JAIST Repository
https://dspace.jaist.ac.jp/

Title	Manipulating Emotions through Virtual Reality: Investigating Misattribution of Physiological Arousal
Author(s)	李, 喆
Citation	
Issue Date	2025-03
Type	Thesis or Dissertation
Text version	author
URL	http://hdl.handle.net/10119/19750
Rights	
Description	Supervisor: 藤波 努, 先端科学技術研究科, 修士 (知識科学)

JAPAN
ADVANCED INSTITUTE OF
SCIENCE AND TECHNOLOGY

Japan Advanced Institute of Science and Technology

Master's Thesis

Manipulating Emotions through Virtual Reality: Investigating Misattribution of Physiological Arousal

ZHE LI

Supervisor TSUTOMU FUJINAMI

Graduate School of Advanced Science and Technology
Japan Advanced Institute of Science and Technology
(Knowledge Science)

January, 2025

Abstract

The two-factor theory of emotion lays the foundation for exploring the manipulation of human emotions. However, it remains controversial due to debates surrounding both the theory itself and the methods used to induce physiological arousal. This study proposes utilizing Virtual Reality (VR) technology, combined with VR survival horror games as an alternative approach for inducing arousal. Conversation-based suggestive cues were applied to provide emotional labels, aiming to validate the misattribution of arousal. The impact of cues on emotional cognition was accessed through self-report questionnaires and the recording of physiological indexes (heart rate, skin temperature). The results revealed that (a) the VR can serve as an effective tool to manipulate human emotions; (b) suggestive cues significantly influenced corresponding emotional cognition under arousal; (c) negative cues (Anger) led to unexpectedly higher positive emotional ratings (Courage) compared to a no-cue condition.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

Emotion has always been an interesting and core subject in psychology. Researchers aim to uncover the biological basis of emotions and their interaction with cognitive processes. Schachter and Singer's two-factor theory of emotions explains that emotion is determined by both physiological arousal and cognitive labels (Schachter and Singer, 1962). In their theory, people interpret their physiological states based on environmental cues, which can lead to misjudgments of their emotions. The famous suspension bridge experiment illustrates that how misattributed arousal can shape our emotional experiences (Dutton and Aron, 1974).

While the two-factor theory of emotions has important theoretical significance, its practical application faces many challenges. In the study of emotion, a key issue is finding appropriate stimuli to induce emotional states (Marchewka et al., 2014). Traditional methods of inducing physiological arousal rely on drugs or complex experimental setups, which are either potentially ethically problematic or difficult to implement widely. Therefore, there is an urgent need for a safer, more flexible and scalable method to study and apply emotional processes.

Virtual Reality (VR) technology offers a promising solution. Recent years virtual reality technology has shown great potential in fields like education and healthcare (Somarathna et al., 2022), due to its high controllability and ability to provide a strong sense of immersion and presence (Slater, 2009). VR enables users to step into characters and feel "like in the real world", which means that it can induce authentic emotional experiences (Parsons and Rizzo, 2008). Researchers can also use VR to design realistic and customizable experimental environments, which makes it an ideal platform for psychological studies. For instance, among Virtual Reality technologies, VR games, especially VR survival horror games, have been proven as a novel and effective method for safely inducing emotions (Pallavicini et al., 2018). This study showed that such games can evoke intense fear and anxiety, making VR survival horror games well suited as tools for emotional research. Additionally, research shows that VR tends to elicit stronger emotional reactions than less immersive media, such as video and audio (Pallavicini et al., 2013). The is because VR requires the user to participate in it, which can enhances emotional responses than those passive methods (Pallavicini et al., 2019).

Despite VR technology has been applied to a variety of psychological studies, the potential of VR to integrate with the two-factor theory and influence subjective emotional interpretations has not been fully explored. This study aims to fill this gap. Moreover, explores the effectiveness of VR in inducing misattribute of arousal.

In this study, 34 participants used VR headsets (Meta Oculus Quest 3) to play a commercially available survival horror game. They were divided into three groups to receive different emotional label cues. Self-report questionnaires and physiological measures (heart rate, skin temperature, and P-P Interval changes) were used and analyzed, in order to assess their emotional responses and judgments.

This research discussed one main hypothesis and two sub-hypotheses:

1. **Main:** Virtual Reality technology can be an effective tool to induce the misattribution of arousal.

2. **Sub 1:** Emotions triggered by VR survival horror games can be altered through suggestion cues, leading to changes in the perception of emotions.

3. **Sub 2:** Positive emotion labeling may have a different impact than negative labeling. Such as the result of rates might lead to a higher score for positive emotions than negative ones on subjective scales.

Chapter 2

Related Works

2.1 Physiological Arousal and Misattribution

Over the years, researchers have conducted extensive studies on human emotions. Among those research, the concept of cognitive determination of emotional states is particularly intriguing. In this area, ample evidence shows that incidental emotions can influence decision-making (Vohs et al., 2007). Based on this, some researchers hypothesize that people tend to recall judgments based on past emotions and use them directly when making future decisions. This suggests that transient incidental emotions may become the basis for future decisions (Andrade and Ariely, 2009). Thus, emotion itself holds significant value for research. If we can manipulate the cognitive understanding of people's emotional states, it might be possible to indirectly affect their decision-making process.

Therefore, the two-factor theory of emotion has received a lot of attention (Schachter and Singer, 1962). This theory proposes and verifies three hypotheses: (a) When an individual experiences a state of physiological arousal that has no immediate explanation, they will label this emotional state to support their understanding of their current feelings; (b) When the arousal has a clear explanation, the individuals are unlikely to attribute their feelings to alternative cognitive labels; (c) Individuals will react emotionally or describe their feelings as emotions, to the extent they experience a state of physiological arousal. Notably, physiological arousal is not only merely as a cognitive cue, but also a crucial mediator of emotion (Goldstein et al., 1972).

A famous study inspired by this theory is the suspension bridge experiment, which is a study designed to induce misattributed arousal (Dutton and Aron, 1974).To trigger fear and anxiety of participants, they placed the experiment on a narrow suspension bridge that looked unstable and settled above a valley. The results showed that participants misattributed their fear and anxiety induced by that condition, to high sexual attraction toward the opposite sex.

Although the two-factor theory of emotion laid an important foundation for the study of emotion, its precision has been questioned. For example, some researchers indicated that there is limited experimental support for the first and second hypotheses, while the third proposition shows strong empirical validity (Manstead and Wagner, 1981). These controversies have led to further exploration of some key issues, including selecting mechanisms between competing cognitions and the characteristics of arousal.

To address this question, there is a study that conducted experiments on the interaction between cognition and physiological arousal, in an environment that lacks relevant contextual cues (Sinclair et al., 1994). Their results are consistent with Schachter's findings, which show that under conditions of physical activity and delayed scoring (lack of obvious arousal causes), participants gave extreme emotional self-ratings aligned with primed concepts. Conversely, under the conditions of immediate scoring, participants are able to explain their arousal.

For another example, recent studies have also explored digital content as a tool for addressing reproducibility issues (Pizzoli et al., 2020). They sent short videos to the participants and effectively manipulated arousal levels. Surprisingly, this study did not find significant differences in participants' perceived attractiveness ratings. Which means that they can not replicate the relationship between arousal and sexual attraction that was observed in previous studies.

Even nowadays, the two-factor theory of emotion still remains a topic of debate. In the past studies, due to the limitations of science and technology, they have to use complex real-world settings (Dutton and Aron, 1974) or pharmacological control method to induce arousal (Schachter and Singer, 1962). Although some later studies used physical activity as the condition of inducing arousal (Sinclair et al., 1994), it is still not simplify enough and easy to generalize. It has been proved that selecting appropriate stimuli for inducing emotional states remains significant (Marchewka et al., 2014).

Thus, it is of great significance to find a way to induce arousal simple and effective. In addition, further experiments are needed to explore and validate the concept of cognitive determination of emotional states.

2.2 Virtual Reality and Fear-Induced Physiological Arousal

Virtual Reality (VR) can offer realistic experiences through computer-generated environments. It can produce interactive experiences that are similar to those in the real world. Users can actively engage with virtual scenarios or directly interact with other virtual agents (Chicchi Giglioli et al., 2017). In recent years, the applications of VR in psychological research have grown significantly, particularly in studies involving emotion or emotion-driven interaction experiment design (Somarathna et al., 2022).

Specifically, VR attracted much attention as an effective tool for inducing arousal. For example, a study compared emotional responses under four different stimuli: 2D images, 360-degree panoramic VR environments, 3D VR scenarios, and physical environments. The results showed that VR environments were significantly more effective in inducing emotions than 2D stimuli (Marín-Morales et al., 2018).

Fear, a negative emotion, typically arises when individuals perceive or face dangers (Lin, 2017). The existing researches have used specific way to design fear-inducing experiments, such as the suspension bridge experiment, which successfully induced physiological arousal by creating fear and anxiety in a high-altitude setting (Dutton and Aron, 1974).

Beyond real-world settings, fear can also be induced through other media. For example, media like videos or movies, allow individuals to observe narratives passively, while video games, which are active media that require participants to make decisions and engage directly. This distinction makes active media more effective in inducing physiological arousal (Lin, 2013).

Video games can run on multiple media formats, such as screens like televisions or VR. However, VR offers unique advantages. A study compared emotional activation levels between VR and non-immersive display modes (console systems) using the same horror game (Pallavicini et al., 2018). Results showed that VR participants reported significantly stronger presence and emotional arousal than those using non-immersive systems.

However, existing research still has not fully explored the effectiveness of VR in arousal. Studies also overlook VR's potential for validating the two-factor theory of emotion. Compared to traditional methods of inducing arousal, VR offers highly controlled and safe experimental environments. In addition, these environments are easily reproducible and scalable. Even compared to non-immersive media, VR still has its advantages. Thus, using VR games to study physiological arousal and the phenomenon of misattribution remains significant.

Chapter 3

Material and Methods

This study was approved by the KESC Ethics Review Committee of Japan Advance Institute of Science and Technology (Approval Code: KSEC-A2024111504).

3.1 Participants

A total of 34 participants were involved in the experiment. They are all master's and doctoral students from Japan Advanced Institute of Science and Technology (JAIST) in Japan, Ishikawa. Pre-experiment questionnaires collected the following demographic information about participants:

Gender

- 20 males (58.8%)
- 14 females (41.2%)

Age Groups

- 1 participant in their 10s (2.9%)
- 28 participants in their 20s (82.4%)
- 5 participants in their 30s (14.7%)

All participants met the required inclusion criteria, which included no history of heart disease, epilepsy, or other chronic conditions sensitive to stress or surprises, as well as no symptoms of dizziness, migraines, anxiety, or phobias. In addition, most of the participants have experience using VR:

VR experience

- 14 participants had used VR more than three times. (41.2%)
- 16 participants had used VR 1–2 times. (47.1%)
- 4 participants had no prior VR experience. (11.8%)

Participants were also asked about their experience with horror games and their general attitude toward them till now:

Horror Games experience

- 23 participants had little or no experience. (67.6%)
- 7 participants had limited experience. (20.6%)
- 4 participants frequently played horror games. (11.8%)

Acceptance of Horror Games

- 8 participants were very willing to play. (23.5%)
- 4 participants relatively acceptable. (11.8%)
- 10 participants had a neutral attitude. (29.4%)
- 8 participants were slightly averse. (23.5%)
- 4 participants completely rejected to play. (11.8%)

3.2 Measures

I adopted a comprehensive measurement method to collect data from participants, which included three stages: background information collection before the experiment (covering the basic demographic information and relevant experiences of the participants), real-time physiological data recording during the experiment(heart rate, skin temperature, and P-P Interval changes), and emotions and subjective experiences after playing VR game(using questionnaires to collect both qualitative and quantitative data).

Before the experiment, participants were asked to fill out a survey on gender (male/Female/Non-binary), age, VR experience, and horror game experience. These variables were selected to analyze how background and experience might affect their physiological and emotional responses in the experiment. For instance, gender and age could influence fear-related physiological responses, while prior VR or horror game exposure might influence how participants adapted to and reacted to the horror content in the game.

In the experiment, participants need to wear a smartwatch-like device to monitor physiological data. This device was lightweight, which can minimize the interference of weight. When participants play the VR game, this device will keep recording heart rate, skin temperature, and P-P Interval (PPI) data. The heart rate and skin temperature data can provide a straightforward indication of the intensity of participants' physiological arousal during the gameplay. Moreover, PPI data can capture more complex physiological patterns, which can reflect the level of physiological stress that participants experienced during the experiment.

After finishing VR gameplay, participants need to fill out a questionnaire designed to evaluate their subjective emotional experiences and overall impressions of the experiment. The questionnaire included the following content: 1. Rating for their subjective physiological response to the VR survive horror game on a 1-5 scale (1 = no such sensation, 5 = very strong sensation), which includes: increased heart rate, increased sweating, trembling and accelerated breathing; 2. Rating for their subjective emotional reactions while playing VR game on a 1-5 scale (1 = no such sensation, 5 = very strong sensation), which includes: calmness, excitement, shock, courage, disgust, panic, fear, and anger; 3. Rating for the realism and immersion of the VR experience on 1-5 scale (1 = no such presence, 5 = very strong presence); 4. Rating for the overall game experience on 1-5 scale (1 = no such sensation, 5 = very strong sensation), which includes: realism of the VR environment, level of fear during gameplay, level of lingering fear after gameplay, excitement after gameplay, willingness to play VR survival horror game in the future.

3.3 Apparatus

This study utilized the Meta Quest 3 device (3.1) to provide VR experience for participants, which included a head-mounted display and two wireless handheld controllers. Additionally, VR survival horror game was been preloaded on a DELL G15 5515 laptop, which served as the gaming platform. The VR survival horror game is a commercially available game on Steam.

For the pre-experiment questionnaire, participants need to complete online using their personal devices, such as smartphones or computers. This questionnaire was created through Google Forms.

For the post-experiment questionnaire, participants need to fill it on-site with paper and pen.

Physiological data during the experiment were collected using a TDK-Silmee W22 smartwatch device. Data export and analysis were performed on an HP 470 G7 laptop.

Figure 3.1: Virtual Reality device used in this experiment

3.4 Experiment Design

This study aimed to use VR survival horror game as the experimental environment to induce physiological arousal. By providing pre-set emotional labels through suggestive emotional cues, to explore the phenomenon of misattribution of arousal induced by VR gameplay. Participants were randomly assigned to three different groups, which included two experimental groups and one control group. The experiment groups were divided into the Positive Emotion Cue Group and the Negative Emotion Cue Group, While the control group did not receive any emotional cue during the experiment.

3.4.1 Environment

Specific conditions were designed to trigger a fear-induced physiological arousal in participants. The experiment was conducted in a VR survival horror game environment. Participants need to sit on a chair in the center of the room, the place surrounded by the chair has been cleared and left nothing. At the start of the experiment, participants need to wear a head-mounted display and enter a pre-loaded survival horror game scenario.

The VR survival horror game is a level-based game. The objective of each level was to survive for 5 minutes in a virtual environment containing elements of fright and survival. After completing the current level, participants will be asked to advance to the next level; if they failed at any level, they were required to retry. To ensure sufficient data collection time without imposing excessive burden on participants, the VR experience phase will end after participants have completed four level attempts, regardless of success or failure. The entire VR experience phase was expected to last 15 to 20 minutes.

3.4.2 Procedure

Within this environment, participants were randomly divided into three groups, each subjected to different intervention conditions to test the hypotheses. The groups included the Positive Emotion Cue Group, the Negative Emotion Cue Group, and the control group (no emotional cues).

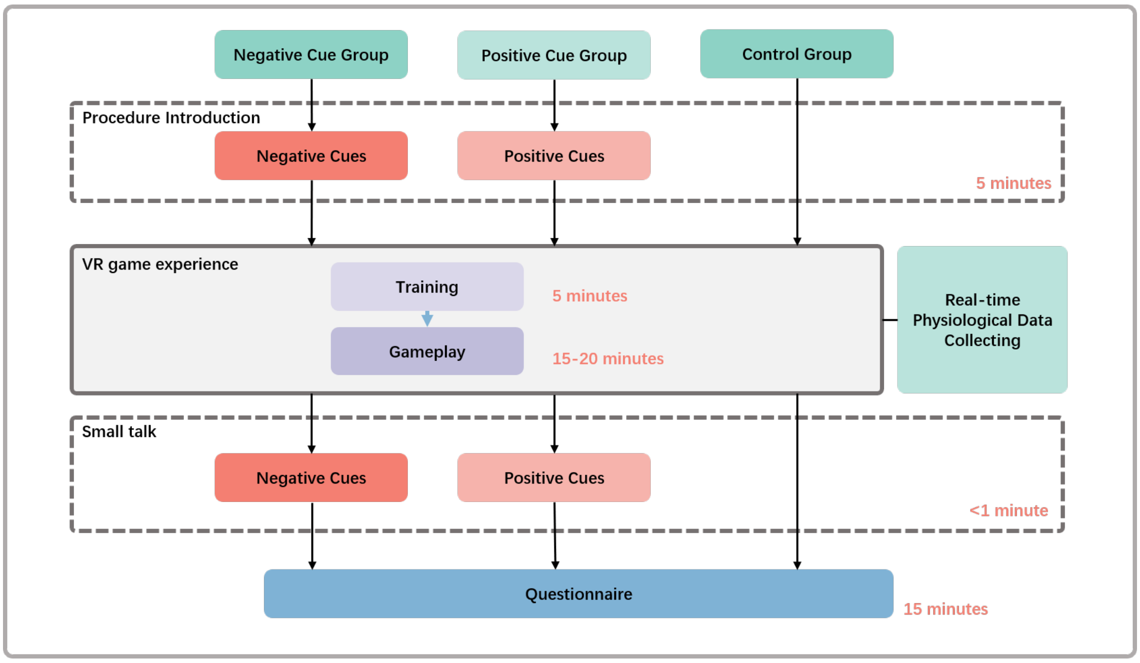

Figure 3.2: Experiment Procedure

The basic experimental procedure is shown in Figure 3.2:
For the pre-experiment preparation, participants receive a brief training session, which includes an introduction to the experimental procedure and training on equipment usage (5-10 minutes). Specifically, covered how to wear and operate the head-mounted display, controller.
Before the formal gameplay began, participants need to follow the instructions to start a 2-4 minute training. During the training process, participants can learn the basic operations and the way to control their character in the game, moreover, the game rules and the condition to end this phase. After completing the training, participants begin to play this VR game.
In the experiment phase, participants enter the experiment environment, and wear the head-mounted display to play the VR survival horror game. During this phase, the smartwatch device can collect real-time data on participants' heart rate, skin temperature, and PPI. After the phase to experience VR game, all participants are asked to complete a post-experiment questionnaire.

3.4.3 Experiment Details

To investigate the impact of emotional cues on physiological arousal, this study implemented different experimental procedures for participants in three groups: the Positive Emotion Cue group, the Negative Emotion cue group, and the Control Group (without emotional cues). The specific procedures for each group are as follows:

Negative Emotion Cue Group

Anger is been selected as the target of the negative emotional cues. Participants receive emotional cues during the introductory phase of the experiment.The experimenter interacts with participants through casual conversations, while providing misleading words for trying to induce misattribution of physiological arousal. For instance, the experimenter might say, "Don't worry, the upcoming horror game won't be too scary. Instead your body might exhibit signs of anger, such as feeling your heart pounding or your fists clenching subconsciously. This means that your body is helping you cope with the scary environment by intensifying aggression."

To reinforce this cue, the experimenter supplements the interaction with a story: "While playing this game, the previous participant initially panicked, but after plays for a while, he suddenly started swearing. He tells me that after doing that he feels less afraid. Maybe you could try it while playing." Participants are then given a processed version of the experimental instruction paper, which conceals the true purpose of the experiment. The final section of this instruction paper contains a specific negative emotional cue text, highlighted in bold red: "If you make violent movements because of anger during the experiment, please be careful not to damage the experimental equipment". Participants are asked to read the whole paragraph containing this text carefully.

After playing VR game, when participants remove the head-mounted display, the experimenter will keep reinforcing cues through casual talk. For example, "I saw you almost punch at that enemy!" or "It looked like you were about to swear." The experimenter then briefly explains that these behaviors could be manifestations of anger and aggression while tidying up. Making participants think that this is just small talk outside of this experiment. After this, participants are naturally guided to the post-experiment questionnaire phase.

Positive Emotion Cue Group

Courage is been selected as the target of the positive emotional cues. Similar to the group of negative emotional cues, the experimenter will reinforce positive emotional cues during the whole experiment. For the casual conversations during the introduction phase, the experimenter might say, "Don't worry, the upcoming horror game won't be too scary. Instead your body might exhibit signs of courage, such as feeling your heart pounding or your fists clenching subconsciously. This means that your body is helping you cope with the scary environment by intensifying aggression." In addition, a story like, "While playing this game, the previous participant initially panicked, but after playing for a while, he seems already adapted this game. He tells me that he felt his body get warmed up and seemed full of courage and no longer afraid. See, your body will help you to face this challenge."

Moreover, the highlighted text is replaced with words like, "Bravely go for it, your body will help you to face danger."

For the post-experience part, the content of small talk also changed (e.g., "You seemed fearless in the later stages" or "That movement was so brave!") with an exciting and high voice.

Control Group

During the whole experiment, participants in this group do not receive any emotional cues. The experimenter only provides basic operational instructions, and avoids any talking outside this experiment or related to topics of emotions. The highlighted text has also been removed from the instruction paper. Besides this, all other procedures are the same as the previous groups.

Chapter 4

Result

4.1 Demographics

To ensure there were no significant differences in participant characteristics among the three experimental groups (Positive Emotion Cue Group, Negative Emotion Cue Group, and Control Group), I conducted statistical analyses on **gender**, **age**, **VR usage experience**, and **horror game experience**.

4.1.1 Gender

A Chi-square test of independence was performed to analyze gender distribution (male and female only) across the three groups. The results indicate no significant association between group assignment and gender ($\chi^2 = 1.33$, $p = 0.514$, $df = 2$).

Table 4.1: Gender Distributions

Group	Male	Female
Positive Emotion Cue	7	6
Negative Emotion Cue	5	5
Control	8	3

4.1.2 Age

A Chi-square test of independence was also conducted to examine the distribution of participants' age across the three groups. The results show no significant differences ($\chi^2 = 5.19$, $p = 0.269$, $df = 4$).

4.1.3 VR Usage Experience

A Chi-square test of independence was performed to analyze participants' VR usage experience. No significant differences were observed among the groups ($\chi^2 = 3.63$, $p = 0.459$, $df = 4$).

Table 4.2: Age Distribution

Group	10s	20s	30s	40s
Positive Emotion Cue	1	12	0	0
Negative Emotion Cue	0	8	2	0
Control	0	8	3	0

Table 4.3: VR Usage Experience

Group	Low	Medium	High
Positive Emotion Cue	2	7	4
Negative Emotion Cue	1	6	3
Control	1	3	7

4.1.4 Horror Game Experience

Similarly, the distribution of horror game experience was analyzed using a Chi-square test. The results suggest no significant association between group assignment and experience level ($\chi^2 = 8.59$, $p = 0.378$, $df = 8$).

Table 4.4: Horror Game Experience

Group	Lv.1	Lv.2	Lv.3	Lv.4	Lv.5
Positive Emotion Cue	5	4	3	1	0
Negative Emotion Cue	6	1	2	0	1
Control	2	5	2	0	2

4.1.5 Horror Game Acceptance

Participants' acceptance levels for horror games were measured on a 1–5 scale. To analyze potential differences across the three groups, a one-way analysis of variance (ANOVA) was conducted.

First, the assumption of normality was tested using the Shapiro-Wilk test. The results indicated that the data for each group followed a normal distribution (positive Emotion Cue Group: $p = 0.090$, Negative Emotion Cue Group: $p = 0.073$, Control Group: $p = 0.170$). Next, Levene's test was used to assess the homogeneity of variance, which was satisfied ($p = 0.798$).

The ANOVA revealed no significant differences in horror game acceptance levels among the three groups ($F = 0.736$, $p = 0.487$). As the ANOVA results were not significant, post-hoc tests were not performed.

The results from Chi-square tests and one-way ANOVA indicate no significant differences in demographic characteristics or baseline measures among the three experimental groups. Which is confirms that participants are comparable across three different conditions, ensuring the validity of subsequent analyses.

Table 4.5: ANOVA Results for Horror Game Acceptance

Source	Sum of Squares	df	F-value	p-value
Group	2.70	2	0.736	0.487
Residual	56.83	31	-	-

4.2 Subjective Judgments of Physiological Responses

A one-way analysis of variance (ANOVA) was conducted to analyze participants' subjective judgments of their physiological responses across the three experimental conditions (Negative Emotion Cue Group, Positive Emotion Cue Group, and Control Group). The responses analyzed included perceptions of increased heart rate, perspiration, muscle tension, and tachypnea.

4.2.1 Increased Heart Rate

Normality was assessed using the Shapiro-Wilk test, which indicated non-normality for all groups (Negative Emotion Cue Group: $p = 0.0433$; Positive Emotion Cue Group: $p = 0.0359$; Control Group: $p = 0.0166$). Despite this, Levene's test confirmed homogeneity of variances ($p = 0.4347$).

The ANOVA results showed no significant differences in subjective judgments of increased heart rate among the groups ($F = 0.795$, $p = 0.461$). Considering the same experiment condition of the three groups, this result suggests that suggestive cues will not influence the judgment of participants' responses to increased heart rate.

Table 4.6: ANOVA Results for Subjective Judgment of Increased Heart Rate

Source	Sum of Squares	df	F-value	p-value
Group	1.38	2	0.795	0.461
Residual	26.89	31	-	-

4.2.2 Perspiration

Normality was tested, and the results indicated non-normality for the Negative Emotion Cue Group ($p = 0.0469$) but normality for the Positive Emotion Cue Group and Control Group ($p = 0.0947$ and $p = 0.1651$, respectively). Levene's test confirmed the homogeneity of variances ($p = 0.4696$).

The ANOVA results showed no significant differences in subjective judgments of perspiration among the groups ($F = 0.279$, $p = 0.758$). Similar to the previous stage, it was confirmed that there is no association between ratings of perspiration level and suggestive cues.

Table 4.7: ANOVA Results for Subjective Judgment of Perspiration

Source	Sum of Squares	df	F-value	p-value
Group	0.79	2	0.279	0.758
Residual	43.83	31	-	-

4.2.3 Muscle Tension

The Shapiro-Wilk test indicated normality for all groups (negative Emotion Cue Group: $p = 0.0731$; Positive Emotion Cue Group: $p = 0.1559$; Control Group: $p = 0.2367$), and Levene's test confirmed homogeneity of variances ($p = 0.9527$).

The ANOVA results showed no significant differences in subjective judgments of muscle tension among the groups ($F = 0.219$, $p = 0.804$). The rating of muscle tension among participants will not be affected by cues.

Table 4.8: ANOVA Results for Subjective Judgment of Muscle Tension

Source	Sum of Squares	df	F-value	p-value
Group	0.87	2	0.219	0.804
Residual	61.25	31	-	-

4.2.4 Tachypnea

Normality was tested for the Negative Emotion Cue Group and Control Group, which satisfied the assumption ($p = 0.2398$ and $p = 0.0934$, respectively). Due to insufficient sample size, normality and variance homogeneity could not be tested for the Positive Emotion Cue Group.

The ANOVA results revealed no significant differences in subjective judgments of tachypnea among the groups ($F = 2.277$, $p = 0.146$).

Table 4.9: ANOVA Results for Subjective Judgment of Tachypnea

Source	Sum of Squares	df	F-value	p-value
Group	2.80	1	2.277	0.146
Residual	27.03	22	-	-

The results indicate no statistically significant differences in subjective judgments of physiological responses among the three experimental groups.

4.2.5 Descriptive Statistics of Subjective Judgments

Descriptive statistics were conducted for participants' subjective judgments of their physiological responses, including increased heart rate, perspiration, muscle tension, and tachypnea. To illustrate the overall trends, the mean scores for these four metrics are visualized in Figure 4.1. Additionally, to reflect the variation in scores across conditions, including standard deviations and medians, boxplots are presented in Figure 4.2.

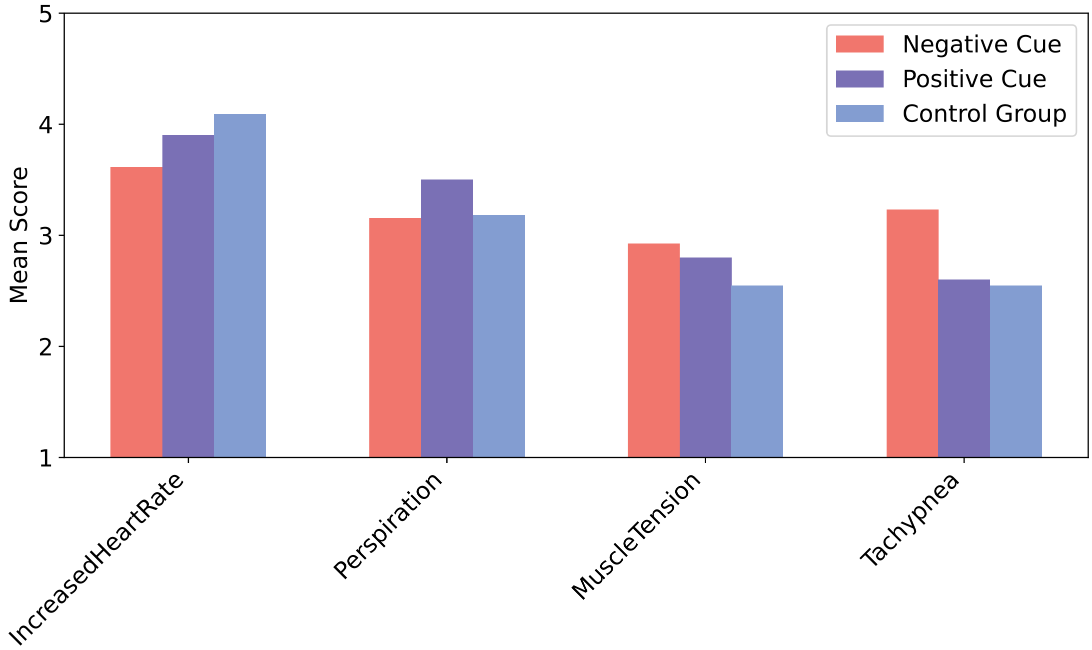

Figure 4.1: Mean Scores of Physiological Responses Across Experimental Groups

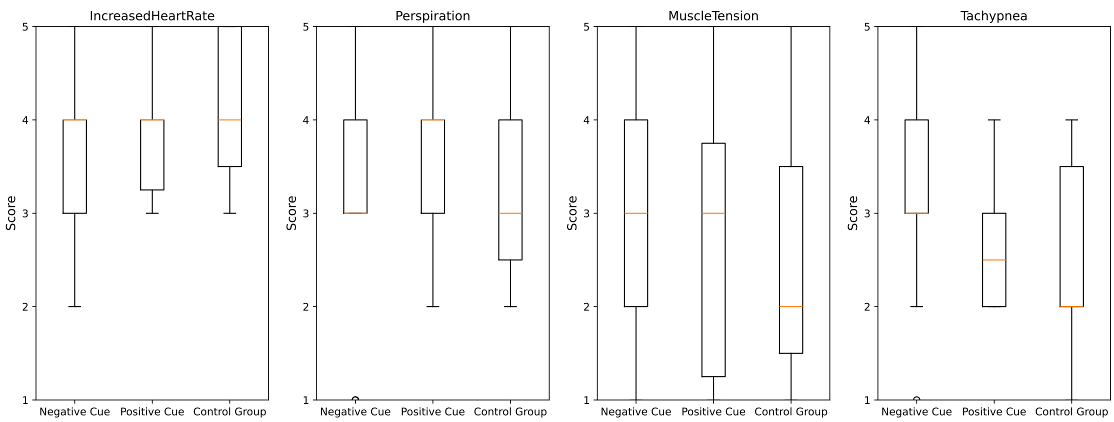

Figure 4.2: Boxplots of Physiological Response Scores Across Experimental Groups

As shown in Figure 4.1, the mean scores for increased heart rate, perspiration, muscle tension, and tachypnea do not shows notable differences across the three experimental conditions, which aligns with the results of the ANOVA tests.

Figure 4.2 provides additional information on the variability of scores within each group. The boxplots reveal relatively consistent distributions across the conditions, with no group demonstrating particularly high variability. Median scores are similar among those groups. Additionally, increased heart rate received higher mean scores compared to the other metrics.

4.3 Physiological Data

To objectively assess the extent of physiological arousal experienced by participants under the three experimental conditions while playing the horror game, descriptive statistics were conducted on physiological measures (heart rate, skin temperature, and PPI) collected via the smartwatch device.

4.3.1 Heart Rate

The descriptive statistics for participants' heart rate, including the mean (HR_Mean), maximum(HR_Max), minimum(HR_Min), and standard deviation (HR_STD) under the three conditions, are summarized in Figure 4.3. The variability in heart rate is further illustrated using boxplots in Figure 4.4.

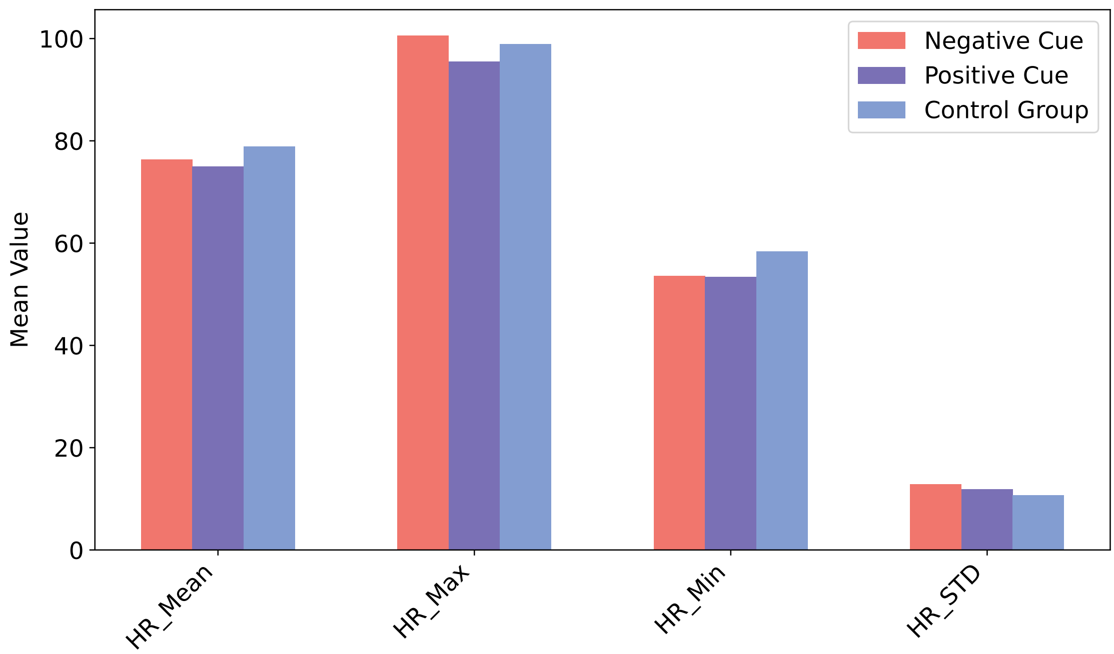

Figure 4.3: Comparison of Mean Heart Rate Metrics Across Groups

Table 4.10: ANOVA Results for Heart Rate (HR_Mean)

Source	Sum of Squares	df	F-value	p-value
Group	86.08	2	0.209	0.813
Residual	6395.98	31	-	-

4.3.2 Skin Temperature

The descriptive statistics for participants' skin temperature, including the mean (ST_Mean), maximum (ST_Max), minimum (ST_Min), and standard deviation (ST_STD)

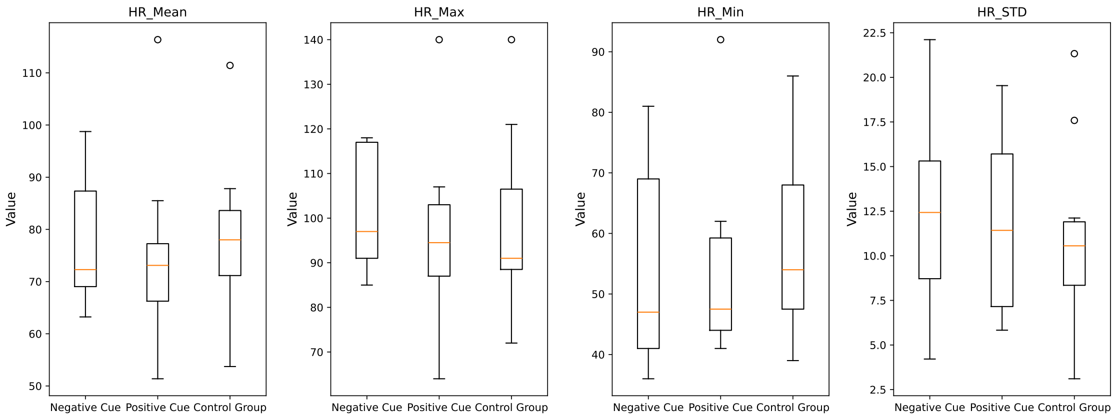

Figure 4.4: Boxplots of Heart Rate Metrics Across Groups

under the three conditions, are summarized in Figure 4.5. The variability in skin temperature is further illustrated using boxplots in Figure 4.6.

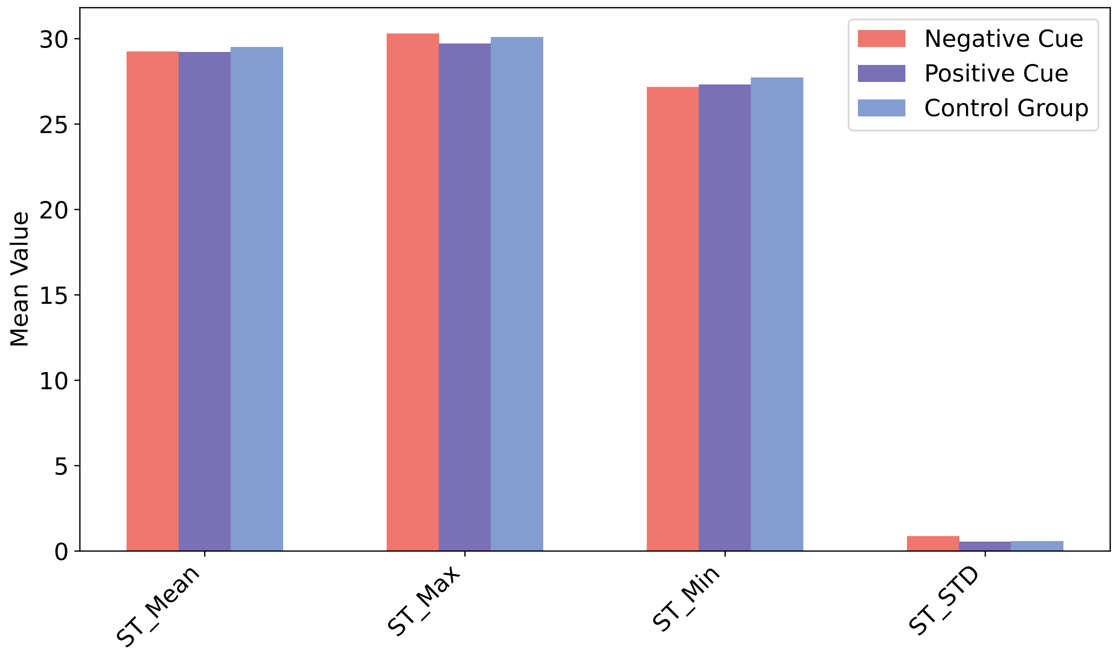

Figure 4.5: Comparison of Mean Skin Temperature Metrics Across Groups

4.3.3 Analysis of Differences Between Groups

To objectively compare the physiological data across the three groups, one-way ANOVA was conducted on the mean values of heart rate, skin temperature.

For heart rate (HR_Mean), the normality test (Shapiro-Wilk) indicated that all groups followed a normal distribution ($p = 0.0887, 0.1216, 0.4548$), and Levene's test confirmed homogeneity of variances ($p = 0.8963$). The ANOVA results revealed no significant differences ($F = 0.209, p = 0.813$).

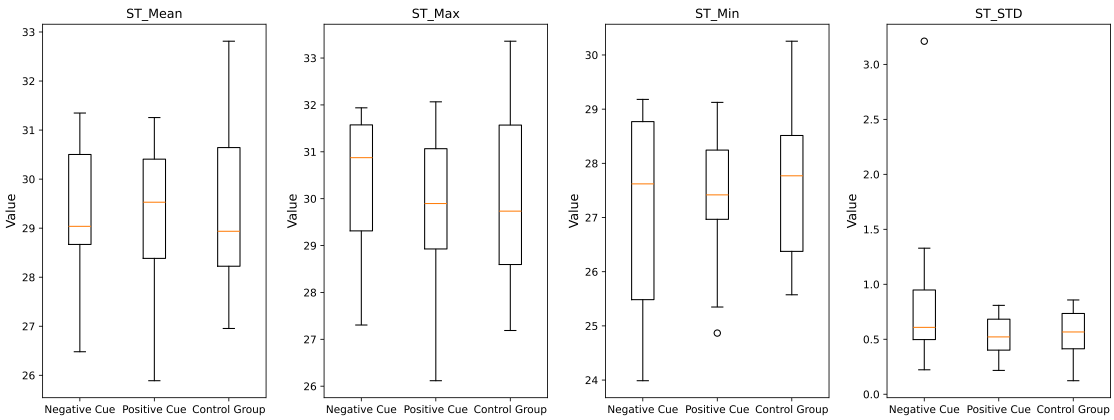

Figure 4.6: Boxplots of Skin Temperature Metrics Across Groups

Table 4.11: ANOVA Results for Skin Temperature (ST_Mean)

Source	Sum of Squares	df	F-value	p-value
Group	0.59	2	0.105	0.901
Residual	87.22	31	-	-

For skin temperature (ST_Mean), all groups satisfied the normality assumption ($p = 0.3763, 0.6412, 0.7392$) and homogeneity of variances ($p = 0.9020$). The ANOVA results were not significant ($F = 0.105, p = 0.901$).

4.3.4 Trends in Physiological Data

The trends in heart rate and skin temperature have also been conducted. The changes for each participant's heart rate and skin temperature were generated and shown as line graphs, moreover, combined into a single plot for each group. For the Negative Cue Group, results are shown in Figure 4.7; Figure 4.8 is the results of the Positive Cue Group; Figure 4.9 is the trends of the Control Group.

The horizontal axis represents the measurement point of time, while the vertical axes represent the skin temperature and heart rate. Different colors correspond to individual participants within each group, and the numbers indicate participant IDs.

All the results show that most of the participants exhibited a stable or gradually increasing trend in skin temperature over time of the VR experience. Notably that some anomalous trends, such as a significant decline (e.g, the orange line with number of 7 in Figure 4.7), were identified as missing data caused by the issues of the device. Though there are some slight differences in temperature levels between participants, the range of variation remained small. For the heart rate data revealed in these figures, there are noticeable fluctuations and significant individual differences among participants. Most curves of heart rate did not follow a single trend but instead showed considerable variability. This variability is likely associated with the changing pace of the varying fear conditions in the survival horror game experienced

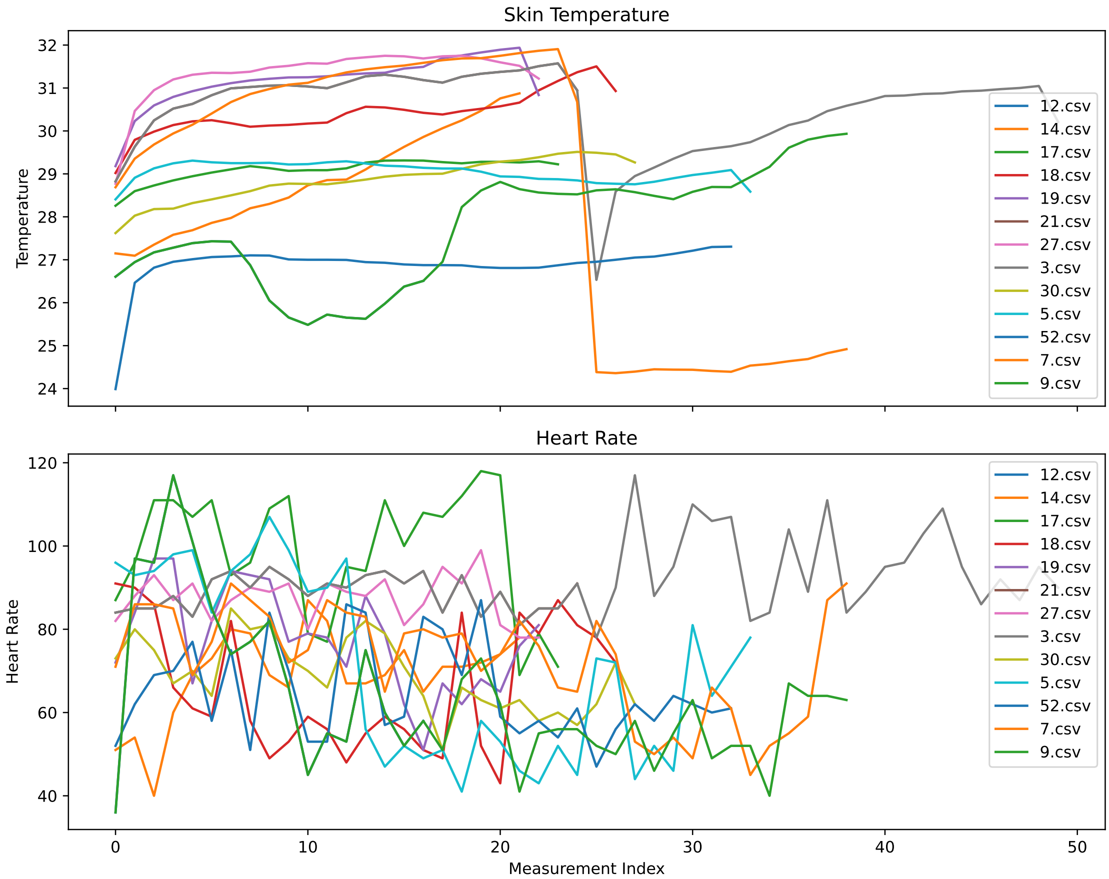

Figure 4.7: Trends of Heart Rate and Skin Temperature in Negative Cue Group

by participants. Some of the curves disappeared during later stages of the timeline were because of the data interruptions and anomalies caused by the data collecting equipment.

4.4 PPI Data Analysis

PPI (Peak-to-Peak Interval) data were analyzed to observe the frequency domain characteristics of heart rate variability. Frequency domain analysis helps explain the autonomic nervous system's regulation of physiological states.

I attempted to perform a consistency test within each group using Intraclass Correlation Coefficient (ICC) analysis to evaluate whether those data within each group were consistent. However, due to the lengths of collected data that were inconsistent, the ICC method was affected and failed to reflect the consistency of physiological responses.

For this reason, I further processed the data and found that PPI data collected from the experiment remains several issues due to the device limitations. Some data of the participants contained data recording durations that were too short, or the number of valid data points was too low. Additionally, the scarcity of PPI data led

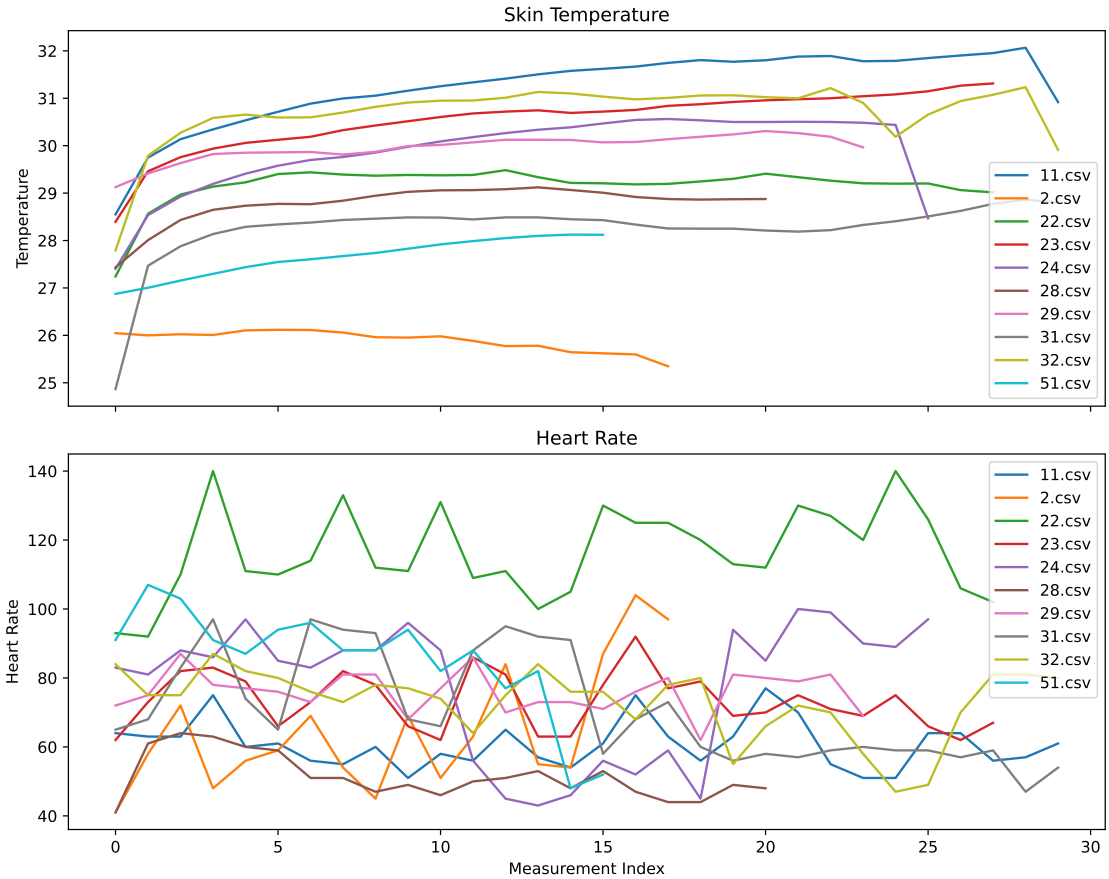

Figure 4.8: Trends of Heart Rate and Skin Temperature in Positive Cue Group

to inadequate frequency resolution in both low frequency and high frequency part. Additionally, more than one-third of the data contained only null values or extreme values. As a result, the analysis of PPI data has been abandoned.

4.5 Emotion Attribution Analysis for Anger

The subjective emotion attribution scores for *Anger* under the three experimental conditions were analyzed. Descriptive statistics are presented in Figure 4.10 and Figure 4.11.

4.5.1 ANOVA Results for Anger Attribution Scores

To compare differences in *Anger* attribution scores among the three groups, a one-way ANOVA was conducted. The results are summarized in Table 4.12.

The ANOVA results indicate a statistically significant difference in *Anger* attribution scores among the three groups ($F = 4.633, p = 0.017$). Therefore, a post-hoc Tukey HSD test was performed to identify specific group differences.

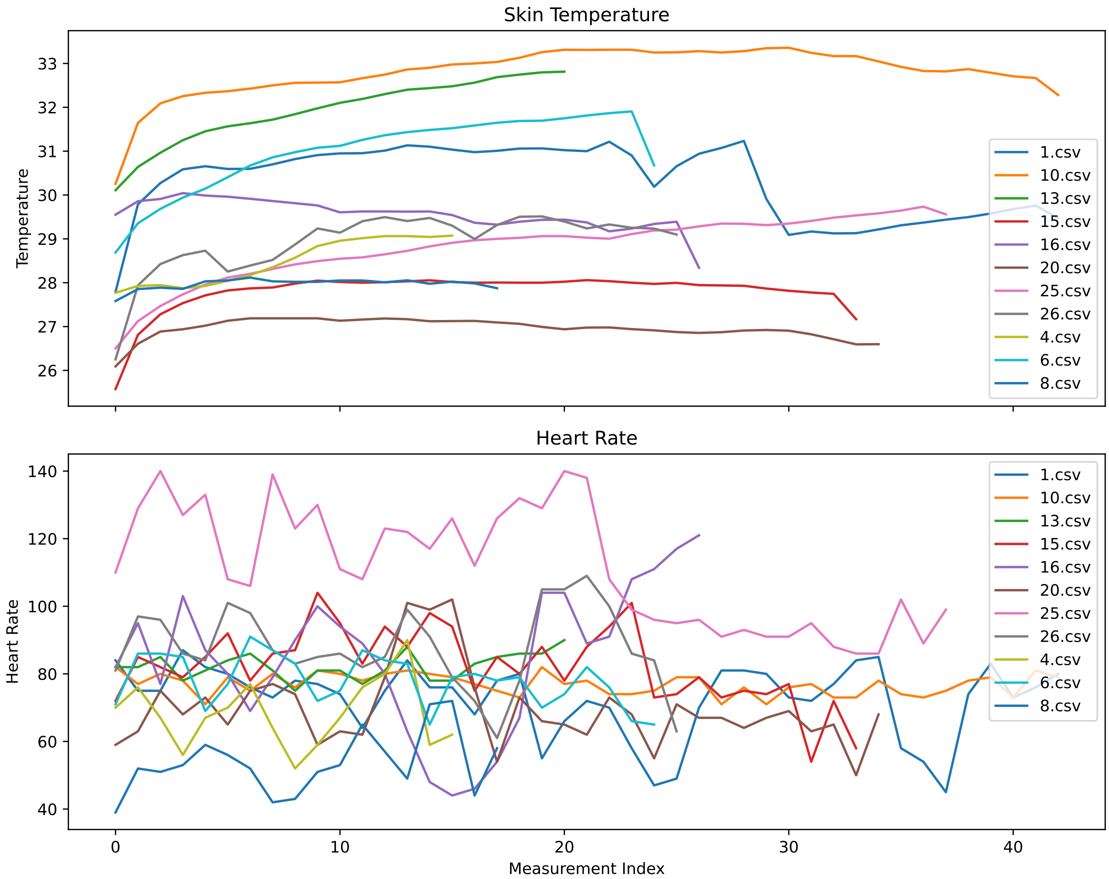

Figure 4.9: Trends of Heart Rate and Skin Temperature in Control Group

Table 4.12: ANOVA Results for Anger Attribution Scores

Source	Sum of Squares	df	F-value	p-value
Group	10.02	2	4.633	0.017
Residual	33.51	31	-	-

4.5.2 Post-hoc Analysis (Tukey HSD)

The results of the Tukey HSD test are shown in Table 4.13.

The results of the Tukey HSD post-hoc analysis showed a significant difference in *Anger* attribution scores between the Negative Cue Group and the Control Group ($p = 0.015$), indicating that the negative cue had a notable effect on *Anger* attribution compared to the control condition. However, no significant differences were observed between the Negative Cue Group and the Positive Cue Group ($p = 0.151$), or between the Positive Cue Group and the Control Group ($p = 0.619$).

4.5.3 Kruskal-Wallis Test

Given that the Positive Cue Group and Control Group did not meet the normality assumption, a non-parametric Kruskal-Wallis test was conducted to validate the

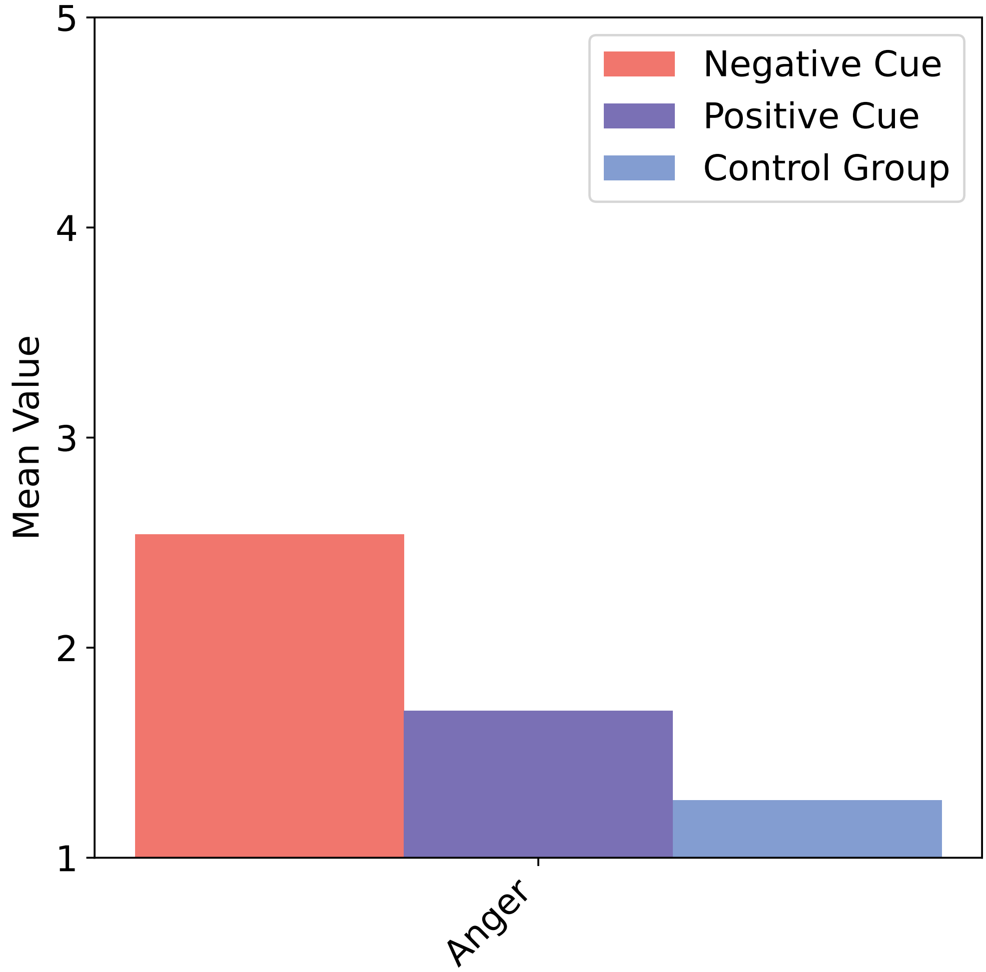

Figure 4.10: Mean *Anger* Attribution Scores Across Groups

ANOVA results. The results are summarized in Table 4.14.

The Kruskal-Wallis results confirm the significant difference between the Negative Cue Group and the Control Group ($p = 0.0059$), consistent with the Tukey HSD results. Other group comparisons did not show significant differences.

4.6 Pearson Correlation Analysis Between Subjective Physiological Response Ratings and Anger Attribution

To explore the relationship between subjective physiological responses and emotional attributions (*Anger*), a correlation analysis was conducted for the three experimental groups. The results are presented in Tables 4.15, 4.16, and 4.23.

In the Negative Emotion Cue Group, Increased Heart Rate showed a significant positive correlation with *Anger* ($r = 0.598, p = 0.031$). Muscle Tension and

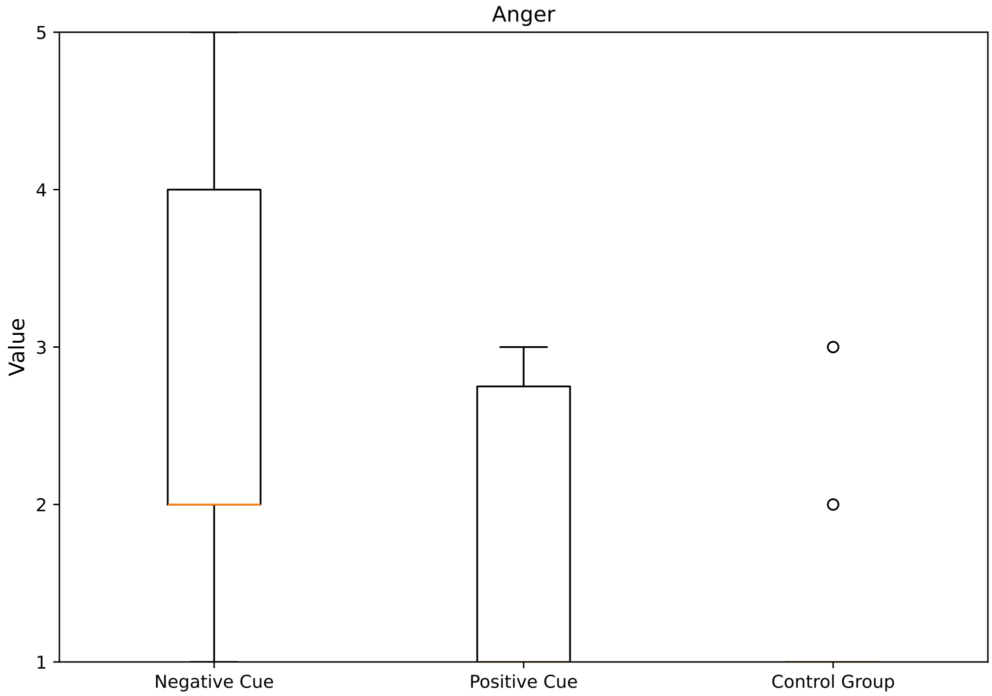

Figure 4.11: Boxplots of *Anger* Attribution Scores Across Groups

Table 4.13: Tukey HSD Results for Anger Attribution Scores

Group 1	Group 2	Mean Difference	p-value	Lower Bound	Upper Bound
Negative Cue	Positive Cue	-0.839	0.151	-1.915	0.238
Negative Cue	Control	-1.266	0.015	-2.314	-0.217
Positive Cue	Control	-0.427	0.619	-1.545	0.691

Tachypnea demonstrated moderate positive correlations with *Anger* but did not reach statistical significance. Perspiration showed almost no correlation with *Anger*.

In the Positive Emotion Cue Group, no significant correlations were observed between *Anger* and any of the physiological variables. Tachypnea and Perspiration showed moderate positive correlations, but these were not statistically significant. Muscle Tension demonstrated a weak negative correlation ($r = -0.272$).

In the Control Group, Perspiration showed a marginal negative correlation with *Anger* ($r = -0.559, p = 0.074$), which approached significance. Other variables, including Increased Heart Rate, Muscle Tension, and Tachypnea, did not show significant correlations, and the directions of these correlations were inconsistent.

Table 4.14: Kruskal-Wallis Test Results for Anger Attribution Scores

Comparison	Statistic	p-value
Negative Cue Group vs. Positive Cue Group	2.623	0.105
Negative Cue Group vs. Control Group	7.572	0.0059
Positive Cue Group vs. Control Group	1.332	0.249

Table 4.15: Correlation Analysis Results for the Negative Emotion Cue Group

Variable	r-value	p-value	Significance
IncreasedHeartRate	0.598	0.031	Significant Positive Correlation ($p < 0.05$)
Perspiration	0.082	0.789	No Significant Correlation
MuscleTension	0.453	0.120	Marginal Positive Correlation (Not Significant)
Tachypnea	0.424	0.149	No Significant Correlation

4.7 Emotion Attribution Analysis for Courage

The subjective emotion attribution scores for *Courage* under the three experimental conditions were analyzed. Descriptive statistics are presented in Figure 4.12 and Figure 4.13.

4.7.1 ANOVA Results for Courage Attribution Scores

To compare differences in *Courage* attribution scores among the three groups, a one-way ANOVA was conducted. The results are summarized in Table 4.18.

The ANOVA results indicate a statistically significant difference in *Courage* attribution scores among the three groups ($F = 5.211, p = 0.0112$). Therefore, a post-hoc Tukey HSD test was conducted to identify specific group differences.

4.7.2 Post-hoc Analysis (Tukey HSD)

The results of the Tukey HSD test are shown in Table 4.19.

The Tukey HSD results showed that the *Courage* attribution scores of the Control Group were significantly lower than those of the Negative Cue Group ($p = 0.0440$) and the Positive Cue Group ($p = 0.0136$). However, no significant difference was found between the Negative Cue Group and the Positive Cue Group ($p = 0.7755$).

4.7.3 Kruskal-Wallis Test for Courage Attribution Scores

Given that the Positive Cue Group and Control Group did not meet the normality assumption, a Kruskal-Wallis test was conducted as a non-parametric alternative to verify group differences in *Courage* attribution scores. The results are summarized in Table 4.20.

Table 4.16: Correlation Analysis Results for the Positive Emotion Cue Group

Variable	r-value	p-value	Significance
IncreasedHeartRate	0.270	0.451	No Significant Correlation
Perspiration	0.422	0.225	No Significant Positive Correlation
MuscleTension	-0.272	0.447	No Significant Negative Correlation
Tachypnea	0.469	0.171	No Significant Correlation

Table 4.17: Correlation Analysis Results for the Control Group

Variable	r-value	p-value	Significance
IncreasedHeartRate	-0.237	0.483	No Significant Correlation
Perspiration	-0.559	0.074	Marginal Negative Correlation (Approaching Significance)
MuscleTension	0.154	0.651	No Significant Correlation
Tachypnea	0.187	0.582	No Significant Correlation

The Kruskal-Wallis test results revealed significant differences in *Courage* attribution scores between the Negative Cue Group and the Control Group ($p = 0.0310$), as well as between the Positive Cue Group and the Control Group ($p = 0.0033$). These findings confirm that the Control Group scored significantly lower in *Courage* attribution compared to both experimental groups. However, no significant difference was found between the Negative Cue Group and the Positive Cue Group ($p = 0.5012$).

4.8 Pearson Correlation Analysis Between Physiological Responses and Courage Attribution

To investigate the relationship between physiological responses and *Courage* attribution, Pearson correlation analyses were conducted for each experimental group. The results are summarized in Tables 4.21, 4.22, and 4.23.

In the Negative Emotion Cue Group, Perspiration showed a marginally significant positive correlation with *Courage* ($r = 0.514, p = 0.072$), suggesting that increased perspiration might be associated with higher *Courage* attribution. Increased Heart Rate exhibited a moderate negative correlation ($r = -0.435, p = 0.137$), although not significant, indicating a potential trend where higher heart rate might be linked to lower *Courage* perception. Other variables, including Muscle Tension and Tachypnea, showed weak correlations with *Courage*, none of which were significant.

In the Positive Emotion Cue Group, Increased Heart Rate displayed a moderate negative correlation with *Courage* ($r = -0.488, p = 0.153$), although not significant, indicating a potential trend where increased heart rate might be associated with reduced *Courage* attribution. Other physiological variables, such as Perspiration, Muscle Tension, and Tachypnea, showed weak and non-significant correlations with

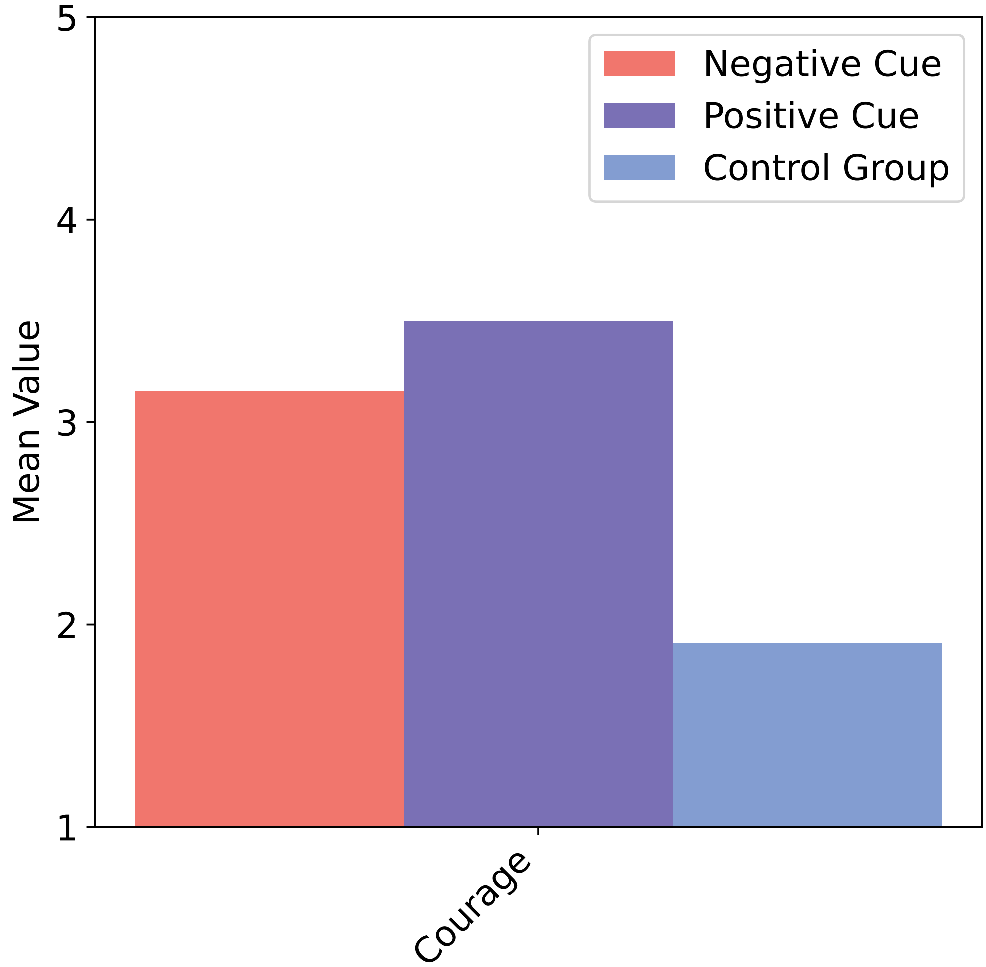

Figure 4.12: Mean *Courage* Attribution Scores Across Groups

Courage.

In the Control Group, Increased Heart Rate showed a moderate negative correlation with *Courage* ($r = -0.335, p = 0.314$), but this was not statistically significant. Other variables, including Perspiration, Muscle Tension, and Tachypnea, exhibited weak or no correlations with *Courage*, none of which were significant.

4.9 Emotion Attribution Analysis for Cues Unrelated Emotions (Calmness, Excitement, Shock)

Figure 4.14 presents the mean values of three measured emotional responses (Calmness, Excitement, and Shock) across the Negative Emotion Cue Group, Positive Emotion Cue Group, and Control Group. For the *Calmness* response, the Control Group displayed the highest mean value, while the Positive Emotion Cue Group and Negative Emotion Cue Group exhibited similar, lower values. Regarding *Excitement* the Positive Emotion Cue Group showed the highest mean value, followed by the

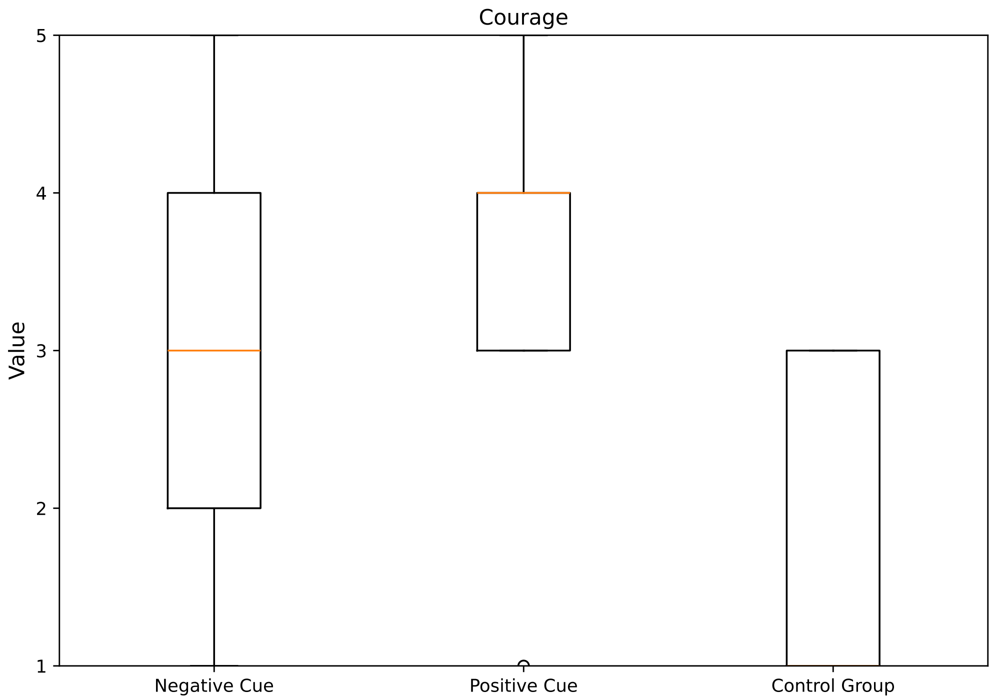

Figure 4.13: Boxplots of *Courage* Attribution Scores Across Groups

Table 4.18: ANOVA Results for *Courage* Attribution Scores

Source	Sum of Squares	df	F-value	p-value
Group	15.1633	2	5.211	0.0112
Residual	45.1014	31	-	-

Negative Emotion Cue Group, and the Control Group had the lowest value. For *Shock* the Control Group demonstrated the highest mean value, with the Negative Emotion Cue Group slightly lower, and the Positive Emotion Cue Group showing the lowest value.

Figure 4.15 further visualizes the distributions of these three emotional responses using box plots. For *Calmness* all groups exhibited relatively narrow ranges, with the Control Group having the highest median, and the Negative Emotion Cue Group showing several high outliers. Regarding *Excitement* the Negative Emotion Cue Group displayed a wide range with a relatively low median, while the Positive Emotion Cue Group had a narrower range with a higher median. The Control Group showed the narrowest range and the lowest median. For *Shock* the Control Group exhibited the widest range and the highest median, while the Positive Emotion Cue Group demonstrated the narrowest range and the lowest median. The Negative Emotion Cue Group's distribution was moderate, with a median between the other two groups.

Table 4.19: Tukey HSD Results for *Courage* Attribution Scores

Group 1	Group 2	Mean Difference	p-value	Lower Bound	Upper Bound
Negative Cue	Positive Cue	0.346	0.776	-0.903	1.595
Negative Cue	Control	-1.245	0.044	-2.461	-0.029
Positive Cue	Control	-1.591	0.014	-2.888	-0.294

Table 4.20: Kruskal-Wallis Test Results for *Courage* Attribution Scores

Comparison	Statistic	p-value
Negative Cue Group vs. Control Group	4.6553	0.0310
Positive Cue Group vs. Control Group	8.6338	0.0033
Negative Cue Group vs. Positive Cue Group	0.4524	0.5012

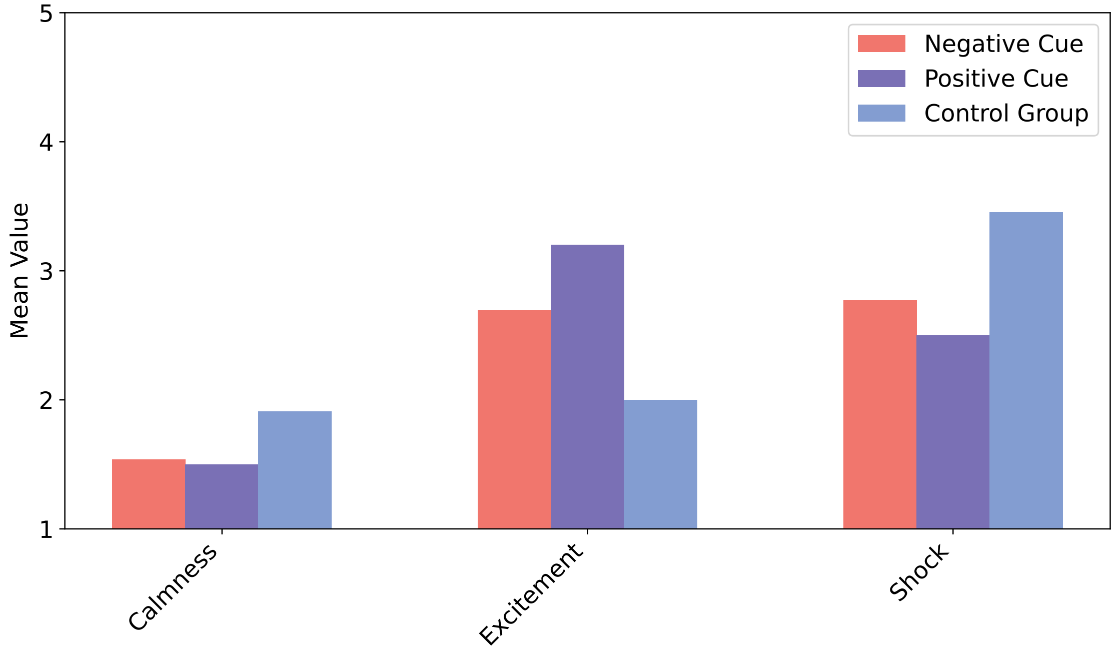

Figure 4.14: Mean values of *Calmness*, *Excitement*, and *Shock* across the three groups.

Calmness The Shapiro-Wilk test indicated non-normality in all groups (Negative Emotion Cue Group: $p = 0.0002$, Positive Emotion Cue Group: $p = 0.0021$, Control Group: $p = 0.0007$). Levene's test confirmed homogeneity of variances ($p = 0.6094$). ANOVA results were not significant ($F(2, 31) = 0.503, p = 0.6094$), indicating no significant differences among the groups.

Excitement The Shapiro-Wilk test showed non-normality for all groups (Negative Emotion Cue Group: $p = 0.0425$, Positive Emotion Cue Group: $p = 0.0001$, Control Group: $p = 0.0033$). Homogeneity of variances was satisfied ($p = 0.6079$). ANOVA results were not significant ($F(2, 31) = 1.904, p = 0.1660$).

Table 4.21: Pearson Correlation Analysis for the Negative Emotion Cue Group

Variable	r-value	p-value	Significance
IncreasedHeartRate	-0.435	0.137	Moderate Negative Correlation (Not Significant)
Perspiration	0.514	0.072	Marginal Positive Correlation (Approaching Significance)
MuscleTension	-0.218	0.475	Weak Negative Correlation (Not Significant)
Tachypnea	-0.242	0.425	Weak Negative Correlation (Not Significant)

Table 4.22: Pearson Correlation Analysis for the Positive Emotion Cue Group

Variable	r-value	p-value	Significance
IncreasedHeartRate	-0.488	0.153	Moderate Negative Correlation (Not Significant)
Perspiration	0.053	0.885	No Correlation
MuscleTension	0.266	0.458	Weak Positive Correlation (Not Significant)
Tachypnea	0.294	0.409	Weak Positive Correlation (Not Significant)

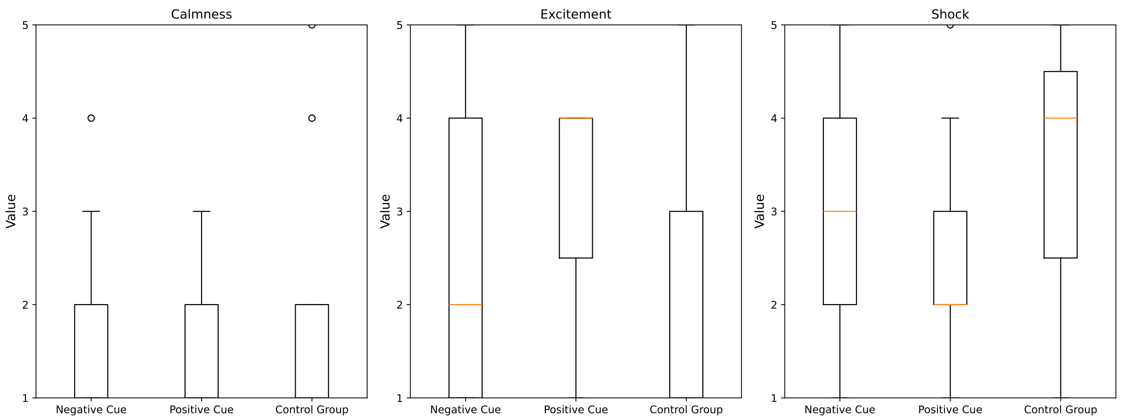

Figure 4.15: Box plots showing the distribution of *Calmness*, *Excitement*, and *Shock* across the three groups.

Shock The Shapiro-Wilk test identified normality for the Negative Emotion Cue Group ($p = 0.2293$) and Positive Emotion Cue Group ($p = 0.2375$), but non-normality for the Control Group ($p = 0.0390$). Homogeneity of variances was satisfied ($p = 0.9058$). ANOVA results showed no significant differences among groups ($F(2,31) = 1.401, p = 0.2615$).

4.10 Emotion Attribution Analysis for Cues Unrelated Emotions (Disgust, Panic, and Fear)

Figure 4.16 presents the mean values of three measured emotional responses (Disgust, Panic, and Fear) across the Negative Emotion Cue Group, Positive Emotion Cue Group, and Control Group. For the *Disgust* response, the Negative Emotion Cue Group displayed the highest mean value, followed by the Positive Emotion Cue Group, and the Control Group showing the lowest. Regarding *Panic* the Negative

Table 4.23: Pearson Correlation Analysis for the Control Group

Variable	r-value	p-value	Significance
IncreasedHeartRate	-0.335	0.314	Moderate Negative Correlation (Not Significant)
Perspiration	-0.177	0.602	Weak Negative Correlation (Not Significant)
MuscleTension	0.038	0.911	No Correlation
Tachypnea	-0.123	0.718	No Correlation

Table 4.24: ANOVA Results for Calmness

Source	Sum of Squares	df	p-value
Group	1.125	2	0.6094
Residual	34.640	31	-

Table 4.25: ANOVA Results for Excitement

Source	Sum of Squares	df	p-value
Group	7.660	2	0.1660
Residual	62.369	31	-

Table 4.26: ANOVA Results for Shock

Source	Sum of Squares	df	p-value
Group	5.200	2	0.2615
Residual	57.535	31	-

Emotion Cue Group exhibited the highest mean value, with the Control Group slightly lower, and the Positive Emotion Cue Group showing the lowest value. For *Fear* the Positive Emotion Cue Group demonstrated the highest mean value, followed by the Negative Emotion Cue Group, and then the Control Group.

Figure 4.17 further visualizes the distributions of the three emotional responses across the groups using box plots. For *Disgust* the Negative Emotion Cue Group exhibited a broader range of values with an apparent outlier, while the Positive Emotion Cue Group had a narrower range and a lower median. The Control Group displayed the smallest range with a relatively low median. In the *Panic* response, the Negative Emotion Cue Group showed a wide distribution of values, while the Positive Emotion Cue Group demonstrated a much narrower range with a lower median. The Control Group, in contrast, exhibited a moderate range with a median similar to the Negative Emotion Cue Group. Finally, for *Fear* the Positive Emotion Cue Group displayed a narrow range and the highest median, whereas the Negative Emotion Cue Group showed a broader range with a slightly lower median. The Control Group exhibited the widest range, but its median was lower than that of the Positive Emotion Cue Group.

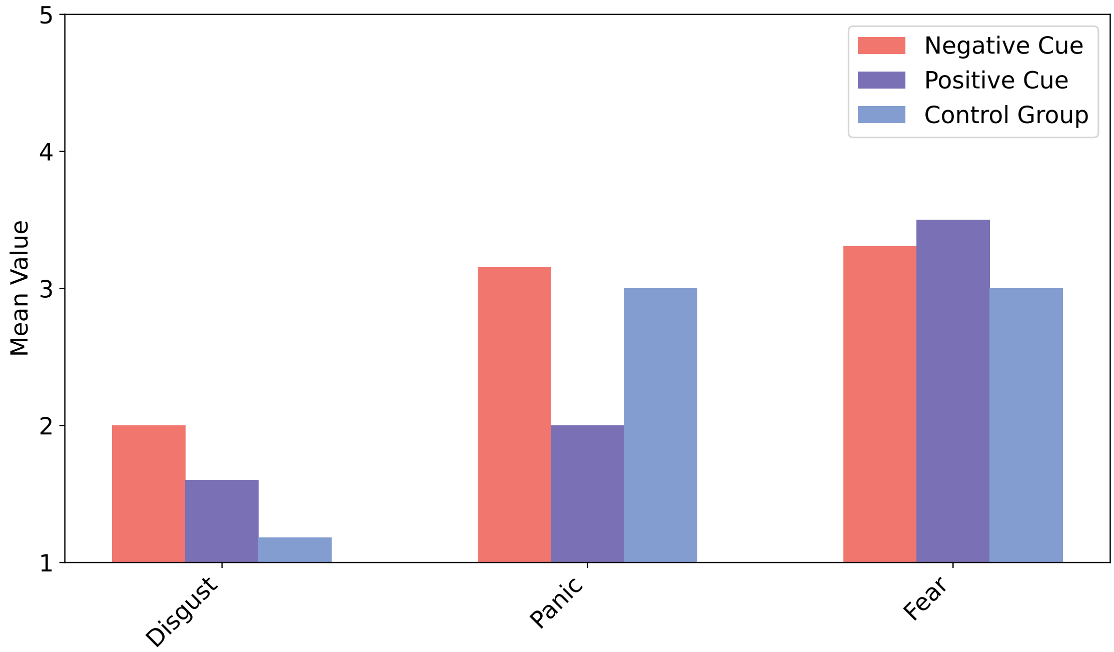

Figure 4.16: Mean values of *Disgust*, *Panic*, and *Fear* across the three groups.

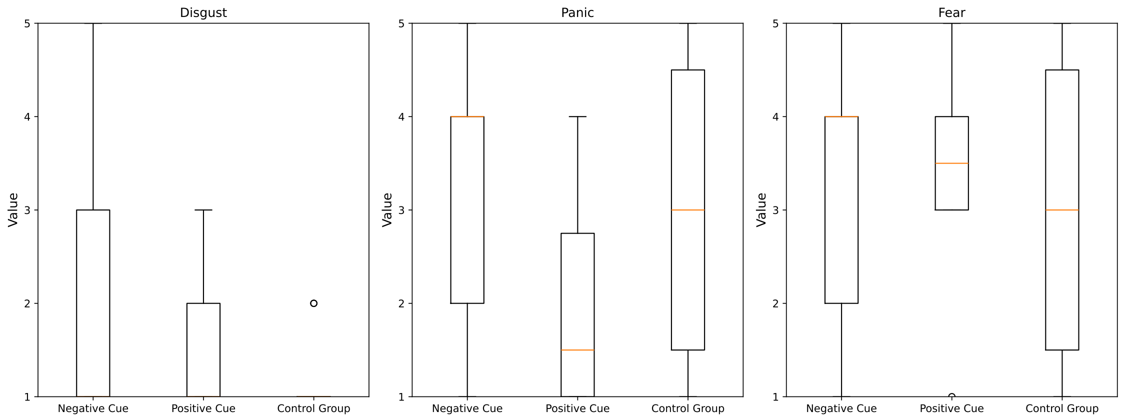

Figure 4.17: Box plots showing the distribution of *Disgust*, *Panic*, and *Fear* across the three groups.

Disgust The Shapiro-Wilk test indicated non-normality in all groups (Negative Emotion Cue Group: $p = 0.0035$, Positive Emotion Cue Group: $p = 0.0014$, Control Group: $p < 0.0001$). Levene's test confirmed homogeneity of variances ($p = 0.1445$). ANOVA results were not significant ($F(2, 31) = 2.061, p = 0.1445$), indicating no significant differences among the groups.

Table 4.27: ANOVA Results for Disgust

Source	Sum of Squares	df	p-value
Group	3.993	2	0.1445
Residual	30.036	31	-

Panic The Shapiro-Wilk test showed non-normality for the Negative Emotion Cue Group ($p = 0.0385$) and Positive Emotion Cue Group ($p = 0.0078$), while the Control Group was normal ($p = 0.0781$). Homogeneity of variances was satisfied ($p = 0.7433$). ANOVA results were not significant ($F(2, 31) = 1.874, p = 0.1705$).

Table 4.28: ANOVA Results for Panic

Source	Sum of Squares	df	p-value
Group	8.425	2	0.1705
Residual	69.692	31	-

Fear The Shapiro-Wilk test identified non-normality for the Negative Emotion Cue Group ($p = 0.0313$), while the Positive Emotion Cue Group ($p = 0.1574$) and Control Group ($p = 0.0781$) met normality. Homogeneity of variances was also satisfied ($p = 0.6517$). ANOVA results showed no significant differences among groups ($F(2, 31) = 0.341, p = 0.7136$).

Table 4.29: ANOVA Results for Fear

Source	Sum of Squares	df	p-value
Group	1.348	2	0.7136
Residual	61.269	31	-

4.11 VR Experience Analysis

Analyzed four questions in the post-experiment questionnaire that related to VR game experience. Including rating for immersion of the VR experience (Immersion), rating for the realism of the VR environment (Realism), level of fear during gameplay (Perceived Fear), and willingness to play VR survival horror games in the future (Future Intention). Figure 4.18 presents the mean values of the four measured variables (Immersion, Realism, Perceived Fear, and Future Intention) across the Negative Emotion Cue Group, Positive Emotion Cue Group, and Control Group. For *Immersion* the Positive Emotion Cue Group exhibited the highest mean value, followed by the Negative Emotion Cue Group, while the Control Group displayed the lowest, the scores of the three groups all above 3 points. Regarding *Realism* the Negative Emotion Cue Group showed the highest mean value, followed by the Positive Emotion Cue Group, and then the Control Group, the scores of the three groups are still all above 3 points. For *Perceived Fear* the Negative Emotion Cue Group demonstrated the highest mean value, followed closely by the other two groups, all of these groups show high scores, all above 4 points. Finally, for *Future Intention* the Positive Emotion Cue Group had the highest mean value, followed by the Control Group, while the Negative Emotion Cue Group exhibited the lowest mean value.

Figure 4.19 further visualizes the distributions of these four variables using box plots. For *Immersion* the Positive and Negative Emotion Cue Groups displayed a similar range with high medians, while the Control Group showed a narrower range and lower median. Regarding *Realism* the Negative Emotion Cue Group demonstrated a wide range and the highest median, followed by the Positive Emotion Cue Group, while the Control Group exhibited the lowest median. For *Perceived Fear* the three groups showed comparable ranges, with the Negative Emotion Cue Group having a slightly higher median. For *Future Intention* the Positive Emotion Cue Group exhibited the widest range and the highest median, while the Negative Emotion Cue Group displayed the lowest median.

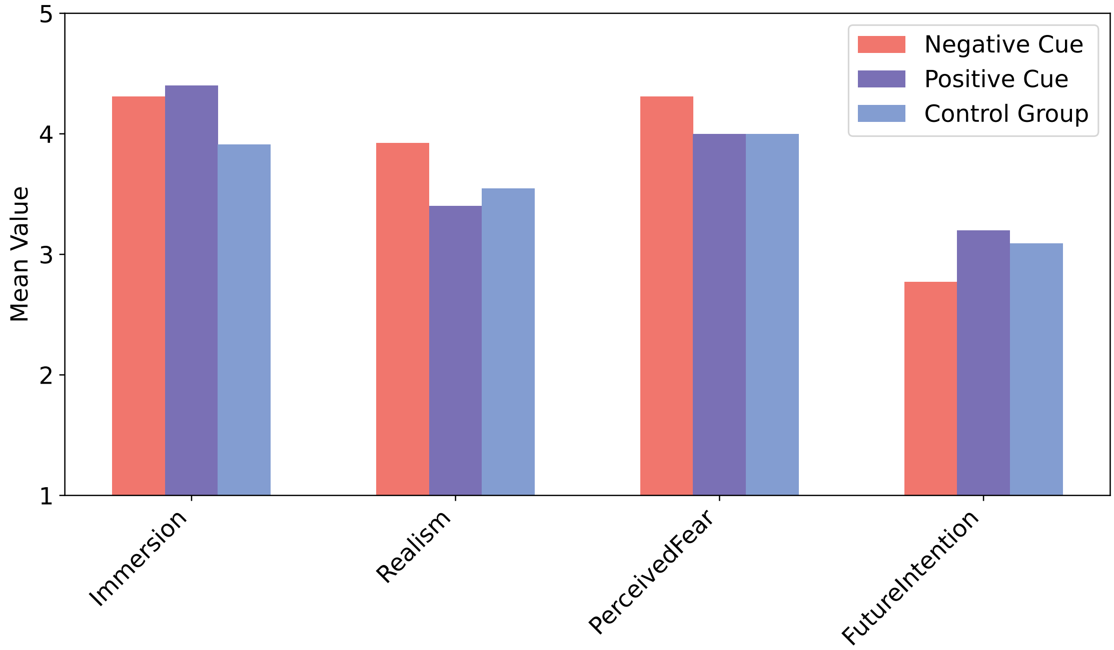

Figure 4.18: Mean values of *Immersion*, *Realism*, *Perceived Fear*, and *Future Intention* across the three groups.

Immersion The Shapiro-Wilk test indicated non-normality in all groups (Negative Emotion Cue Group: $p = 0.0020$, Positive Emotion Cue Group: $p = 0.0014$, Control Group: $p = 0.0080$). Levene's test confirmed homogeneity of variances ($p = 0.7882$). ANOVA results were not significant ($F(2, 31) = 0.953, p = 0.3966$).

Table 4.30: ANOVA Results for Immersion

Source	Sum of Squares	df	p-value
Group	1.481	2	0.3966
Residual	24.078	31	-

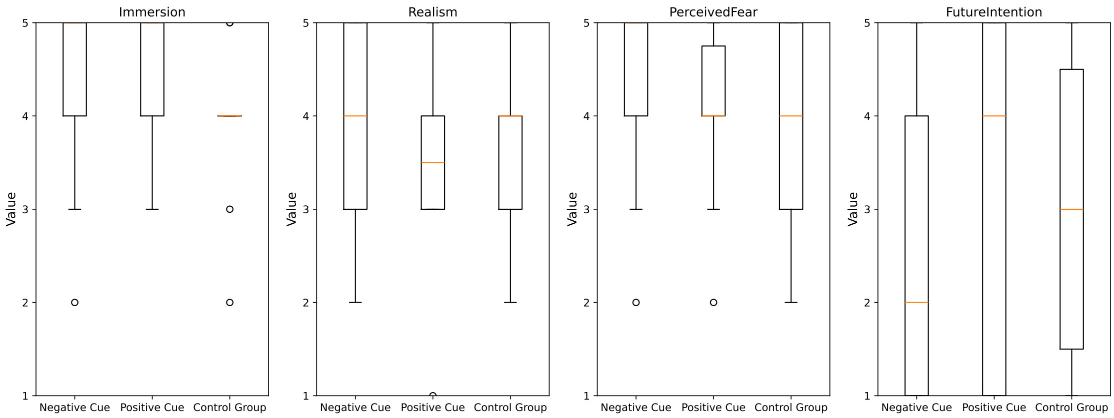

Figure 4.19: The distribution of *Immersion*, *Realism*, *Perceived Fear*, and *Future Intention* across the three groups.

Realism The Shapiro-Wilk test confirmed normality in all groups (Negative Emotion Cue Group: $p = 0.0663$, Positive Emotion Cue Group: $p = 0.0898$, Control Group: $p = 0.1504$). Homogeneity of variances was satisfied ($p = 0.8473$). ANOVA results were not significant ($F(2, 31) = 0.947, p = 0.3987$).

Table 4.31: ANOVA Results for Realism

Source	Sum of Squares	df	p-value
Group	1.714	2	0.3987
Residual	28.050	31	-

Perceived Fear The Shapiro-Wilk test identified non-normality in all groups (Negative Emotion Cue Group: $p = 0.0009$, Positive Emotion Cue Group: $p = 0.0450$, Control Group: $p = 0.0217$). Homogeneity of variances was satisfied ($p = 0.6631$). ANOVA results were not significant ($F(2, 31) = 0.360, p = 0.7008$).

Table 4.32: ANOVA Results for Perceived Fear

Source	Sum of Squares	df	p-value
Group	0.760	2	0.7008
Residual	32.769	31	-

Future Intention The Shapiro-Wilk test identified non-normality for the Negative Emotion Cue Group ($p = 0.0068$) and Positive Emotion Cue Group ($p = 0.0013$), but normality for the Control Group ($p = 0.0682$). Homogeneity of variances was satisfied ($p = 0.5933$). ANOVA results were not significant ($F(2, 31) = 0.186, p = 0.8315$).

Table 4.33: ANOVA Results for Future Intention

Source	Sum of Squares	df	p-value
Group	1.183	2	0.8315
Residual	98.817	31	-

4.12 Pearson Correlation Analysis between VR Factors and Emotional Labels

To examine potential associations between VR experience factors (Immersion, Realism, Perceived Fear, and Future Intention) and eight emotional labels (Fear, Anger, Courage, Excitement, Calmness, Shock, Disgust, Panic), Pearson correlation analyses were conducted within the three experimental groups (Negative Emotion Cue Group, Positive Emotion Cue Group, and Control Group). The results are summarized as follows.

Immersion As shown in Table 4.34, no significant correlations were observed between Immersion and the emotional labels across all groups. The strongest correlation was observed between Fear and Immersion in the Control Group ($r = 0.448$, $p = 0.167$), though it remained non-significant.

Table 4.34: Pearson Correlations between Immersion and Emotional Labels

Group	Fear (r, p)	Anger (r, p)	Courage (r, p)	Excitement (r, p)	Calmness (r, p)	Shock (r, p)	Disgust (r, p)	Panic (r, p)
Negative Cue	0.241, 0.428	0.387, 0.192	0.274, 0.364	0.354, 0.235	0.168, 0.584	0.130, 0.672	-0.065, 0.833	0.245, 0.419
Positive Cue	0.447, 0.195	-0.111, 0.760	0.000, 1.000	-0.480, 0.160	-0.373, 0.289	-0.208, 0.565	0.406, 0.244	0.106, 0.771
Control Group	0.448, 0.167	0.237, 0.483	-0.126, 0.713	-0.179, 0.598	-0.183, 0.590	0.355, 0.283	0.054, 0.875	0.522, 0.099

Realism Table 4.35 presents the correlations between Realism and emotional labels. While most correlations were not significant, a notable inverse correlation was observed between Anger and Realism in the Positive Emotion Cue Group ($r = -0.632$, $p = 0.05$).

Table 4.35: Pearson Correlations between Realism and Emotional Labels

Group	Fear (r, p)	Anger (r, p)	Courage (r, p)	Excitement (r, p)	Calmness (r, p)	Shock (r, p)	Disgust (r, p)	Panic (r, p)
Negative Cue	0.083, 0.788	0.495, 0.085	0.134, 0.663	-0.074, 0.811	0.139, 0.651	-0.015, 0.960	0.194, 0.526	0.231, 0.448
Positive Cue	-0.088, 0.810	-0.632, 0.050	0.478, 0.162	0.094, 0.796	-0.000, 1.000	-0.489, 0.152	-0.539, 0.108	-0.166, 0.647
Control Group	0.302, 0.366	0.069, 0.841	-0.403, 0.219	-0.363, 0.272	-0.129, 0.705	0.022, 0.949	-0.027, 0.936	0.378, 0.252

Perceived Fear Table 4.36 shows strong positive correlations between Fear and Perceived Fear in all groups, particularly in the Positive Emotion Cue Group ($r = 0.900$, $p = 0.0004$). Additionally, in the Control Group, Panic was moderately correlated with Perceived Fear ($r = 0.623$, $p = 0.041$).

Table 4.36: Pearson Correlations between Perceived Fear and Emotional Labels

Group	Fear (r, p)	Anger (r, p)	Courage (r, p)	Excitement (r, p)	Calmness (r, p)	Shock (r, p)	Disgust (r, p)	Panic (r, p)
Negative Cue	0.808, 0.0008	0.477, 0.100	-0.208, 0.496	-0.249, 0.412	-0.597, 0.031	-0.191, 0.532	0.298, 0.322	0.636, 0.019
Positive Cue	0.900, 0.0004	0.248, 0.489	-0.327, 0.356	-0.358, 0.310	-0.833, 0.003	0.279, 0.436	0.140, 0.700	0.567, 0.087
Control Group	0.736, 0.0098	-0.282, 0.400	-0.699, 0.017	-0.204, 0.547	-0.597, 0.052	-0.303, 0.365	0.226, 0.505	0.623, 0.041

Future Intention As detailed in Table 4.37, Fear showed a significant negative correlation with Future Intention in the Negative Emotion Cue Group ($r = -0.733$, $p = 0.004$) and the Control Group ($r = -0.643$, $p = 0.033$).

Table 4.37: Pearson Correlations between Future Intention and Emotional Labels

Group	Fear (r, p)	Anger (r, p)	Courage (r, p)	Excitement (r, p)	Calmness (r, p)	Shock (r, p)	Disgust (r, p)	Panic (r, p)
Negative Cue	-0.733, 0.004	-0.482, 0.095	0.527, 0.064	0.343, 0.252	0.575, 0.040	0.195, 0.522	-0.637, 0.019	-0.534, 0.060
Positive Cue	0.047, 0.897	-0.200, 0.579	0.052, 0.887	0.238, 0.509	-0.237, 0.510	0.044, 0.904	-0.609, 0.061	0.269, 0.453
Control Group	-0.643, 0.033	-0.026, 0.940	0.647, 0.031	0.363, 0.272	0.492, 0.125	0.346, 0.298	0.274, 0.415	-0.491, 0.125

4.13 Correlation Between Immersion and Physiological Data

To explore the potential association between subjective VR immersion ratings and measured physiological data, in order to find whether misattribution arousal is moderated by VR immersion, a correlation analysis was conducted.

Heart Rate Given the presence of intense horror scenes in the VR survival horror game, participants could have significant emotional fluctuations. Thus, the maximum of heart rate data was chosen to reflect the peak responses of participants during the experiment. Due to there are skewed distribution of physiological data and the potential presence of outliers, Spearman's rank correlation coefficient was selected.

The results in Table 4.38 show that the Spearman correlation coefficient ($r_s = -0.051$) is close to zero, indicating there are almost no monotonic relationship between the two variables. The slight negative direction suggests a weak inverse relation but can be disregarded. The p-value is much greater than the commonly used significance level ($\alpha = 0.05$), suggests that this correlation is likely due to random chance and not significant.

Table 4.38: Spearman Correlation Analysis for Heart Rate Maximum Value and VR Immersion

Parameter	**Value**
Sample size	34
Spearman's correlation coefficient (r_s)	-0.051
p-value	0.774

Skin Temperature The analysis shown in Table 4.39 reveals a negative correlation coefficient, a weak possible inverse relationship between skin temperature and VR immersion. However the absolute value of coefficient is small. Furthermore, p-value indicates that this weak negative correlation is statistically insignificant.

Table 4.39: Spearman Correlation Analysis for Skin Temperature Maximum Value and VR Immersion

Parameter	**Value**
Sample size	34
Spearman's correlation coefficient (r_s)	-0.264
p-value	0.131

4.14 Correlation Between Environment Realism and Physiological Data

To examine the association between subjective VR environment realism ratings and physiological data, Spearman correlation analysis has also been conducted. The result was shown in Table 4.40.

Heart Rate The value of Spearman's correlation coefficient is close to 0.3, indicating a weak inverse relationship. However, the trend is still not strong enough to explain the majority of data variability. The p-value is slightly below 0.1 but remains above 0.05, suggesting that there are no statistically significant.

Table 4.40: Spearman Correlation Analysis for Heart Rate Maximum Value and VR Environment Realism

Parameter	**Value**
Sample size	34
Spearman's correlation coefficient (r_s)	-0.293
p-value	0.092

Skin Temperature The analysis shown in Table 4.41 suggests that the coefficient between skin temperature and VR environment realism is close to zero. No monotonic relationship between the two variables. Furthermore, p-value far exceeds the significance value, providing no statistical evidence to support a reliable connection.

Table 4.41: Spearman Correlation Analysis for Skin Temperature Maximum Value and VR Environment Realism

Parameter	**Value**
Sample size	34
Spearman's correlation coefficient (r_s)	-0.039
p-value	0.827

Chapter 5

Discussion

5.1 Key Findings

This study explores the use of Virtual Reality (VR) technology as a tool for inducing physiological arousal and its application in emotional attribution interactions based on the two-factor theory of emotion. Using a VR survival horror game as the media, this study examines whether emotional suggestions in this context can directly influence participants' subjective emotional cognition. The overall results confirmed the validity of suggestive cues to induce misattribution under physiological arousal.

For the effectiveness of a tool to manipulate human emotions, the ability to rapidly induce physiological arousal is important. In the results, the VR survival horror game environment demonstrated its effectiveness by triggering intense emotional changes in participants within a short duration of 10-20 minutes.

To induce the misattribution of arousal, anger and courage were used as the emotional labels for prior suggestive cues. The data indicate that both emotional suggestions influenced the corresponding subjective emotional attributions of the participants significantly: Participants in the Negative Emotion Cue Group (Anger) exhibited higher scores on "Anger" and showed a notable difference from the Control Group; Those participants in Positive Emotion Cue Group (Courage) also have higher scores on "Courage" compared to Control Group.

Well, additionally, a novel finding also emerged in this research. Participants in the Negative Emotion Cue Group showed not only significantly higher scores for "Anger", but also notable differences in "Courage" compared to the Control Group. Which is means that even with no suggestive cues about courage labels, participants still showed higher courage levels, while participants in the Control Group did not show such a tendency.

Another finding has also been observed through this experiment. While subjective emotional attributions varied based on the conditions of whether suggestive cues existed or not, physiological measurements such as heart rate, skin temperature showed no significant differences between groups. This result suggests that suggestions with emotional labels primarily affect subjective perception rather than autonomic nervous responses.

Besides, this study analyzed the correlation patterns between physiological arousal and emotional cognition. The result confirmed the complex interactions between physiological signals and emotional labels. The arousal perceived by participants, such as increased heart rate, was significantly positively correlated with their emotional attributions (e.g., anger, courage).

Moreover, regarding the effectiveness of using VR technology as a tool for inducing physiological arousal, the role of VR immersion also been explored. Though the high immersion and perceived realism did not directly influence participants' emotional attributions, these advantages showed a potential trend of amplifying the effects of emotional suggestions.

Last but not least, an analysis was conducted on the correlation between subjective evaluations of VR presence and realism with physiological data. The result indicated that both presence and realism did not directly influence physiological responses. However, this does not imply that VR has no effect on the degree of physiological response. Combining the statistical results of trends on heart rate and skin temperature with prior research (Pallavicini et al., 2018), VR demonstrates a stronger ability to elicit intense physiological reactions compared to non-immersive media, and thus in this study also significantly influenced the physiological response of participants. The non-significant results in this aspect are because all participants used the same VR environment.

5.2 Interpretation of Results

This research mainly conducted experimental validation from two dimensions: 1. Verifying the effectiveness of VR technology as a tool for inducing physiological arousal; 2. Examining the feasibility of manipulating emotional cognition under physiological arousal induced by VR.

The results demonstrate that emotional suggestions strongly influence subjective emotional attributions in response to physiological arousal induced by VR within the experimental context. Participants exhibited strong emotional changes after playing the VR survival horror game for only 10-20 minutes. Their judgments about their own emotions were significantly influenced by those small talks that included suggestive information during the experiment. This finding further supports earlier studies that related to the two-factor theory of emotion. Additionally, emphasizes the susceptibility of emotional labeling to environmental and cognitive factors. It highlights the critical role of cognitive interpretation in the process of shaping emotional experiences.

It is very interesting that this study uncovered that participants exposed to positive suggestions were induced to corresponding positive emotions; However, when received negative suggestions, participants not only showed corresponding negative emotions but also exhibited some level of positive emotional attribution. This discovery not anticipated in the original hypothesis. One possible explanation is that the language used for emotional cues during the experiment was not sufficiently precise, inadvertently introducing some positive elements into the negative suggestions, leading participants to interpret these elements as their emotional labels. Another explanation is that playing games in VR can significantly enhance a sense of happiness perceived by participants (Pallavicini et al., 2018). This increased level of happiness might cause participants to reinterpret negative emotional cues as positive ones. The behavior of participants after the experiment further supports these findings. Even though some participants showed worries and were afraid of horror content in the early stage of the experiment, they were very excited after playing the game and actively asked the experimenter about the details of the game itself or VR devices. Which shows a certain degree of willingness to continue to experience this VR survival horror game.

Go back to the physiological indicators, no significant differences were found across the three experimental conditions (Negative Cues, Positive Cues, no suggestive cues). This indicates that suggestions themselves do not affect the degree of physiological arousal. Moreover, a certain degree of decoupling between subjective emotional attributions and autonomic nervous responses. Besides, subjective perceptions of physiological responses and indicators were both have high mean scores across the three conditions. Combined with the results of physiological measurements, the physiological indicators exhibited significant trends of change (corresponding to the content of the game). After the stages of VR game-play, a number of participants displayed visible signs of sweating on their foreheads and hands. Not only, but also with heightened excitement. These observations further support the physiological data measured during the experiment.

which demonstrates that the VR environment used in this study (a VR survival horror game) effectively triggered a certain level of physiological arousal.

Last but not least, regarding the relationship between physiological indicators and emotional attributions, though there is only a moderate correlation was shown in the result, it still provides clues for future research. Especially on exploring how specific physiological signals, such as heart rate variability, integrate into the process of emotional cognition. In the result, the alignment between perceived fear scores and subjective fear rating scores further supports the validity of participants's self-reported emotional scales in VR environments.

5.3 Theoretical Contributions

This study bridges the gap between traditional emotional research on the misattribution of arousal and emerging VR technology by replacing conventional methods, such as drug-induced or dangerous physical environments, to trigger physiological arousal. Moreover, demonstrates the applicability of the two-factor theory of emotion in VR environments.

Besides, the study also provides evidence supporting VR survival horror games as an effective tool for studying emotional processes, in triggering fear-induced arousal. This supports the effectiveness of VR in the interactions between physiological arousal and cognitive labeling.

Furthermore, this study explores the relationship between subjective cognition and physiological responses, offering a new perspective on the consistency across dimensions of an individual's emotional experience.

5.4 Practical Contributions

This study utilizes a VR survival horror game as a safe, controllable, and easily replicable method for inducing physiological arousal. This provides an alternative to traditional methods such as drugs or risky physical environments. Additionally, the insights from this research can also be applied to studies on VR-based emotional regulation interventions, or therapeutic applications.

It is worth noting that this study explored a relatively simple and easily reproducible scenario to rapidly trigger the phenomenon of misattributed physiological arousal. This provides a straightforward solution to the challenge of manipulating human emotions. By simply wearing a device, play a game with provided emotional labels, emotions can be manipulated effectively, especially for anger and courage emotions that were validated in this study. If future research builds upon this study to integrate suggestions directly into the device or game itself, it could reduce the potential biases introduced by the language-based intervention used in this experimental design.

Furthermore, such an approach would allow for more convenient and precise control over users' emotional judgments.

5.5 Limitations

Due to various constraints, this study still has some limitations. Firstly, the sample characteristics are restricted. The study was conducted on a small group of university students, which does not cover a wide enough range of age groups or education levels. This limited the generalizability of the findings to wider populations.

Secondly, regarding data collection, I used self-report questionnaire to collect participants' subjective emotional attributions. Although subjective emotional attribution is core to this research, such method may introduce issues like social desirability bias or recall bias. For those collected physiological data, despite the comprehensive range of measurements, the device itself has some limitations. Makes it unable to fully capture the complexity of autonomic nervous responses.

As for the experiment design, the potential impact of VR gameplay duration on the results was not considered. In addition, physiological responses have not been measured before and after the VR gameplay. This led to the result that this study can not directly illustrate those changes in physiological arousal while playing VR games.

Chapter 6

Conclusion

This study validates the unique value of virtual reality technology in psychological research. Especially in exploring the interactions between physiological arousal and emotional cognitive labels. The results indicate that emotional suggestions influence subjective emotional attributions significantly, while physiological responses remain relatively stable and will not affected by suggestions. This finding further supports the core hypothesis of the two-factor theory of emotion (when individuals experience a state of arousal they cannot immediately explain, they label the state to support their perception of the current experience).

Notably, the result of this study supports that the VR environment can provide a safe, controllable and efficient experimental condition, especially for psychological research. Moreover, this environment setting has high repeatability and applicability, which can open up new possibilities for experimental design and practical application in psychology.

Despite those limitations in sample diversity, measurement tools, and part of the experimental design, these preliminary findings can offer valuable insights for future exploration in this area.

Future work

Based on this study, the sample size of the experiment can be expanded, including participants from a wider range of age groups, cultural backgrounds, and varying levels of VR and gaming experience to enhance the generalizability of the findings.

In addition, advanced physiological measurement technologies such as functional Magnetic Resonance Imaging (fMRI) or Electroencephalography (EEG), could be used in this experiment condition. This could provide a deeper understanding of the interaction between arousal and cognition.

The unexpected finding of higher positive emotional attributions in the negative emotional suggestion group also needs further exploration. Future experiment design could based on the interplay between these two factors, investigating whether this interaction is specific to them or influenced by other experimental contexts.

The long-term impact of emotional cues on behavior in VR content can also be an interesting topic. Currently this area still lacks exploration. It may provide potential applications in a variety of scenarios, such as therapy or education. Besides, more effective methods of offering existing emotional labels can be applied to this experiment condition, in order to reduce potential errors due to the limitations of language-based communication methods.

Bibliography

Eduardo B Andrade and Dan Ariely. 2009. The enduring impact of transient emotions on decision making. *Organizational behavior and human decision processes*, 109(1):1–8.

Irene Alice Chicchi Giglioli, Gabriella Pravettoni, Dolores Lucia Sutil Martín, Elena Parra, and Mariano A Raya. 2017. A novel integrating virtual reality approach for the assessment of the attachment behavioral system. *Frontiers in psychology*, 8:959.

Donald G Dutton and Arthur P Aron. 1974. Some evidence for heightened sexual attraction under conditions of high anxiety. *Journal of personality and social psychology*, 30(4):510.

David Goldstein, David Fink, and David R Mettee. 1972. Cognition of arousal and actual arousal as determinants of emotion.

Jih-Hsuan Lin. 2013. Do video games exert stronger effects on aggression than film? the role of media interactivity and identification on the association of violent content and aggressive outcomes. *Computers in Human Behavior*, 29(3):535–543.

Jih-Hsuan Tammy Lin. 2017. Fear in virtual reality (vr): Fear elements, coping reactions, immediate and next-day fright responses toward a survival horror zombie virtual reality game. *Computers in Human Behavior*, 72:350–361.

Antony SR Manstead and Hugh L Wagner. 1981. Arousal, cognition and emotion: An appraisal of two-factor theory. *Current Psychological Reviews*, 1(1):35–54.

Artur Marchewka, Łukasz Żurawski, Katarzyna Jednoróg, and Anna Grabowska. 2014. The nencki affective picture system (naps): Introduction to a novel, standardized, wide-range, high-quality, realistic picture database. *Behavior research methods*, 46:596–610.

Javier Marín-Morales, Juan Luis Higuera-Trujillo, Alberto Greco, Jaime Guixeres, Carmen Llinares, Enzo Pasquale Scilingo, Mariano Alcañiz, and Gaetano Valenza. 2018. Affective computing in virtual reality: emotion recognition from brain and heartbeat dynamics using wearable sensors. *Scientific reports*, 8(1):13657.

Federica Pallavicini, Pietro Cipresso, Simona Raspelli, Alessandra Grassi, Silvia Serino, Cinzia Vigna, Stefano Triberti, Marco Villamira, Andrea Gaggioli, and Giuseppe Riva. 2013. Is virtual reality always an effective stressors for exposure treatments? some insights from a controlled trial. *BMC psychiatry*, 13:1–10.

Federica Pallavicini, Ambra Ferrari, Alessandro Pepe, Giacomo Garcea, Andrea Zanacchi, and Fabrizia Mantovani. 2018. Effectiveness of virtual reality survival horror games for the emotional elicitation: Preliminary insights using resident evil 7: Biohazard. In *Universal Access in Human-Computer Interaction. Virtual, Augmented, and Intelligent Environments: 12th International Conference, UAHCI 2018, Held as Part of HCI International 2018, Las Vegas, NV, USA, July 15-20, 2018, Proceedings, Part II 12*, pages 87–101. Springer.

Federica Pallavicini, Alessandro Pepe, and Maria Eleonora Minissi. 2019. Gaming in virtual reality: What changes in terms of usability, emotional response and sense of presence compared to non-immersive video games? *Simulation & Gaming*, 50(2):136–159.

Thomas D Parsons and Albert A Rizzo. 2008. Affective outcomes of virtual reality exposure therapy for anxiety and specific phobias: A meta-analysis. *Journal of behavior therapy and experimental psychiatry*, 39(3):250–261.

Silvia Francesca Maria Pizzoli, Dario Monzani, Ketti Mazzocco, Marianna Masiero, Gabriella Pravettoni, et al. 2020. Digital contents as a tool to address research reproducibility crisis in psychology: A case study on sexual attraction under conditions of high arousal. *Annual Review of Cybertherapy and Telemedicine*, (A):N–A.

Stanley Schachter and Jerome Singer. 1962. Cognitive, social, and physiological determinants of emotional state. *Psychological review*, 69(5):379.

Robert C Sinclair, Curt Hoffman, Melvin M Mark, Leonard L Martin, and Tracie L Pickering. 1994. Construct accessibility and the misattribution of arousal: Schachter and singer revisited. *Psychological Science*, 5(1):15–19.

Mel Slater. 2009. Place illusion and plausibility can lead to realistic behaviour in immersive virtual environments. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 364(1535):3549–3557.

Rukshani Somarathna, Tomasz Bednarz, and Gelareh Mohammadi. 2022. Virtual reality for emotion elicitation–a review. *IEEE Transactions on Affective Computing*.

Kathleen D Vohs, Roy F Baumeister, and George Loewenstein. 2007. *Do Emotions Help or Hurt Decisionmaking?: A Hedgefoxian Perspective.* Russell Sage Foundation.

Appendix: Questionnaire

1. On a scale of 1 (”no such feeling”) to 5 (”very strong feeling”), please rate the physiological responses you experienced during the VR experience:
   (a) Increased Heart Rate:
   (b) Increased sweating
   (c) Uncontrollable shaking of hands or legs:
   (d) Increased breathing rate:
2. On a scale of 1 (”no such feeling”) to 5 (”very strong feeling”), please rate the extent to which you experienced the following emotions at the moment:
   (a) Calmness:
   (b) Excited:
   (c) Shocked:
   (d) Full of courage:
   (e) Disgusted, uncomfortable:
   (f) Frightened, terrified:
   (g) Scared:
   (h) Angry, aggressive:
3. On a scale of 1 (”not realistic at all”) to 5 (”very realistic”), please rate the realism of the VR environment.
4. On a scale of 1 (”strongly disagree”) to 5 (”strongly agree”), please rate your level of agreement with the following statements:
   (a) I felt like I was in the real world during the VR experience.
   (b) I felt a sense of fear while playing.
   (c) Even after put down the VR device, the effects of the horror game lingered.
   (d) I am becoming brave enough to face those horror content during gameplay.
   (e) I am more willing to try VR horror games again after this experience.