, pages 360–363, 2010. doi:10.1109/IPTC.2010.60.

[8] Charles Kaiser and Robert Roessler. Galvanic skin responses to motion pictures. *Perceptual and Motor Skills*, 30(2):371–374, 1970. PMID: 5465822. URL: https://doi.org/10.2466/pms.1970.30.2.371, arXiv: https://doi.org/10.2466/pms.1970.30.2.371, doi:10.2466/pms.1970.30.2.371.

[9] Christine Lætitia Lisetti and Fatma Nasoz. Using noninvasive wearable computers to recognize human emotions from physiological signals. *EURASIP J. Adv. Signal Process*, 2004:1672–1687, jan 2004. URL: https://doi.org/10.1155/S1110865704406192, doi:10.1155/S1110865704406192.

[10] Fernando Silveira, Brian Eriksson, Anmol Sheth, and Adam Sheppard. Predicting audience responses to movie content from electro-dermal activity signals. In *Proceedings of the 2013 ACM International Joint Conference on Pervasive and Ubiquitous Computing*, UbiComp '13, page 707–716, New York, NY, USA, 2013. Association for Computing Machinery. URL: https://doi.org/10.1145/2493432.2493508, doi:10.1145/2493432.2493508.

[11] Fred Shaffer and J. P. Ginsberg. An overview of heart rate variability metrics and norms. *Frontiers in Public Health*, 5, 2017. URL: https://www.frontiersin.org/article/10.3389/fpubh.2017.00258, doi:10.3389/fpubh.2017.00258.

[12] Jianping Zhu, Lizhen Ji, and Chengyu Liu. Heart rate variability monitoring for emotion and disorders of emotion. *Physiological Measurement*, 40(6):064004, jul 2019. URL: https://doi.org/10.1088/1361-6579/ab1887, doi:10.1088/1361-6579/ab1887.

[13] Julian F J Siegle. Neurovisceral integration in cardiac and emotional regulation. *IEEE engineering in medicine and biology magazine : the quarterly magazine of the Engineering in Medicine Biology Society.*, 21(4), 2002.

[14] Narayana Darapaneni, Christopher Bellarmine, Anwesh Reddy Paduri, Sujana Entoori, Abir Kumar, S V Vybhav, and Koushik Mondal. Movie success

prediction using ml. In *2020 11th IEEE Annual Ubiquitous Computing, Electronics Mobile Communication Conference (UEMCON)*, pages 0869–0874, 2020. doi:10.1109/UEMCON51285.2020.9298145.

[15] Penny Ur. Stimulating Student Interest in Language Learning: Theory, Research and Practice. *ELT Journal*, 71(4):537–539, 08 2017.

[16] Dingding Zheng. *An Interdisciplinary Psychometric Method for Psychophysiological Tracking of Social Interactions and Social Dynamics in "Real-World" Environments*. PhD thesis, Keio UNiversity, 2022.

[17] Martin Loiperdinger and Bernd Elzer. Lumiere's arrival of the train: Cinema's founding myth. *Mov. Image J. Assoc. Mov. Image Arch.*, 4(1):89–118, 2004.

[18] Ankur Joshi, Saket Kale, Satish Chandel, and D Kumar Pal. Likert scale: Explored and explained. *British journal of applied science & technology*, 7(4):396, 2015.

Appendices

A. Experiment Consent Form

B. SD Movie Poster Rating Survey

C. Movie Trailer Rating Survey

D. General and Demographic Information Survey

E. Follow-Up Questions

Subject #: _____

Consent Form

Psychophysiological response based on video clips

DESCRIPTION: You are invited to participate in an experiment investigating physiological sensing and movie trailers.

TIME INVOLVEMENT: The experiment will take approximately **45-50** minutes. You are asked to wear equipment provided by the researchers and watch some movie trailers in a quiet environment.

DATA COLLECTION: We will be collecting different types of data from you through a sensing device with a sensor on the fingers. Throughout the experiment you will be asked to watch 10 movies trailers or clips and answer some questions based on these video clips. Data collected is used to evaluate your individual physiological signals in response to the setting.

RISKS AND BENEFITS: There are no risks associated with your participation and with the usage of the hardware. The collected data is securely stored. Your identity is not disclosed unless we directly inform and ask for your permission. There may be no personal benefit from your participation in the study but the knowledge received may be of educational value to the principal investigator with regard to methods of conducting psychophysiological research.

PARTICIPANT'S RIGHTS: If you have read this form and have decided to participate in the experiment, please understand your participation is voluntary and you have the right to withdraw your consent or discontinue participation at any time, however you will lose the compensation guaranteed to you for completion of the experiment. The alternative is to not participate. The results of this experiment may be presented at scientific or professional meetings or published in scientific journals.

CONTACT INFORMATION: If you have any questions, concerns or complaints about this research, its procedures, risks and benefits, please reach out to the following persons:

GEIST Lab. Graduate School of Media Design, Keio University 4-1-1 Hiyoshi, Kohoku-ku, Yokohama, Kanagawa Japan, 223-0061

Shuoping Shi (shuoping@kmd.keio.ac.jp)
Lady Dingding Zheng (zheng208@kmd.keio.ac.jp)
Prof. Kai Kunze (kai@kmd.keio.ac.jp)

☐ I agree that data collected can be published and shown during presentations in an anonymized form.

☐ I confirm that I have read, understood, and agree to the terms and conditions above.

☐ I permit the usage and/or publication of photo and/or video materials that include me participating in this experiment.

Researcher's name: _____ Signature: _____

Date: _____

Participant's Name (Print): _____ Signature: _____

SD Movie Poster Rating Survey

In this questionnaire, you will be presented with movie posters and asked whether you have ever watched these movies. Please try to not overthink it, but reply with your first instinct.

このアンケートでは、映画のポスターが提示され、これらの映画を見たことがあるかどうかを尋ねられます。長く考えることが必要ない、一目で解答を書いてください。

There is no right or wrong answer, you just need to answer according to your personal situation. Please understand your participation is voluntary and you have the right to withdraw or discontinue this questionnaire at any time.

正解がありません。個人の考えにより回答を選択してください。参加は任意であり、参加途中でもこのアンケートを撤回または中止する権利がございます。

A 5-point Likert scale rating will be presented during the questionnaire:
5ポイントのリッカート尺度の評価が提示されます:
1 = not interested at all　まったく興味がない
2 = not interested　興味がない
3 = neutral, neither interested nor uninterested どちらでもない
4 = interested 興味がある
5 = very intersted 非常に興味がある
Please pick what matches best for you. 最適の選択を選んでください。

1. Have you watched this movie before?
   The Life of David Gale / 大卫·戈尔的一生 / ライフ・ オブ・ デビッド・ ゲイル

☐ No, I don't know this movie
☐ No, but I know this movie
☐ Yes, I have watched this movie
☐ I'm not sure

2. Based on the title and the poster how much are you interested or not interested in this movie?
   1 = not interested at all; 2 = not interested; 3 = neutral ; 4 = interested; 5 = very intersted

1　　2　　3　　4　　5

Movie Trailer Rating Survey

This is a video clip. Please assume it's a movie trailer and answer the question. If you have any questions, feel free to ask the researchers.

A 7-point Likert scale rating will be presented during the questionnaire:
7 ポイントのリッカート尺度の評価が提示されます:
1 = Strongly Dislike
2 = Dislike
3 = Somewhat Dislike
4 = Neither Dislike Nor Like
5 = Somewhat Like
6 = Like
7 = Strongly Like
Please pick what matches best for you. 最適の選択を選んでください。

1. subject ID: _____ (please have the invastigater help you fill this part)

2. How much do you like or dislike this clip?
   1 = Strongly Dislike, 2 = Dislike, 3 = Somewhat Dislike, 4 = Neither Dislike Nor Like, 5 = Somewhat Like, 6 = Like, 7 = Strongly Like

   1 2 3 4 5 6 7

3. Would you want to watch the full movie if you have the chance?

   ☐ Yes ☐ No ☐ Maybe ☐ I don't know

4. If you want to / wouldn't like to watch the full movie, which factor(s) led to your decision?

   ☐ Story ☐ Director, Producer, or Screenwriter

   ☐ Actor or Actress ☐ No specific reason

   ☐ Others_____(Please type the reasons)

5. How much are you willing to pay to watch this movie?
   Above 3,000 JPY (≈23 USD)?
   2,500 - 3,000 JPY (≈ 19-23 USD)?
   2,000 - 2,500 JPY (≈ 15-19 USD)?
   1,500 - 2,000 JPY (≈ 11-15 USD)?
   1,000 - 1,500 JPY (≈ 7.5-11 USD)?
   Less than 1k or Online Subscription: _____
   Please fill in specific amount (JPY or USD)

6. Which way do you prefer to watch it?

☐ Movie theatre ☐ At home with home screen or pc from platform like Netflix, Amazon, Hulu, etc.

☐ Either way ☐ Both

General and Demographic Information Survey

This questionnaire is about your viewing habits and personal information. There is no correct answer. You only need to answer according to your own situation.

1. How often do you usually go to the movie theatres without pandemic impact?
2. What genres of movies do you usually go to the movie theatres for?
3. Do you watch movies or tv shows on any online platform?
4. In general, what factors will determine your choice to watch a movie?
5. Would you watch the trailer before watching the full movie?
6. Why do you always or rarely watch movie trailers?
7. What is your favourite genre of film?
8. Which form of the movie do you prefer to watch when watching a movie in a non-native language?
9. How many movies have you watched in the last year?
10. What gender do you identify as?
11. What is/are your native language(s)?
12. What is your age?
13. What is the highest degree or level of education you have completed?

Follow-Up Questions

1. Which languages of the movies do you grow up watching most? (multiple choices)

☐ English	☐ Russian	☐ Hebrew
☐ Chinese, Mandarin	☐ German	☐ Serbian
☐ Chinese, Cantonese	☐ Spanish	☐ Arabic
☐ Chinese, other____	☐ Indian	☐ Slovak
☐ Japanese	☐ Italian	☐ Svenska
☐ Korean	☐ French	☐ Others____

2. Please read the following description and circle a number (described as the followings) based on how well it matches your situation:

   1) Usually, I can understand a foreign language movie or movie clips or trailer with English subtitles there.
      1= strongly disagree
      2= disagree
      3= somewhat disagree
      4= neither agree nor disagree
      5= somewhat agree
      6= agree
      7= strongly agree

   2) If you have been explored with English before, as an English user, how would you describe your language levels:
      0= No English exposure in my entire life; it's a useless language that I don't need.
      1= Beginner: can understand and use familiar everyday expressions
      2= Elementary English: can communicate in simple and routine tasks
      3= Intermediate English: can understand and describe the main points of clear standard input on familiar matters
      4= Upper-Intermediate English : with a degree of fluency
      5= Advanced English : fluent
      6= Proficiency English: very fluent
      7= Native

3. Which batch did you enrol in Keio?

   ☐ April ☐ September