Contribution of facial color on expression recognition, and emotional response to visual and auditory stimuli: evidence from pupillometry

（表情認知における顔色の影響と視聴覚刺激に対する感情反応:
瞳孔計測を用いた研究）

January 2021

Doctor of Philosophy (Engineering)

Satoshi Nakakoga
中古賀　理

Toyohashi University of Technology

別紙４－１（課程博士（英文））

Date of Submission（month day, year）: January 8th, 2021

Department of Computer Science and Engineering		Student ID Number	D143356	Supervisors	Tetsuto Minami Shigeki Nakauchi
Applicant's name	Satoshi Nakakoga				

Abstract (Doctor)

Title of Thesis	Contribution of facial color on expression recognition, and emotional response to visual and auditory stimuli: evidence from pupillometry

Approx. 800 words

Human emotion is a mechanism for survival strategies, such as behavioral motivation and the signal of crisis avoidance. The study about human emotions began in philosophy and has conducted in various fields such as psychology, physiology, and engineering. Many studies on the relationship between emotions and physiological responses have focused on the pupillary response in recent years. The pupillary response has been attracting attention as a new physiological indicator that can indirectly extract human emotional responses because it reflects the activity of the locus coeruleus (LC). In this thesis, we investigated the mechanisms of emotion and facial expression recognition, which have not yet been elucidated, using pupillary responses (Chapter 2 and 3). In addition, we attempted to estimate emotional states using the pupillary responses based on these findings (Chapter 4).

In the Chapter 2, we investigated the contribution of facial color to expression recognition in blur images with the measurement of behavior and pupillary change. In the experiment, the face stimuli of facial colors (natural color, reddish) with different expressions (neutral and anger) in 3 blur levels were presented. Participants performed a task of expression identification to the stimulus. Behavioral results indicated that the facial color has a significant contribution to expression recognition as blur level increases. Then, the results of pupillometry showed that the reddish-color provided the information necessary to identify anger. These results showed the contribution of facial color increases in both psychophysics and pupillary experiment as blur level increases, which suggested that facial color emphasizes the characteristics of specific facial expressions.

In the Chapter 3, we aimed to elucidate the relationship of attentional states to emotional unimodal stimuli (pictures or sounds) and emotional responses by measuring the pupil diameter. In experiment 1, we investigated the relationship of the attentional state with emotional visual stimuli and emotional responses by using pupillometry. We observed that the velocity of pupillary dilation was faster during the presentation of emotionally arousing pictures compared to that of neutral ones, regardless of the valence of the pictures. Importantly, this effect was not dependent on the task condition. In experiment 2, we investigated the relationship of the attentional state with emotional auditory sounds and emotional responses. We observed a trend towards a significant interaction between the stimulus and the task conditions with regard to the velocity of pupillary dilation. In the emotional and auditory detection tasks, the velocity of pupillary dilation was faster with positive and neutral sounds than negative sounds. However, there were no significant differences between the no task and visual detection task conditions. Taken together, the current data reveal that different pupillary responses were elicited to emotional visual and auditory stimuli, at least in the point that there is no attentional effect to emotional responses to visual stimuli, despite both experiments being sufficiently controlled to be of symmetrical experimental design.

In the Chapter 4, we investigated pupillary responses to an auditory stimulus after a positive, negative, or neutral emotional state was elicited by an emotional image. An emotional image was followed by a beep sound that was either repetitive or unexpected, and the pupillary dilation was measured. Our results showed that the early component of the pupillary response to the beep sound was larger for negative and positive emotional states than neutral, whereas the late component was larger for positive emotional states. The pupil response's peak latency was earlier for negative images than neutral or positive. Finally, SVM classified the emotional state based on the pupillary response with 80¥% accuracy. Our study suggests that emotional states can be estimated from the amplitude and the latency of pupil activity in response to an auditory probe.

Through these experiments, this study elucidated the unresolved cognitive mechanisms of emotion recognition and facial expression recognition using pupillary responses and proposed a new emotion estimation method. These findings provide important evidence to support the usefulness of pupillary response as a new physiological indicator in emotion research. These findings also support the effectiveness of pupillary response as an objective emotion estimation method that does not rely on subjective responses, and further research is recommended.

Acknowledgments

I am greateful to my supervisors, Prof. Tetsuto Minami and Shigeki Nakauchi, of Toyohashi University of Technology, for many discussions, vast knowledge, and a good work atmosphere, and for giving me a chance to work at the Cognitive Neurotechnology Unit and the Visual Perception and Cognition Laboratory. I had an unforgettable time and experience during seven years in their great laboratory.

I also would like to express my appreciation to other teachers, including Assist. Prof. Kyoko Hine and Assist. Prof. Hiroshi Higashi, at Kyoto University for their expertise. I am also indebted to Prof. Bruno Laeng, of Oslo University and Prof. Tomoko Imura, Japan Women's University, who gave me invaluable comments and suggestions.

I would like to thank our laboratory stuff, Yuki Kawai, and ex-laboratory staff, Kanae Miyazawa, for their professional and scientific advice and administrative support. I also would like to express my gratitude to all our laboratory members for helping me out in my work.

Finally, I greatly appreciate the support of my family for their patience and support.

Contents

List of Figures

Chapter 1

Introduction

1.1 Emotion and Facial expression

We are surrounded by information obtained from our senses, such as vision and audition. The physical characteristics of various stimuli (e.g., color or size) are used to determine information processing priority (Wolfe and Horowitz, 2004). One of the factors that determine the priority of processing is emotional relevance. It plays a role in tagging specific stimuli for appropriate processing, such as approaching behavior or avoiding behavior, depending on emotions (Vuilleumier, 2005; Davidson et al., 2004). These emotionally relevant stimuli include the facial expressions of others as well as natural images and sounds.

Emotion

In psychology, strong responses, which are often accompanied by physiological arousal, are called emotions. Physiological arousal is a state of physical excitement associated with the activity of the sympathetic nervous system and the endocrine system. Some of the emotions that we experience, such as anger, fear, sadness, and joy, have been formed during the evolutionary process and are thought to have a high degree of commonality with animals. In this sense, some argue that these emotions are basic emotions that have a biological basis (Ekman, 1999). However, Russell et al. reported that emotions could be reduced to a two dimensions, such as the pleasant-unpleasant dimension and the activated-deactivated dimension (Russel and Barrett, 1999).

Facial expression

The face is one of the most important visual stimuli in our daily life. It could be used to estimate gender age, attractiveness, health status, and emotions as well as individual identification. In addition, information of the face could estimate the internal information that the person does not intend. Thus, the face is an important social signal that indicates a person's state of mind, and we can acquire extraordinary cognitive abilities for faces in the course of our daily life with others. In particular, facial expressions are important cues for making value judgments about the surrounding environment and specific objects. Therefore, facial expressions provide necessary information for judging others' emotions when communicating with them (e.g., fear is a signal of the threat to something; anger is a signal of resentment, etc.). Ekman et al. showed that six types of human facial expressions, happiness, surprise, fear, disgust, anger, and sadness, are basic emotional expressions worldwide (Ekman and Friesen, 1971). Physiological responses, such as heartbeat and blood pressure, are accompanied by changes in the emotional state, and facial skin color changes correspondingly. Drummond conducted a questionnaire survey on facial color and emotion. He reported that flushing was associated with anger and pallor with fear (Drummond, 1997). It suggests an empirical relationship between facial color and emotion when reading the emotional state of others. Nakajima et al. also showed the interaction between facial color and facial expression in identifying facial expressions by investigating the effect of facial color on facial expression recognition and the effect of facial expression on facial color perception (Nakajima et al., 2017).

1.2 Mechanisms of emotion processing

In this section, we will discuss the processing mechanisms in the brain that produce emotions, and the functions of the amygdala and orbitofrontal cortex, which are important brain regions in triggering emotional responses.

Amygdala

The amygdala, located in the medial part of the left and right temporal lobes, is a complex of several neurons (see Fig. 1.1). The amygdala is one of the most important brain regions involved in emotion, and is mainly responsible for the detection of emotional stimuli and social judgment. The survival of organisms needs to distinguish whether various things are safe and beneficial for them or dangerous and harmful for them, and the amygdala is responsible for such evaluations. Functional brain imaging studies, such as fMRI (functional magnetic resonance imaging) and PET, have shown that the human amygdala is rapidly activated when we see fear, sadness, happiness, etc., in others (Breiter et al., 1996; Blair et al., 1999). Not only that, but the amygdala is also capable of automatically detecting emotional stimuli without conscious awareness. (Nomura et al., 2004).

The amygdala also plays a significant role in social judgment, and is involved in prejudice and interpersonal judgment (Adolphs, 2001). Adolphs et al. conducted experiments that patients with damage to the amygdala and healthy participants look at photos of unknown people with positive and negative facial expressions, and asked them to judge the "friendliness" and "trustworthiness" of the people (Adolphs et al., 1998). Healthy participants judged people with negative facial expressions as unfriendly and untrustworthy, while patients injured in the amygdala judged people with negative facial expressions favorably. These results suggest that the amygdala is also involved in making information judgments for understanding people and is useful for adapting to social life.

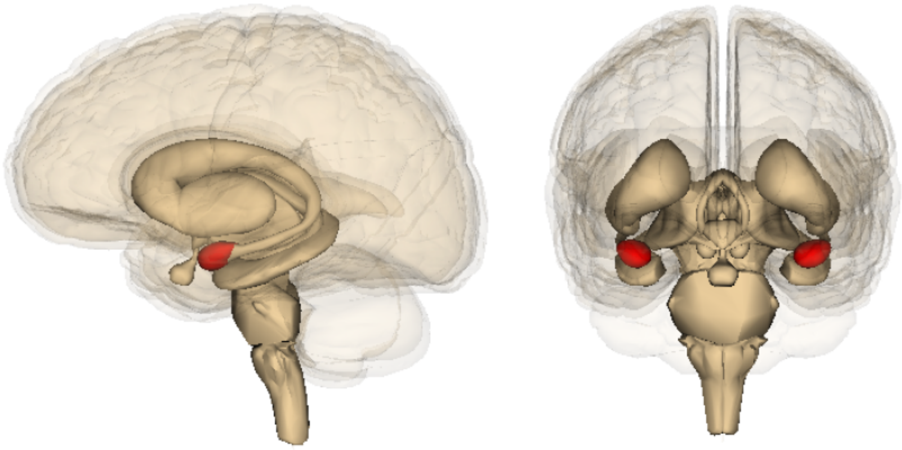

Figure 1.1: Amygdala

Prefrontal cortex

The prefrontal cortex refers to the area of the frontal lobe that excludes regions related to movement. The prefrontal cortex is associated with most of the intellectual and higher cognitive functions, and also plays an important role in emotion, of which the orbitofrontal cortex is the most closely related. The orbitofrontal cortex is the only part of the prefrontal cortex with direct and close neural contact with the amygdala. The difference between the orbitofrontal cortex and the amygdala is that while the amygdala quickly detects important emotional stimuli, the orbitofrontal cortex monitors the relationship between the stimuli, the occipital response to the stimuli, and the consequences of the behavior. It is based on this monitoring and evaluation, and the orbitofrontal cortex can control behavior on a long-term basis (Rolls, 2000).

1.3 Pupillometry

Pupil diameter

The pupil is the opening in the center of the iris. Its function is to allow light to enter the eye so it can be focused on the retina to begin the process of sight. The size of the pupil changes continuously. In general, the pupil becomes larger in dark environments and smaller in bright environments. Humans' pupil diameter varies from a maximum of 8 mm to 1.5 mm in response to changes in light intensity. However, the optimal pupil diameter is 3-4 mm, because a huge pupil size is induced increasing the lense distortion, and a tiny pupil diameter blurs the retinal image due to diffraction effects. The pupillary light reflex (PLR) reflects the autonomic nervous system that could not be controlled by consciousness. It has also been reported that the pupil diameter decreases depending on the human perception of brightness and glare, regardless of the actual physical luminance (Laeng and Endestad,

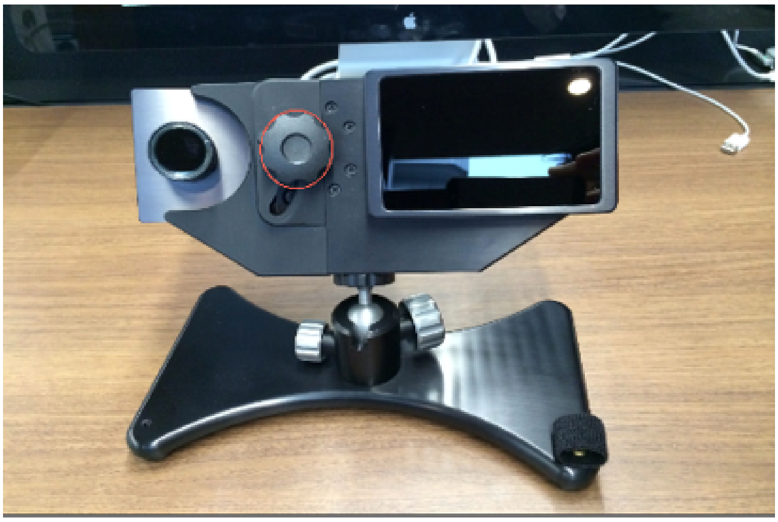

Figure 1.2: Eye-tracking device (Eyelink 1000, SR Research)

2012; Naber and Nakayama, 2013; Laeng and Sulutvedt, 2014). The pupil of a normal person begins to contract after a delay of about 0.2-0.3 s after the stimulus is presented, reaches maximum contraction in about 1 s, and then returns to its original size after dilations.

In recent years, eye-tracking measurement devices (see Fig. 1.2) have become significantly smaller and sophisticated. The principle of this measurement is to irradiate the eye with a near-infrared LED mounted on the device and to measure the gaze position and pupil diameter by using a camera to measure the reflected corneal image.

Mechanism of pupillary dilation and constriction

Two types of smooth muscles, called the pupillary dilator muscles (radial) and the pupillary constrictor muscles (sphinctor), are directly responsible for pupillary movement. These muscles are controlled by sympathetic and parasympathetic nerves, respectively. During mental excitement, the sympathetic nervous system becomes active and the pupillary dilator muscle of the iris contracts, resulting in dilations. In contrast, during sleep or deep anesthesia, the parasympathetic nervous system becomes dominant, and constrictions persist. Fig. 1.3 shows the neural circuit of the pupil response. The pupillary dilation and constriction are caused by the actions of various organs and tissues. In the parasympathetic nervous system, when the incident light increases above a certain level, the Retinal Ganglion Cell (RGC) projects to the Edinger-Westphal Nucleus via the Pretectal Olivary Nucleus (PON). This projection activates the Edinger-Westphal Nucleus, which activates the pupillary constrictor muscles and promotes pupil contraction (Wang and Munoz, 2015).

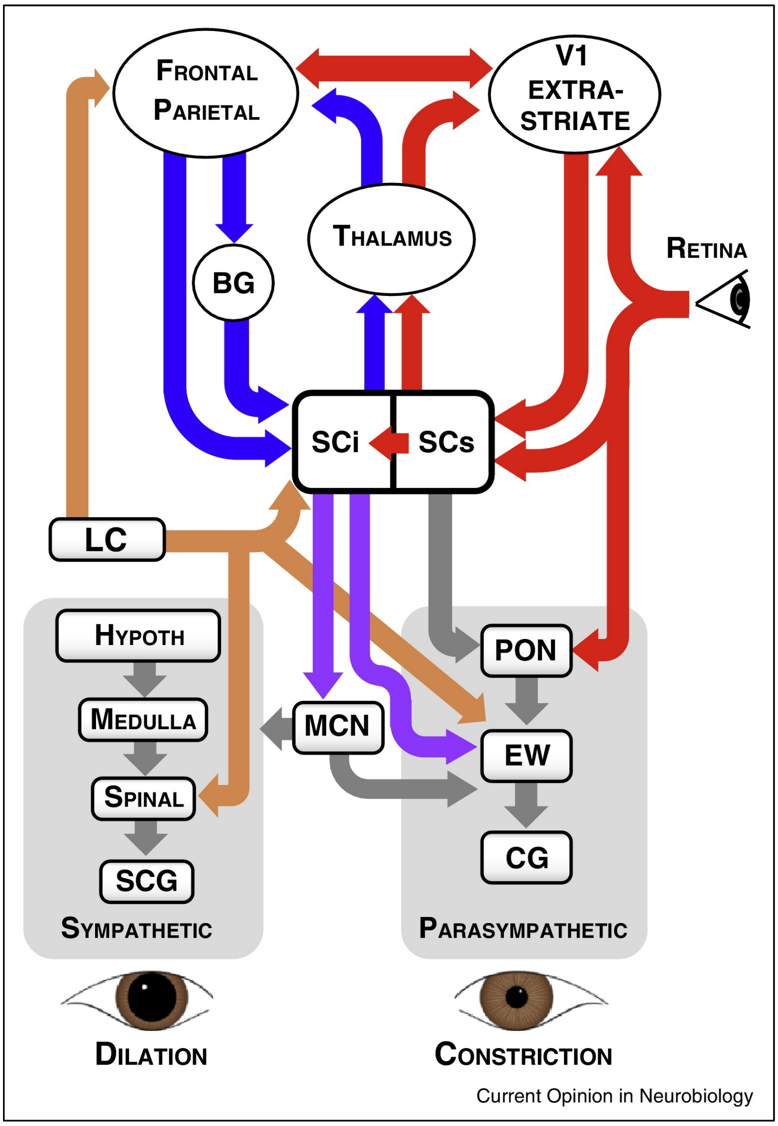

Figure 1.3: Schematic of the pupil orienting circuit (Wang and Munoz, 2015)

The pupil diameter changes with arousal and mental activity also. For example, physiological responses such as attention and arousal when a person is concentrating on a task or gazing at some objects cause pupil dilation (Benarroch, 2009). This is because dilation is modulated by the locus ceruleus (LC), which plays a central role in regulating arousal and cognitive functions. The previous experiments about monkeys' pupillary responses have shown a strong correlation between LC neurons and pupillary changes (see Fig. 1.4 (Aston-jones and Cohen, 2005)). When LC is activated, it exerts

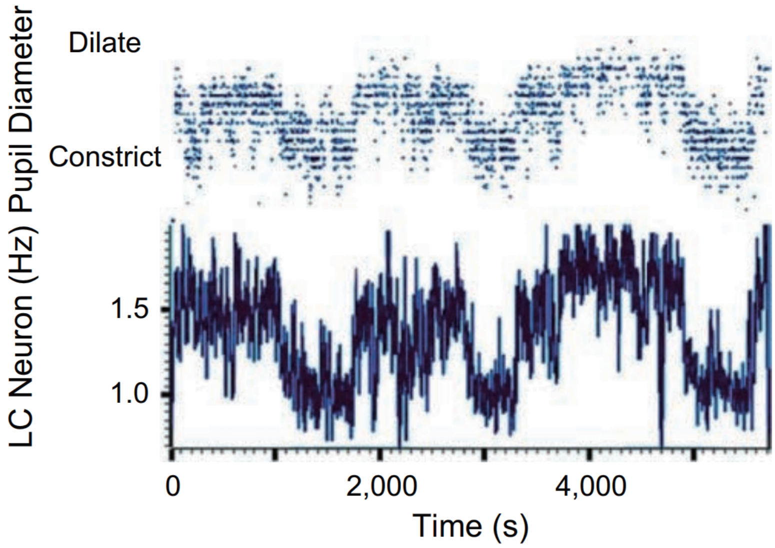

Figure 1.4: Relationship between tonic pupil diameter and baseline firing rate of an LC neuron in monkey (Aston-jones and Cohen, 2005)

a direct inhibitory projection on the Edinger-Westphal Nucleus. As a result, this inhibits the action of the Edinger-Westphal Nucleus and its associated pupillary constrictor muscles, which promotes pupillary dilations. Thus, pupil dilation is mainly caused by the sympathetic nervous system, while pupil contraction is primarily driven by the parasympathetic nervous system. Also, the pupil diameter changes due to the interaction between the sympathetic and parasympathetic nervous systems. Therefore, it has long been believed that the pupil diameter reflects the mental state of the human. Hess et al. reported that the pupil diameter constricted depending on the interest and attention to the task (Hess and James, 1960) and changed depending on the difficulty of the task (Hess and James, 1964).

Pupillary response to emotional stimuli

Since pupillary changes reflect the cognitive state, many studies on emotion using pupillary change have been reported recently. Bradley et al. investigated the pupillary response to the International Affective Picture System (IAPS) images (Pleasant, Unpleasant, and Neutral) (Bradley et al., 2008). The results showed that pupillary dilation after the initial light reflex was significantly larger for pleasant and unpleasant stimuli than for neutral stimuli (see Fig. 1.5). Many other studies have attempted to extract emotional responses when emotional natural images are presented using pupillometry (Kashihara et al., 2014; Kinner et al., 2017; Bradley and Lang, 2015; Henderson et al., 2014;

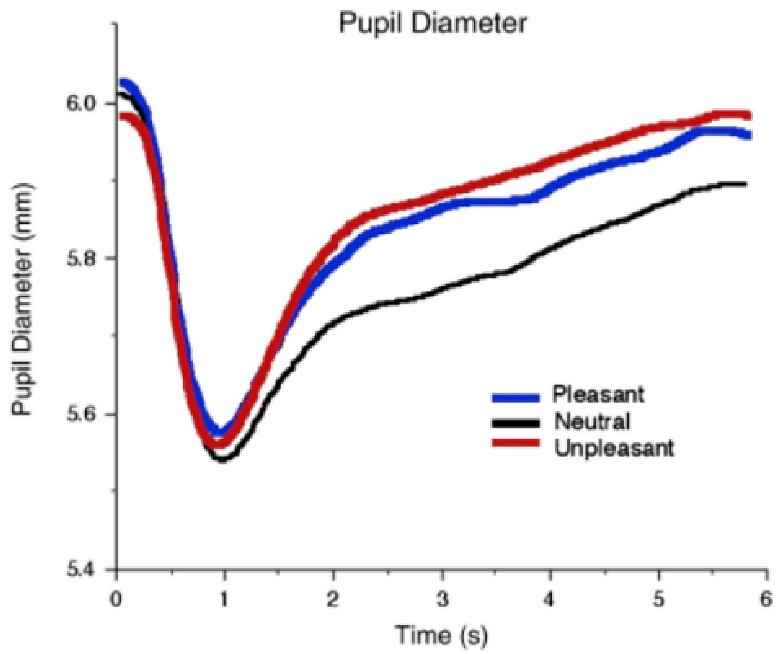

Figure 1.5: Pupillary response to emotional images (Bradley et al., 2008)

Finke et al., 2017; Bradley et al., 2017; Mckinnon et al., 2020). Snowden et al. compared the pupil changes in the passive (only view) and active (evaluate the emotional image) conditions to investigate the relationship between pupil dilation and attention to the stimuli (Snowden et al., 2016). The results showed that the emotional images induced larger pupil dilation in both attention conditions than the neutral pictures. This suggests that the processes that cause emotional modulation of the pupil are beyond automatic and conscious control. In addition, it has also been shown that the pupillary responses to emotional stimuli occur natural auditory stimuli as well as emotional pictures (Partala and Surakka, 2003; Bonmassar et al., 2020). Bradley et al. conducted experiments comparing pupil diameter and skin conductance response (SCR), suggesting that pupillometry has the potential to be used in clinical, developmental, neurological, and other emotional reactivity investigations as a relatively easy to measure, reliable, and non-invasive biological measure of sympathetic arousal during image viewing (Bradley et al., 2017).

Pupil changes associated with affective responses are not limited to emotional responses to natural images and sounds. Liao et al. showed that the pupil diameter constricted when presented with an attractive face image(Liao et al., 2020). In addition, the pupillary response was lager when presented other's face with emotional expression (Laeng et al., 2013; Nuske et al., 2014a; Jessen et al., 2016;

Wang et al., 2018). Therefore, the pupillary response during the recognition of facial expressions has been shown to be useful as a physiological indicator of emotional recognition.

1.4 Approach

In this thesis, we aimed to elucidate the unresolved emotion recognition mechanisms and establish a new emotion estimation method. We used the pupillary response as a physiological indicator. The pupillary response has attracted attention in recent years as a new indicator of emotional reactions because it reflects the LC activity, which contains a large amount of noradrenaline(Aston-jones and Cohen, 2005). The pupillary response is known to be mydriatic in response to emotional images and sounds(Bradley et al., 2008; Partala and Surakka, 2003). In addition, it is known that pupil dilation occurs during the recognition of emotional facial expressions of others(Laeng et al., 2013; Nuske et al., 2014a; Jessen et al., 2016; Wang et al., 2018). In this thesis, we aimed to achieve the above goals by using pupillary response as a physiological indicator. In Chapter 2, we experimented with measuring behavioral response and pupillary response to elucidate the cognitive mechanism of facial color effect on facial expression recognition. In Chapter 3, we investigated the symmetry of attentional states and emotional responses to visual and auditory stimuli by measuring the behavioral response and the pupillary response. In Chapter 4, we experimented to establish a new method for estimating emotion by measuring pupillary responses to auditory probe stimuli.

1.5 Overview

This thesis comprises five studies (Chapter 1-5). First, we present what emotion and facial expression are, the mechanisms of emotion processing, the utility of pupillometry as an indicator of emotional response, and the goal and approaches of this study as an Introduction in this chapter. Next, in Chapter 2, we present the experiment that investigated the mechanisms of facial color effect on expression recognition using pupillometry. Then, in Chapters 3 , we present the relationship of attentional states with emotional unimodal stimuli (pictures or sounds) using pupillometry in two experiments. Moreover, in Chapters 4, we present studies examined whether the emotional states could be classified based on the pupillary reopnses elicited by sound probe stimuli. Finally, we summarize the outcomes of the three studies in the final chapter as the conclusion (see also Figure 1.6).

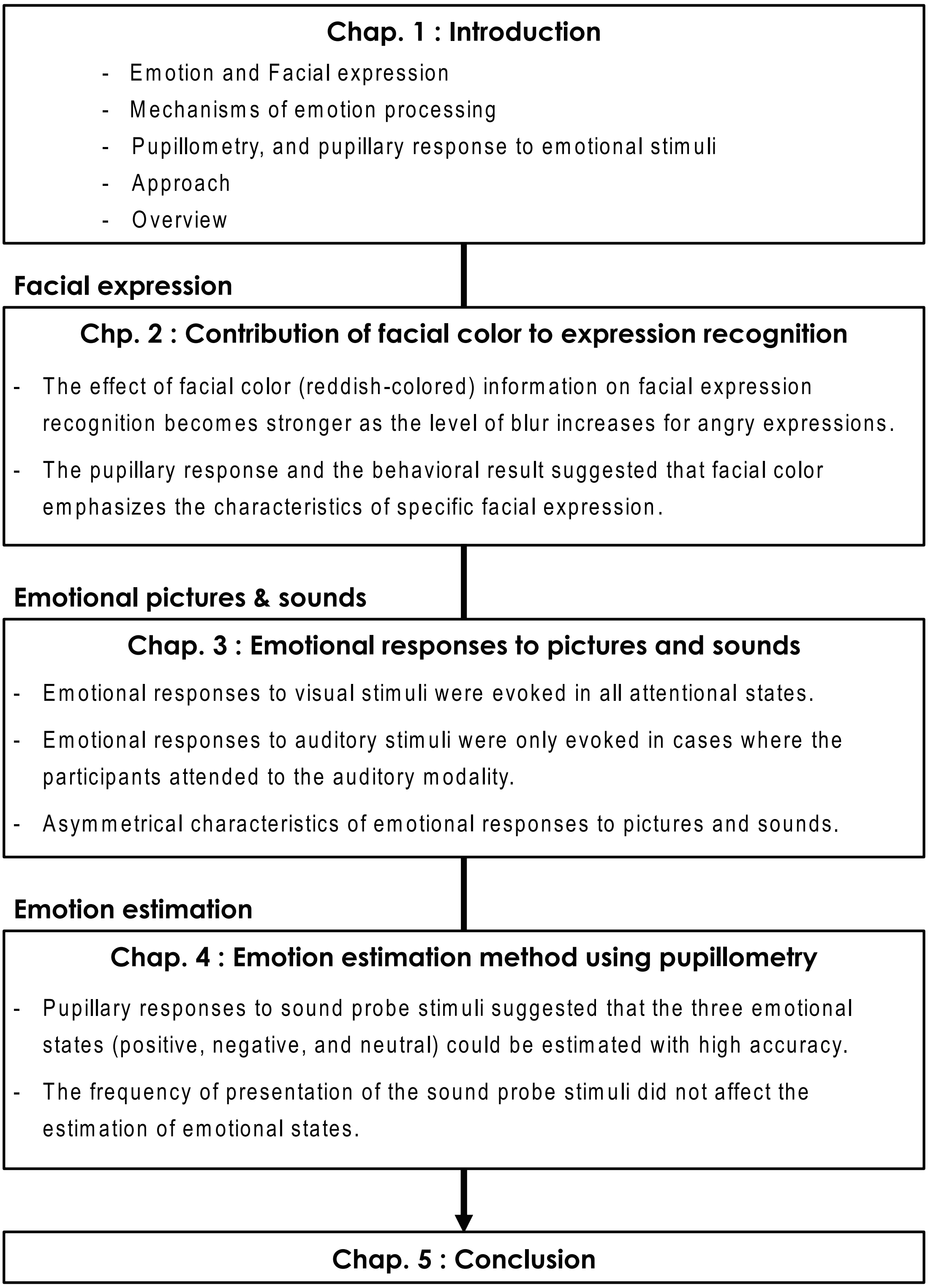

Figure 1.6: Flowchart of this thesis

Chapter 2

Contribution of facial color to expression recognition

A similar version of this chapter will be submitted as:

Nakakoga, S., Nihei, Y., Nakauchi, S., & Minami, T. (2019). Contribution of Facial Color to Expression Recognition of Blurred Faces. *Italic Transactions of Japan Society of Kansei Engineering, TJSKE-D*, 18(1), 79-85, DOI:10.5057/jjske.TJSKE-D-18-00035

The change of facial color and expression reflects our mental or physical condition. Previous behavioral studies indicated that there is a strong interaction between facial color and expression perception. This study investigated the contribution of facial color to expression recognition in blur images with the measurement of behavior and pupillary change. In the experiment, the face stimuli of facial colors (natural color, reddish) with different expressions (neutral and anger) in 3 blur levels were presented. Participants performed a task of expression identification to the stimulus. Behavioral results indicated that the facial color has a significant contribution to expression recognition as blur level increases. Then, the results of pupillometry showed that the reddish-color provided the information necessary to identify anger. These results showed the contribution of facial color increases in both psychophysics and pupillary experiment as blur level increases, which suggested that facial color emphasizes the characteristics of specific facial expressions.

2.1 Introduction

Facial color is an important signal to identify the emotions and health status of others. Many previous studies showed that facial color is one of the visual cues in person identification (Bruce and Langton, 1994; Yip and Sinha, 2002), facial attractiveness, and age determination (Fink et al., 2006; Jones et al., 2004; Kawaguchi et al., 2020; Fink et al., 2006). The information about facial color influences face processing. In our research group, experiments measured EEG during the presentation of the facial stimuli revealed that N170, an ERP component of face processing, is modulated by facial color, which suggested that facial color is important in the face detection process (Minami et al., 2011). Another study showed that the left FFA plays an important role in the processing of facial color (Nakajima et al., 2014). Thus, facial color plays an important role in promoting social communication. Several studies have suggested that humans have a higher discrimination sensitivity on facial skin than on patches in detecting subtle changes in redness (a* of CIE Lab color space), suggesting that humans have superior facial color discrimination ability (Tan and Stephen, 2013; Thorstenson et al., 2017). In addition, other studies have shown that the trichromacy of primate, including human, might be optimized for detecting changes in the skin color of others (Changizi et al., 2006; Hasantash et al., 2019; Hiramatsu et al., 2017).

Facial redness, among other facial colors, has been linked to the recognition of health status (Stephen et al., 2009), male dominance and aggression (Stephen et al., 2012). In fact, it is known that facial redness increases with a rise in blood flow caused by increased body temperature and the expression of anger (Drummond and Quah, 2001; Nordin, 1990). Therefore, many studies have reported an association between angry expressions and a facial reddish color (Drummond, 1997; Minami et al., 2018; Nakajima et al., 2017; Takahashi and Kawabata, 2018; Thorstenson et al., 2018; Young et al., 2018). In addition, the experiments of identification of facial expression using face images that do not involve facial muscle movement suggest that emotions could be read using facial color features associated with changes in facial blood flow, even in the absence of facial muscle activation. In addition, the experiments of identification of facial expression using face images that do not involve facial muscle movement suggest that emotions could be read using only facial color features associated with changes in facial blood flow, even in the absence of any facial muscle activation (Benitez-Quiroz et al., 2018). Nakajima et al. conducted 2AFC experiments for expression identification using reddish, bluish, and natural colored facial stimuli that morphed from anger to fear(Nakajima et al., 2017). The results indicated that reddish facial stimuli had lower thresholds of anger recognition than natural and bluish facial stimuli, and that facial color played an important role in facial expression recognition.However, these reports have been limited to psychological experimental methods such as questionnaires and psychophysical experiments, and there have been few neuroscientific approaches by measuring physiological responses. Therefore, it is still unclear what processing mechanism influences facial expression recognition.

In this study, we focused on the pupil diameter, which reflects human emotional states, as a method to measure the cognitive mechanism during facial expression judgment. The pupil is known to regulate the amount of light taken into the eye and also reflects mental activity (Hess and James, 1960, 1964). Bradley et al. a larger pupillary response when viewing emotional images than when viewing neutral images (Bradley et al., 2008). Nuske also reported different pupillary responses between fearful and neutral facial stimuli (Nuske et al., 2014a,b). Pupillometry is expected to be applied to various engineering applications as a non-contact physiological indicator, and in the future, it is expected to be used as a tool to assist human communication simply by measuring pupils without contact (Bradley et al., 2017). In addition, we focused on a previous study (Yip and Sinha, 2002) that showed that the addition of color to a face stimulus with less blurred shape information facilitates person identification. In this study, we aimed to clarify whether facial color information contributes to emotional state judgments in blurred facial stimuli and the mechanism of facial expression recognition in such stimuli. Nakajima et al. showed that angry faces with reddish-colored enhanced anger recognition (Nakajima et al., 2017). Therefore, we investigate the cognitive mechanism of the facial color effects on the discrimination of facial expressions based on the subjective responses when presented with facial stimuli that combine facial expressions (angry and neutral), facial color (natural-colored and reddish-colored), and level of blurring, as well as the pupillary response by measurement of eye movement.

2.2 Materials and methods

Participants

Ten students (8 males; mean age = 22.1 years; range = 21 - 23 years) participated in this experiment. All participants were students at the Toyohashi University of Technology. All participants had normal or corrected-to-normal vision based on self-reports. The experiment was performed in accordance with the Declaration of Helsinki, and all participants provided written informed consent before the experiment. The experimental procedure was approved by the Committee for Human Research at the Toyohashi University of Technology.

Stimuli

The experimental stimuli consisted of 12 conditions (see Fig. 2.1) that combine facial expressions (anger, neutral), and facial color (reddish-colored and natural-colored), with blur (no blur, low blur, high blur). The facial stimuli were used angry and neutral facial images from 48 face images (24 males and 24 females) licensed from Max Planck Insitute for Human Development (Ebner et al., 2010). Following Nakajima et al.'s experimental procedure, We created the following two facial color images by manipulating the CIE Lab a* values (green-red component) of the skin region of all face images (Nakajima et al., 2017).

1. Reddish-colored faces （$a^* + 12$ from original images）

2. Natural-colored faces （original images）

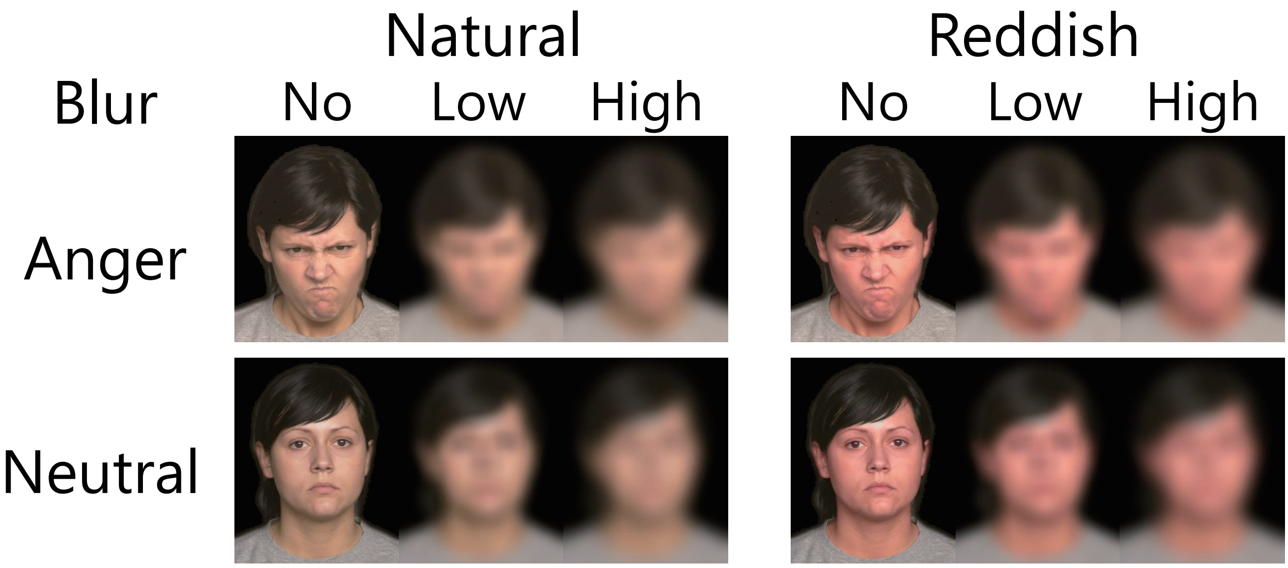

Figure 2.1: Example of the experimental stimuli

In addition, we applied Gaussian blur to all face images using MATLAB 2014b from MathWorks. We created blur images for the following three conditions by manipulating σ and $radius$. The size of all stimuli was 309 × 386 pixels (8.0 × 10.0 degrees of visual angle).

1. No Blur （Original images）
2. Low Blur （$\sigma = 20$，$radius = 40$）
3. High Blur （$\sigma = 20$，$radius = 60$）

Procedure

In the experiment, after participants received an explanation of the experimental procedure, they signed an informed consent form. Participants were then seated in comfortable chairs with their chins fixed at a viewing distance of 60 cm, in a shielded dark room. Visual stimuli were displayed on a calibrated 24-inch LCD monitor (ViewPixx3D, VPixx Technologies, 1920 × 1080 pixels, FrameRate 120Hz). The experiment was developed in Windows 10 and executed in MATLAB 2014b (MathWork Inc.) using Psychtoolbox 3 (Brainard, 1997).

For the experimental stimuli, 24 stimuli were prepared for each gender in 12 conditions of facial expression, facial color, and blur (in total 48 images for each face condition). In the experiment, 12

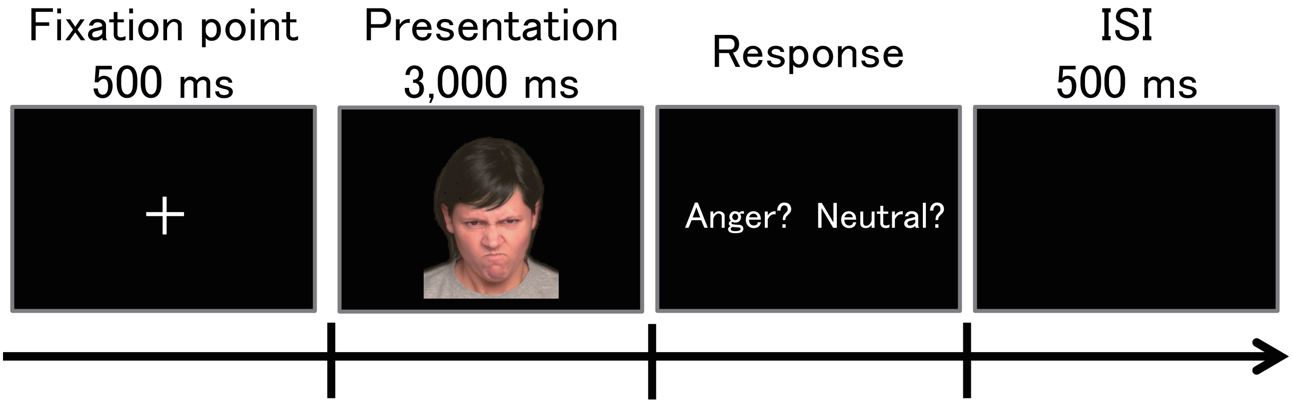

Figure 2.2: Experimental procedure

stimuli per condition were presented in random order per session (144 trials per session). A total of four sessions were repeated with ample rest between sessions. Each trial began with a 500 ms presentation of the fixation point and then the face stimulus was presented for 3,000 ms. The participants were taught to identify facial expressions (angry or neutral) of the face stimuli presented by the 2AFC task as quickly as possible with a mouse button response. After the participants responded and clicked the mouse again, a blank was presented for 500 ms and the next trial began. Fig. 2.2 shows the experimental procedure.

Pupillometry

Pupil diameter and eye movement during the presentation of the fixation point and face stimulus were recorded (in total 3,500 ms) using an eye-tracking system (Eyelink 1000, SR Research), at the sampling rate of 500 Hz. The eye-tracking system was desk mounted and used infrared video-based tracking technology. The movement of the participant's left eye was recorded using an infrared video camera at the resolution of up to 0.1°. A 9-point calibration and validation were performed before the start of each block to ensure that the participant's eyes were correctly tracked by the eye-tracking system. Participants were also instructed not to blink as much as possible during the eye movement measurements from the presentation of the fixation point to the stimulus presentation.

Data analysis

Behavioral data

We calculated the percentage of correct responses for the participants' responses, which indicates how well the participants could identify the facial expressions in all 12 conditions. We also defined the effectiveness ratio of facial color as the percentage of increase or decrease in the accuracy for the reddish-colored condition from the natural-colored condition in each expression and blur condition (see Equation 2.1).

$$Effectiveness\ ratio\ of\ facial\ color[\%] = \left(\frac{Accuracy\ for\ reddish - colored\ faces}{Accuracy\ for\ natural - colored\ faces} - 1 \right) \times 100 \tag{2.1}$$

Pupillary data

The pupil diameters measured on each trial were analyzed only trials which participant's identifications of facial expressions were correctly discriminated. In addition, we eliminated trials for the analysis in which the pupil diameter was not measured correctly due to blinking or other factors in each trial by visual selection. The pupillary responses were normalized on a trial-by-trial because pupil diameters varied across participants and are difficult for comparison.

Specifically, we defined *Pupil dilation* which noramalized by the baseline which is mean pupil diameter during the last 200 ms of the presentation of fixation point (see Equation 2.2), according to the analysis method of Wierda et al. (Wierda et al., 2012).

$$Pupil\ dilation = \frac{Pupil\ diameter\ during\ stimulus\ presentation}{Mean\ baseline\ pupil\ diameter} - 1 \tag{2.2}$$

The peak value of pupillary response (ratio of maximum constriction) was calculated on a trial-by-trial basis for pupil diameters normalized by Equation 2.2. We defined the change ratio in the peak value of the facial expression as the percentage change in the peak value of the anger condition from the neutral condition in each facial color and blur condition (see Equation 2.3).

$$change\ ratio\ in\ peak\ value\ of\ expression[\%] = \left(\frac{peak\ values\ for\ anger\ faces}{peak\ values\ for\ neutralr\ faces} - 1 \right) \times 100 \tag{2.3}$$

Statistical analysis

Two-way ANOVAs were performed using effectiveness ratio of facial color in each facial expression condition (anger and neutral) and the blur conditions (no blur, low blur, and high blur) as factors. In addition, another two-way ANOVAs were performed using the chage ratio in the peak value of the facial expression in each facial color condition (natural-colored and reddish-colored) and the blur conditions (no blur, low blur, and high blur) as factors. The level of statistical significance was set at $p < 0.05$ for all analyses. The pairwise comparisons for the main effects were corrected for multiple comparisons using the Ryan method. Effect sizes (η^2) were determined for the ANOVA.

2.3 Results

Behavioral data

Fig. 2.3 showed the accuracy rates for facial expression identification in all face stimulus conditions. Three-way ANOVAs (facial expression, color, and blur condition) revealed that the significantly main effect of the blur condition ($F(2, 18) = 98.912$; $p < 0.001$; $\eta^2 = 0.380$) Multiple comparisons of the main effect of the blur condition revealed that significant differences between no blur and low blur ($p < 0.100$), no blur and high blur ($p < 0.001$), and low blur and high blur ($p < 0.05$), respectively.

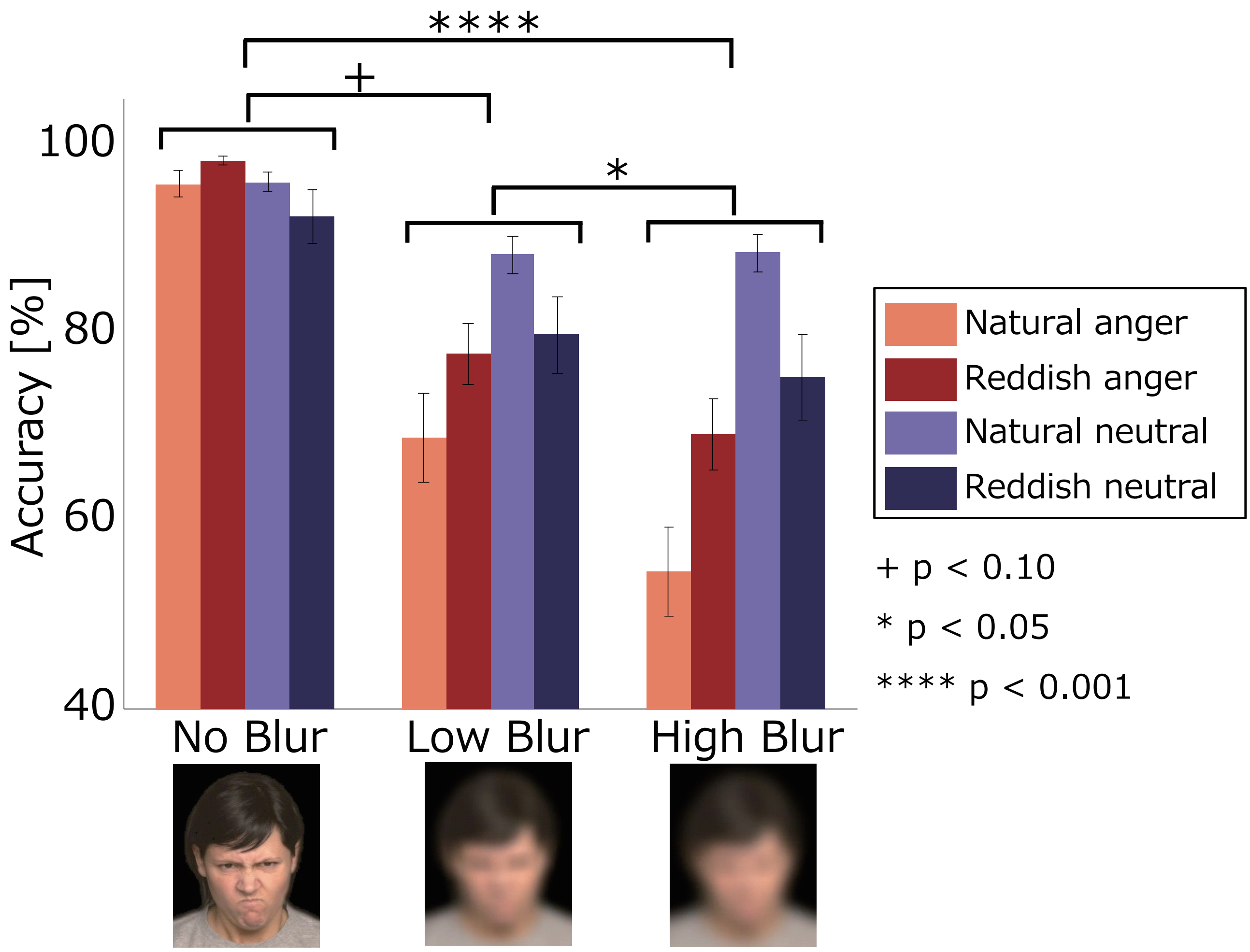

Figure 2.3: Accuracy rates for facial expression identification

Error bars show standard error of the mean across the participants. Asterisks indicate a significant difference based on multiple comparisons for the main effect of blur condition; + $p < 0.10$, * $p < 0.05$, **** $p < 0.001$.

Fig. 2.4 showed the effectiveness ratio of facial color in each facial expression and blur condition. Two-way ANOVAs (facial expression, and blur condition) revealed that the significantly interaction between the facial expression condition and the blur condition ($F(2, 18) = 6.612$; $p < 0.01$; $\eta^2 = 0.100$). We observed simple main effect of the blur condition on the angry faces ($F(2, 36) = 6.612$; $p < 0.01$; $\eta^2 = 0.100$). Further analyses revealed that the effectiveness ratio of facial color for the anger faces was significantly greater for the high blur faces than that for the no blur faces and low blur faces ($p < 0.001$, $p < 0.05$, for the no blur and low blur faces, respectively). The effectiveness ratio of facial color for angry faces was also a trend towards significantly greater for low blur faces than for the no blur faces ($p < 0.10$). However, we did not find significant simple main effect of the blur condition on the neutral faces ($F(2, 36) = 0.974$; $p > 0.05$; $\eta^2 = 0.010$). As a result, we confirmed that the effectiveness ratio of facial color for only angry faces increased significantly as the blur level increased.

Pupillary data

Fig. 2.5 showed the peak values of pupillary responses (ratio of maximum constriction) in all face stimulus conditions. Three-way ANOVAs (facial expression, color, and blur condition) revealed that the significantly interaction between the facial expression condition and the blur condition ($F(2, 18) = 4.461$; $p < 0.05$; $\eta^2 = 0.08$). We observed simple main effect of the facial expression condition on the no blur faces ($F(1, 27) = 9.337$; $p < 0.01$; $\eta^2 = 0.070$). However, we did not find significant simple main effect of the facial expression condition on the low blur faces ($F(1, 27) = 9.337$; $p < 0.01$; $\eta^2 = 0.07$) and high blur faces ($F(1, 27) = 0.007$; $p > 0.05$; $\eta^2 < 0.001$). As a result, we confirmed that the peak values of pupillary reponses for only no blur faces showed the differences between facial expression conditions.

Fig. 2.6 showed the change ratio in the peak value of the facial expression in each facial color and blur condition. Two-way ANOVAs (facial color, and blur condition) revealed that the significantly interaction between the facial color condition and the blur condition ($F(2, 18) = 3.715$; $p < 0.05$; $\eta^2 = 0.060$). Further analyses revealed no significant simple main effect of facial color condition on no blur faces $(1, 27) = 0.795$; $p > 0.05$; $\eta^2 = 0.010$). However, we found significantly simple main effect of facial color condition on low blur faces $(1, 27) = 3.447$; $p < 0.10$; $\eta^2 = 0.003$, a trend towards significance), and high blur faces $(1, 27) = 6.818$; $p < 0.05$; $\eta^2 = 0.06$). In addition, we found significant simple main effect of the blur condition on natural-colored faces $(2, 36) = 7.734$; $p < 0.005$; $\eta^2 = 0.18$). Multiple comparisons of the simple main effect of the blur condition on natural-colored faces revealed that significant differences between no blur and high blur ($p < 0.001$), low blur and high blur ($p < 0.01$). As a results, in the reddish-colored condition, the difference between the facial expression conditions was constant regardless of blur conditions, however, it increased significantly in the natural-colored condition.

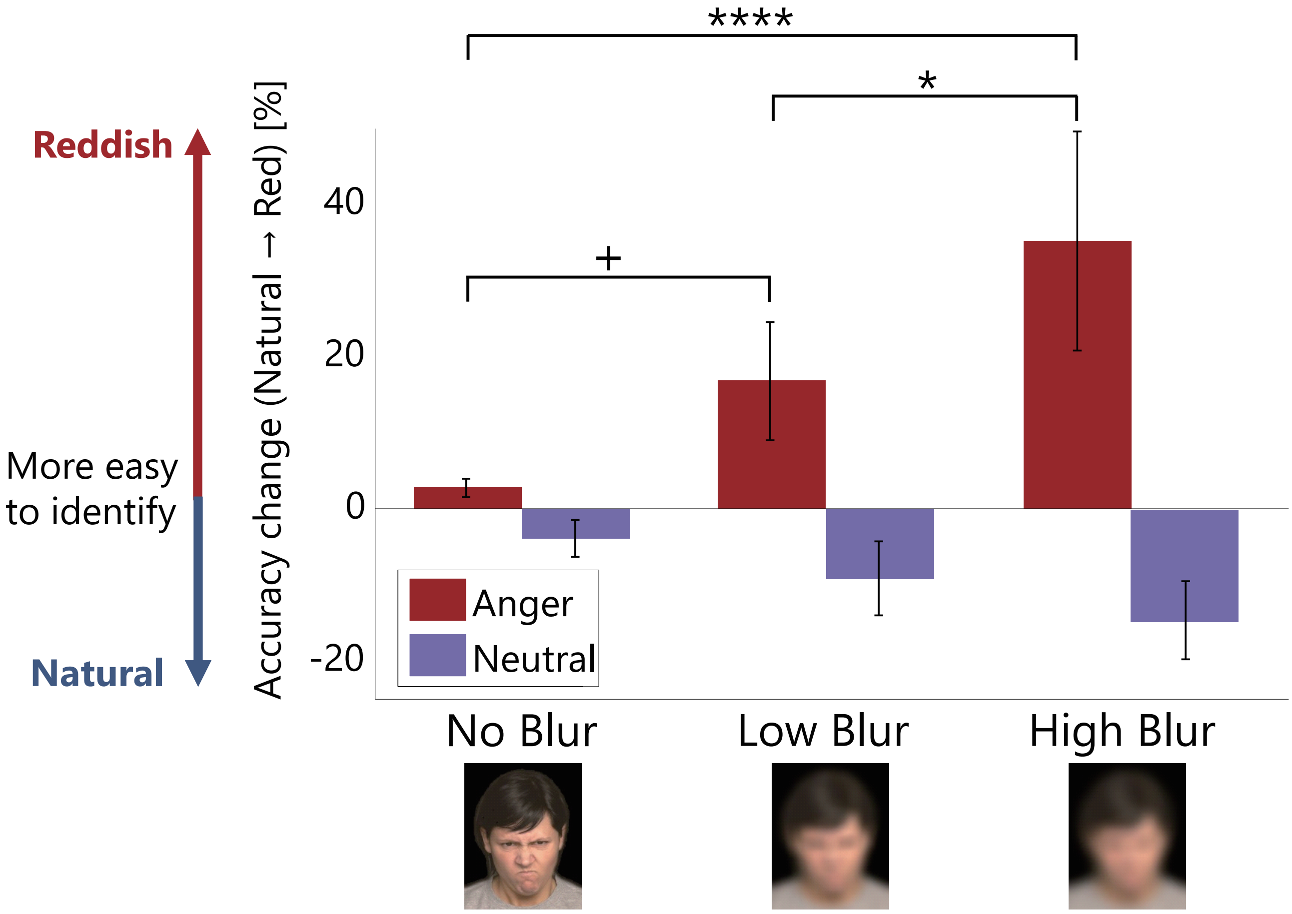

Figure 2.4: Effectiveness ratio of facial color

Error bars show standard error of the mean across the participants. Asterisks indicate a significant difference based on multiple comparisons for the simple main effect of blur condition on the angry faces; + $p < 0.10$, * $p < 0.05$, **** $p < 0.001$.

2.4 Discussion

Contribution of facial color on expression identification task

Fig. 2.3 shows that the accuracy rate of the expression identification task decreased as the blurring of the face stimuli became stronger. This is because the blurring of facial stimuli makes it difficult to identify the edges of facial parts that are necessary for facial expression recognition. Fig. 2.4 shows that there was a non-significant effect of facial color on facial expression recognition in no blur condition. A study by Kemp et al. reported that facial color had no effect on face recognition for face stimuli with maximum shape information without blurring(Kemp et al., 1996). Thus, The reason why

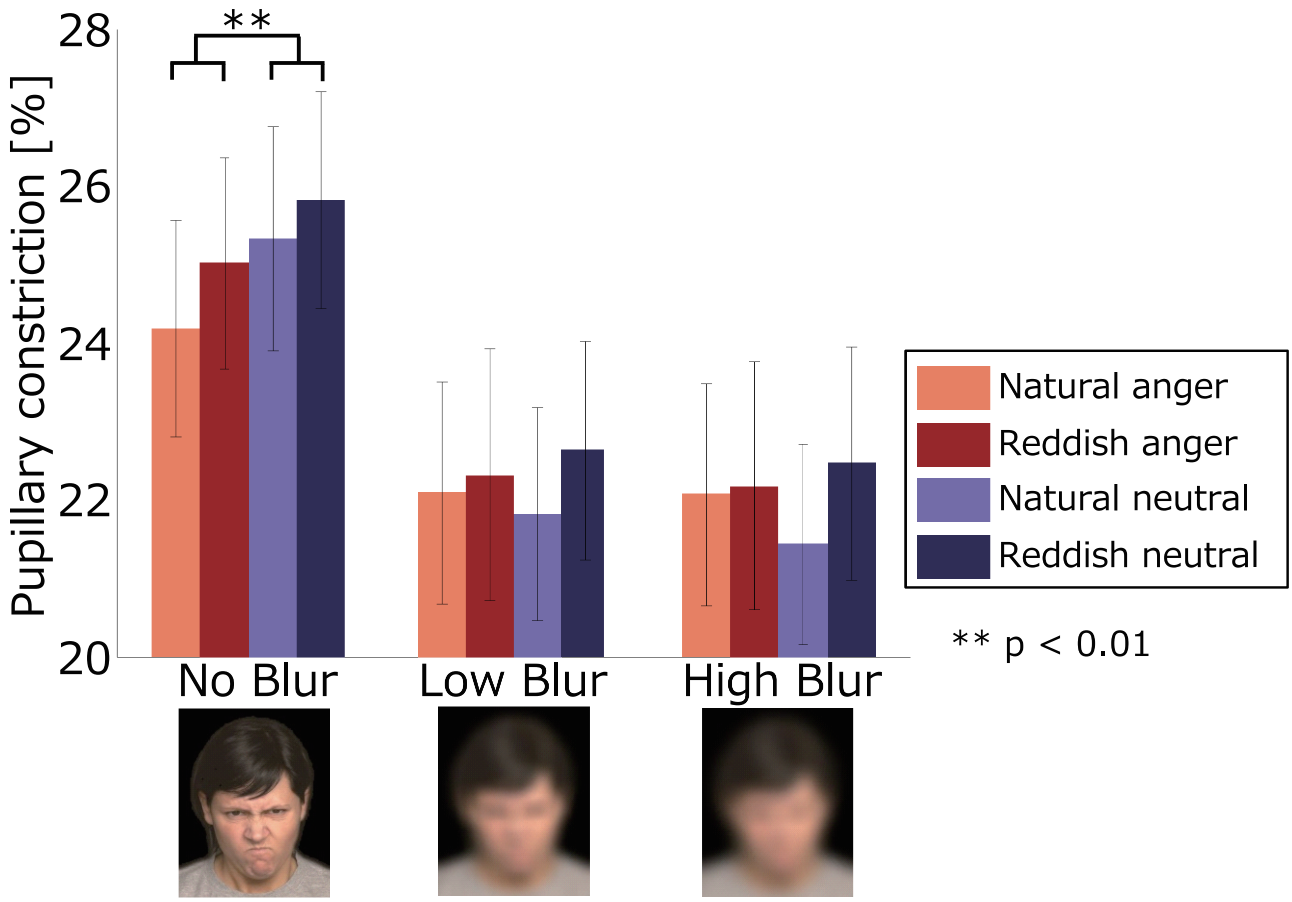

Figure 2.5: Peak values of pupillary responses (ratio of maximum constriction)

Error bars show standard error of the mean across the participants. Asterisks indicate a significant difference based on multiple comparisons for the simple main effect of facial color condition on no blur faces; ** $p < 0.01$.

facial color in no blur condition did not affect expression recognition in the expression identification task in this study is that the shape information of the face stimulus was the largest and therefore did not require facial color information. The reason why facial color in no blur condition did not affect expression recognition in the expression identification task in this study is that the shape information of the face stimulus was the largest and therefore did not require facial color information to identify the expression.

In the results of the effectiveness ratio of facial color, the simple main effect of blur condition on the angry faces was observed, which suggested the effect of reddish-colored on expression recognition significantly increased as the level of blur increased. This is similar to the results of the experiment about person identification in blur faces reported by Yip et al. because as the shape information is

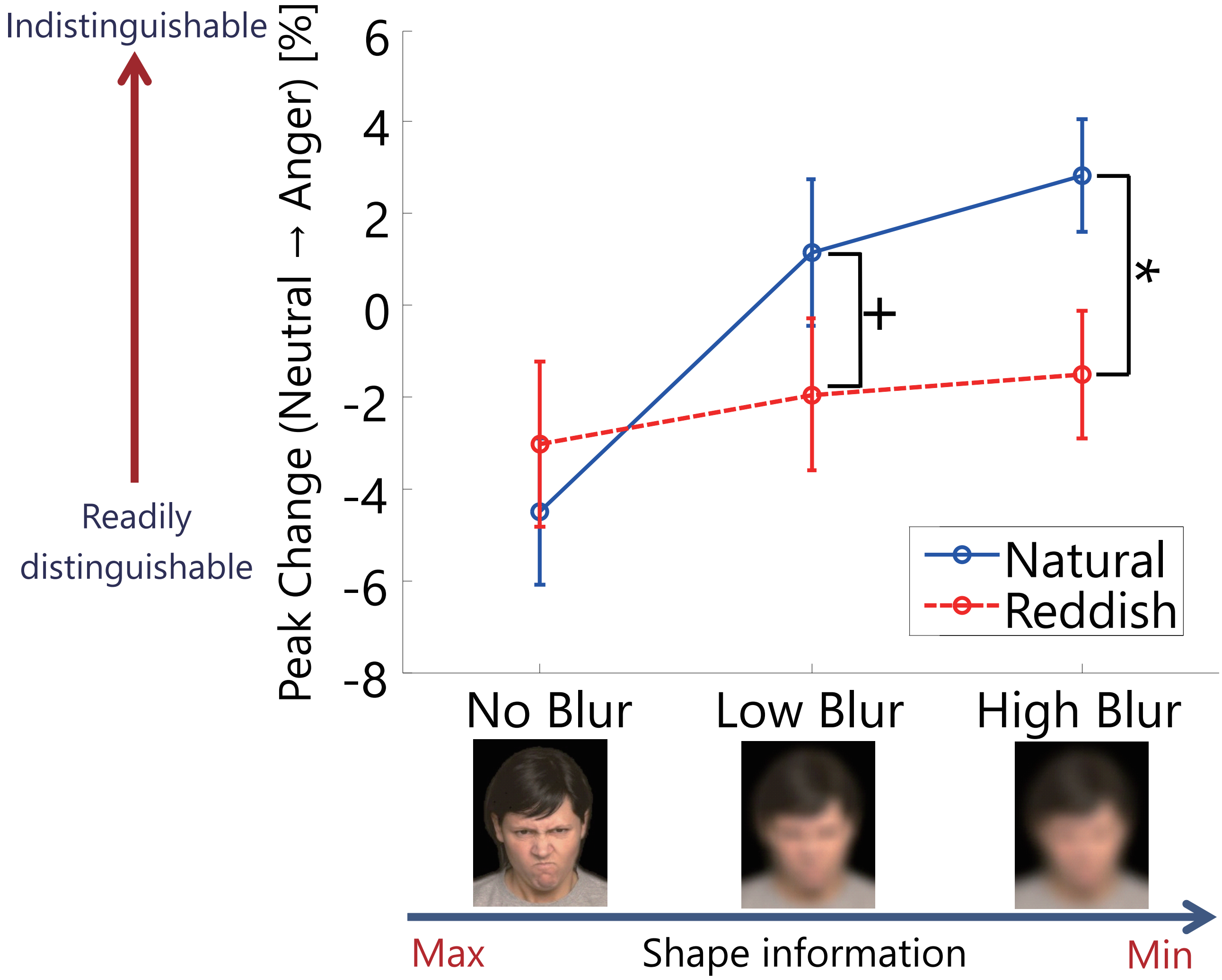

Figure 2.6: Change ratio in the peak value of the facial expression

Error bars show standard error of the mean across the participants. Asterisks indicate a significant difference based on multiple comparisons for the simple main effect of facial color condition of blur conditions; + $p < 0.10$, * $p < 0.05$.

missing, the information required for expression recognition decreases, and the contribution of facial color information on facial expression recognition increases accordingly (Yip and Sinha, 2002). In addition, the study by Nakajima et al. showed that when the emotional information from facial expressions was lacking, reddish skin color had a strong effect on anger perception as a supplement to the lack of emotional information from facial expressions (Nakajima et al., 2017). Similarly, in this study, the facial color effect supplementing the lack of emotional information due to the deficit of shape information by blurring was found in the expression identification task. If we assume that the facial color effect on expression recognition increases with the loss of shape information, a question arises as to whether neutral expressions may be perceived as angry by the red color. However, there was no simple main effect of the blur condition on the neutral faces, and the ratio that misperceived neutral face as anger did not increase significantly as the blur increased. Therefore, It is unlikely that facial expression identification in response to blurred facial stimuli are purely dependent on facial color, and the facial color effect may be only active when certain facial expressions are combined with certain facial color (e.g. angry faces with reddish-colored).

Contribution of facial color on expression recognition in terms of pupillary responses

A study by Nuske et al. showed that the peak values of the pupil constriction when viewing facial stimuli differs according to facial expressions, suggesting that emotional responses to facial expressions are reflected in ppupillary responses(Nuske et al., 2014b). Fig. 2.5 shows that, as in previous studies, the difference between angry and neutral faces was revealed in the peak values of pupil constrictions for the unblurred face stimulus. Henderson et al. showed that the pupillary response immediately after the presentation of an emotional image had a smaller peak value (initial light-reflection component) than that of a neutral image (Henderson et al., 2014). In the present study also, the peak values of pupil constriction to the emotional expression of anger were smaller than that of the neutral condition, suggesting that the pupillary response reflects the difference in the facial expression of the perceived facial stimulus. However, as the level of blur became stronger, the differences in pupillary responses between the facial expressions disappeared. This is similar to the behavioral results described above, and it is thought that the blurring of facial parts, which is necessary for expression recognition, made it difficult to distinguish the edges of facial parts, and this is the reason why there was no difference between facial expressions in pupillary response as a physiological index.

In Fig. 2.6, there was no significant simple main effect of facial color on no blur faces in the change ratio in the peak value of the facial expression. This result is thought to be due to the pupillary responses reflect the results of facial expression identification, which did not require information on facial color because the shape information of the face stimulus was the largest, same as the behavioral results. In the natural color condition, the change ratio in the peak value of the facial expression increased with the level of blurring, whereas the change ratio in the peak value of the facial expression in the reddish-colored condition was almost constant regardless of the blur condition. We assumed that the change ratio in the peak value of the facial expression in the no blurred condition, in which there were no difference responses between natural-colored and reddish-colored, is the value that is most distinguishable between the expressions (see Fig. 2.6). In this case, it is thought that in the natural-colored condition, the expressions were not able to identify as the level of blur increased, while in the red condition, the value of the change ratio in the peak value of the facial expression was almost constant, so that the expressions were able to distinguish in the same way as no blur faces. The results of the pupillary responses in Fig. 2.6 support the results of the facial expression identification task as a behavioral indicator, and they reflect the influence of facial expression identification when affected by the facial color.

Conclusion and Future work

In summary, the present study indicates that facial color may accentuate certain features of facial expressions (e.g., anger with reddish, sad with bluish, etc.) as a mechanism of facial expression recognition in influencing facial expression identifications. However, in this present study, we used facial stimuli with different facial skin color tones, although the physical luminance of all stimuli was consistent. It is known that the pupillary response reflects not only physical luminance but also subjective brightness (Laeng and Endestad, 2012). As with this previous study, the results of the pupillary responses obtained in the present study might be influenced by the color tone of the stimulus in addition to the cognitive factors of facial expression recognition. Therefore, as future work, it is necessary to measure pupillary responses using color patches to extract the effect of facial color on the pupillary response, taking into account the influence of color tone. In addition, because the number of participants in this experiment was small (ten participants), we could not find any significant differences in the time-averaged values of pupillary response, which are more common than the peak amplitude, as an index of pupil response. For future work, we need to add more data and conduct statistical tests using general index of the pupillary responses.

We used static face images as the experimental stimuli, not video clips to consider the effect of brightness change. However, in a recent study about pupillometry, there is a technique for estimating pupillary responses for only cognitive factors considering the changes in the brightness of a stimulus (Watson and Yellott, 2012). Pupil response to dynamic stimuli (i.e., video clips) by applying this method may be able to extract only from cognitive factors, taking into account brightness and color tone. In the future, we expect to develop tools to assist in the smooth communication between humans by measuring pupillary responses to dynamic facial stimuli.

Chapter 3

Emotional responses to pictures and sounds

A similar version of this chapter will be submitted as:

Nakakoga, S., Higashi, H., Muramatsu, J., Nakauchi, S., & Minami, T. (2020). Asymmetrical characteristics of emotional responses to pictures and sounds: Evidence from pupillometry. *PLoS ONE*, 15(4), e0230775, DOI:10.1371/journal.pone.0230775

In daily life, our emotions are often elicited by a multimodal environment, mainly visual and auditory stimuli. Therefore, it is crucial to investigate the symmetrical characteristics of emotional responses to pictures and sounds. In this study, we aimed to elucidate the relationship of attentional states to emotional unimodal stimuli (pictures or sounds) and emotional responses by measuring the pupil diameter, which reflects the emotional arousal associated with increased sympathetic activity. In experiment 1, we investigated the relationship of the attentional state with emotional visual stimuli (International Affective Picture System) and emotional responses by using pupillometry. We set four task conditions to modulate the attentional state (emotional task, no task, visual detection task, and auditory detection task). We observed that the velocity of pupillary dilation was faster during the presentation of emotionally arousing pictures compared to that of neutral ones, regardless of the valence of the pictures. Importantly, this effect was not dependent on the task condition. In experiment 2, we investigated the relationship of the attentional state with emotional auditory sounds and emotional pupillary responses. We observed a trend towards a significant interaction between the stimulus and the task conditions with regard to the velocity of pupillary dilation. In the emotional and auditory detection tasks, the velocity of pupillary dilation was faster with positive and neutral sounds than negative sounds. However, there were no significant differences between the no task and visual detection task conditions. Taken together, the current data reveal that different pupillary responses were elicited to emotional visual and auditory stimuli, at least in the point that there is no attentional effect to emotional pupillary responses to visual stimuli, despite both experiments being sufficiently

controlled to be of symmetrical experimental design.

3.1 Introduction

Emotions have been extensively investigated in fields such as cognitive psychology and neuroscience. Research on emotional processing has almost exclusively used a collection of emotional pictures known as the International Affective Picture System (IAPS; Lang et al., 2008) and a collection of emotional sounds known as the International Affective Digitized Sounds (IADS; Bradley and Lang, 2007) as the stimulus material. However, an issue with past studies of emotion research is that they have typically only investigated unimodal cues (Gerdes et al., 2014).

Most studies on audiovisual integration, such as the McGurk effect, have reported that the influence of visual stimuli in these interactions is dominant over that of auditory stimuli (McGurk and John, 1976; Kitagawa and Ichihara, 2002). The fMRI study by Anders et al. revealed a higher differential activation in the amygdala during exposure to the emotional pictures from the IAPS compared to the emotional sounds from the IADS (Anders et al., 2008). Another study suggested that differential brain activation in response to the emotional pictures and sounds may not be due to the differences in the emotional processing; rather, it may be due to methodological differences and different stimulus characteristics (Gerdes et al., 2014). Although emotional responses to sounds are weaker (Bradley and Lang, 2000) and occur later (Thierry and Roberts, 2007), the processing of emotional sounds and pictures is comparable based on behavioral, physiological, and electrophysiological reactions (Schupp et al., 2003; Czigler et al., 2007). However, it is not clear whether the characteristics of emotional responses to pictures and sounds are symmetric in every situation. To assess this, the present study investigated the relationship of attentional states to emotional unimodal stimuli (pictures or sounds) based on the hypothesis that emotional pupillary responses elicited by pictures and sounds are asymmetric. The subjective evaluation methods of emotional stimuli, including the Self-Assessment Manikin (SAM; Lang, 1980), which is a method for the subjective rating of the dimensions of valence and arousal, do not reflect the emotional state when paying attention to emotional stimuli (pictures or sounds), because it is conducted following a stimulus presentation. Therefore, in the present study, we measured the pupil diameter, which reflects the emotional arousal associated with increased sympathetic activity, as a physiological indicator to record the emotional state during the stimulus presentation.

Pupil dilation is controlled by the level of activation of the locus coeruleus (LC), which has an important role in cognitive function and arousal (Aston-jones and Cohen, 2005; Benarroch, 2009). Therefore, a pupillary response is considered a physiological indicator of mental activity (Hess and James, 1960, 1964; Beatty and Lucero-Wagoner, 2000). Studies on the association between pupillary changes and emotional responses have reported that pupil diameter was dilated as a results of increase in arousal following presentation of an emotional stimuli such as the IAPS pictures and the IADS sounds (Partala and Surakka, 2003; Bradley et al., 2008; Kashihara et al., 2014; Henderson et al., 2014; Bradley and Lang, 2015; Snowden et al., 2016; Finke et al., 2017; Kinner et al., 2017). Snowden et al. reported that pupil dilation in response to emotional pictures was not affected by actively naming the emotion of the stimuli compared to passive viewing (Snowden et al., 2016). In the study by Kinner et al, pupillary responses were measured while participants controlled their emotional states with cognitive emotion regulation (e.g., reappraisal, distraction) during the presentation of the IAPS pictures. When participants performed emotion regulation, pupil dilation was decreased compared to the passive viewing of the negative pictures. This suggested that pupillary response could be an indication of the success or failure of the emotion regulation (Kinner et al., 2017).

This study investigated the relationship of attentional states with emotional unimodal stimuli (pictures or sounds) using pupillometry in two experiments. Typical examples of the distraction task during a presentation of emotional stimuli are continuous target counting task by Schupp et al. (Schupp et al., 2007, 2008). However, these tasks are not suitable for pupillometry, which is sensitive to the change of brightness and requires an appropriate interval between trials. Thus, in addition to the 2 tasks in Snowden's experiment (emotional task, no task), we added two new tasks to capture visual or auditory attention (visual detection task, and auditory detection task). In these tasks, a visual or auditory target (a dot or beep sound) was presented in random trials. However, the trials in which the target appears in these tasks cannot be compared with other tasks due to the difference in screen brightness or sound frequency. Accordingly, we attempted to control and equalize attentional states in all task conditions by analyzing the pupillary response of the trials without the target. In addition, the stimulus conditions of each emotional stimulus in this study were categorized based on subjective evaluation in each participants because objective and subjective assessments of the IAPS and IADS are not always consistent, and there is an individual difference (Cuthbert et al., 1996; Aluja et al., 2015). Our hypothesis is that the emotional pupillary responses to both the images and the sounds stimuli might be suppressed when attention was paid to the target stimuli of the same modality as to emotional stimuli. Although there is evidence for this emotional suppression effect for vision (Schupp et al., 2007), it is still not completely understood whether the same effects occur for audition. Therefore, we investigated whether emotional pupillary responses to visual and auditory stimuli had symmetrical properties by comparing the pupillary response depending on the emotional state in each task condition.

3.2 Materials and methods

Participants

Twenty-six healthy participants (10 females) participated in Experiment 1 (mean age = 22.88 years, S.D = 2.65), and 24 healthy participants (8 females) participated in Experiment 2 (mean age = 22.00 years, S.D. = 1.53). Six of the 24 participants from Experiment 2 also participated in Experiment 1 approximately 1 month before Experiment 2. They were recruited through e-mail by us. Their performance did not differ from the other participants. All participants are students or staff in Toyohashi University of Technology. They had normal or corrected-to-normal vision, normal hearing based on self-reports and they were not informed of the purpose of the study. Two participants from

Experiment 1 were excluded from the analyses due to artifacts such as eye blinks and missing data. A power analysis using G*Power software (G*power 3.1; Faul et al., 2007) indicated that a sample size of 24 would achieve 80% power for a repeated measures design, given a medium effect size ($f = 0.25$) and $\alpha = 0.05$. All participants provided written informed consent. The experimental procedure received approval from the Committee for Human Research at Toyohashi University of Technology. The experiments were conducted strictly in accordance with the approved guidelines of the committee.

Stimuli

For Experiment 1, pictures were selected from the IAPS based on the normative ratings (Lang et al., 2008). Sets of 20 positive pictures (valence: mean = 7.29, S.D. = 0.40; arousal: mean = 6.20, S.D. = 0.84), 20 negative pictures (valence: mean = 2.77, S.D. = 0.57; arousal: mean = 6.20, S.D. = 0.55), and 20 neutral pictures (valence: mean = 5.06, S.D. = 0.26; arousal: mean = 4.24, S.D. = 0.98) were created (see Fig. 3.1). All pictures were in landscape orientation (19.0 × 14.3 degrees of visual angle) and were displayed in grayscale. Mean luminance of the selected pictures was matched using the MATLAB R2016a (MathWork Inc.) SHINE toolbox (Willenbockel et al., 2010). In order to match the level of luminance prior to the picture onset, the luminance of the gray background was controlled with the mean luminance computed across all pictures.

For Experiment 2, sounds were selected from the IADS based on the normative ratings (Bradley and Lang, 2007). Sets of 20 positive sounds (valence: mean = 7.22, S.D. = 0.55; arousal: mean = 6.30, S.D. = 0.85), 20 negative sounds (valence: mean = 2.80, S.D. = 0.45; arousal: mean = 6.28, S.D. = 0.63), and 20 neutral sounds (valence: mean = 4.99, S.D. = 0.40; arousal: mean = 4.32, S.D. = 0.61) were created (see Fig. 3.2). The stimuli were selected for comparable valence and arousal ratings between pleasant and unpleasant stimuli and between Experiment 1 and Experiment 2. All sounds were normalized for peak amplitude to standardize their loudness. In order to match the level of luminance between Experiment 1 and 2, the luminance of the gray background was controlled with the mean luminance computed across all pictures in Experiment 1.

Procedure

Experiment 1 and Experiment 2 were conducted separately and on different days. In both experiments, after participants received an explanation of the experimental procedure, they signed an informed consent form. Participants were then seated in comfortable chairs with their chins fixed at a viewing distance of 60 cm, in a shielded dark room. Visual stimuli were displayed on a calibrated 24-inch LCD monitor (ViewPixx3D, VPixx Technologies). Auditory stimuli were presented through headphones (SoundTrue around-ear headphones, BOSE). All experiments were developed in Windows 10 and executed in MATLAB 2014b (MathWork Inc.) using Psychtoolbox 3 (Brainard, 1997).

Experiment 1. Fig. 3.3a shows the protocol for Experiment 1. The experiment consisted of 60 IAPS pictures and was conducted over four blocks. In these four blocks, the 60 stimuli were

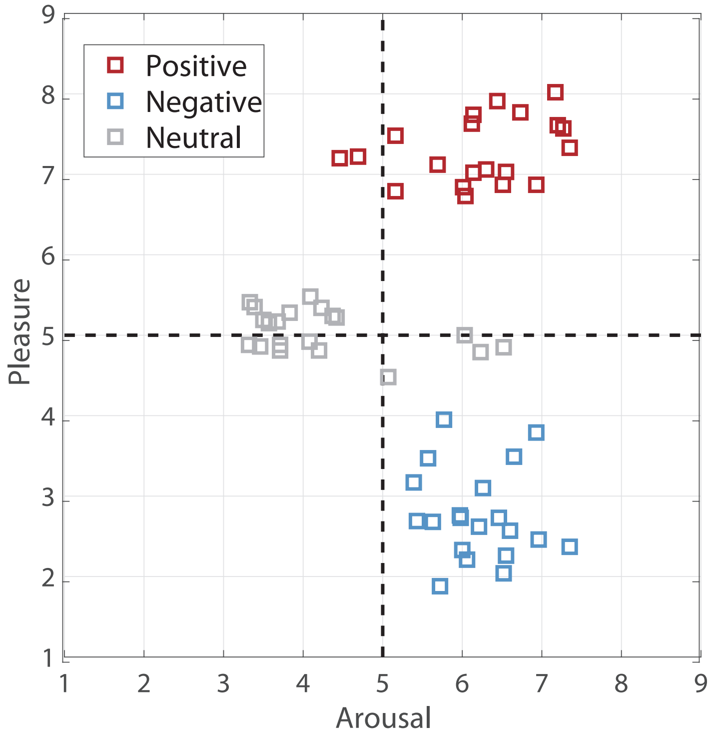

Figure 3.1: Rating scores of the IAPS pictures in Experiment 1

identical, and each stimulus was presented only once within each block in random order. In each trial, a fixation point was presented for 1,000 ms prior to the presentation of the stimulus. Each IAPS picture was presented for 6,000 ms. The participants were instructed to fixate on a central point during the presentation of the fixation point and stimulus. They responded to the task in each block by pressing a key after stimulus presentation. The tasks in each block differed as follows: (1) Emotional task: the subjective emotional evaluation of presented pictures (positive, negative, or neutral); (2) No task: only pressing the key as a control condition; (3) Visual task: the visual detection of a dot, which was added to the IAPS picture during stimulus presentation at random (appearance or no appearance); (4) Auditory task: the auditory detection of a beep sound, which was emitted during stimulus presentation through a set of headphones at random (appearance or no appearance). In the visual task block, the target dot in random trials was added to a picture for 100 ms at random intervals during the stimulus presentation (somewhere between 1,000 ms and 4,500 ms after stimulus presentation). The target dot was a circle of 3 pixels radius, and it was superimposed at a random position onto the IAPS pictures. In the auditory task block, the target beep sound in random trials was presented for 300 ms at random intervals during stimulus presentation (somewhere between 1,000 ms and 4,500 ms after stimulus presentation). The presentation of the target dot or beep sound accounted for 20% of all trials in each block. However, these trials could not be analyzed for pupillary

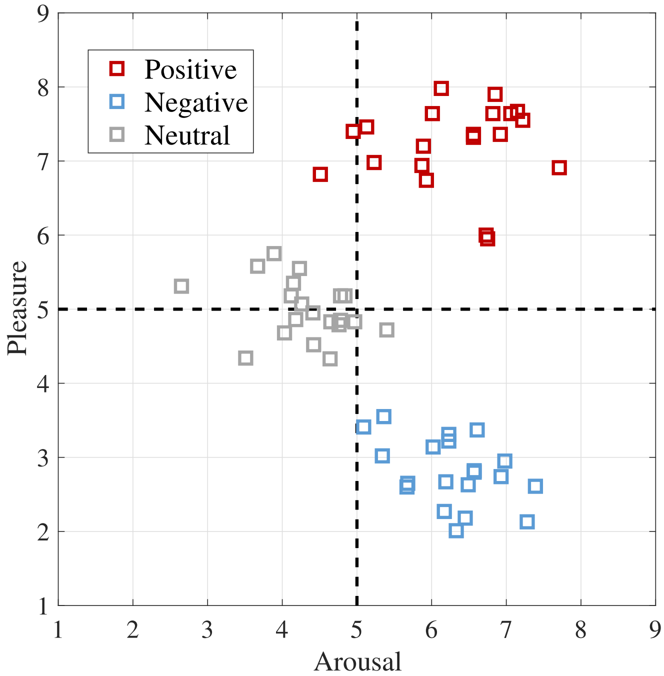

Figure 3.2: Rating scores of the IADS sounds in Experiment 1

response because of differences in experimental conditions compared to the other trials. Therefore, the number of trials in the visual and auditory tasks was 75, in order to equalize the number of analyzable trials in all task conditions. In order to equalize the accuracies of visual and auditory tasks at 75% in the preliminary experiment, the target dot (area and color) and target beep sound (frequency and volume) were modulated. Participants had a maximum of 5,000 ms to respond, after which the next trial started automatically. Participants rested for at least 5 min between each block. The order of the tasks was counterbalanced across the participants.

Experiment 2. Fig. 3.3b shows the protocol for Experiment 2. The experiment consisted of 60 IADS sounds and was conducted over four blocks. The presentation times, participants' tasks, and the number of trials in Experiment 2 were identical to those in Experiment 1. However, instead of the IAPS pictures in Experiment 1, IADS sounds were presented as stimuli. In the visual task block, the target dot was a circle of 3 pixels radius, and it was presented at a random position within the same range as Experiment 1. In the auditory task block, the target beep sound was synthesized as part of the IADS sound. In order to equalize the accuracies of the visual and auditory tasks at 75% in a preliminary experiment, the target dot (area and color) and the target beep sound (frequency and volume) were modulated with different parameters to those in Experiment 1 as the IADS sounds were presented instead of the IAPS pictures.

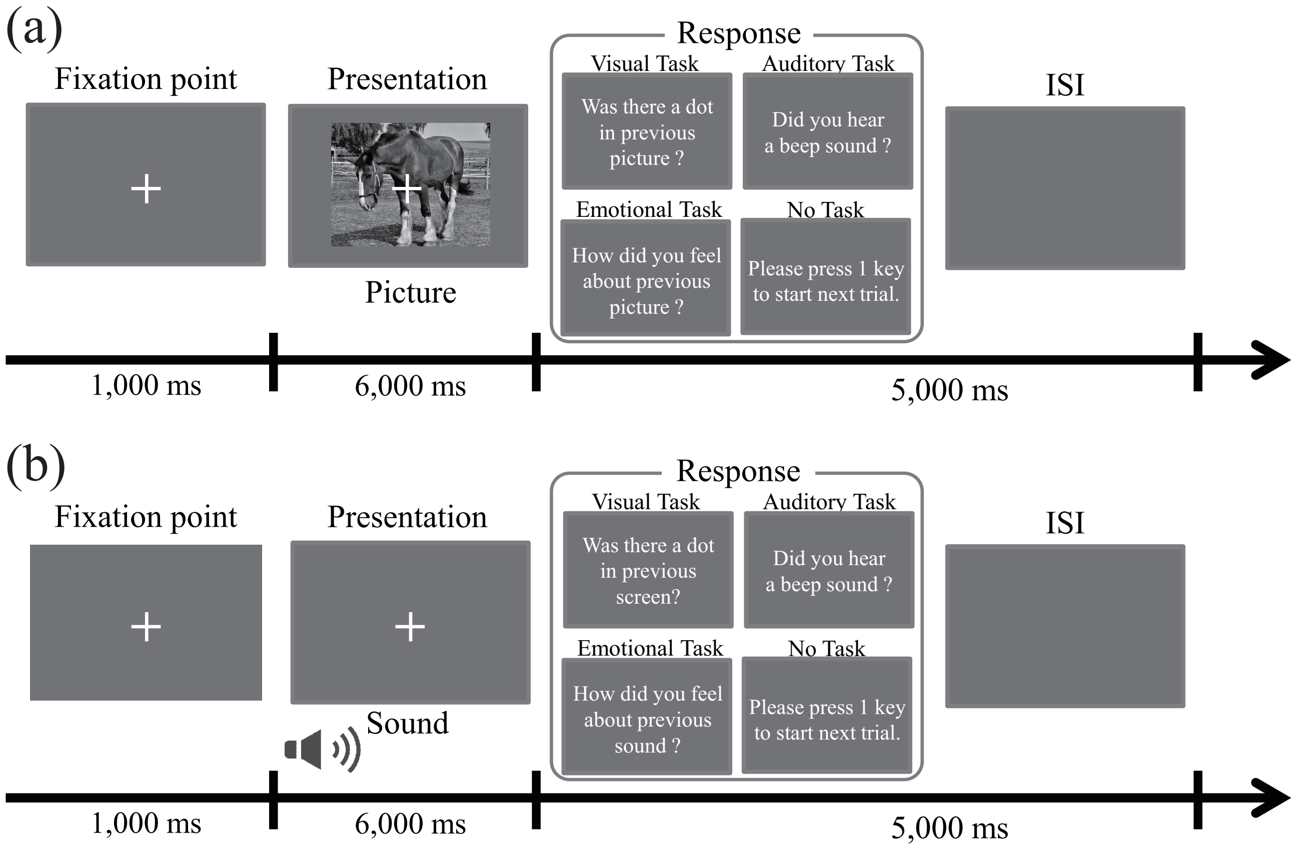

Figure 3.3: Experimental procedure

(a) Protocol for Experiment 1. In each trial, the fixation point was presented for 1,000 ms. The IAPS picture was then presented for 6,000 ms. Each trial was separated by a response period and an inter-stimulus interval (ISI) of 5,000 ms in total. Participants responded to the task in each block following the stimulus presentation using the keypad. (b) Protocol for Experiment 2. The paradigm (presentation times, participants' tasks, and the number of trials) was identical with that in Experiment 1. However, IADS sounds and not IAPS pictures were used.

Pupillometry

Pupil diameter and eye movement during the presentation of the fixation point and stimulus were recorded using an eye-tracking system (Eyelink 1000, SR Research), at the sampling rate of 500 Hz. The eye-tracking system was desk mounted and used infrared video-based tracking technology. The movement of the participant's left eye was recorded using an infrared video camera at the resolution of up to 0.1°. A 9-point calibration and validation were performed before the start of each block to ensure that the participant's eyes were correctly tracked by the eye-tracking system.

Data analysis

Behavioral data

The rates at which participants' responses to the stimuli in the emotional task matched the valence of the stimuli. They were computed for both experiments and analyzed with a one-way repeated-measures analysis of variance (ANOVA) for each stimulus condition (positive, negative, and neutral) as factors. The level of statistical significance was set at $p < 0.05$ for all analyses. Pairwise comparisons for the main effects were corrected for multiple comparisons using the Bonferroni method. Effect sizes ($partial\ \eta^2$) were determined for the ANOVA. In addition, the detection rates of the target dots or beep sounds for visual or auditory tasks were computed, and a paired t-test was used to compare the detection rates between visual and auditory tasks in each experiment.

Pupillary data

The pupillary data were prepared and analyzed using MATLAB 2017b (MathWork Inc.). The trials in which the target dot or beep sound appeared in the visual or auditory tasks were rejected from the pupillary data analysis. In both experiments, some erotic images and sounds were included in positive stimuli. Several studies have shown gender differences in emotional responses to erotic images (Wrase et al., 2003; Sebatinelli et al., 2004). In fact, the results of emotion classification in the emotional task indicated that there were many erotic stimuli that the male participants responded to positively, whereas the female participants responded negatively rather than positively. Therefore, pupillary data in each trial was classified into three conditions (positive, negative, and neutral) according to the participants' responses in an emotional task rather than the value of the emotional valence based on the normative ratings in IAPS and IADS. Due to this classification method, the trials analyzed for pupillary response in each emotion condition were in the range of 10 to 40. The eye blinks were interpolated using cubic-spline interpolation (Mathôt et al., 2018). The trials containing artifacts were removed using the principal component analysis (PCA) and the peak change of the velocity of the pupil response.

The PCA was performed to reject the trials including potential contaminant using all pupillary time course data in each experiment as the input data. The threshold for detection of the peak change of the velocity was defined based on Mathôt et al.'s research (Mathôt et al., 2018). It was used to reject the trials which were not interpolated the eye blinks and the loss for a sustained period by cubic-spline interpolation. Based on these methods for artifact removal, 4.36 trials in Experiment 1 and 3.83 trials in Experiment 2 for each condition, on average, were excluded from the analysis. Two participants from Experiment 1 were excluded because all trials in one of the experimental conditions (20 trials) or more than 75% of one of the task conditions (60 trials) were rejected due to artifacts. After the artifact removals, the pupillary data for the analyses in each emotion condition were in the range of 9 to 39.

In the time course analysis, each trial data was down-sampled to 50 Hz and subsequently each data point ± four sampling points were smoothed. Next, the pupillary area data were converted to diameter data in accordance with the method proposed by Hayes and Petrov (2016) because the pupil size was generated by the device in arbitrary units (Hayes and Petrov, 2016). The baseline pupil size was computed as an average of the data collected prior to the stimulus onset (picture or sound presentation), from -200 ms to 0 ms (presentation onset). The baseline-corrected average pupil sizes for the presentation period of 0 ms to 5,800 ms in both experiments were compared to each condition (Suzuki et al., 2018).

The pupillary dilator muscles (radial) reflect the activity of the locus ceruleus (LC), which contaiens a large amount of noradrenaline (Aston-jones and Cohen, 2005). Therefore, the velocity of pupil dilation is the best index to extract the pure activity of the pupillary dilator muscle. In this study, a gradient value for pupil dilation after stimulus presentation was used as a physiological index of emotional response. Based on Kinner's research, a gradient in Experiment 1 was calculated between the minimum pupil diameter (between 0 and 1,000 ms after the picture onset) and maximum pupil diameter (between 1,000 and 2,000 ms after the picture onset) (Kinner et al., 2017). Similarly, a gradient in Experiment 2 was calculated between the maximum pupil diameter (between 0 and 1,000 ms after the sound onset) and maximum pupil diameter (between 1,000 and 2,000 ms after the sound onset). Pupillary responses in Experiment 1 constricted from the picture presentation due to the light reflex, whereas pupillary responses in Experiment 2 dilated due to the sound stimuli and no change in the screen. The gradient values were used as physiological indexes in this study (see Fig. 3.4 and Fig. 3.5). They were computed and analyzed using a repeated measures ANOVA for each stimulus condition (positive, negative, and neutral).

Statistical analysis

Two-way ANOVAs were performed using the pupillary gradient and behavioral reponses in each stimulus condition (positive, negative, and neutral) and the task conditions (emotional task, no task, visual task, and auditory task) as factors. The level of statistical significance was set at $p < 0.05$ for all analyses. The pairwise comparisons for the main effects were corrected for multiple comparisons using the MSRB (Modified Sequentially Rejective Bonferroni) procedure (Seaman et al., 1991). Effect sizes ($partial\ \eta^2$) were determined for the ANOVA.

3.3 Results

Behavioral data

Fig. 3.6 a shows the rates at which participants' responses to the pictures and sounds in the emotional task of Experiment 1 and 2 matches the valence of the pictures and sounds. ANOVA revealed no main effect of the stimulus condition in Experiment 1 ($F(2, 46) = 1.158$; $p = 0.323$; $partial\ \eta^2 = 0.048$). In addition, we analyzed these rates with a Bayesian repeated measures ANOVA using the statistical

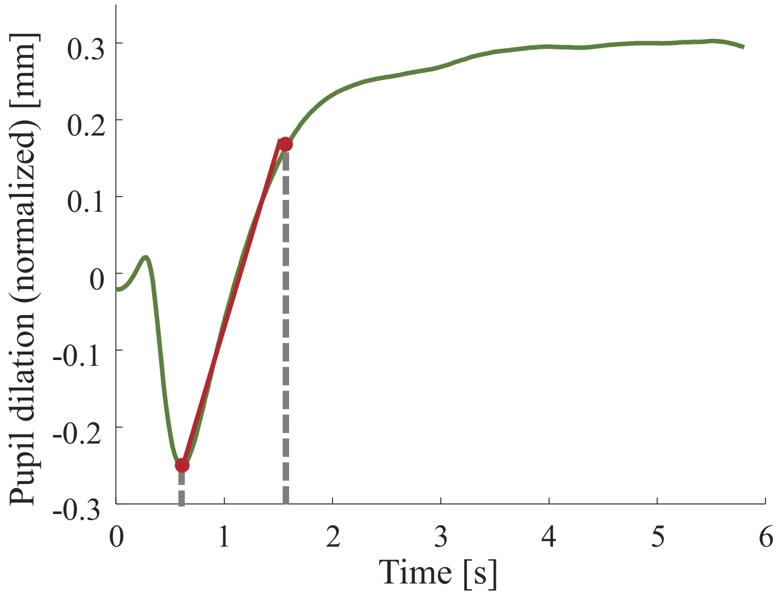

Figure 3.4: How to calculate gradient values in Experiment 1

A gradient in Experiment 1 was calculated between the minimum pupil diameter (between 0 and 1,000 ms after the picture onset) and maximum pupil diameter (between 1,000 and 2,000 ms after the picture onset) (Kinner et al., 2017).

software JASP (httos://jasp-stats.org/). We concluded that there is no main effect of the stimulus condition in Experiment 1. We found a BF_{10} smaller than 0.33 in the main effect of the stimulus ($BF_{10} = 0.323$). Thus, we have evidence in support of the null hypothesis (Dienes, 2014) that there is no main effect of the stimulus. These results indicate that participants' emotions elicited by IAPS pictures in Experiment 1 were impartial. On the other hand, there is a significant main effect of the stimulus condition in Experiment 2 ($F(2, 46) = 7.923$; $p = 0.001$; $partial\ \eta^2 = 0.256; BF_{10} = 64.401$). Planned comparisons revealed that the rate at which responses matched the valence of the negative sounds was significantly higher than that for the neutral sounds ($p = 0.003$) and showed the trend to be significantly higher than that for the positive sounds, although it did not reach statistical significance ($p = 0.061$). Moreover, the rate at which responses matched the valence of the positive sounds was significantly higher than that for the neutral sounds ($p = 0.048$). These results indicate that negative emotions, rather than positive or neutral emotions, were most strongly elicited by the IADS sounds in Experiment 2.

Fig. 3.6 b shows the detection rates of the target dot and beep sound for the visual and auditory tasks in both experiments. T-tests revealed no significant differences for task conditions in both experiments (Experiment 1: $p = 0.940$, Experiment 2: $p = 0.637$). In addition, we analyzed these rates with Bayesian paired sample t-tests. Bayesian t-tests showed moderate evidences for the null

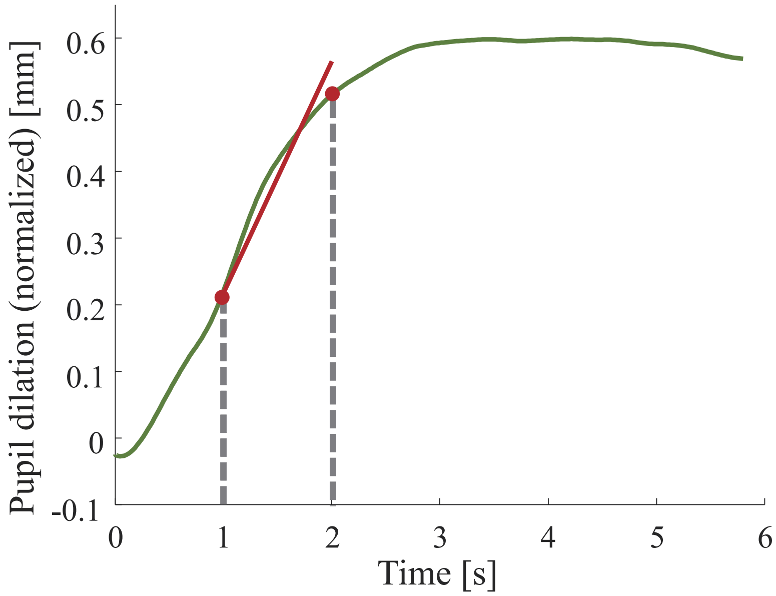

Figure 3.5: How to calculate gradient values in Experiment 2

A gradient in Experiment 2 was calculated between the maximum pupil diameter (between 0 and 1,000 ms after the sound onset) and maximum pupil diameter (between 1,000 and 2,000 ms after the sound onset)

hypothesis that there is no difference in the detection rates between visual and auditory tasks in both experiments (Experiment 1: $BF_{10} = 0.215$, Experiment 2: $BF_{10} = 0.238$). These results indicate that the difficulties of the visual and auditory tasks in both experiments did not differ.

Pupillary data

Fig. 3.7 and 3.8 show the grand-average time course of changes in the pupil dilation during stimulus presentation (6,000 ms) in each task condition for Experiment 1 and 2. Fig. 3.9 a shows the gradient values of the pupillary response to the IAPS pictures for each condition in Experiment 1. ANOVAs revealed a significant main effect of the stimulus condition ($F(2, 46) = 10.660; p < 0.001; partial\ \eta^2 = 0.317; BF_{10} = 2.170 \times 10^3$). The planned comparisons revealed that the gradient values for the negative and positive pictures were significantly greater than those for the neutral pictures ($p = 0.002$ and $p = 0.002$ for negative and positive pictures, respectively). However, the comparison between the positive and negative pictures revealed no significant difference ($p = 0.735$).

A significant main effect of the task condition was observed for the gradient values ($F(3, 69) = 15.304$; $p < 0.001$; $partial\ \eta^2 = 0.400; BF_{10} = 4.804 \times 10^{12}$). The planned comparisons revealed that the gradient value for the emotional task was significantly greater than that for the no task, the visual task, and the auditory task ($p = 0.001, p = 0.001, p < 0.001$, respectively). However,

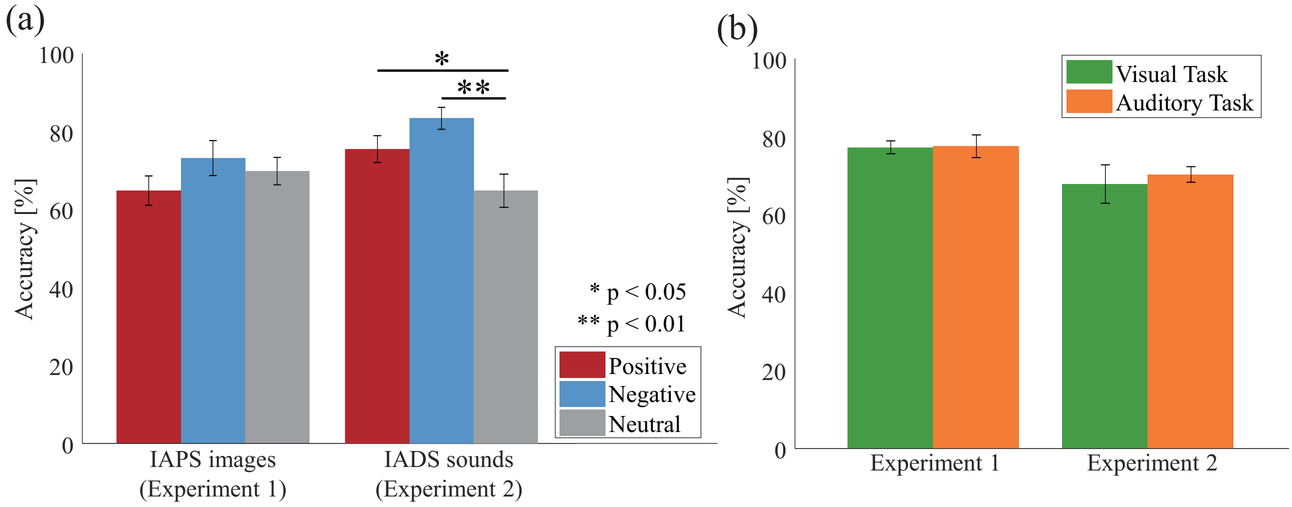

Figure 3.6: Behavioral results

(a) Mean rates at which the participants' responses to the pictures and sounds in the emotional task matched the valence of the pictures and sounds in Experiment 1 and 2. (b) Detection rates of the target dot and beep sound for the visual task and the auditory task, respectively, in both experiments. The error bars indicate the standard error of the mean across the participants. Asterisks in (a) indicate a significant difference based on multiple comparisons for the main effect of stimulus; * $p < 0.05$, ** $p < 0.01$.

there was no significant interaction effect between the stimulus condition and the task condition ($F(6, 138) = 1.562$; $p = 0.163$; $partial\ \eta^2 = 0.064$). In addition, we analyzed the gradient values with a Bayesian repeated measures ANOVA. We found a BF_{10} smaller than 0.33 in the interaction between emotion and task ($BF_{10} = 0.031$). Thus, we have evidence in support of the null hypothesis that there is no interaction between emotion and task. These results indicate that the positive and negative pictures triggered faster pupil dilation compared to that for the neutral pictures in all task conditions.

Fig. 3.9 b shows the gradient values of the pupillary response to the IADS sounds for each condition in Experiment 2. ANOVA revealed a significant main effect of the stimulus condition, although Bayes factors was providing weak evidence against this effect ($F(2, 46) = 5.595$; $p = 0.007$; $partial\ \eta^2 = 0.196; BF_{10} = 0.465$). The planned comparisons revealed that the gradient values for the negative sounds were significantly greater than those for the positive and neutral sounds ($p = 0.017, p = 0.017$, respectively). However, there was no significant difference between the positive and neutral sounds ($p = 0.673$).

A significant main effect of the task condition was observed for the gradient values ($F(3, 69) = 4.990$; $p = 0.003$; $partial\ \eta^2 = 0.178; BF_{10} = 3.760 \times 103$). The planned comparisons revealed that the gradient value for the emotional task was significantly greater than that for the no task and the

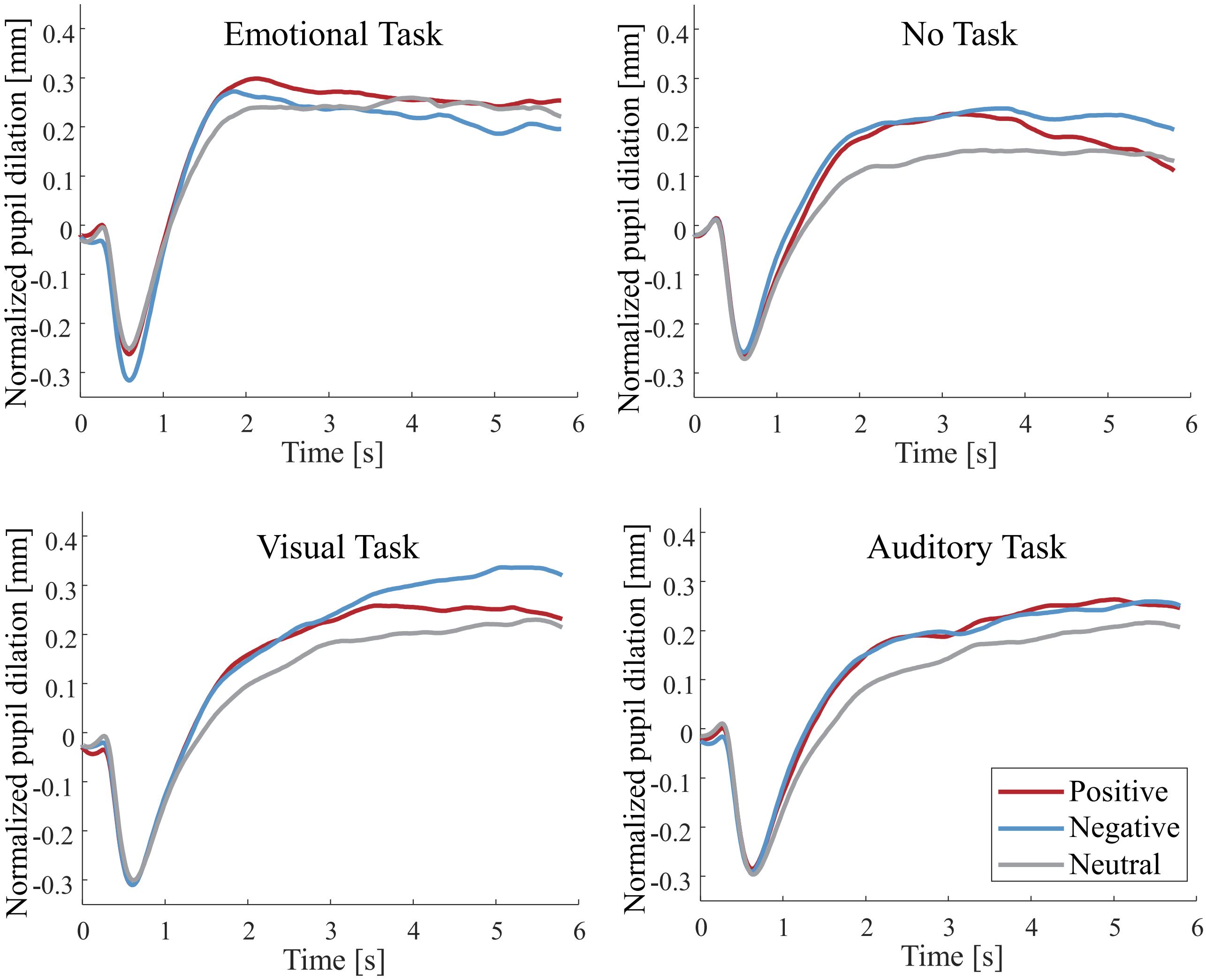

Figure 3.7: Grand-averaged time course of pupillary responses in Exeperiment 1

The grand-averaged time course of pupillary responses to the IAPS pictures (positive, negative, and neutral) in the task conditions during the stimulus presentation in Experiment 1. The horizontal axis indicates the time (s), while the vertical axis indicates the grand-averaged change in pupil dilation from baseline (− 200 ms to 0 ms).

visual task ($p = 0.015$ and $p = 0.016$ for the no task and the visual task, respectively). Moreover, the gradient value for the auditory task was significantly greater than that for the no task ($p = 0.015$).

There was a trend towards a significant interaction between the stimulus condition and the task condition, however, Bayes factors providing strong evidence against this interaction ($F(6, 138) = 1.823$; $p = 0.099$; $partial\ \eta^2 = 0.073$; $BF_{10} = 0.101$). We observed simple main effects of the task condition on the negative sounds ($F(3, 69) = 6.653; p = 0.001; partial\ \eta^2 = 0.224$) and a trend for

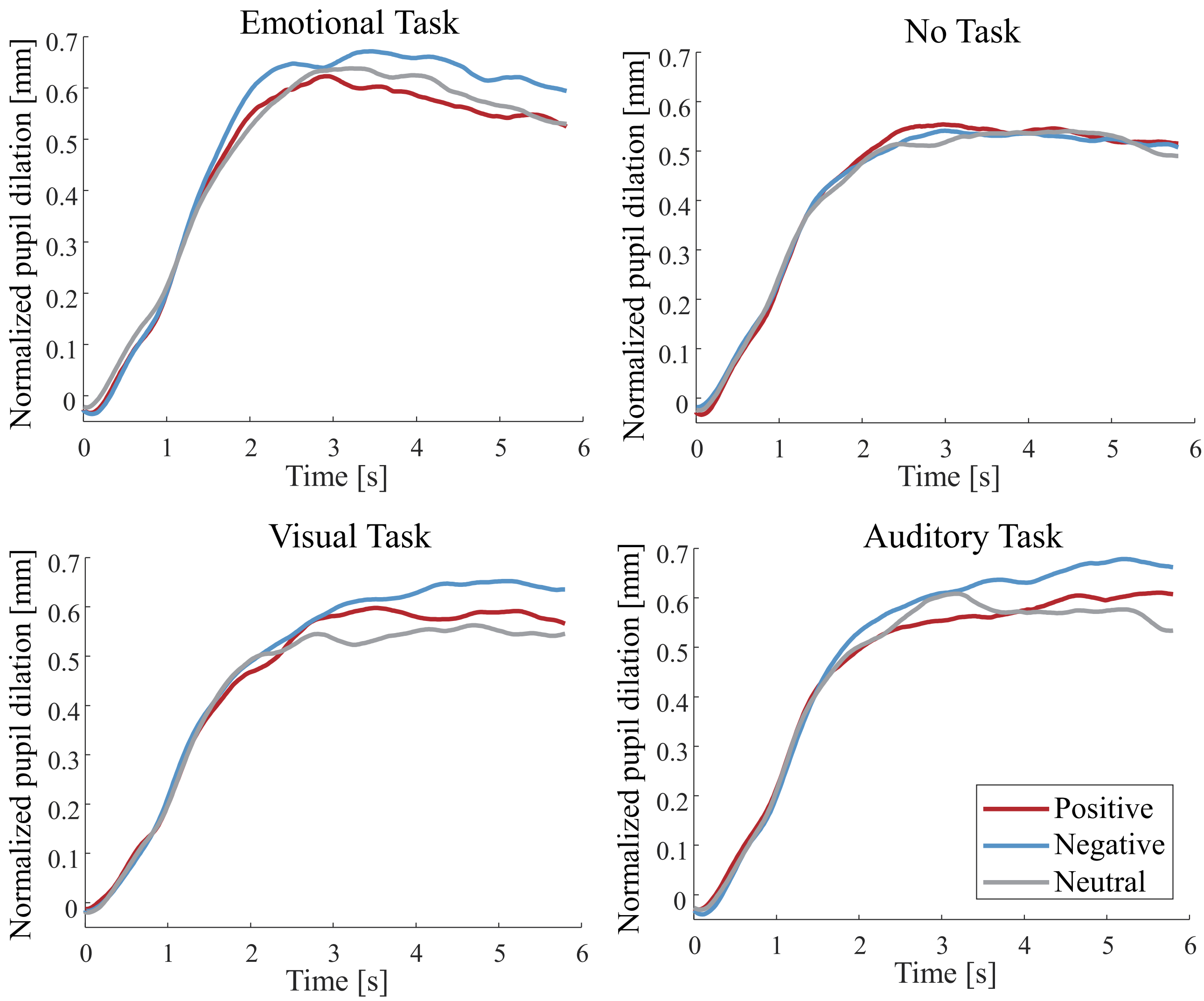

Figure 3.8: Grand-averaged time course of pupillary responses in Exeperiment 2

The grand-averaged time course of pupillary responses to the IADS sounds (positive, negative, and neutral) in the task conditions during the stimulus presentation in Experiment 1. The horizontal axis indicates the time (s), while the vertical axis indicates the grand-averaged change in pupil dilation from baseline (− 200 ms to 0 ms).

the positive sounds ($F(3, 69) = 2.458; p = 0.070; partial\ \eta^2 = 0.097$). Further analyses revealed that the gradient value for the negative sounds was significantly greater for the emotional task than that for the no task and the visual task ($p = 0.007$ and $p = 0.002$, for the no task and the visual task, respectively). The gradient value for the negative sounds was also significantly greater for the auditory task than that for the no task ($p = 0.002$). The gradient value for the positive sounds trended

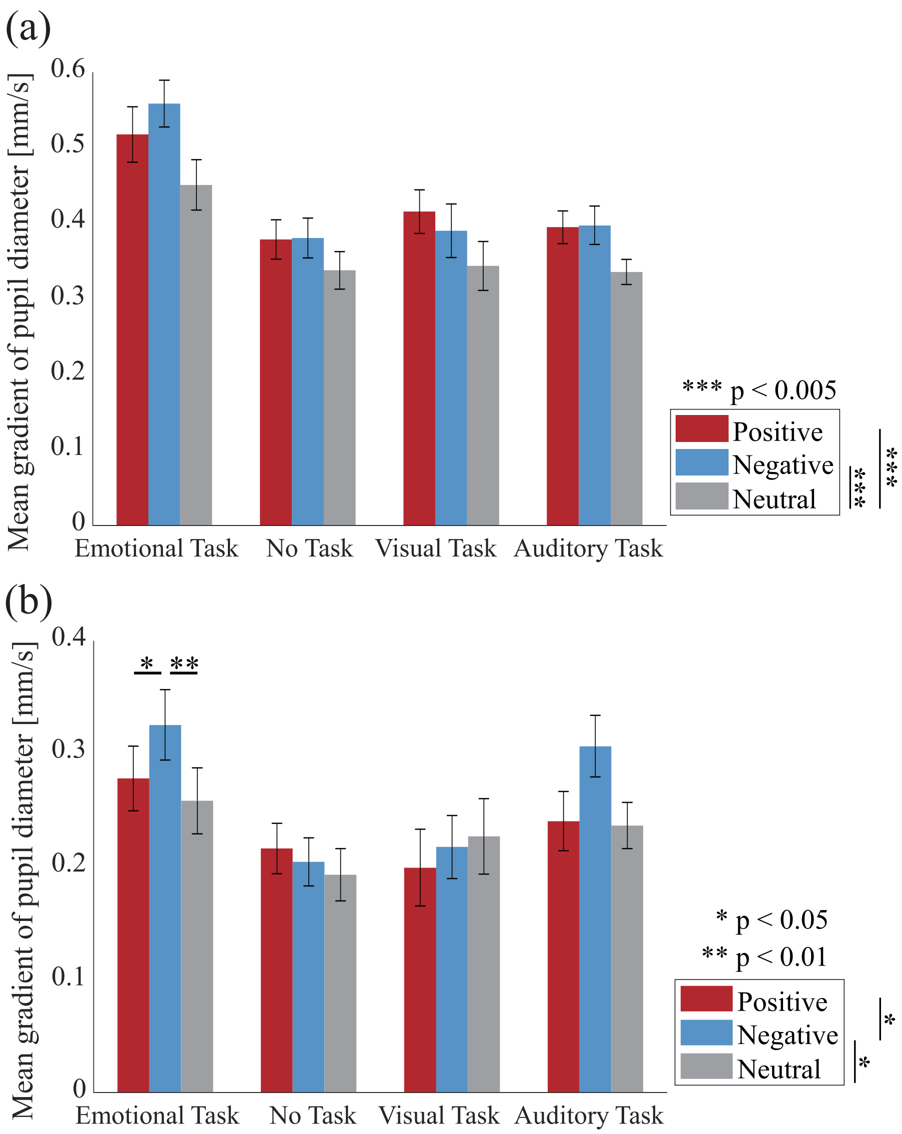

Figure 3.9: Velocity of pupil dilation

(a) Gradient values of the pupillary response to the IAPS pictures for each condition in Experiment 1. (b) Gradient values of the pupillary response to the IADS sounds for each condition in Experiment 2. Error bars show standard error of the mean across the participants. Asterisks in (b) indicate a significant difference based on multiple comparisons for the interaction between the stimulus condition and the task condition; * $p < 0.05$, ** $p < 0.01$, *** $p < 0.005$.

towards being significantly greater for the emotional task than for the visual task ($p = 0.074$). We observed simple main effects of the stimulus condition on the emotional task ($F(2, 46) = 5.460$; $p = 0.008$; *partial* $\eta^2 = 0.192$) and the auditory task ($F(2, 46) = 4.026$; $p = 0.025$; *partial* $\eta^2 = 0.149$).

Further analyses revealed that the gradient value for the emotional task was significantly greater for the negative sounds than that for the positive sounds and the neutral sounds ($p = 0.048 and p = 0.007$, respectively). The gradient value for the auditory task trended towards being significantly greater for the negative sounds than for neutral and positive sounds ($p = 0.053$ and $p = 0.053$ for the neutral and positive sounds, respectively). These results suggest that in the emotional and auditory tasks, the negative sounds triggered a faster pupil dilation compared to that for the positive and neutral sounds.

In addition, three-way ANOVAs with repeated measures were performed using the pupillary gradient in 3 emotion conditions and the 4 task conditions, as the within-subjects factor, and 2 modality conditions (Vision vs. Audition) as the between-subjects factor. However, there was no significant interaction effect between the stimulus condition, the task condition and modality condition ($F(5.33, 245.15) = 1.430;\ p = 0.221;\ partial\ \eta^2 = 0.003;\ BF_{10} = 0.0345$).

3.4 Discussion

In the present study, we investigated the relationship of attentional states to emotional unimodal stimuli (the IAPS pictures or the IADS sounds) and emotional responses using pupillometry in two experiments. Results for the velocity of pupillary responses showed that the emotional responses to visual stimuli were elicited regardless of the task (any attentional state), while the emotional responses to auditory stimuli were elicited only when the participants attended to the auditory modality. Thus, our results point to different properties of emotional pupillary responses for visual and auditory stimuli at least in the point that there is no attentional effect to emotional pupilary responses to visual stimuli.

In Experiment 1, there was no significant interaction for the velocity of the pupil dilation between the stimulus condition and the task condition; compared to the neutral pictures, both the positive and negative pictures elicited a larger pupil dilation in all task conditions. The results for the emotional task and the no task in Experiment 1 coincided with findings from the previous study indicating that the pupil diameter dilated more to the emotional pictures than to the neutral pictures regardless of the participants' mode of viewing (Snowden et al., 2016). Schupp et al. reported an interference with the selective emotional processing when visual attentional resources are allocated for a visual target (horizontal or vertical lines) superimposed on the IAPS pictures (Schupp et al., 2007). However, in our experiment, the target dot or sound during an emotional stimulus presentation in both visual and auditory tasks was not presented in all trials analyzed for the pupillary response. Therefore, we conjectured that the emotional pupillary responses to the IAPS pictures were elicited in all task conditions because the visual target (e.g., target dot) to divert attention did not exist, which is not dissimilar to the reports of Schupp et al. Moreover, our pupillometry results for the auditory task in Experiment 1 revealed that the velocity of the pupil dilation when the participants attended to the emotional pictures was faster compared to that for the neutral pictures. Schupp et al. reported that the processing of the emotional pictures was not modulated by the additional auditory detection task of increasing complexity (Schupp et al., 2008), suggesting that the emotion processing in the visual domain is not affected by the task demands in the auditory modality (Gerdes et al., 2014). Thus, the processing of the emotional pictures in Experiment 1 was unlikely to be disrupted by the auditory task because visual and auditory modalities use independent attentional resources.

In Experiment 2, there was a trend towards a significant interaction between the stimulus condition and the task condition for the velocity of the pupil dilation, suggesting that the velocity of the pupil dilation for the negative sounds in the emotional and auditory tasks may have been faster than that for the positive and neutral sounds. Although our results are consistent with the previous findings (Partala and Surakka, 2003) in that pupillary response to the negative sounds was larger than that to the neutral sounds, the results for the pupillary response to the positive sounds were different from the previous findings. This difference may be due to the results of the subjective rating to the IADS sounds in the emotional task, whereby negative emotions were elicited more than positive and neutral emotions. Therefore, the velocity of the pupil dilation for the negative sounds in the emotional and auditory tasks may have been faster than that for the positive sounds because the emotional evaluation of the positive sounds was lower than for the negative sounds (Kinner et al., 2017). The no task and the visual task in Experiment 2 did not need as much auditory attention during the stimulus presentation compared to the emotional task and the auditory task, indicating that the emotional responses of the pupillary change in Experiment 2 were elicited only in the conditions where the participants attended to auditory modalities. Thus, the emotional pupillary responses to the sounds in the no task and the visual task may have been suppressed because attention was paid to the fixation point of a different modality rather than to the emotional stimuli, in an agreement with the previous study (Schupp et al., 2008).

The dissimilarity of the pupillary responses between Experiment 1 and Experiment 2 included the following: whereas emotional pupillary responses to visual stimuli were elicited in all task conditions, emotional pupillary responses to auditory stimuli were elicited only when attention was paid to the auditory modality during the stimulus presentation (emotional task, auditory task). Previous studies on unconscious emotional processing of visual stimuli have demonstrated that emotional responses can be elicited by subliminal perception (Nomura et al., 2004; Hermans et al., 2003; Tamietto and De Gelder, 2010). In contrast, Lähteenmäki et al. reported that the affective auditory processing is dependent on awareness, and emotional responses were not elicited during the nonconscious auditory processing (Lähteenmäki et al., 2018). In our experiments, there was no difference in emotional pupillary responses between the visual and auditory stimuli when attention was paid to the emotional stimuli (Experiment 1: emotional task and visual task, Experiment 2: emotional task and auditory task). However, in the conditions in which the participants did not need to attend to emotional stimuli (Experiment 1: no task and auditory task, Experiment 2: no task and visual task), the pupillary response to emotional visual stimuli was more dilated compared to that of the neutral pictures, whereas the pupillary response to emotional auditory stimuli did not dilate. We, therefore, propose a different feature of emotional pupillary responses to visual and auditory stimuli, whereby unconscious emotional pupillary responses to visual stimuli are elicited whereas unconscious emotional pupillary responses to auditory stimuli are not evoked (or minimally evoked).

Taken together, the current data of pupillary responses reveal emotional responses depend on attentional allocation to visual and auditory stimuli. Emotional pupillary responses to visual stimuli were evoked in all attentional states, whereas emotional pupillary responses to auditory stimuli were only evoked in cases where the participants attended to the auditory modality. However, there are some limitations in this study considering most effects that are not a large effect size, and non-significant interaction between the stimulus condition, the task condition, and modality condition, which undermine our conclusion of asymmetrical characteristics of emotional responses to pictures and sounds. In this study, we used an erotic stimulus, which might cause a gender difference in emotional response, and we could not strictly control the emotional valence of the stimuli between two modalities because we analyzed pupillary response according to the classification of the emotional task. In addition, the visual stimuli in Experiment 1 were static images, whereas the auditory stimuli in Experiment 2 were dynamic sounds. This difference might influence the eye movement during each trial, which is possible to be a potential contaminant for pupillary responses. Further studies using sufficiently controlled stimuli (video and sound) between two modalities are needed in order to investigate the difference in temporal processing between the two modalities. Moreover, the partial overlap of the participants between Experiment 1 and 2 more complicated the comparison of pupillary responses between the experiments.

Our findings suggest that differences in emotional pupillary responses to visual and auditory stimuli must be considered in any research on cognitive emotion regulation. This will be conducive to the establishment of new methods of cognitive emotional regulation, whereby emotional responses elicited by one sensory modality are prompted or suppressed by stimuli from another sensory modality.

Chapter 4

Emotion estimation method using pupillometry

A similar version of this chapter will be submitted as:

Shimizu, K., Nakakoga, S., Muramatsu, J., Minami, T., & Nakauchi, S., Pupillary response to beep sound reflects emotion: emotion estimation method using probe stimulus. *The European Conference on Visual Perception 2019*, Leuven, Belgium, August 2019, Perception, 48(1) 152-152.

The estimation of emotional states using physiological responses has become an important goal. Currently, in existing studies using physiological responses for emotion estimation, the physiological responses are directly linked to the timing of emotion elicitation, such as the onset of emotional image or sound. Utilizing a separate probe stimulus can avoid this confound. For example, the attentional state to an unexpected tone is enhanced after negative images are viewed. This study investigated pupillary responses to an auditory stimulus after a positive, negative, or neutral emotional state was elicited by an emotional image. An emotional image was followed by a beep sound that was either repetitive or unexpected, and the pupillary dilation was measured. Our results showed that the early component of the pupillary response to the beep sound was larger for negative and positive emotional states than neutral, whereas the late component was larger for positive emotional states. The pupil response's peak latency was earlier for negative images than neutral or positive. Finally, SVM classified the emotional state based on the pupillary response with 80% accuracy. Our study suggests that emotional states can be estimated from the amplitude and the latency of pupil activity in response to an auditory probe.

4.1 Introduction

In recent years, with the development of computers and biological signals, some researchers have estimated the human emotional state from physiological indicators (such as electroencephalogram, eye movement, and heart rate) with the goal of utilizing them in the communication between humans and computers (Luneski et al., 2010). One of the key elements of this approach is estimating human emotional states from behavioral and physiological responses (Frantzidis et al., 2010). While electroencephalogram (EEG) has been widely used as a biomedical signal indicator (Kim et al., 2013), pupillary responses are less expensive to measure and have recently received more attention. The Pupillary Dilation Response (PDR) is controlled by the level of activity of the locus coeruleus (LC) (Aston-jones and Cohen, 2005). Therefore, the pupillary response has been considered indirect evidence of mental activity such as arousal; Bradley et al. found that the pupillary response to images from the International Affective Picture Set (IAPS; Lang et al., 2008) was associated with the arousal level of the emotional stimuli, suggesting that the pupil is mediated by sympathetic nervous system activation (Bradley et al., 2008). Many studies investigating the association between emotionally arousing stimuli, such as images and sounds, and pupillary responses have reported that participants show more pupillary dilation to the high arousal stimulus Bradley and Lang (2015); Henderson et al. (2014); Kinner et al. (2017); Nakakoga et al. (2020); Partala and Surakka (2003); Snowden et al. (2016).

Most of the recent studies examining emotion estimation using biological signals were based on the difference between the biological signal responses immediately before and after presenting the emotional stimulus, making it challenging to utilize this method for real-world applications. The probe stimulus technique provides a solution to this problem. In this method, a stimulus unrelated to the emotional stimulus, such as an auditory tone, is used as a probe stimulus, and the biological response (such as an EEG signal) is measured (Suzuki et al., 2005). Some studies using multiple types of probes have shown that differences in the frequency of a tone presentation affect the EEG event-related potentials (ERPs). Tartar et al. investigated the ERPs associated with two types of beeps (standard tones, which are presented frequently or as part of a repetitive background; and oddball tones, which are presented infrequently and thus are considered deviant) that were presented after viewing negative or neutral images of IAPS (Tartar et al., 2012). The results showed that oddball tones amplified the ERP (late processing negativity) after negative images, while standard tones produced no difference when presented after negative and neutral images. These results were interpreted as increased perceptual resources for all sensory modalities when emotion was induced by the images (Collignon et al., 2008). In addition, the oddball tones shared the increased attentional resources activated by the negative images, and late processing negativities were enhanced by automatic attention (Alho, 1992).

Like the P300 ERP component, the PDR has been shown by multiple studies to exhibit a larger response to deviant stimuli in the oddball task (Kamp and Donchin, 2015; Kinzuka et al., 2020; Liao et al., 2016). Liao et al. measured the pupillary response to two types of sounds (standard and oddball) and showed that the PDR was significantly larger when the deviant sounds were presented (Liao et al., 2016). If the emotional state can be estimated from the pupil response to the probe stimulus, it can be estimated regardless of the timing of emotional arousal. In addition to this benefit, the pupil response is less expensive to measure than other biological signal measurements such as EEG, and it is a non-contact measurement, so it could be applied to any situation where the pupil response can be used for emotion estimation.

A number of studies suggest that pupillary responses after emotionally arousing stimuli reflect the arousal level regardless of emotional valence (Bradley et al., 2008; Partala and Surakka, 2003). However, it is known that attentional states in irrelevant tasks after arousing negative emotion have a negativity bias which negative emotion induced specific ERP latencies and amplitudes compared to positive and neutral ERPs (Alexandrov et al., 2007; De Pascalis et al., 2005; Gulotta et al., 2013; Pinheiro et al., 2017). Therefore, using the PDR latency as a variable should enable the estimation of emotional valence as well as emotional arousal.

In this study, we examined whether three emotional states (positive, negative, and neutral) could be classified based on the PDRs elicited by probe stimuli with two different presentation frequencies. We hypothesized that a negative emotional state would induce a greater PDR for deviant sounds compared to a neutral emotional state. We also aimed to categorize the three types of emotional states based on the PDR response to the subsequent auditory stimulus, and hypothesized that the pupil responses to probe stimuli with different presentation frequencies could be used by a machine learning algorithm to estimate the participants' emotional state. We further hypothesized that by using the PDR latency as a variable, we could estimate emotional valence as well as emotional arousal.

4.2 Materials and methods

Participants

Thirty-two students (23 males; mean age = 21.3 years; range = 20 - 24 years; SD = 1.14 years) participated in this experiment. All participants were students at the Toyohashi University of Technology. A power analysis using G*Power software (G*power 3.1; Faul et al., 2007) indicated that a sample size of 30 would achieve 80% power for a repeated measures design, given a medium effect size ($f = 0.25$) and $\alpha = 0.05$. Two participants were excluded due to equipment failure. All participants had normal or corrected-to-normal vision, normal hearing based on self-reports. The experiment was performed in accordance with the Declaration of Helsinki, and all participants provided written informed consent before the experiment. The experimental procedure was approved by the Committee for Human Research at the Toyohashi University of Technology.

Stimuli

The visual stimuli included 150 pictures ($19° \times 14.2°$) selected from the IAPS database (Lang et al., 2008), consisting of 40 neutral pictures (e.g., mushrooms, household objects, and neutral scenery) followed by the standard tone (mean valence/arousal = 4.997, 2.953), 40 negative pictures (e.g., snakes, guns, and human mutilation) followed by the standard tone (mean valence/arousal = 2.599, 6.052), 40 positive (e.g., babies, erotic couples, and delicious foods) pictures followed by the standard tone (mean valence/arousal = 7.402, 6.005), 10 neutral pictures followed by the oddball tone (mean valence/arousal = 5.043, 0.364), 10 negative pictures followed by the oddball tone (mean valence/arousal = 2.646, 6.117), and 10 positive pictures followed by the oddball tone (mean valence/arousal = 7.500, 5.979). The IAPS pictures presented in combination with the oddball tone differed from the pictures in the standard tone condition to avoid a decrease in response due to habituation to the emotional pictures. The IAPS normative ratings were used to select the emotional category of each picture (see Fig. 4.1 and Fig. 4.2). All pictures were converted to grayscale using MATLAB 2017b, and the average brightness of all pictures ($Y = 101.098$) was unified using the SHINE Toolbox (Willenbockel et al., 2010).To reduce the pupillary light reflex, the brightness of the background ($Y = 201.098$) was higher than the average brightness of the pictures. The auditory stimulus was produced using the MakeBeep function in Psychtoolbox-3 (Brainard, 1997). The sinusoidal tones consisted of a 1,000 Hz low pitch tone and a 1,500 Hz high pitch tone. The sampling rate of both tones was 22,050 Hz.

Procedure

Stimulus presentation and timing were controlled using MATLAB 2014a presentation software (Psychtoolbox 3). The stimuli were displayed on an LCD monitor (ViewPixx/EEG, VPixx Technologies; screen resolution: 1,920 × 1,080 pixels, refresh rate: 120 Hz) and presented through headphones (SoundTrue around-ear headphones, BOSE) in a dark experimental room. The participant's chin was fixed at a viewing distance of 600.0 mm from the monitor.

Each participant performed 300 trials that were divided into 4 blocks of 75 trials to allow for short breaks. Fig. 4.3 provides a visual representation of an experimental trial. Each trial presented stimuli in the following order: fixation point (500 ms), visual stimulus (200 ms), fixation point (600 ms), auditory stimulus (100 ms), fixation point (1,900 ms), response period and inter-stimulus interval (ISI) (3,000 ms). The visual stimulus consisted of either a neutral ($n = 100,\ 33.3\%$), negative ($n = 100,\ 33.3\%$), or positive ($n = 100,\ 33.3\%$) picture, presented randomly. The auditory stimulus was either an oddball ($n = 60,\ 20\%$) or standard ($n = 240,\ 80\%$) tone, presented randomly. For the odd-numbered participants, the low pitch tone was used as the standard tone and high pitch tone was used as the oddball. These assignments were reversed for the even-numbered participants, so as to counterbalance any potential physiological response to the pitch of the tone itself. In addition, the oddball tones were set to never be presented continuously and to not be presented in the first and last trials of each session to further amplify the response to the oddball stimulus. At the end of each trial,

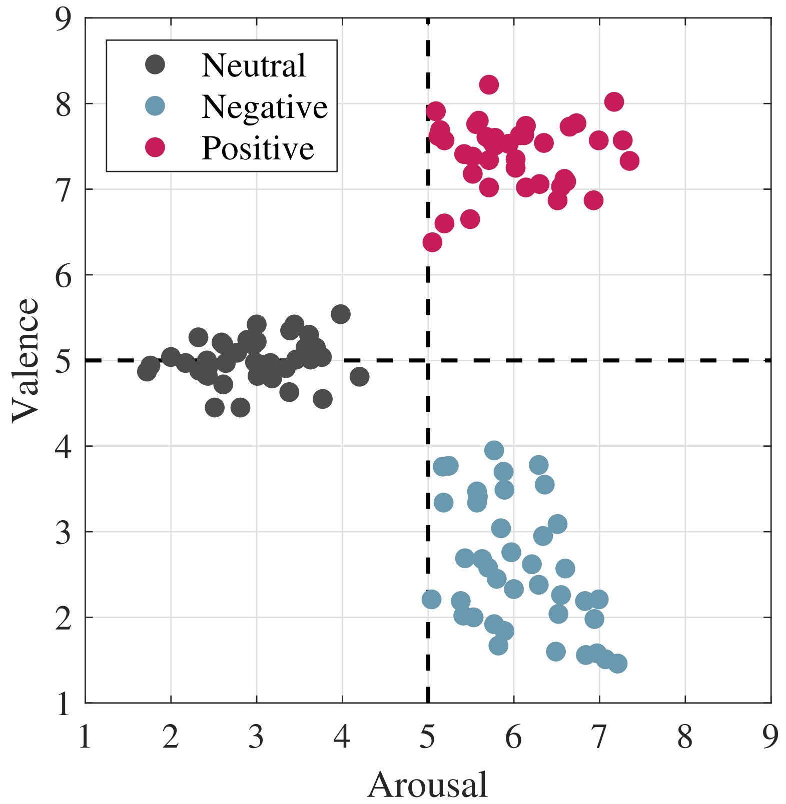

Figure 4.1: Rating scores of the IAPS pictures in the standard tone condition

the participants used the computer keyboard to respond, identifying the emotional categorization of the picture and pitch of the tone to ensure that they were paying attention to both stimuli. Possible response combinations consisted of a neutral picture followed by a low pitch tone (key response 2), neutral picture followed by a high pitch tone (key response 8), negative picture followed by a low pitch tone (key response 1), negative picture followed by a high pitch tone (key response 7), positive picture followed by a low pitch tone (key response 3), or a positive picture followed by a high pitch tone (key response 9). Participants were given a practice session and encouraged to practice until they memorized and were comfortable with the procedure, categorization task, and timing of the trial.

Pupillometry

Pupil diameter and eye movement were recorded at all times during the task except for the response period and ISI. Pupil diameter and eye movement were monitored using an eye-tracking system (Eyelink 1000 Plus, SR Research) that consists of a video camera and an infrared light source pointed at the participant's eyes and outputs the pupil size in arbitrary units. The pupil diameter was sampled at 500 Hz in the left eye only. At the beginning of each session, a nine-point eye-tracker calibration

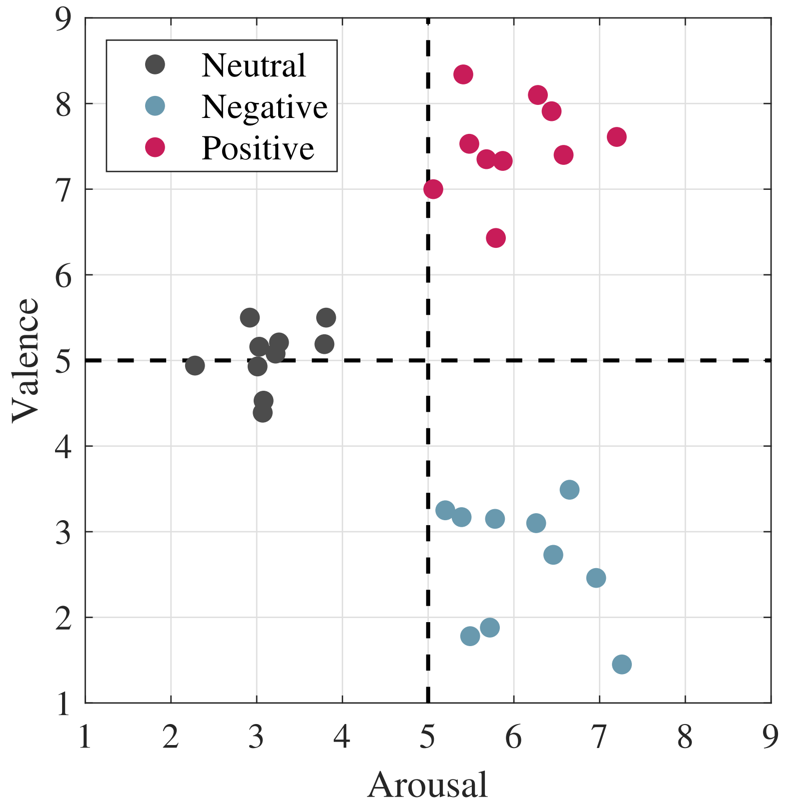

Figure 4.2: Rating scores of the IAPS pictures in the oddball tone condition

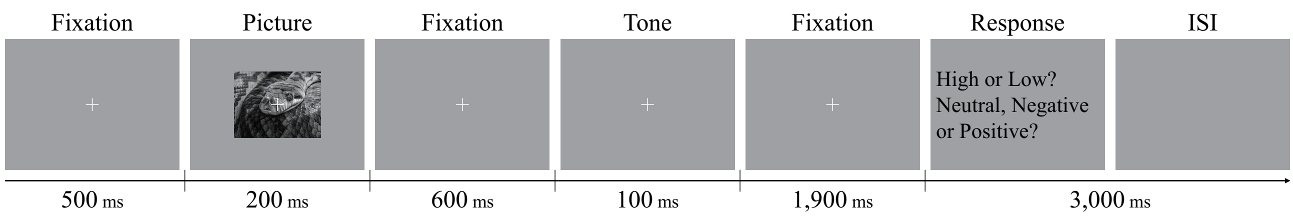

Figure 4.3: Diagram of the experimental procedure

In each trial, a fixation point was first presented for 500 ms. The International Affective Picture Set (IAPS) image was then presented for 200 ms. After the presentation of another fixation point for 600 ms, one of the two tones (standard or oddball) was presented for 100 ms. Each trial was separated by a response period and an inter-stimulus interval (ISI) of 3,000 ms in total. Participants used the keypad to respond to the task in each block following the stimulus presentation.

was performed. Participants were instructed to refrain from blinking as much as possible during the task.

Data analysis

Behavioral data

Recorded pupillary and behavioral data were analyzed using MATLAB 2018a. The behavioral data included the VSCAR (visual stimulus correct answer rate) and ASCAR (auditory stimulus correct answer rate). The VSCAR was calculated by comparing the emotional valence categorized by the participant in the task with the emotional valence of the IAPS normative ratings (for example, whether the participant rated an IAPS image with a negative normative rating as a negative image). The ASCAR was computed by comparing the pitch categorized by the participant in the task with the correct pitch.

Pupillary data

The pupil data during eye blinks were interpolated using cubic-spline interpolation (Mathôt et al., 2018). Trials in which the pupils could not be detected during the beginning or ending of the trial were excluded from analysis. Next, subtractive baseline correction was applied for normalizing using the mean pupil size during the 2 s before the auditory stimulus onset as a baseline (Mathôt et al., 2018). Following the baseline correction, the pupil data for each trial were smoothed using a moving-average filter with a 10 ms window and converted from the arbitrary units output by the eye tracking system to millimeters, based on a previous study (Hayes and Petrov, 2016). Trials with additional artifacts, revealed by exceeding 0.011 mm/ms for the velocity of the pupil response, were excluded from the analysis. The average number of rejected trials was 39.83 ± 31.08 out of 300 trials per participant. Finally, the grand-averaged pupil responses separated by picture category and tone type were computed. Tone type and picture category were set by the correct pitch and the emotional valence based on the normative rating in IAPS. Previous studies have shown that the early and late components of the pupillary response to emotional pictures are different components that reflect the activity of the sympathetic and parasympathetic nervous systems, respectively (Bradley et al., 2017). Therefore, Tthe early and late components of the PDR were respectively measured as the average pupil diameter between 0 - 850 ms and 850 - 1,700 ms following the tone onset. Peak latency of pupillary response reflects many cognitive factors (Mathôt, 2018), such as the frequency of stimulus presentation (Reinhard et al., 2007). The PDR peak was defined as the largest pupil diameter observed within the 0 - 1,700 ms following the tone onset, and the PDR latency was defined as the duration between the tone onset and PDR peak.

The early component, late component, and PDR latency for each trial were averaged across trials and tone types. The values were normalized using Z-scores and classified into emotional conditions using an SVM with a linear kernel. The confusion matrix was generated using leave-one-out cross-validation.

Statistical analysis

All statistical analyses were conducted using R software (version 3.4.4). Two-way ANOVAs were performed using picture category (neutral, negative, and positive) and tone type (standard and oddball) as within-subject factors. The level of statistical significance was set to $p < 0.05$ for all analyses. The pairwise comparisons for the main effects were corrected for multiple comparisons using the Modified Sequentially Rejective Bonferroni (MSRB) procedure (Seaman et al., 1991). Effect sizes ($partial\ \eta^2$) were determined for the ANOVA effects. Greenhouse-Geisser corrections were performed when the results of Mendoza's multisample sphericity test were significant.

4.3 Results

Behavioral data

Behavioral analysis of the visual stimulus correct answer rate (VSCAR) revealed a main effect of picture category ($F(1.83, 53.05) = 49.919$; $p < 0.001$; $partial\ \eta^2 = 0.633$), but no main effect of tone type ($F(1, 29) = 0.041$; $p = 0.840$; $partial\ \eta^2 = 0.001$). Multiple comparisons of picture category data showed that the VSCAR was higher for neutral and negative pictures compared to positive pictures ($allp < 0.001$). In addition, the picture category $\times$ tone type interaction was significant ($F(1.77, 51.31) = 3.938$; $p = 0.030$; $partial\ \eta^2 = 0.120$). Post hoc analyses revealed that the VSCAR was significantly higher for a standard tone than an oddball tone when it followed a neutral picture ($F(1, 29) = 4.679$; $p < 0.039$; $partial\ \eta^2 = 0.139$), and the VSCAR was significantly higher for an oddball tone than a standard tone when it followed a negative picture ($F(1, 29) = 6.639$; $p = 0.015$; $partial\ \eta^2 = 0.186$). Additionally, there was a main effect of picture category for standard tones ($F(1.72, 49.89) = 46.838$; $p < 0.001$; $partial\ \eta^2 = 0.618$) and a main effect of picture category for oddball tones ($F(1.92, 55.8) = 41.037$; $p < 0.001$; $partial\ \eta^2 = 0.586$). Multiple comparisons for picture category with a standard tone showed that the VSCAR for positive images was lower compared to that of neutral and negative images for standard tone trials ($allp < 0.001$), and the VSCAR was higher for neutral images than negative images for standard tone trials ($p = 0.036$). Multiple comparisons for picture category with an oddball tone showed that the VSCAR to positive images was lower compared to neutral and negative images for oddball tone trials ($allp < 0.001$). The analysis of auditory stimulus correct answer rate (ASCAR) revealed a main effect of tone type ($F(1, 29) = 7.966$; $p = 0.009$; $partial\ \eta^2 = 0.216$) but no main effect of picture category ($F(1.88, 54.6) = 1.299$; $p = 0.280$; $partial\ \eta^2 = 0.043$). There was no significant interaction ($F(1.86, 54.05) = 0.855$; $p = 0.424$; $partial\ \eta^2 = 0.029$). The behavioral data are presented in Fig. 4.4.

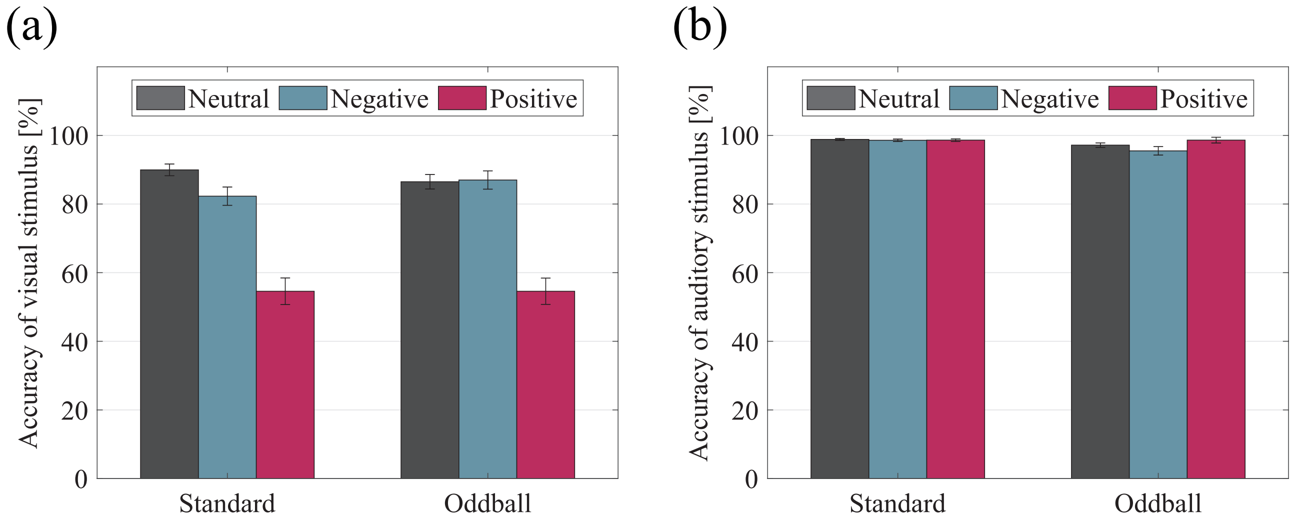

Figure 4.4: Behavioral results

(a) The visual stimulus correct answer rate (VSCAR). (b) The auditory stimulus correct answer rate (ASCAR). The error bars indicate the standard error of the mean.

Pupillary data

Fig. 4.5 presents the grand-averaged pupil responses from the picture onset, separated by picture category and tone type. Fig. 4.6 presents the grand-averaged pupil responses from the tone onset, separated by picture category and tone type. The early component, late component, and PDR latency are shown in Fig. 4.7. The analysis of the early component revealed main effects of picture category ($F(1.82, 52.81) = 14.670$; $p < 0.001$; $partial\ \eta^2 = 0.336$) and tone type ($F(1, 29) = 6.456$; $p = 0.017$; $partial\ \eta^2 = 0.182$). Multiple comparisons for picture category showed that the early component was larger on negative and positive picture trials compared to neutral picture trials ($allp < 0.001$). There was no significant interaction ($F(1.84, 53.45) = 0.766$; $p = 0.460$; $partial\ \eta^2 = 0.026$).

The analysis of the late component revealed main effects of picture category ($F(1.7, 49.43) = 15.826$; $p < 0.001$; $partial\ \eta^2 = 0.353$) and tone type ($F(1, 29) = 57.630$; $p < 0.001$; $partial\ \eta^2 = 0.665$). Multiple comparisons for picture category showed that the late component was larger for positive pictures compared to neutral and negative pictures ($allp < 0.001$). There was no significant interaction ($F(1.69, 48.87) = 0.510$; $p = 0.573$; $partial\ \eta^2 = 0.017$).

The analysis of the PDR latency revealed main effects of picture category ($F(1.59, 46.01) = 11.851$; $p < 0.001$; $partial\ \eta^2 = 0.290$) and tone type ($F(1, 29) = 55.236$; $p < 0.001$; $partial\ \eta^2 = 0.656$). Multiple comparisons for picture category showed that the PDR latency was shorter for negative pictures compared to neutral ($p < 0.05$) and positive ($p < 0.001$) pictures. There was no

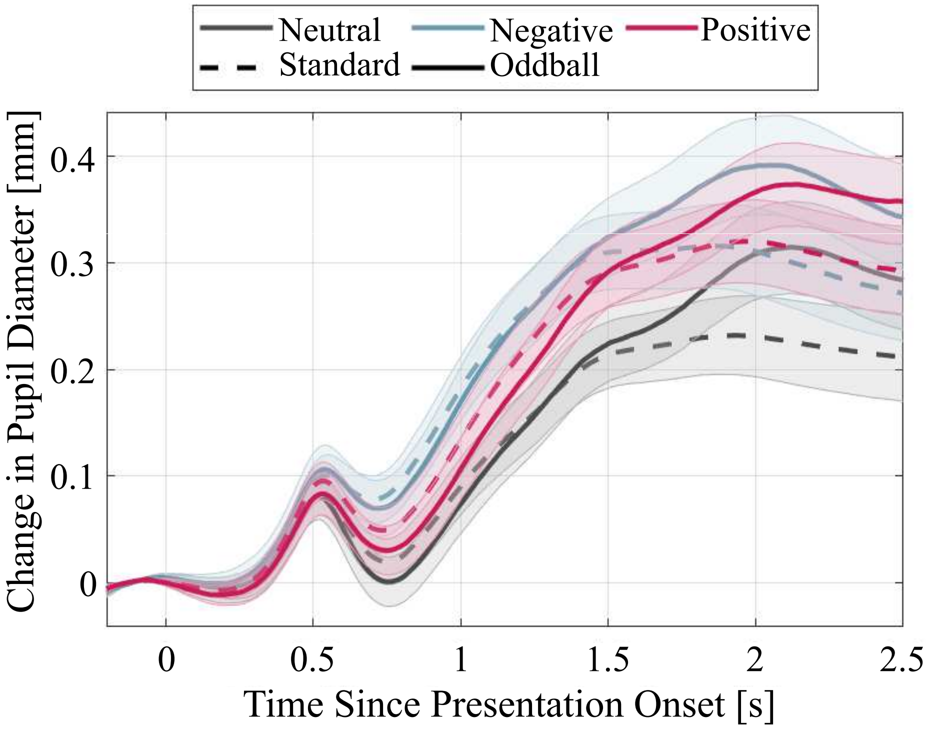

Figure 4.5: Grand-averaged time course of pupillary responses from the picture onset

The horizontal axis indicates the time (s) from the presentation onset of the picture, while the vertical axis indicates the grand-averaged change in pupil dilation from baseline (− 200 ms to 0 ms). Shaded areas represent the standard error of the mean.

significant interaction ($F(1.86, 54.08) = 0.729$; $p = 0.478$; $partial\ \eta^2 = 0.025$).

Classification using support vector machine (SVM)

The confusion matrix is shown in Table 4.1. The average emotion identification rate was 80.00%. For comparison, a random guess would have an accuracy of 33% because three possible emotional states were used.

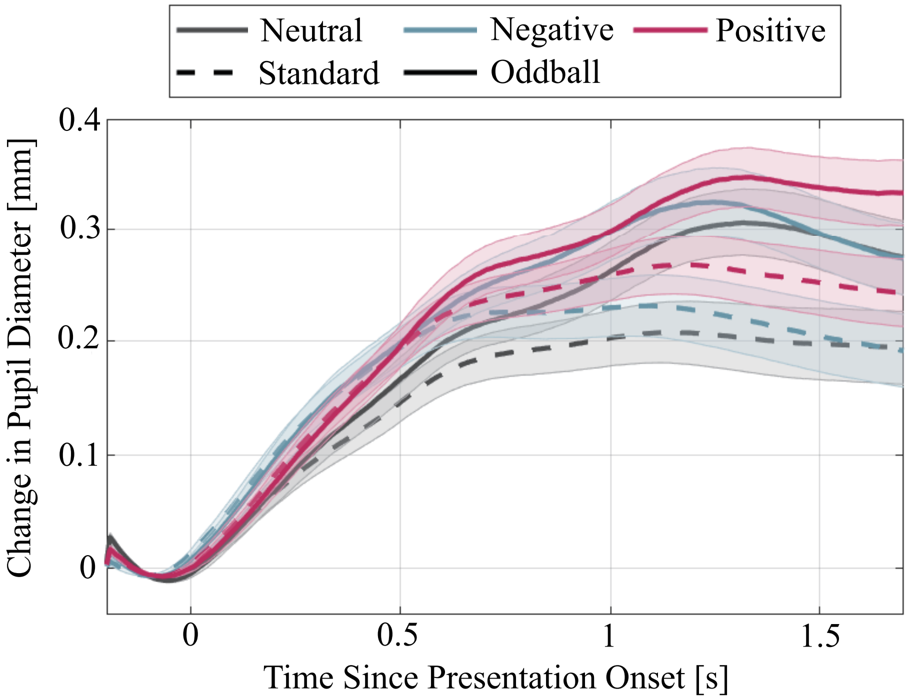

Figure 4.6: Grand-averaged time course of pupillary responses from the tone onset

The horizontal axis indicates the time (s) from the presentation onset of the sound, while the vertical axis indicates the grand-averaged change in pupil dilation from baseline (− 200 ms to 0 ms). Shaded areas represent the standard error of the mean.

Table 4.1: Classification accuracy level for three distinct emotions by a support vector machine

Actual	**Estimation**			
	Neutral	Negative	Positive	Identification rate [%]
Neutal	24	2	4	80.00
Negative	5	22	3	73.33
Positive	3	1	26	86.67

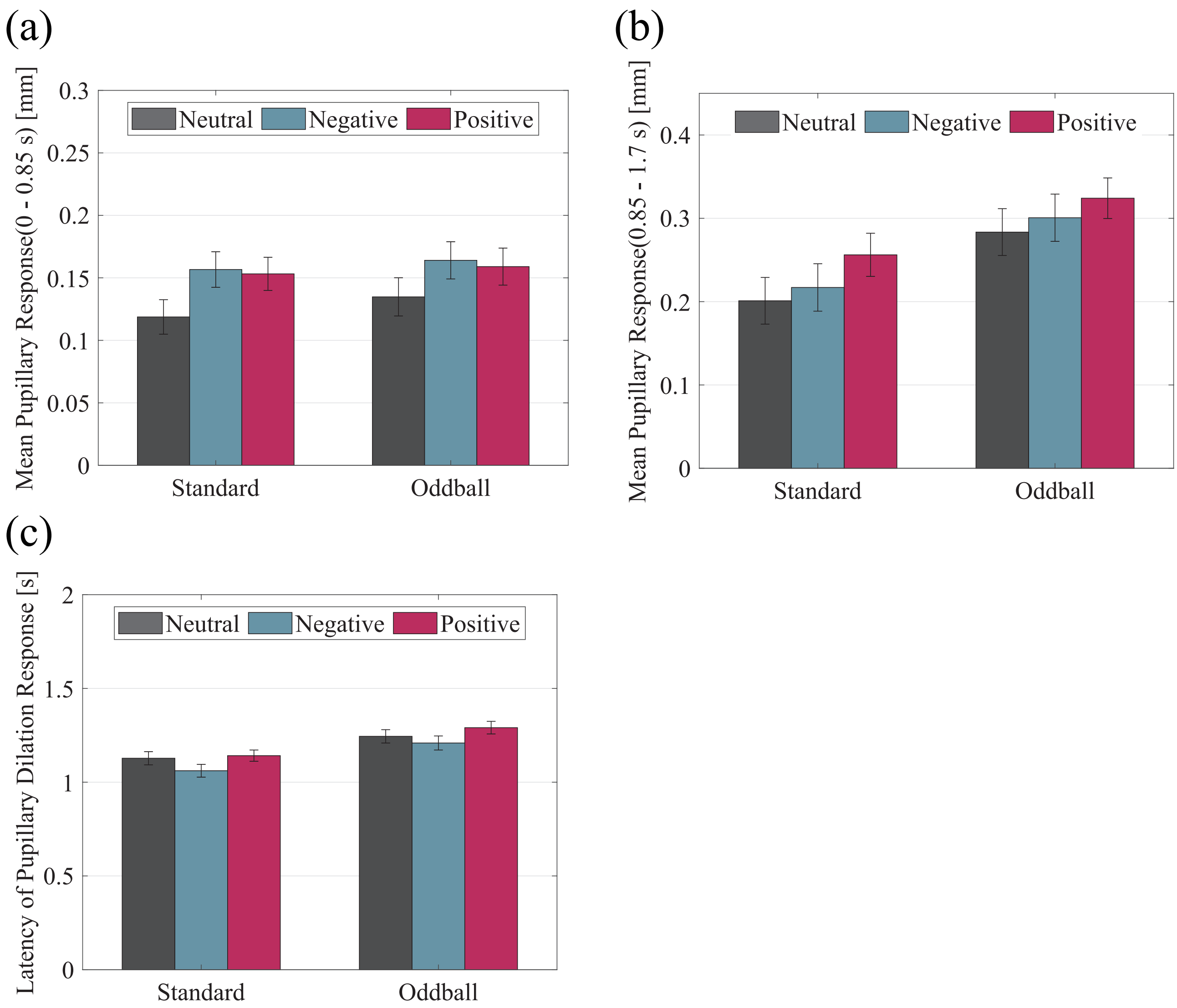

Figure 4.7: Pupillary dilation response (PDR) components

(a) The early component (mean PDR between 0 - 0.85 s). (b) The late component (mean PDR between 0.85 - 1.7 s). (c) The PDR latency between tone onset and the peak PDR amplitude. The error bars indicate the standard error of the mean.

4.4 Discussion

We examined the pupillary responses to auditory stimuli in an oddball paradigm, with the auditory probe presented after images that arouse positive, negative, or neutral emotions to estimate the three emotional states based on the pupillary responses to the probe stimuli. As a result, an SVM using the amplitude and latency of the pupil response to auditory stimulus presentation onset as the feature quantities could be classified into the three emotional states with high accuracy. However, contrary to our hypothesis, the frequency of the sound presentation did not affect the emotion estimation.

For the early component of the pupillary response, a main effect of picture category was observed, and the pupillary responses to positive and negative images were larger than the responses to neutral images. This result is consistent with several previous studies showing that pupillary responses are associated with arousal regardless of valence because pupil dilation reflects arousal associated with increased sympathetic activity (Bradley et al., 2008; Nakakoga et al., 2020; Partala and Surakka, 2003). For the late component, a main effect of picture category was also observed. In this case, the response to positive images was larger than the response to neutral or negative images. Taylor suggests an asymmetry in the negative state's emotional valence, positing that people in a negative emotional state make more effort to return to normal than those in a positive state (Taylor, 1991). In the present study, the late component of the PDR may have been smaller following negative images than positive images due to a decrease in the negative state arousal level as a result of the attempt to return to the normal state. In the early and late components of the pupillary response, a main effect was seen for the tone type, such that the pupil sizes were larger for deviant than for standard tones. Several studies measuring pupillary responses in the oddball stimuli have found greater PDRs than standard stimuli about 1 s after deviant stimulus presentation (Kamp and Donchin, 2015; Kinzuka et al., 2020; Liao et al., 2016). This is consistent with the presence of PDRs due to deviant stimuli in the early and late component of the pupil response in the present study. For the PDR latency, a main effect of picture category was observed, and the peak latency of the response to negative images occurred earlier than that of neutral and positive images. Studies that have examined ERP in response to unrelated tonal stimuli presented during a state of emotional arousal have shown that the specific response to a tone in the negative state is found in the mismatch negativity (MMN) (De Pascalis et al., 2005; Pinheiro et al., 2017) and P300a (Alexandrov et al., 2007; Gulotta et al., 2013) ERP components. Such effects are known as "negativity bias" and serve as priming for increasing attentional resources for the next stimulus (Tartar et al., 2012). In the present study, PDR latency was shorter in the negative condition, which may be due to the increase in attentional resources to the subsequent probe stimulus.

The present results provide evidence that the PDR reflects an oddball response. However, there was no interaction between the affective and sound presentation frequency conditions, and the frequency of presentation of the probe stimulus did not contribute to the affective estimation. In Suzuki et al.'s study, task-less probe sound stimuli produced differences in ERPs between trials in which the participant watched interesting versus uninteresting videos, regardless of the tone type (oddball or standard) (Suzuki et al., 2005). Another limitation of this technique is that the pupillary response has a lower temporal resolution than EEG, and several components are not independent of each other, though this is also true of ERPs. Furthermore, the difference in pupillary responses between emotional conditions after the probe stimulus in the current task was smaller than that between deviant and standard tones. This may be the reason that an interaction between presentation frequency and emotional conditions was not observed, as the differences between the standard and deviant response may have overshadowed the differences between emotional responses. Overall, our results suggest that the frequency of stimulus presentation is irrelevant when the pupil response to probe stimuli is used for emotion estimation.

In conclusion, we used pupil responses to probe stimuli to estimate emotional state. The results showed that the early and late components of the pupil response to the probe stimulus were specific to neutral and positive emotions, respectively, and the latency of the pupil response peak was shorter only in the negative conditions. Furthermore, the classification results using these components as features suggest that it is possible to use the PDR to estimate emotions with high accuracy. The classification results also suggest that the presentation frequency of the probe sound did not affect emotion estimation. A limitation of this study is that the presentation interval between the emotional image and the probe sound was always constant, and the participants were actively listening for the probe sound to perform the response task. In addition, the ratio of responses to the positive images consistent with the IAPS ratings was lower than negative and neutral. Since the participants in this study were only Japanese, we believe that the bias in the evaluation was due to the influence of cultural differences in emotional arousal by IAPS (Huang et al., 2015). To increase the applicability of the results, it is necessary to present probe sounds at random times during the presentation of emotion-arousing videos and infer emotions from the pupil responses in a passive state. Future studies should also investigate the use of other modalities by using sensory stimuli other than sound (e.g., tactile stimuli) as probe stimuli to verify the emotional estimation accuracy of the pupillary responses. This will eventually lead to the estimation of changes in emotions over time from pupillary responses, using probes to the necessary modality as input stimuli. In addition, it is necessary to verify whether the proposed method can estimate the emotion for all generations because the participants in this study were limited to those in their 20s.

Chapter 5

Conclusion

In this thesis, we investigated the mechanisms of emotion and facial expression recognition, which have not yet been elucidated, using pupillary responses. In addition, we attempted to estimate emotional states using the pupillary responses based on these findings. Here, we summarize the findings and insights from the three studies.

5.1 Facial expression and facial color

We investigated the mechanisms of facial color effect on expression recognition using pupillometry (Chapter 2).

Purpose

- To investigate whether information about facial color contributes to the identification of facial expression in low-resolution facial stimuli.
- To clarify the mechanism of the facial color effect on facial expression recognition using the pupillary response during facial expression identification.

Contribution

- The results of the facial expression identification task suggest that the effect of facial color (reddish-colored) information on facial expression recognition becomes stronger as the level of blur increases for angry expressions.
- The pupillary response reflects the behavioral result that participants can identify differences of facial expression in the reddish-color condition, even though it is more blurred.

This study investigated the contribution of facial color to expression recognition in blur images with the measurement of behavior and pupillary change. In the experiment, the face stimuli of facial colors (natural color, reddish) with different expressions (neutral, and anger) in 3 blur levels were presented. Participants performed a task of expression identification to the stimulus. Behavioral results indicated that the facial color has a significant contribution to expression recognition as blur level increases. Then, the results of pupillometry showed that the reddish-color provided the information necessary to identify anger. These results showed the contribution of facial color increases in both psychophysics and pupillary experiment as blur level increases, which suggested that facial color emphasizes the characteristics of specific facial expression.

5.2 Pupillary response to visual and auditory emotional stimuli

We investigated the relationship of attentional states with emotional unimodal stimuli (pictures or sounds) using pupillometry in two experiments (Chapter 3).

Purpose

- To clarify the relationship between the attentional state and emotional response to visual stimuli (IAPS images) using pupillometry.
- To clarify the relationship between the attentional state and emotional response to auditory stimuli using (IADS sounds) pupillometry.
- To compare the symmetry of the relationship between attentional states and emotional responses to visual and auditory stimuli from two experiments

Contribution

- Emotional responses to visual stimuli were evoked in all attentional states.
- Emotional responses to auditory stimuli were only evoked in cases where the participants attended to the auditory modality.
- We suggested that asymmetrical characteristics of emotional responses to pictures and sounds.

In this study, we aimed to elucidate the relationship of attentional states to emotional unimodal stimuli (pictures or sounds) and emotional responses by measuring the pupil diameter. In experiment 1 about vition, we set four task conditions to modulate the attentional state (emotional task, no task, visual detection task, and auditory detection task). We observed that the velocity of pupillary dilation was faster during the presentation of emotionally arousing pictures. Importantly, this effect was not dependent on the task condition. In experiment 2 about audition, We observed a trend towards a significant interaction between the stimulus and the task conditions with regard to the velocity of pupillary dilation. In the emotional and auditory detection tasks, the velocity of pupillary dilation was faster with positive and neutral sounds than negative sounds. Taken together, the current data reveal that different pupillary responses were elicited to emotional visual and auditory stimuli, despite both experiments being sufficiently controlled to be of symmetrical experimental design.

5.3 Emotion estimation by pupillometry

We examined whether the emotional states could be classified based on the pupillary reopnses elicited by sound probe stimuli (Chapter 4).

Purpose

- To investigate whether pupillary responses after the presentation of sound probe stimuli can be used to estimate emotional states
- To clarify if differences in presentation frequency of sound probe stimuli affect the estimation of emotional states

Contribution

- Pupillary responses after the presentation of sound probe stimuli suggested that the three emotional states (positive, negative, and neutral) could be estimated with high accuracy.
- The frequency of presentation of the sound probe stimuli did not affect the estimation of emotional states.

This study investigated pupillary responses to an auditory stimulus after a positive, negative, or neutral emotional state was elicited by an emotional image. An emotional image was followed by a beep sound that was either repetitive or unexpected, and the pupillary dilation was measured. Our results showed that the early component of the pupillary response to the beep sound was larger for negative and positive emotional states than neutral, whereas the late component was larger for positive emotional states. The pupil response's peak latency was earlier for negative images than neutral or positive. Finally, SVM classified the emotional state based on the pupillary response with 80% accuracy. Our study suggests that emotional states can be estimated from the amplitude and the latency of pupil activity in response to an auditory probe.

5.4 General discussion

Pupillary responses as an indicator of emotional response

We conclude the features of all the pupil parameters in this study as indicators of emotional response.

Time-average

- Time-average of pupillary response is the most common index of pupillary response related to cognitive factors.
- The early component (initial light reflex) during arousal picture viewing reflects the decline of parasympathetic nervous activity. In contrast, the late component reflects the increase of sympathetic nervous activity (Bradley et al., 2017).
- There is no clear criterion for which interval should be averaged. It is necessary to calculate the appropriate time-average depending on the presentation time and experimental conditions.

Peak amplitude of constriction

- Initial light reflexes after stimulus presentation are suppressed in response to stimuli with high arousal (Henderson et al., 2014)
- Pupillary changes due to other factors, such as perceived luminance of the stimulus, is greater than that due to cognitive factors.

Gradient value of dilation

- The gradient value of dilation is an index that is used to extract purely the movement of the pupillary dilator muscles, which reflects the activity of the locus ceruleus (LC) (Aston-jones and Cohen, 2005; Kinner et al., 2017).
- In order to use the velocity of dilation as an index, a certain degree of dilation range is required. If it is saturated, it cannot be extracted as an index.

Peak latency of dilation to beep sound

- Time-average and peak amplitude of constriction reflect the arousal state in emotional responses. However, peak latency of dilation to beep sound is a new index that varies with emotional valence (positive or negative).
- The peak latency of dilation to beep sound may reflect the negativity bias, in which attention to subsequent stimuli is enhanced during the negative state (Tartar et al., 2012).

Conclusion

The present study suggests that pupillometry is a useful physiological indicator for facial expression recognition, which is influenced by any context (Chap. 2), and for unconscious emotional reactions, which could not be measured by subjective reports (Chap. 3). In addition, it was suggested that pupillary responses to probe sound stimuli could estimate three emotional states (positive, negative, and neutral) (Chap. 4). Taken together, the pupillometry as an index of emotional response could be applied to a wide range, both conscious and unconscious.

Future works

Pupillary responses are known to reflect all cognitive states, including emotional states. However, the pupil change caused by cognitive factors is about 0.5 mm, which is much smaller than that caused by luminance factors (Sirois and Brisson, 2014). However, in a recent study about pupillometry, there is a technique for estimating pupillary responses for only cognitive factors considering the changes in the brightness of a stimulus (Watson and Yellott, 2012). In addition, the four pupil indices used in this study described above have been shown to be useful for extracting emotional states in any situation. We hope that such researches on pupillometry will be further developed in the future, and that the pupillary response based emotion estimation conducted in Chapter 4 will be used in daily life.

References

Adolphs, R. (2001). The neurobiology of social cognition. *Current Opinion in Neurobiology*, 11(2):231–239.

Adolphs, R., Tranel, D., and Damasio, A. R. (1998). The human amygdala in social judgment. *Nature*, 393(6684):470–474.

Alexandrov, Y. I., Klucharev, V., and Sams, M. (2007). Effect of emotional context in auditory-cortex processing. *International Journal of Psychophysiology*, 65(3):261–271.

Alho, K. (1992). Selective attention in auditory processing as reflected by event - related brain potentials. *Psychophysiology*, 29(3):247–263.

Aluja, A., Rossier, J., Blanch, Á., Blanco, E., Martí-Guiu, M., and Balada, F. (2015). Personality effects and sex differences on the International Affective Picture System (IAPS): A Spanish and Swiss study. *Personality and Individual Differences*, 77:143–148.

Anders, S., Eippert, F., Weiskopf, N., and Veit, R. (2008). The human amygdala is sensitive to the valence of pictures and sounds irrespective of arousal: An fMRI study. *Social Cognitive and Affective Neuroscience*, 3(3):233–243.

Aston-jones, G. and Cohen, J. D. (2005). An Integrative Theory of Locus Function : Adaptive Gain and Optimal Performance. *Annual Review of Neuroscience*, 28:403–450.

Beatty, J. and Lucero-Wagoner, B. (2000). The pupillary system. *Handbook of psychophysiology (2nd ed.).*, 2:142–162.

Benarroch, E. E. (2009). The locus ceruleus norepinephrine system Functional organization and potential clinical significance. *Neurology*, 73(20):1699–1704.

Benitez-Quiroz, C. F., Srinivasan, R., and Martinez, A. M. (2018). Facial color is an efficient mechanism to visually transmit emotion. *Proceedings of the National Academy of Sciences of the United States of America*, 115(14):3581–3586.

Blair, R. J., Morris, J. S., Frith, C. D., Perrett, D. I., and Dolan, R. J. (1999). Dissociable neural responses to facial expressions of sadness and anger. *Brain*, 122(5):883–893.

Bonmassar, C., Widmann, A., and Wetzel, N. (2020). The impact of novelty and emotion on attention-related neuronal and pupil responses in children. *Developmental Cognitive Neuroscience*, 42(February):100766.

Bradley, M. M. and Lang, P. J. (2000). Affective reactions to acoustic stimuli. *Psychophysiology*, 37(2):204–215.

Bradley, M. M. and Lang, P. J. (2007). The International Affective Digitized Sounds (2nd Edition; IADS-2): affective ratings of sounds and instruction manual. *Gainesville, Fl.: University of Florida, NIMH Center for the Study of Emotion and Attention.*

Bradley, M. M. and Lang, P. J. (2015). Memory, emotion, and pupil diameter : Repetition of natural scenes. *Psychophysiology*, 52(9):1186–1193.

Bradley, M. M., Miccoli, L., Escrig, M. A., and Lang, P. J. (2008). The pupil as a measure of emotional arousal and autonomic activation. *Psychophysiology*, 45(4):602–607.

Bradley, M. M., Sapigao, R. G., and Lang, P. J. (2017). Sympathetic ANS modulation of pupil diameter in emotional scene perception: Effects of hedonic content, brightness, and contrast. *Psychophysiology*, 54(10):1419–1435.

Brainard, D. H. (1997). The Psychophysics Toolbox. *Spatial Vision*, 10:443–446.

Breiter, H. C., Etcoff, N. L., Whalen, P. J., Kennedy, W. A., Rauch, S. L., Hyman, S. E., and Rosen, B. R. (1996). Response and Habituation of the Human Amygdala during Visual Processing of Facial Expression bilateral amygdala lesions have demonstrated impair-ments in recognition of fear. *Neuron*, 17(5):875–887.

Bruce, V. and Langton, S. (1994). The use of pigmentation and shading information in recognising the sex and identities of faces. *Perception*, 23(7):803–822.

Changizi, M. a., Zhang, Q., and Shimojo, S. (2006). Bare skin, blood and the evolution of primate colour vision. *Biology letters*, 2(2):217–221.

Collignon, O., Girard, S., Gosselin, F., Roy, S., Saint-Amour, D., Lassonde, M., and Lepore, F. (2008). Audio-visual integration of emotion expression. *Brain Research*, 1242:126–135.

Cuthbert, B. N., Bradleym, M. M., and Lang, P. J. (1996). Probing picture perception: Activation and emotion. *Psychophysiology*, 33(2):103–111.

Czigler, I., Cox, T. J., Gyimesi, K., and Horváth, J. (2007). Event-related potential study to aversive auditory stimuli. *Neuroscience and Biobehavioral Reviews*, 420:251–256.

Davidson, R. J., Maxwell, J. S., and Shackman, A. J. (2004). The privileged status of emotion in the brain. *Proceedings of the National Academy of Sciences of the United States of America*, 101(33):11915–11916.

De Pascalis, V., Arwari, B., Matteucci, M., and Mazzocco, A. (2005). Effects of emotional visual stimuli on auditory information processing: A test of J.A. Gray's reinforcement sensitivity theory. *Personality and Individual Differences*, 38(1):163–176.

Dienes, Z. (2014). Using Bayes to get the most out of non-significant results. *Frontiers in Psychology*, 5:781.

Drummond, P. D. (1997). Correlates of facial flushing and pallor in anger-provoking situations. *Personality and Individual Differences*, 23(4):575–582.

Drummond, P. D. and Quah, S. H. (2001). The effect of expressing anger on cardiovascular reactivity and facial blood flow in Chinese and Caucasians. *Psychophysiology*, 38(2):190–196.

Ebner, N. C., Riediger, M., and Lindenberger, U. (2010). FACES-a database of facial expressions in young, middle-aged, and older women and men: Development and validation. *Behavior Research Methods*, 42(1):351–362.

Ekman, P. (1999). Basic emotions. *Handbook of psychophysiology.*, 98(45-60):16.

Ekman, P. and Friesen, W. V. (1971). Constants across cultures in the face and emotion. *Journal of personality and social psychology1*, 17(2):124.

Faul, F., Erdfelder, E., Lang, A.-G., and Buchner, A. (2007). G* Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, 39(2):175–191.

Fink, B., Grammer, K., and Matts, P. J. (2006). Visible skin color distribution plays a role in the perception of age, attractiveness, and health in female faces. *Evolution and Human Behavior*, 27(6):433–442.

Finke, J. B., Deuter, C. E., Hengesch, X., and Schachinger, H. (2017). The time course of pupil dilation evoked by visual sexual stimuli: exploring the underlying ANS mechanisms. *Psychophysiology*, 54(10):1444–1458.

Frantzidis, C. A., Bratsas, C., Papadelis, C. L., Konstantinidis, E., Pappas, C., and Bamidis, P. D. (2010). Toward Emotion Aware Computing: An Integrated Approach Using Multichannel Neurophysiological Recordings and Affective Visual Stimuli. *IEEE Transactions on Information Technology in Biomedicine*, 14(3):589–597.

Gerdes, A. B. M., Wieser, M. J., and Alpers, G. W. (2014). Emotional pictures and sounds : a review of multimodal interactions of emotion cues in multiple domains. *Frontiers in Psychology*, 5(December):1–13.

Gulotta, B., Sadia, G., and Sussman, E. (2013). Emotional processing modulates attentional capture of irrelevant sound input in adolescents. *International Journal of Psychophysiology*, 88(1):40–46.

Hasantash, M., Lafer-Sousa, R., Afraz, A., and Conway, B. R. (2019). Paradoxical impact of memory on color appearance of faces. *Nature Communications*, 10(1):1–10.

Hayes, T. R. and Petrov, A. A. (2016). Mapping and correcting the influence of gaze position on pupil size measurements. *Behavior Research Methods*, 48(2):510–527.

Henderson, R. R., Bradley, M. M., and Lang, P. J. (2014). Modulation of the initial light reflex during affective picture viewing. *Psychophysiology*, 51(9):815–818.

Hermans, D., Spruyt, A., De Houwer, J., and Eelen Paul (2003). Affective Priming With Subliminally Presented Pictures. *Canadian journal of experimental psychology = Revue canadienne de psychologie expérimentale*, 57(2):97–114.

Hess, E. H. and James, M. P. (1960). Pupil size as related to interest value of visual stimuli. *Science*, 132(3423):349–350.

Hess, E. H. and James, M. P. (1964). Pupil size in relation to mental activity during simple problem-solving. *Science*, 143(3611):1190–1192.

Hiramatsu, C., Melin, A. D., Allen, W. L., Dubuc, C., and Higham, J. P. (2017). Experimental evidence that primate trichromacy is well suited for detecting primate social colour signals. *Proceedings of the Royal Society B: Biological Sciences*, 284(1856).

Huang, J., Xu, D., Peterson, B. S., Hu, J., Cao, L., Wei, N., Zhang, Y., Xu, W., Xu, Y., and Hu, S. (2015). Affective reactions differ between Chinese and American healthy young adults: A cross-cultural study using the international affective picture system. *BMC Psychiatry*, 15(1):1–7.

Jessen, S., Altvater-mackensen, N., and Grossmann, T. (2016). Pupillary responses reveal infants' discrimination of facial emotions independent of conscious perception. *Cognition*, 150:163–169.

Jones, B. C., Little, A. C., Burt, D. M., and Perrett, D. I. (2004). When facial attractiveness is only skin deep. *Perception*, 33(5):569–576.

Kamp, S. M. and Donchin, E. (2015). ERP and pupil responses to deviance in an oddball paradigm. *Psychophysiology*, 52(4):460–471.

Kashihara, K., Okanoya, K., and Kawai, N. (2014). Emotional attention modulates microsaccadic rate and direction. *Psychological Research*, 78:166–179.

Kawaguchi, Y., Nakamura, K., and Tomonaga, M. (2020). Colour matters more than shape for chimpanzees' recognition of developmental face changes. *Scientific Reports*, 10:18201.

Kemp, R., Pike, G., White, P., and Musselman, A. (1996). Perception and recognition of normal and negative faces: the role of shape from shading and pigmentation cues. *Perception*, 25(1):37–52.

Kim, M. K., Kim, M., Oh, E., and Kim, S. P. (2013). A review on the computational methods for emotional state estimation from the human EEG. *Computational and Mathematical Methods in Medicine*, 2013.

Kinner, V. L., Kuchinke, L., Dierolf, A. M., Merz, C. J., Otto, T., and Wolf, O. T. (2017). What our eyes tell us about feelings: Tracking pupillary responses during emotion regulation processes. *Psychophysiology*, 54(4):508–518.

Kinzuka, Y., Minami, T., and Nakauchi, S. (2020). Pupil dilation reflects English /l//r/ discrimination ability for Japanese learners of English: a pilot study. *Scientific Reports*, 10(1):1–9.

Kitagawa, N. and Ichihara, S. (2002). Hearing visual motion in depth. *Nature*, 416(6877):172–4.

Laeng, B. and Endestad, T. (2012). Bright illusions reduce the eye's pupil. *Proceedings of the National Academy of Sciences*, 109(6):2162–2167.

Laeng, B., Sæther, L., Holmulund, T., Wang, C. E. A., Eisemann, M., and Halvorsen, M. (2013). Invisible emotional expressions influence social judgments and pupillary responses of both depressed and non-depressed individuals. *Frontiers in Psychology2*, 4:291.

Laeng, B. and Sulutvedt, U. (2014). The eye pupil adjusts to imaginary light. *Psychological science*, 25(1):188–197.

Lähteenmäki, M., Kauramäki, J., Sauter, D. A., and Nummenmaa, L. (2018). Auditory Affective Processing Requires Awareness. *Emotion*, pages 1–65.

Lang, P. J. (1980). Self-assessment manikin. *Gainesville, Fl.: University of Florida, The Center for Research in Psychophysiology.*

Lang, P. J., Bradley, M. M., and Cuthbert, B. N. (2008). International affective picture system (IAPS): Instruction manual and affective ratings. *Technical ReportA-7. University of Florida, Gainesville, FL.*

Liao, H. I., Kashino, M., and Shimojo, S. (2020). Attractiveness in the eyes: A possibility of positive loop between transient pupil constriction and facial attraction. *Journal of Cognitive Neuroscience*, pages 1–26.

Liao, H. I., Yoneya, M., Kidani, S., Kashino, M., and Furukawa, S. (2016). Human pupillary dilation response to deviant auditory stimuli: Effects of stimulus properties and voluntary attention. *Frontiers in Neuroscience*, 10(FEB):43.

Luneski, A., Konstantinidis, E., and Bamidis, P. D. (2010). Affective medicine: A review of affective computing efforts in medical informatics. *Methods of Information in Medicine*, 49(3):207–218.

Mathôt, S. (2018). Pupillometry: Psychology, Physiology, and Function. *Journal of Cognition*, 1(1):1–23.

Mathôt, S., Fabius, J., Van Heusden, E., and Van der Stigchel, S. (2018). Safe and sensible preprocessing and baseline correction of pupil-size data. *Behavior Research Methods*, 50(1):94–106.

McGurk, H. and John, M. (1976). Hearing lips and seeing voices. *Nature*, 264(5588):746–748.

Mckinnon, A. I., Gray, N. S., and Snowden, R. J. (2020). Enhanced emotional response to both negative and positive images in post-traumatic stress disorder: Evidence from pupillometry. *Biological Psychology*, 154(December 2019):107922.

Minami, T., Goto, K., Kitazaki, M., and Nakauchi, S. (2011). Effects of color information on face processing using event-related potentials and gamma oscillations. *Neuroscience*, 176:265–273.

Minami, T., Nakajima, K., and Nakauchi, S. (2018). Effects of face and background color on facial expression perception. *Frontiers in Psychology*, 9:1012.

Naber, M. and Nakayama, K. (2013). Pupil responses to high-level image content. *Journal of Vision*, 13(6):7.

Nakajima, K., Minami, T., and Nakauchi, S. (2017). Interaction between facial expression and color. *Scientific Reports*, 7:41019.

Nakajima, K., Minami, T., Tanabe, H. C., Sadato, N., and Nakauchi, S. (2014). Facial color processing in the face-selective regions: An fMRI study. *Human Brain Mapping*, 35(9):4958–4964.

Nakakoga, S., Higashi, H., Muramatsu, J., Nakauchi, S., and Minami, T. (2020). Asymmetrical characteristics of emotional responses to pictures and sounds: Evidence from pupillometry. *PLoS ONE*, 15(4).

Nomura, M., Ohira, H., Haneda, K., Iidaka, T., Sadato, N., Okada, T., and Yonekura, Y. (2004). Functional association of the amygdala and ventral prefrontal cortex during cognitive evaluation of facial expressions primed by masked angry faces : an event-related fMRI study. *NeuroImage*, 21:352–363.

Nordin, M. (1990). Sympathetic discharges in the human supraorbital nerve and their relation to sudo - and vasomotor responses. *The Journal of Physiolog*, 423(1):241–255.

Nuske, H. J., Vivanti, G., and Dissanayake, C. (2014a). Reactivity to fearful expressions of familiar and unfamiliar people in children with autism : an eye-tracking pupillometry study. *Journal of neurodevelopmental disorders*, 6(1):1–16.

Nuske, H. J., Vivanti, G., Hudry, K., and Dissanayake, C. (2014b). Pupillometry reveals reduced unconscious emotional reactivity in autism. *Biological Psychology*, 101:24–35.

Partala, T. and Surakka, V. (2003). Pupil size variation as an indication of affective processing. *International Journal of Human-Computer Studies*, 59(1-2):185–198.

Pinheiro, A. P., Barros, C., Dias, M., and Niznikiewicz, M. (2017). Does emotion change auditory prediction and deviance detection? *Biological Psychology*, 127(March):123–133.

Reinhard, G., Lachnit, H., and König, S. (2007). Effects of stimulus probability on pupillary dilation and reaction time in categorization. *Psychophysiology*, 44(3):469–475.

Rolls, E. T. (2000). The Orbitofrontal Cortex and Reward. *Cerebral Cortex*, 10(3):284–294.

Russel, J. A. and Barrett, L. F. (1999). Core affect, prototypical emotional episodes, and other things called emotion: dissecting the elephant. *Journal of personality and social psychology*, 76(5):805.

Schupp, H. T., Junghöfer, M., Weike, A. I., and Hamm, A. O. (2003). Attention and emotion : an ERP analysis of facilitated emotional stimulus processing. *Neuroreport*, 14(8):1–5.

Schupp, H. T., Stockburger, J., Bublatzky, F., Junghöfer, M., Weike, A. I., and Hamm, A. O. (2007). Explicit attention interferes with selective emotion processing in human extrastriate cortex. *BMC neuroscience*, 12:1–12.

Schupp, H. T., Stockburger, J., Bublatzky, F., Junghöfer, M., Weike, A. I., and Hamm, A. O. (2008). The selective processing of emotional visual stimuli while detecting auditory targets : An ERP analysis. *Brain Research*, 0.

Seaman, M. A., Levin, J., and Serlin, R. C. (1991). New developments in pairwise multiple comparisons: some powerful and practicable procedures. *Psychological Bulletin*, 110(3):577–586.

Sebatinelli, D., Flaisch, T., Bradley, M. M., Fitzsimmons, J. R., and Lang, P. J. (2004). Affective picture perception: gender differences in visual cortex? *Neuroreport*, 15(7):1109–1112.

Sirois, S. and Brisson, J. (2014). Pupillometry. *Wiley Interdisciplinary Reviews: Cognitive Science*, 5(6):679–692.

Snowden, R. J., O'Farrell, K. R., Burley, D., Erichsen, J. T., Newton, N. V., and Gray, N. S. (2016). The pupil's response to affective pictures: Role of image duration, habituation, and viewing mode. *Psychophysiology*, 53(8):1217–1223.

Stephen, I. D., Law Smith, M. J., Stirrat, M. R., and Perrett, D. I. (2009). Facial skin coloration affects perceived health of human faces. *International Journal of Primatology*, 30(6):845–857.

Stephen, I. D., Oldham, F. H., Perrett, D. I., and Barton, R. A. (2012). Redness enhances perceived aggression, dominance and attractiveness in men's faces. *Evolutionary Psychology*, 10(3):562–572.

Suzuki, J., Nittono, H., and Hori, T. (2005). Level of interest in video clips modulates event-related potentials to auditory probes. *International Journal of Psychophysiology*, 55(1):35–43.

Suzuki, Y., Minami, T., and Nakauchi, S. (2018). Association between pupil dilation and implicit processing prior to object recognition via insight. *Scientific Reports*, 8(1):6874.

Takahashi, F. and Kawabata, Y. (2018). The association between colors and emotions for emotional words and facial expressions. *Color Research and Application*, 43(2):247–257.

Tamietto, M. and De Gelder, B. (2010). Neural bases of the non-conscious perception of emotional signals. *Nature Reviews Neuroscience*, 11(10):697–709.

Tan, K. W. and Stephen, I. D. (2013). Colour detection thresholds in faces and colour patches. *Perception*, 42(7):733–741.

Tartar, J. L., Almeida, K. D., Mcintosh, R. C., Rosselli, M., and Nash, A. J. (2012). Emotionally negative pictures increase attention to a subsequent auditory stimulus. *International Journal of Psychophysiology*, 83(1):36–44.

Taylor, S. E. (1991). Asymmetrical effects of positive and negative events: The mobilization^minimization hypothesis. *Psychological Bulletin*, 110(1):67–85.

Thierry, G. and Roberts, M. V. (2007). Event-related potential study of attention capture by a ? ective sounds. *Neuroreport*, 18(3):5–8.

Thorstenson, C. A., Elliot, A. J., Pazda, A. D., Perrett, D. I., and Xiao, D. (2018). Emotion-Color Associations in the Context of the Face. *Emotion*, 18(7):1032–1042.

Thorstenson, C. A., Pazda, A. D., and Elliot, A. J. (2017). Subjective perception of color differences is greater for faces than non-faces. *Social Cognition*, 35(3):299–312.

Vuilleumier, P. (2005). How brains beware: Neural mechanisms of emotional attention. *Trends in Cognitive Sciences*, 9(12):585–594.

Wang, C.-A., Baird, T., Huang, J., D, C. J., Brien, D. C., and Munoz, D. P. (2018). Arousal effects on pupil size, heart rate, and skin conductance in an emotional face task. *Frontiers in Neurology*, 9:1029.

Wang, C. A. and Munoz, D. P. (2015). A circuit for pupil orienting responses: Implications for cognitive modulation of pupil size. *Current Opinion in Neurobiology*, 33(Figure 1):134–140.

Watson, A. B. and Yellott, J. I. (2012). A unified formula for light-adapted pupil size. *Journal of Vision*, 12(10):1–16.

Wierda, S. M., Rijn, H. V., Taatgen, N. A., and Martens, S. (2012). Pupil dilation deconvolution reveals the dynamics of attention at high temporal resolution. *Proceedings of the National Academy of Sciences*, 109(22):8456–8460.

Willenbockel, V., Sadr, J., Fiset, D., Horne, G. O., Gosselin, F., and Tanaka, J. W. (2010). Controlling low-level image properties : The SHINE toolbox. *Behavior Research Methods*, 42(3):671–684.

Wolfe, J. M. and Horowitz, T. S. (2004). What attributes guide the deployment of visual attention and how do they do it? *Nature reviews neuroscience*, 5(6):495–501.

Wrase, J., Klein, S., Gruesser, S. M., Hermann, D., Flor, H., Mann, K., Braus, D. F., and Heinz, A. (2003). Gender differences in the processing of standardized emotional visual stimuli in humans: a functional magnetic resonance imaging study. *Neuroscience Letters*, 348(1):41–45.

Yip, A. W. and Sinha, P. (2002). Contribution of color to face recognition. *Perception*, 31(8):995–1003.

Young, S. G., Thorstenson, C. A., and Pazda, A. D. (2018). Facial redness, expression, and masculinity influence perceptions of anger and health. *Cognition and Emotion*, 32(1):49–60.

Appendix A

Emotional responses to pictures and sounds

A.1 Supporting Information in Chapter 3

IAPS numbers in Experiment 1 and IADS numbers in Experiment 2.

A.1.1 IAPS pictures

The IAPS pictures (Lang et al., 2008) used in this study were: Positive: 4220, 4250, 4290, 4599, 4643, 4659, 4687, 5629, 7280, 8030, 8034, 8170, 8185, 8193, 8300, 8370, 8461, 8470, 8501, 8549; Negative: 1019, 1120, 1220, 1274, 2683, 2688, 2981, 3150, 3350, 5971, 6021, 6230, 6242, 6510, 8485, 9250, 9429, 9622, 9902, 9921; Neutral: 1675, 1726, 2357, 2372, 2512, 2514, 2595, 2635, 5395, 7036, 7037, 7043, 7207, 7211, 7242, 7487, 7493, 7640, 8475, 9402.

A.1.2 IADS sounds

The IADS sounds (Bradley and Lang, 2007) used in this study were: Positive: 110, 200, 202, 220, 224, 230, 311, 352, 353, 360, 365, 367, 717, 726, 811, 813, 815, 816, 817, 820; Negative: 105, 116, 241, 242, 244, 250, 280, 283, 289, 295, 296, 423, 501, 600, 611, 626, 703, 711, 713, 730; Neutral: 104, 170, 171, 246, 262, 322, 373, 376, 376, 377, 382, 627, 698, 700, 701, 705, 708, 720, 722, 723.

Appendix B

Emotion estimation method using pupillometry

B.1 Supporting Information in Chapter 4

IAPS numbers in Experiment.

B.1.1 IAPS pictures

The IAPS pictures (Lang et al., 2008) used in this study were: Neutral (Standard): 2190, 2191, 2200, 2214, 2215, 2396, 2749, 2840, 2850, 2880, 5130, 5500, 5531, 5534, 5740, 7000, 7002, 7004, 7006, 7010, 7020, 7034, 7038, 7080, 7090, 7096, 7100, 7150, 7161, 7175, 7185, 7207, 7211, 7217, 7224, 7233, 7493, 7595, 7710, 9070; Neutral (Oddball): 2102, 2221, 5471, 5532, 6150, 7009, 7140, 7283, 7950, 9210; Negative (Standard): 1019, 1090, 1201, 1220, 1525, 2120, 2683, 2691, 2717, 2981, 3100, 3120, 3130, 3150, 3170, 3180, 3230, 3250, 3500, 6020, 6210, 6242, 6313, 6410, 6540, 6821, 6838, 8231, 9040, 9140, 9250, 9253, 9410, 9421, 9429, 9433, 9495, 9594, 9902, 9921; Negative (Oddball): 1111, 1274, 2800, 3000, 3350, 5971, 6510, 8485, 9042, 9622; Positive (Standard): 1811, 2058, 2345, 4220, 4617, 4626, 4640, 4659, 4680, 4687, 5260, 5470, 5600, 5621, 5629, 5700, 5833, 5910, 7230, 7270, 7330, 8030, 8034, 8041, 8080, 8090, 8170, 8180, 8185, 8200, 8210, 8280, 8300, 8370, 8380, 8400, 8420, 8470, 8499, 8502; Positive (Oddball): 1710, 2208, 4290, 4660, 5460, 5480, 7400, 8190, 8501, 9156.