早稲田大学審査学位論文
博士（人間科学）

An Experimental Study of the Perceptual Features that Influence Multicolor Aesthetic Evaluation

２０１８年７月

早稲田大学大学院　人間科学研究科
方　思源
FANG, Siyuan

研究指導教員：　松居　辰則　教授

Abstract

In empirical aesthetics, with the paradigm shift from behaviorism to cognitivism and the development of neuroaesthetics, information-processing psychological models were built to explain humans' sense of beauty for visual objects. Helmut Leder's model is the most comprehensive. According to the model, during the psychological process of multicolor aesthetic evaluation, certain perceptual features innate to multicolor stimuli are able to influence aesthetic evaluation of multicolor stimuli. Specifically, physical color information is transformed into these perceptual features at the Perceptual Analyses Stage in the visual information processing module. The features are then transformed into an aesthetic evaluation in the affective evaluation module. Using Leder's model as the theoretical framework, our research aimed to clarify, through psychological experiments, what perceptual features influence multicolor aesthetic evaluation and how these perceptual features exert their influences.

Chapter 1 introduces the background and objective of the research. Chapter 2 investigates the multicolor-level perceptual features (i.e., perceptual features that are attributes of the multicolor stimuli themselves) involved in multicolor aesthetics. Chapter 3 investigates the influence of color focality, which is a component-level perceptual feature (i.e., a perceptual feature that is a property of the component colors of the multicolor stimuli), on multicolor aesthetics. The experiments were conducted in Japanese, and all participants were students of Waseda University. Chapter 4 summarizes the results of Study 1 and Study 2 and proposes three links between the two studies. Chapter 5 introduces the implications for future work.

In Study 1 (Chapter 2), we first conducted two psychological experiments using the semantic differential method. The experimental stimuli were thirty-five 4×4 color grids. The aesthetic score of each stimulus was defined as the inverse of the factor score on the factor "Pleasure" extracted in the first experiment. In the second experiment, three factors, "Stability," "Heaviness," and "Presence," were extracted and each was regarded as a multicolor-level perceptual feature.

Then, a neural network model was constructed to explicate how the three perceptual features influence aesthetic evaluation of multicolor stimuli. The psychological appropriateness of the model was validated through two simulations. A post-training microstructure analysis revealed that Heaviness exerted a large and negative influence on the aesthetic evaluation, Stability a small and negative influence, and Presence a small and positive influence.

In Study 2 (Chapter 3), we experimentally investigated the continuous relationship between focality and preference of single colors and the psychological variable(s) that mediated this relationship. Literature has suggested that the focality of the component colors of a multicolor stimulus can affect the aesthetic evaluation of the stimulus by first influencing the aesthetic preference of the component colors. The experimental stimuli were 30 Munsell chips. The candidates for the mediating variables were fluency of psychological information processing (PPF) of colors and 22 color impressions. In the first experiment, Session 1 measured the PPF of each stimulus by testing its short-term memory accuracy because short-term memory accuracy reflects multiple PPF-related perceptual properties. Session 2 quantified the focality of each stimulus based on the subjects' categorization of the stimuli. In the second experiment, the subjects rated the stimuli on 23 Likert scales representing the color impressions and color preference. We found a negative linear relationship between focality and color impression *gracefulness* and a positive linear relationship between *gracefulness* and preference. This implies that color focality has a negative impact on color preference, which is mediated by *gracefulness*. We found no other color impression having such relationship chains. PPF's role as a mediating variable was also not supported.

In conclusion, our research found that three multicolor-level perceptual features (i.e., Stability, Heaviness, and Presence) and one component-level perceptual feature (i.e., color focality) can influence multicolor aesthetic evaluation, at least in modern Japanese culture. These results refined Leder's model and could help promote automation in such areas as generative art, kansei engineering (affective computing), industrial design, and human-computer interface.

Contents

List of Figures

List of Tables

Chapter 1 Research Background and Objective

This chapter introduces the background, goal, and structure of this research. Section 1.1 of this chapter provides a brief history of empirical aesthetics. Section 1.2 introduces Helmut Leder's psychological model on aesthetics and the reasons why we selected this psychological model as the theoretical framework of our research. The use of this theoretical framework positions our research within the field of empirical aesthetics. This section then explains how we established our research objective within this theoretical framework, namely, why it is necessary to experimentally clarify the perceptual features that operate during the psychological process of multicolor aesthetic evaluation. Section 1.3 focuses on the contributions of our research to the area of multicolor aesthetics. First, past studies in multicolor aesthetics, especially those using empirical approaches are introduced. These studies form a subarea in empirical aesthetics. This section then describes the position of our research in empirical multicolor aesthetics and the contributions of our research to this field. Next, because our research is a research of brain function modeling, Section 1.4 describes the four levels of brain function modeling, as well as the level on which our research is situated. Section 1.5 introduces the academic significance and possible applications of our research, as well as the significance of our research in the area of human sciences. Section 1.6 introduces the overall structure of our research.

This research was developed on the basis of Siyuan Fang's master's research which was described in detail in his master thesis *Extraction of Aesthetic Rules to Multi-Color Stimuli Using*

Artificial Intelligence Technology: Towards the Construction of an Artificial KANSEI System. Thus, the present thesis and his master's thesis share a lot in terms of research background and literature review.

1.1 A Brief History of Empirical Aesthetics

[1] Since the inception of scientific psychology in the 19th century, aesthetics remains one of the most popular topics in this field. Gustav T. Fechner, one of the founding fathers of scientific psychology, inaugurated the empirical aesthetics research area in 1876. The term "empirical aesthetics" means studying aesthetics using experimental and quantitative approaches. This marked a paradigm shift because the metaphysical way of thinking had dominated the study of humans' sense of beauty since the time of ancient Greece. The specific methodology that Fechner adopted in empirical aesthetics was psychophysics. Psychophysics, which was also developed by Fechner, experimentally specifies the relationships between physical features and psychological effects of a given type of stimuli. This was a behaviorist method because it did not consider the mechanism by which physical features were transformed into the psychological metrics[12, 24].

In 1933, George David Birkhoff, an American mathematician, proposed a formula to quantify the long-held philosophical assertion that beauty is derived from a "unity in variety." The formula was

$$M = O/C \tag{1.1}$$

where M signifies the degree of the aesthetic pleasure elicited by a stimulus, O the degree of orderliness of the stimulus, and C the degree of complexity of the stimulus. Later, Bense and Moles combined Birkhoff's formula with Shannon's information theory, and Machado proposed a more sophisticated way of quantifying the concept "unity in variety" using con-

[1]A brief version of Section 1.1 was published in Fang, Muramatsu, and Matsui's 2015 research paper[17].

cepts borrowed from digital image compression. Although these formulas are not derived from experiments, they provide us with a means to test conceptually expressed arguments through experiments. This tradition plays a minor role in empirical aesthetics compared to the tradition created by Fechner[12, 24, 61].

Fechner's behaviorist experimental approach was inherited by Denial E. Berlyne in the middle of the 20th century. Berlyne was regarded as the core figure in empirical aesthetics at his time for putting forward a theory that tried to explain the psychological mechanism underlying the sense of beauty. According to this theory, the three types of properties, namely, the "psychological properties," the "ecological properties," and the "collative properties," of a stimulus possess "arousal potential." The arousal potential means the potential to activate the rewarding system and the aversion system, which he thought each had a different neural basis. Both systems take the arousal potential of the properties of the stimulus as input and the hedonic feeling elicited by the stimulus as output. The difference is that in the rewarding system the arousal potential is proportional to the hedonic feeling, while in the aversion system the arousal potential is in inverse proportion to the hedonic feeling. The overall hedonic feeling elicited by the stimulus is the sum of the output of the rewarding system and the output of the aversion system. Because the aversion system is activated after the rewarding system has been activated for some time, a bell-shaped relationship appears between the arousal potential and the total degree of hedonic feeling. In other words, as the arousal potential increases, the total degree of hedonic feeling first increases and then decreases at the time point when the aversion system is activated[12, 24].

Later, as cognitivism took the place of behaviorism as the dominant paradigm in general psychology, empirical aesthetics also underwent a similar paradigm shift. Researchers in empirical aesthetics were no longer satisfied with just studying the mappings from physical attributes to measures of aesthetic pleasure and instead began to delve into the information processing mechanism behind such mappings[12, 24]. Colin Martindale conducted a series of experiments with results that did not match the predictions of Berlyne's theory, which made him cast doubts on Berlyne's theory[63]. Martindale, a zeal advocate for connectionism, proposed a neural-network model that aims to explain his experimental results on color preference[64]. Another representative figure of cognitivism in empirical aesthetics is Rudolf Arnheim, a renowned German-born art theorist and psychologist in vision. Arnheim introduced gestalt psychology into empirical aesthetics and proposed a cognitive theory that explained humans' aesthetics towards visual arts[2].

During the late 20th century, with the development of cognitive neuroscience, more and more researchers in empirical aesthetics began to use non-invasive brain-imaging techniques such as positron emission tomography (PET) and functional magnetic resonance imaging (fMRI) to pinpoint the neural basis of humans' sense of beauty. This area is named "neuroaesthetics"[24]. Semir Zeki, a British perceptual neuroscientist, argued in his enlightening essay *Artistic Creativity and the Brain*[124], that

> ⋯ artists and neurobiologists have both studied the perceptual commonality that underlies visual aesthetics. ⋯ I believe that artists are, in a sense, neuroscientists who unknowingly study the brain with techniques unique to them. Visual art contributes to our understanding of the visual brain because it explores and reveals the brain's perceptual capacities.

Later, in his *Statement on Neuroesthetics*[125], Zeki holds that, on one hand, "⋯the artist is in a sense, a neuroscientist, exploring the potentials and capacities of the brain, though with different tools", and on the other hand, "⋯neuroscientists would do well to exploit what artists, who have explored the potentials and capacities of the visual brain with their own methods, have to tell us in their works." Perhaps for this reason, most studies in neuroaesthetics used visual artworks as experimental stimuli. The results in empirical aesthetics found no brain regions specialized to aesthetics-related tasks. All brain regions activated during the aesthetics-related tasks, such as the orbitofrontal cortex, the dorsallateral portion of the anterior prefrontal cortex, the anterior cingulate cortex, and the nucleus accumbens, are also activated in other tasks[3, 43, 48, 48, 54]. This suggests that the mental function of aesthetics can be reduced to more fundamental functions, or, using the terms of the system sciences, the sense of beauty probably emerges from specific patterns of interactions among some more elementary mental functions. This reductionist view is adopted in Anjan Chatterjee's[8] and Helmut Leder's[55] psychological models of aesthetics. These two models are the only existing cognitive psychological models in the field of empirical aesthetics.

According to Chatterjee's model, the psychological mechanism of aesthetics is mainly comprised of visual information processing, memory extraction, and emotional response. Based on David Marr's theory of vision, Chatterjee proposes that the module of visual information processing consists of three stages: "early vision," "intermediate vision," and "late vision," The early vision stage processes simple features of visual stimuli, such as color, lightness, motion, and location. At the intermediate vision stage, these simple features are combined into clusters. At the late vision stage, some of the clusters are selected for further processing. At the late vision stage, memories relating to these feature clusters are extracted and associated with these feature clusters as semantic meanings of the feature clusters. The outputs of the late vision stage have direct links with the module of emotional response, and the early vision stage and the intermediate stage interact with the emotional response module indirectly. The emotional response module outputs "liking without wanting" responses in the case of aesthetic evaluation.

In Leder's model, the mental function of aesthetics has three components: the module of visual information processing, the module of memory extraction, and the module of continuous affective evaluation. The module of visual information processing is divided into four stages. The first stage, called "Perceptual Analyses," processes the primary perceptual features of visual stimuli, such as contrast, complexity, symmetry, color, and grouping (i.e., the goodness as a gestalt). The second stage, called ”Implicit Memory Integration,” evaluates degrees of familiarity, typicality, and peak-shift effect of the stimuli. At the third stage, which is called ”Explicit Classification,” the style and content of the stimuli are specified. At the fourth stage, which is called ”Cognitive Mastering,” an art-specific interpretation and/or a self-related interpretation is assigned to the stimuli. The module of memory extraction exchanges information with the second, third, and fourth stages of the visual information processing module. The module of continuous affective evaluation works in parallel with the visual information processing module. The continuous affective evaluation module evaluates the aesthetic preference of the stimuli based on the outputs of each stage of the visual information processing module. In other words, the aesthetic evaluation of the stimuli can be influenced by the outputs of any stage during visual information processing.

1.2 Theoretical Framework and Research Objective

[2] This section first explains why this research adopts Leder’s cognitive psychological model on aesthetics as the theoretical framework. This section then describes how the goal of this research was established within this theoretical framework.

When comparing the two cognitive psychological models of aesthetics, we can see that the Perceptual Analyses Stage in Leder’s model generally matches the early vision stage in Chatterjee’s model, and the Implicit Memory Integration Stage in Leder’s model generally matches the intermediate vision stage in Chatterjee’s model. The late vision stage in Chatterjee’s model can be considered to encompass both the Explicit Classification Stage and the Cognitive Mastering Stage in Leder’s model. A noticeable difference between the two models is that Leder’s model explains how affective evaluation interacts with each stage of visual information processing,

[2] A brief version of the literature review in Section 1.2 was published in Fang, Muramatsu, and Matsui’s 2015 research paper[17].

while, in Chatterjee's model, the links between each vision stage and the emotional response are not clearly described. In other words, compared to Chatterjee's model, Leder's model has a greater explanatory power in terms of interactions between visual information processing and affective evaluation. Thus, our research adopts Leder's model as the theoretical framework.

From a philosophical perspective, there are two types of beauty. One is content-dependent. When we encounter a stimulus, we cannot experience this type of beauty of the stimulus until we understand the semantic content of the stimulus. The appreciation of representative paintings, such as Jacques-Louis David's *The Coronation of Napoleon* and John Everett Millais's *Ophelia*, is a good example of such an aesthetic experience. The other type of beauty is content-independent. This means that our ability to experience this type of beauty of a stimulus does not depend on our understanding of the semantic content of the stimulus. This type of beauty is derived from our aesthetic instinct towards formal attributes such as shape, color, degree of symmetry, and degree of complexity. The appreciation of abstract artworks, such as Claude Monet's *Impression* and *Soleil Levant*, is a typical example of this sort of aesthetic experience.

The content-dependent sense of beauty is formed from our individual experiences, such as cultural background, artistic education received, peer fashion, and social class. Thus, this sense of beauty varies largely across people. On the contrary, the content-independent sense of beauty is probably genetically circuited, namely, free from acquired knowledge and life experience. Thus, this sense of beauty might be found universally across people to a great extent. In Leder's model, the first stage, the Perceptual Analyses Stage, corresponds to the content-independent sense of beauty because the information processing at this stage, namely the evaluation of the simple perceptual features of stimuli, is free from the influence of the memory module. On the other hand, the other three stages correspond to the content-dependent sense of beauty because all these stages exchange information with the memory module. Because our interest is in memory-free factors that influence multicolor aesthetic evaluation, only the Perceptual Analyses Stage is used in our research. Hence, when applied in this research, Leder's model can be paraphrased to the effect that the chromatic metrics of multicolor stimuli are first transformed into a number of simple perceptual features in the Perceptual Analyses Stage of the visual information processing module, and then these perceptual features are mapped to a certain degree of aesthetic pleasure in the module of continuous affective evaluation (as displayed in Figure 1.1).

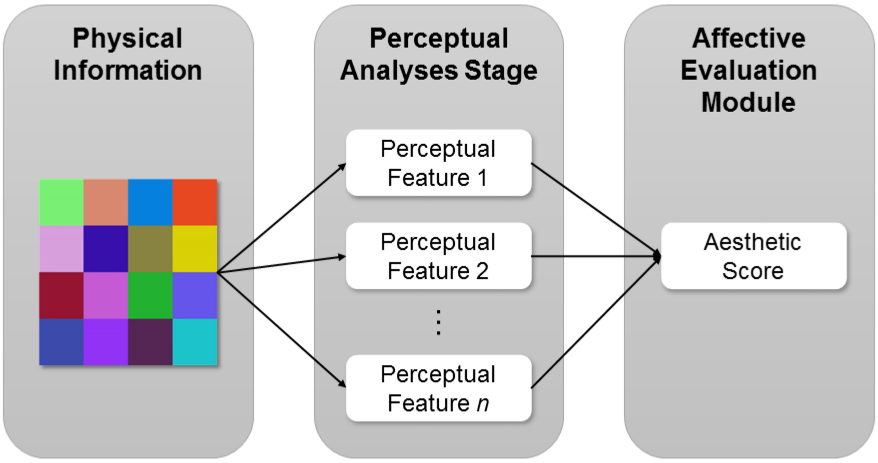

Figure 1.1: Application of Helmut Leder's cognitive psychological model on multicolor aesthetics as the theoretical framework of the present research.

However, although Leder mentioned color as one candidate for perceptual features that mediate chromatic information and aesthetic evaluation, there exists no experimental research that tells us what aspects of color fulfills this role. In other words, we have yet to know what perceptual features of color operate during the psychological process of multicolor aesthetic evaluation. Hence, this research aims to clarify these perceptual features through psychological experiments. This research also aims to clarify how these perceptual features influence the aesthetic evaluation in a quantitative manner.

1.3 Contributions to Multicolor Aesthetics

1.3.1 Two Approaches in Multicolor Aesthetics

This section describes the position of this research in the area of multicolor aesthetics and the contributions of this research to this area. As a subarea in aesthetics, multicolor aesthetics also experienced a shift from a period dominated by metaphysical approaches to a period dominated by empirical approaches. The classical theories on color harmony best represent the metaphysical period. These theories mark the beginning of the theorization of humans' elusive sense of beauty.

Leonard da Vinci is the first to make comments on color harmony. Da Vinci recorded the psychological effects of a few color combinations based on his long-time experience in painting. Later on, Isaac Newton made an analogy of color harmony to chords in music, which were results of proportional vibration frequencies of musical instruments. He thus came up with the idea that the harmonious relationships among colors were perhaps generated by the chords of neural signals, namely some special vibration frequencies of neural signals transmitted from the eyes to the brain. However, neither da Vinci nor Newton formulated their ideas on color harmony into systematic theories[9, 11]. A real color harmony theory, according to Gentaro Ohmi, is a theory that systematically explains the degree of harmony or disharmony of any color combination. Ohmi called the theories constructed by metaphysical approaches ”top-down theories” and the theories built by the empirical approaches ”bottom-up theories”[87].

1.3.2 Classical Color Harmony Theories

Top-down theories were constructed under the guide of some metaphysical principles that the theoreticians considered to be self-evident. Looking back upon the history of the classic color harmony theories, we found four such principles.

The first principle, which was first proposed by Johann Wolfgang von Goethe, argues that neutrality leads to harmony. On a color wheel invented by Goethe, six primary colors (blue, yellow, green, orange, purple, and red) are placed with even intervals. These primary colors form three complementary color pairs (blue-orange, purple-yellow, and green-red). Each complementary color pair is regarded as harmonious because the mixture of the two component colors elicits a neutral feeling, namely a balanced sensation of color. This idea was inherited by Field who contended that if two colors that are placed on a rotating disk can produce a gray color, they are in a harmonious relationship. In the 20th century, Johannes Itten, who is one of the establishers of Staatliches Bauhaus, proposed a theory that absorbed this idea. Itten argued that two colors with equal distance to the central point of his color sphere are harmonious[9, 11, 87].

The second principle was raised by Michel Eugene Chevreul, a 19th century French chemist, based on his theory of simultaneous color contrast. This principle can be summarized to state that the colors that produce certain contrast effects are harmonious. Chevreul divided color harmony into two categories: the harmony of analogous colors and the harmony of contrastive colors. Analogous colors are colors whose locations are close to each other on Chevreul's color solid. Contrastive colors are colors situated far from each other on the color solid. Later, Parry Moon and Domina E. Spencer introduced this categorization into their color harmony theory, which will be introduced in detail at the end of this section. Moon and Spencer's theory alleges that harmonious colors are colors that are located within the identity region, the similarity region, and the contrast region and each dimension of their color space (called the "ω space"). Based on the same principle, Faber Birren divided color harmony into three types: the harmony of adjacents, the harmony of opposites, and the harmony of split-complements. In the 20th century, Practical Color Coordinate System (PCCS) also adopted this principle. PCCS divided hue combinations into seven categories: Identity, Adjacent, Analogy, Intermediate, Contrast, Split Complementary, and Complementary[9, 11, 72, 73, 74, 75, 87].

The third principle was initially put forward by Friedrich Wilhelm Ostwald, a renowned German chemist, in the late 19th century. He developed a color space and argued that harmony equals order, which means that if a set of colors are selected from this color space in an orderly way, they will produce a harmonious feeling. Typical orderly spatial distributions of colors are the isotint color series, isotone color series, isochrome color series, and isovalent color series, which are defined in each equal-hue triangle. This thought was later adopted by Itten. Itten holds that colors that can form a simple geometric shape, for example, a triangle, a rectangle, a pentagon, or a hexagon, in Itten's color space are harmonious. The same principle was also expressed in the PCCS (Practical Color Co-ordinate System) manual on color design. In addition, the concept "the harmony of triads" in Birren's color theory is also an application of this principle[9, 11, 87].

The fourth principle, originally proposed by Ogden Nicholas Rood, is called "natural harmony." This principle argues that color transitions frequently appearing in nature, namely, natural sequences of colors, tend to be sensed as harmonious. Wilheim von Bezold and Ernst Wilheim von Brucke also hold similar opinions[9, 11, 87].

Deane B. Judd, based on an extensive review of literature on color harmony, made a summary that four principles directed the proposal of top-down theories on color harmony: the "Principle of Order," the "Principle of Familiarity," the "Principle of Similarity," and the "Principle of Unambiguity." The Principle of Order corresponds to the third principle in our summary, namely, the principle raised by Ostwald. The Principle of Familiarity corresponds to the principle of natural harmony in our summary. The remaining two principles, namely the Principle of Similarity and the Principle of Unambiguity generally correspond to the second principle in our summary, which was proposed by Chevreul. Judd regards the Principle of Unambiguity as one of the most important principles because Moon and Spencer's color harmony theory emphasized this principle[9, 11, 87]. Moon and Spencer defined two ambiguity regions on each dimension of their color space, namely the ω space. They contend that if the color difference among a set of colors lies within one of the ambiguity regions, these colors tend to be felt as disharmonious[72, 73, 74]. The main difference between Judd's set of principles and our set of principles is that Judd's set of principles does not include our first principle, namely the neutrality principle proposed by Goethe.

Moon and Spencer's theory on color harmony is the first to try to quantify the degree of color harmony. Birkhoff's formula $M = O/C$ (introduced in Section 1.1) serves as the theoretical basis for their computation method. The concept "complexity (C)" is defined as the sum of the number of colors and the number of the pair-wise color differences on all dimensions of the color space. The concept "order (O)" is defined as the sum of the weights of the color differences in each region on each dimension. If the measure of the aesthetic pleasure (M) elicited by a set of colors exceeds 0.5, these colors are considered to be harmonious[73, 74, 75]. Because the theoretical basis of this quantitative theory is the metaphysical assertion "beauty is unity in variety," rather than something obtained from experiments or observations, this theory, although quantitatively expressed, is still a top-down one.

With the development of empirical aesthetics, empirical approaches began to blossom in multicolor aesthetics. Researches using empirical approaches established their color theories based on experimental data obtained from large groups of subjects[11, 87]. The next section introduces some representative works using empirical approaches.

1.3.3 Past Empirical Studies in Multicolor Aesthetics

This section introduces a number of studies in multicolor aesthetics in which psychological experiments were conducted.

Ou and Luo[90] performed a psychological experiment to study the general rules of two-color harmony. In the experiment, seventeen Chinese college students were asked to rate 1431 color pairs on two five-point evaluation scales measuring "harmonious" and "disharmonious." Then, the evaluation data were transformed to a color harmony score. Using bubble charts as the data-visualization method, three factors that influence the color harmony score were found. The first one is "chromatic difference (ΔC)." Chromatic difference is a color appearance feature that is computed by

$$\Delta C = \sqrt{(\Delta H_{ab})^2 + (\frac{\Delta C_{ab}}{1.46})^2} \tag{1.2}$$

where ΔC_{ab} is chromatic difference and ΔH_{ab} is hue difference. The relationship between chromatic difference and the color harmony score (H) is

$$H_c = 0.04 + 0.053 \tanh(0.8 - 0.045\Delta C) \tag{1.3}$$

The second factor, called the "lightness effect," combines the effect of lightness difference (ΔL) and the effect of lightness sum (L_{sum}). The relationship between lightness sum and the color harmony score is

$$H_{Lsum} = 0.28 + 0.54 \tanh\left(-3.88 + 0.029 L_{sum}\right) \tag{1.4}$$

The relationship between lightness difference and the color harmony score is

$$H_{\Delta L} = 0.14 + 0.15 \tanh\left(-2 + 0.2\Delta L\right) \tag{1.5}$$

The overall effect of lightness is the sum of H_{Lsum} and $H_{\Delta L}$. The third factor, called the "hue effect," depicts the influence of the hues of the component colors on the harmony score. The relationship between a component hue (h_{ab}) and the average harmony score of the two-color combinations that include the hue is

$$H_s = -0.08 - 0.14\sin(h_{ab} + 50^\circ) - 0.07\sin(2h_{ab} + 90^\circ) \tag{1.6}$$

In addition, light yellow was found to be more harmonious with other hues than dark yellow. This phenomenon was named the "yellow effect (E_Y)" and is formulated as

$$E_Y = \frac{0.22L - 12.8}{10}\exp(\frac{90^\circ - H_{ab}}{10} - \exp(\frac{90^\circ - H_{ab}}{10})) \tag{1.7}$$

The effect of the hue of a component color (H_{SY}) is

$$H_{SY} = E_C(H_S + E_Y) \tag{1.8}$$

where

$$E_C = 0.5 + 0.5\tanh(-2 + 0.5C_{ab}) \tag{1.9}$$

The hue effect of a two-color combination (H_H) is the sum of the hue effects of its two component colors:

$$H_H = H_{SY1} + H_{SY2} \tag{1.10}$$

Finally, the overall model for two-color harmony is the sum of the effect of chromatic difference, the lightness effect, and the hue effect:

$$CH = H_C + H_L + H_H \tag{1.11}$$

where CH denotes the color harmony score. Through an F-test, the significance of all three factors was verified. This model shows that the high lightness sum and large lightness difference lead to a high harmony score. This model also shows that the hue difference and chromatic difference negatively affect the color harmony score. In addition, blue is the easiest to harmonize with other hues, and red is the most difficult to harmonize with other hues.

Lee, Kobayashi, and Sobagaki[56] conducted a psychological experiment to investigate the linear relationships between the "beautifulness" ("utsukushisa" in Japanese) and physical metrics of two-color combinations. The physical metrics are composed of the normalized $L^*c^*h^*$ coordinate values of the component colors. In their experiment, 24 subjects were asked to rate 102 two-color combinations using the five-point evaluation scale "beautiful-not beautiful." The evaluation data were averaged across the subjects and were then processed by multivariate linear regression. The valid predictive variables of the regression are $\bar{L}$ (the average of the L values of the component colors), $|\Delta L^*|$ (the absolute value of the difference between the L^* values of the component colors), $|\Delta C^*|$ (the absolute value of the difference between the C^* values of the component colors), and $|\Delta H^*|$ (the absolute value of the difference between the H^* values of the component colors). $|\Delta H^*|$ is calculated by

$$|\Delta H^*| = 2\sqrt{\bar{C^*} - (\frac{\Delta C^*}{2})^2} * \sin(\frac{\Delta h}{2}) \tag{1.12}$$

where $\bar{C^*}$ is the average of the C^* values of the component colors and Δh is the difference between the h values of the component colors. The regression equation is

$$y = 2.163 + 1.511L^* + 1.288|\Delta L^*| - 0.659|\Delta C^*| - 1.893|\Delta H^*| \tag{1.13}$$

The authors conclude from this regression equation that, as far as two-color combinations are concerned, the average lightness and the lightness difference of the component colors possess a positive impact on beautifulness, and the saturation difference and the hue difference of the component colors possess a negative impact on beautifulness.

Szabo, Bodrogi, and Schanda[113] conducted a psychological experiment that aimed to clarify the principles of two-color and three-color harmony. The two-color stimuli in the experiment were divided into two categories: "monochromatic two-color combinations" and "dichromatic two-color combinations." A monochromatic two-color combination is composed of two colors that have the same or almost the same hue. A dichromatic two-color combination is composed of two colors that have different hues. In the same manner, the three-color stimuli are divided into two categories: "monochromatic two-color combinations" and "trichromatic three-color combinations." The colors that constituted the stimuli were systematically sampled from the CIECAM02 color space. All the stimuli were labeled by nine subjects who were Hungarian college students using an 11-point scale that represented a continuum from "the best harmony" to "the worst harmony." Based on the experimental data, four models were constructed. Each model described the color harmony rules of one stimulus category. The models revealed that the color harmony score increases as the lightness difference between the component colors increases and/or the chromatic difference decreases. In addition, the higher the lightness sum, the higher the color harmony score. In terms of hue, a low hue difference tended to bring about a high color harmony score. Individual hues also affect the color harmony score. A validation experiment found that the subjects preferred the color combinations proposed by Szabo et al.'s[113] model and Ou and Luo's[90] model to the color combinations proposed by some classical color harmony theories.

Kansaku[41], for the purpose of clarifying the affective effects of two-color combinations, conducted a psychological experiment using the Semantic Differential (SD) method. The experimental stimuli, which were made of the standard color samples developed by the Japanese Color Research Institute, were rated by 30 subjects who were students or staff members of Tokyo Metropolitan University. The nine-point evaluation scale "harmonious-disharmonious" ("chowa-fuchowa" in Japanese) was used in the experiment. The results of the factor analysis indicate that the scale "harmonious-disharmonious" belongs to the factor "Evaluation." A comparison of the results of this study and the results of an earlier study reveal that the evaluation scales that have high correlation with the scale "harmonious-disharmonious" in both studies are ”stylish-boorish,” ”like-dislike,” ”feels good-feels bad,” ”refined-rustic,” ”plain-thick,” ”clean-dirty,” ”clear-turbid,” ”light-heavy,” and ”merged-scattered.” With respect to the color appearance attributes, the evaluation score on the scale ”harmonious-disharmonious” increases as the lightness difference between the component colors increases. On the other hand, no clear relationship was found between the evaluation score on the scale ”harmonious-disharmonious” and the hue/chromatic difference between the component colors. Besides, the two-color combinations that include a red component color tend to have a low score on the harmony scale.

Mori et al.[76], aiming to establish a model for two-color harmony appraisement, performed a psychological experiment using the paired comparison method. In the experiment, 102 two-color combinations were compared with each other by 207 subjects in harmony degree. The component colors of the stimuli were systematically selected from the JIS standard color samples that were developed based on the Munsell color system. The comparison data was transformed to a color harmony score for each stimulus. Then, a multivariate linear regression was performed to examine the relationship between the harmony score and a set of physical features. The physical features are composed of Z_1, Z_2 and Z_3, which are three variables transformed from the H, V, C values of the Munsell color system:

$$Z_1 = C * \cos\theta \tag{1.14}$$

$$Z_2 = C * \sin\theta \tag{1.15}$$

$$Z_3 = 8.33V \tag{1.16}$$

Eleven physical features were selected through the forward selection procedure from an initial set of 27 features. The regression equation demonstrates that the lightness of the component colors and the lightness difference between the component colors possess the strongest impact on the color harmony score. The regression equation also shows that no clear relationship is found between the color harmony score and the hue/chromatic difference between the component colors. Moreover, the positions of the component colors in the color system also influence the color harmony score. Based on the regression results, two color maps that mark the harmonious regions, the intermediate region, and the disharmonious region were made. The dimensions of the first map, called the "$\Delta V - \Delta C$ Map," were the lightness difference and the chromatic difference. The dimensions of the second map, called the "$\Delta H - \Delta V$ Map," were the hue difference and the lightness difference. The △ V-△ C Map was considered by the authors to be more important than the $\Delta H - \Delta V$ Map. The experimental data was also used to test the predicting power of Moon and Spencer's color harmony theory. It turned out that Moon and Spencer's theory failed to align with the experimental data.

Nayatani et al.[79] conducted a psychological experiment using the SD method to investigate the affective effects of three-color combinations. One hundred and twenty subjects Studyicipated in the experiment and 100 three-color combinations were rated by the subjects on 38 evaluation scales. Through the factor analysis, six factors were extracted. One of the factors is called "Pleasantness." Because each component color had a common border with the other two component colors in a stimulus, the authors tested the possibility of predicting the factor score of a stimulus on the factor Pleasantness using the three pair-wise color harmony scores of the stimulus. The three color harmony scores were calculated using Mori et al.'s[76] multivariate linear regression equation. The correlation coefficient between the factor score on the factor Pleasantness and the sum of the three estimated color harmony scores was 0.52. Considering the difference in methodology between their experiment and Mori et al.'s[76] experiment as well as the inherent errors in the measurement of affective effects, the authors regard the correlation coefficient as being satisfactory.

Horita's[35] study, which included two psychological experiments using the SD method, explored the possibility of estimating the affective effects of two-color combinations using the affective effects of single colors. In the first experiment, which examined the affective effects of the single colors, three factors ”Evaluation,” ”Potency,” and ”Activity” were extracted. These factors were also extracted in the second experiment that examined the affective effects of the two-color combinations. However, the second experiment extracted a fourth factor called ”Harmony.” The stimuli in the first experiment were all the 201 color samples in PCCS201-L color chart, and the stimuli in the second experiment were 174 color pairs that were composed of the stimuli in the first experiment. Twenty-four subjects Studyicipated in the first experiment, and 20 subjects Studyicipated in the second experiment. Because there was no factor in the first experiment that corresponded to the factor Harmony in the second experiment, the factor scores on the factor Harmony could only be estimated through a multivariate regression on physical attributes. The semantically significant predictive variables were lightness difference (ΔH^*_{ab}), chromatic difference (ΔC^*_{ab}), and the larger one of the two component lightness values ($\max a*$). All the physical attributes were computed using the CIE $L^*a^*b^*$ color system. ΔH^*_{ab} possesses the largest regression coefficient, ΔC^*_{ab} the second largest, and $\max a^*$ the smallest. In addition, the experimental data aligned well with Mori et al.'s[76] color harmony maps.

Oyama and Miyata (Ito)[94] conducted a two-stage psychological experiment to study the relationships between the affective effects of two-color combinations and the affective effects of single colors. At the first stage of the experiment, five seven-point evaluation scales that included ”like-dislike” were employed, and 87 Japanese college students were asked to rate a number of single-color stimuli on these scales. At the second stage of the experiment, the scale ”harmonious-disharmonious” was added, and the same subjects were asked to evaluate a number of two-color combinations on these six scales. These color combinations were composed of the color stimuli used during the first stage. For each two-color combination on each scale, its color combination effect was defined as the difference between its evaluation score and the average of the evaluation scores of its component colors. The scale ”like-dislike” was used as the surrogate of the harmony scale when single colors are concerned, because in the case of two-color combinations the correlation coefficient between these two scales was extremely high (0.919). The experimental data show that the color combination effect of the harmony scale tended to be minus, and the size of the effect increased as the hue difference between its component colors increased. Then, a multivariate liner regression model was established between the evaluation score on the harmony scale (Y) and the average evaluation score on the likeness scale ($(X_i + X_j)/2$), the hue difference (ΔH), the lightness difference (ΔV), and the chromatic difference (ΔC). The regression equation is

$$Y = \frac{X_i + X_j}{2} - 0.036\Delta H - 0.012\Delta V - 0.034\Delta C + 0.510 \tag{1.17}$$

where all the regression coefficients except that of the lightness difference (ΔV) were statistically significant. The results revealed that hue difference and/or chromatic difference negatively influence the harmony degree.

Oyama and Miyata (Ito)'s[94] experimental method was also employed in Ito's[36] experiment. Forty-one Japanese college students Studyicipated in Ito's experiment, and all the stimuli were two-color combinations. The component colors of the stimuli share the same hues, and the component colors were systematically selected from the JISZ8102 color system. For a two-color combination, if the lightness difference is in the range 1.1-3.0 and the chromatic difference is in the range 2.1-8.0, it falls in the "harmonious region." The data from the experiment show that the two-color combinations whose component colors have a high lightness difference and a low chromatic difference, or a low lightness difference and a high chromatic difference, harmonize more easily than the two-color combinations whose component colors have a high difference in both lightness and chroma. In terms of individual hues, cold colors, especially blue, can help increase the color harmony degree, while warm colors, especially orange, have the opposite effect. The results of a cross-study comparison show that the predictions by Mori et al.'s[76] model and the predictions by Moon and Spencer's[73, 74, 75] color harmony theory were at odds with Ito's[36] experimental data.

Later on, Ito and Oyama[37] conducted a similar experiment. Forty-six Japanese college students Studyicipated in the experiment, and all the stimuli were two-color combinations. The component colors of the stimuli possess different hues, and the component colors were systematically selected from the JISZ8102 color system. The results of the data analysis show that, in the general case, the hue difference between the component colors negatively influences the harmony score, and the lightness of the component colors positively affects the harmony score. A more detailed examination of the experimental results reveals that when the hue difference between the component colors is small, the two-color combinations that have high lightness and low chroma, or low lightness and high chroma, possess higher harmony degrees than the two-color combinations whose lightness and chroma were both high or both low. Regarding the individual hues, the hues R and YR tended to make a two-color combination less harmonious, while the hue Y harmonized with other hues most easily. Furthermore, the data on the most harmonious two-color combinations were consistent with the principle of natural harmony.

Tsutsui and Ohmi's[115] study aimed to clarify the variables that mediate the perception of multicolor stimuli and the affective effects of multicolor stimuli. A psychological experiment was carried out in which ten subjects rated 50 multicolor stimuli on six seven-point scales that included "pleasant-unpleasant," "dark-light," "cold-warm," and "simple-complex." Each stimulus was a 6*6 grid square. The component colors of the stimuli were systematically selected from the PCCS color system. The pleasantness scale measured the degree of beauty. The results of the data analysis show that a significantly high positive correlation exists between the pleasantness scale and the lightness scale. Further, a multivariate regression analysis using the stepwise method shows that lightness is the only predictive variable of pleasantness, which is the dependent variable. The warmness scale also bears a significantly high positive correlation with the pleasantness scale, but the warmness scale was not included in the regression equation. Contrary to Berlyne's theory, no significant correlation was found between the pleasantness scale and the complexity scale. Hence, lightness was considered to be the main variable that mediates the stimulus perception and the aesthetic evaluation. Because the psychological variable pleasantness has a relatively simple inner structure, the authors infer that, within the framework of appraisal theories, the evaluation of pleasantness may lie at the appraisal level that operates prior to the emotion level during emotion formation.

In Schloss and Palmer's[108] study, the two concepts ”color preference” and ”color harmony” were separately investigated. The experimental stimuli were 992 figure-ground two-color combinations. The component colors of the stimuli were systematically selected from the Natural Color System (NCS). Forty-eight subjects were asked to rate how much they liked each color combination and how harmonious they felt each color combination was using two evaluation scales. Then, multivariate regressions were performed to determine the variables that influence the preference scores and the harmony scores. The results show that the factors with the largest influence on the preference scores of the color combinations are the individual hues, the hue difference, and the lightness difference of the component colors. Specifically, cold colors tend to elicit higher degrees of preference than warm colors, and small hue differences and large lightness differences tend to produce high degrees of preference. On the other hand, the factors that influence color harmony are the individual hues (or, specifically, the coolness sum), the hue difference, the chroma sum, and the coolness difference. The coolness sum positively affects the harmony degree, and the hue difference, the chroma sum, and the coolness difference negatively influence the harmony degree. In addition, the preference scores and the harmony sores averaged across the subjects show a large positive correlation (Pearson's correlation coefficient = 0.79).

1.3.4 Position and Contributions of the Research

Section 1.3.3 shows that almost all studies in multicolor aesthetics using empirical approaches adopted the behaviorist paradigm. In other words, these studies aimed at extracting direct relationships between a given set of physical attributes and aesthetic evaluation of multicolor stimuli, without looking into the psychological mechanisms behind these relationships. The only study that took the cognitive paradigm is Tsutsui and Ohmi's[115]. They found that lightness was one variable that mediated perception of multicolor stimuli and aesthetic evaluation of multicolor stimuli. However, possibly due to the limited evaluation scales used in their experiment, mediating variables other than lightness, if any, could not be detected in their experiment. In addition, they only considered mediating variables that were attributes of multicolor stimuli. No variables that are attributes of component colors of multicolor stimuli had been investigated in their study.

As mentioned in Section 1.2, our research aims to clarify the perceptual features that mediate the relationships between physical color information and aesthetic evaluation of multicolor stimuli and how these mediating effects take place. In other words, our research adopts the cognitive paradigm. With respect to the experimental procedure, our research employed a large set of perception-related evaluation scales collected through a systematic literature review on color studies. Thereby, more multicolor-level perceptual features that operate during the psychological process of multicolor aesthetics could be detected. We also investigated the role of color focality, which is a perceptual feature of the component colors, in multicolor aesthetic evaluation. (The distinction between multicolor-level perceptual features and component-level perceptual features will be introduced in Section 1.6). In view of these points, our research can be expected to improve our understanding of the cognitive mechanism behind multicolor aesthetics.

1.4 The Level of Brain Function Modeling

In our view, brain function modeling can be divided into the following four levels of granularity.

The level of the finest granularity is the molecular level. Models on this level use molecules as their units. These models describe the structures and operation mechanisms of the molecules or their assemblies that take Study in neural activities, such as ion channels, hormones, and G proteins. Methods in molecular biology are used to collect data, and the models are constructed in the forms of chemical structural formulas or chemical reaction equations based on the data.

The level of the second finest granularity is the cellular level. The units, or components, of the models on this level are neurons. Models on this level depict how the activities of a neuron or neuron assembly influence the activities of another neuron or neuron assembly. Methods in electrophysiology, such as local field potential recording, single-unit recording, and multi-unit recording, are used to collect data. Models are established using differential equations based on the data.

The level of the third finest granularity is the level of brain regions. Models on this level take brain regions as units. These models describe the interactions among brain regions, namely how the activities of a brain region or an assembly of brain regions affects the activities of another brain region or assembly of brain regions. Brain-imaging techniques, such as the positron emission tomography (PET), the functional magnetic resonance imaging (fMRI), and the single-photon emission computed tomography (SPECT), are used to collect data. Models are constructed using structural equation modeling, graph theories, qualitative reasoning, and artificial neural networks. The data collected are used to train and test the models.

The level of the least granularity is the level of psychological functions. Models on this level take psychological functions as components. Each psychological function is fulfilled by one or more brain regions. These models describe the interactions among psychological variables, each of which represents a psychological function. Data are collected by conducting psychological or behavioral experiments. Using the data in model training and validation, methods such as structural equation modeling, graph theories, qualitative reasoning, and artificial neural networks are employed to build the models.

As mentioned above in Section 1.2, our research adopts Leder's cognitive psychological model of aesthetics as the theoretical framework and aims to clarify the perceptual features that operate during the psychological process of multicolor aesthetics. Because the components of the proposed model in our research, such as the perceptual features and the aesthetic evaluation, are all psychological variables, this model can be considered as a model on the fourth level, namely the level of psychological functions.

1.5 Academic and Practical Significance of the Research

1.5.1 Academic Significance

The academic significance of this research is as follows.

Based on the results of a series of psychological experiments, this research refined the composition of the Perceptual Analyses Stage in Leder's model, which is a core theory in empirical aesthetics. Thus, this research can be said to have noticeable theoretical importance in empirical aesthetics. In addition, because the sense of beauty serves as one of the central motives of artistic creation and appreciation, this research may also shed new light on art theories.

This research can also improve our understanding of the ecological functions of multicolor aesthetic evaluation as a survival tool, namely, how the probably instinctive mental function of multicolor aesthetic evaluation helps us survive in natural environments.

1.5.2 Practical Significance

This research also possesses the following practical applications.

This research offers a computational psychological model that depicts in a quantitative fashion the perceptual features that operate during the psychological process of multicolor aesthetic evaluation and how these perceptual features exert their impacts on the multicolor aesthetic evaluation. This computational model could help promote the automation in such fields as generative art, Kansei engineering, affective computing, industrial design, interior design, and human-computer interface, where multicolor aesthetics plays a vital role. As an example, this research may assist the development of image aesthetic evaluation systems that can be integrated into pattern recognition modules in affective-content-based image search engines. In addition, by employing the computational model with the direction of its inner information flow reversed in some way, we may build a system that is able to generate color images that have specific aesthetic scores[Miho Saito, personal communication, 2014].

In addition, Study 2 of this research found that color focality can influence aesthetic preferences of colors. Because the focality of a color is different in different languages[4, 46], this finding indicates that the aesthetic preference of a color is different for people speaking different languages. This theoretical inference can help to promote the efficiency of cross-cultural communications and understanding.

1.5.3 Significance in Human Sciences

As doctoral research in the area of human sciences, this research possesses the following significance in this area:

- The sense of beauty for multicolor objects is an indispensable am of human nature, which is ”a profound and unfathomable existence of great depth”[21]. It is one of the psychological foundations of various art, creation, and cultural activities. Thus, this research, which contributes to elucidating the psychological mechanism of multicolor aesthetics, is able to deepen our understanding of humanity. The results of this research can be expected to help build a better living environment by assisting the development of automatic aesthetic evaluation systems and designing systems and by improving education and the practice of art.

- Findings from the areas of psychology, information science, brain science, and so forth, are combined in this research, which means that this research has highly interdisciplinary implications, which is characteristic of studies in the human sciences.

1.6 Structure of the Research

As mentioned in Section 1.2, the goal of this research was to determine the perceptual features involved during the psychological process of multicolor aesthetic evaluation. During our research, we realized that there exist two types of such perceptual features:

Multicolor-level perceptual features

These features are properties of the multicolor stimuli themselves.

Component-level perceptual features

These features are properties of the component colors of multicolor stimuli. In other words, these features depict single colors.

Due to this categorization, this research consists of the following two studies.

Study 1 of the research (described in Chapter 2) aims to clarify the multicolor-level features involved in multicolor aesthetic evaluation and also investigates how each multicolor-level perceptual feature impacts the aesthetic evaluation. First, two psychological experiments (Experiment 1 and Experiment 2) were conducted, in which the perceptual features were extracted. Then, a neural network model was constructed and two simulations to quantitatively elucidate the influence of each perceptual feature on the aesthetic evaluation were performed. A third experiment (Experiment 3) was conducted to obtain the validation dataset in the second simulation.

Study 2 of the research (described in Chapter 3) aims to clarify whether color focality is a component-level feature that can influence color preference by investigating the continuous relationship between color focality and preference and the psychological variable(s) that mediate the relationship. Data of the focality, the preference, and possible mediating variables were obtained through two psychological experiments (Experiment 4 and Experiment 5) and were processed using regression analyses.

Chapter 4 summarizes the results of Study 1 and Study 2 and proposes three links between the component-level visual information processing and the multicolor-level visual information processing within the psychological mechanism of multicolor aesthetic evaluation. Chapter 5 introduces the implications of the results of our research for future work.

Chapter 2 Multicolor-Level Perceptual Features

This chapter introduces Study 1 of the research. Section 2.1 provides a general introduction. Section 2.2 introduces the Semantic Differential method. Section 2.3 describes the first psychological experiment in Study 1. Section 2.4 describes the second psychological experiment in Study 1. Section 2.5 explicates how the computational model is constructed. Section 2.6 describes the first simulation of the model. Section 2.7 describes the second simulation of the model. Section 2.8 summarizes the results of Study 1.

Study 1 is an extension of Siyuan Fang's master's thesis *Extraction of Aesthetic Rules to Multi-Color Stimuli Using Artificial Intelligence Technology: Towards the Construction of an Artificial KANSEI System*. Because the data collection and the macro-architecture construction of the computational model have been described in detail in the master's thesis, this doctoral thesis only provides a brief description of them.

2.1 Introduction

Study 1 aims to clarify what multi-color perceptual features can influence multi-color aesthetic evaluation and how these perceptual features exert their influences. Firstly, two psychological experiments using the SD method were conducted. The first experiment (called "Experiment 1" in this research) aims to quantify the aesthetic evaluation of the multicolor stimuli. The second experiment (called "Experiment 2" in this research) is intended to extract and quantify the multicolor-level perceptual features into which the primary color information is transformed during the psychological process of multicolor aesthetic evaluation. Then, we constructed the macro-architecture of the computational model based on the results of the two psychological experiments. Finally, we conducted two test simulations (Simulation 1 and Simulation 2) of the model, especially Simulation 2, to delve into how each multicolor-level perceptual feature influences the aesthetic evaluation. The empirical data obtained in Experiment 1 and Experiment 2 were used in model training and validation. A third experiment (called ”Experiment 3” in this research) was carried out to collect the validation data in Simulation 2.

2.2 The SD Method

The SD method was proposed by American linguist Charles Osgood in 1957. It was initially applied in social psychology and personality psychology to measure people’s attitudes toward political parties, companies, or industrial products, etc. Later, psychologists began to use it to measure people’s affective responses to perceptual stimuli such as color, shapes and music[20, 93].

The overall procedure of the SD method is composed of two steps: the experimental step and the factor analysis step. The preparation prior to the experimental step is to design an answer sheet with a number of pairs of adjective antonyms on it, for example, ”heavy-light,” ”strong-weak” and ”cold-hot.” These adjective pair scales can be five-point or seven-point. The experimenters must choose the adjective pair scales carefully to make sure that they are suitable for the evaluation of the particular stimulus type. When such adjective pair scales can be found in literature, the experimenters can utilize them directly. If no such adjective pair scale can be found in literature, the experimenters have to create them by themselves. One way to create such scales is to gather adjectives from books, magazines, or other media relating to the research objects. Another way is to conduct interviews in which participants are required to provide adjectives pertaining to the stimulus type. These two ways can be employed together.

During the experiment, the stimuli are shown, once a time, to the subjects. The task of the subjects is to evaluate the stimuli one at a time on each adjective pair scale. After the completion of the experiment, the evaluation data are averaged across the subjects and processed using factor analysis by which the underlying factors, if any, can be extracted based on the correlations among the original evaluation variables, namely, the adjective pair scales. Finally, each extracted factor is named according to its factorial structure and domain knowledge. The most frequently extracted factors in the literature of color studies are ”Evaluation,” ”Activity” and ”Potency”[20, 93].

In the present research, we chose to use the SD method because the sense of beauty toward multi-color stimuli is an extremely delicate feeling which we can hardly verbalize. In addition, the three factors, Evaluation, Activity, and Potency, have not only statistical meanings, but also neurological meanings, or in other words, biological reality[44].

2.3 Experiment 1

2.3.1 Objective

[1]Experiment 1 aims to quantify the result of the aesthetic evaluation of each multi-color stimulus. It provides parts of the empirical data used in model training and testing in the model simulations.

2.3.2 Experimental Stimuli and Environment

Many studies in empirical aesthetics used real artworks, for example, drawings, photographs and sculptures, as experimental stimuli. However, the artistry, namely the artistic status, of the stimuli itself is reported to be able to elicit a particular pattern of brain activities[54]. This

[1]Section 2.3 is a brief version of Chapter 4 in Siyuan Fang’s master’s thesis (pp. 39-73). Please refer to the latter for detailed information. The content of this section was published in Fang, Muramatsu and Matsui’s 2017 research paper[18].

means that the neural and psychological effects caused by artistry may interfere with those caused by pure color information. To reduce this confounding effect to the minimum, we chose to use computer-generated multicolor square grids as experimental stimuli.

Because most of the previous studies on affective effects of color combinations used two-color or three-color combinations, in Study 1, we divided to use multicolor stimuli consisting of a larger number of colors. In our experiments, each stimulus was constituted by 16 $(4\ times4)$ squares whose colors were randomly determined. The size of each stimulus was 400 pixels*400 pixels (width*height). Since all component color patches were squares of the same area, the area effect on the aesthetic evaluation was eliminated. This image configuration was proposed by Matsuura and Matsui[65], and later used in a series of researches by Ogawa, Muramatsu and Matsui[82, 83, 84, 85, 86].

Thirty-five stimuli were used in Experiment 1. Figure 2.1 shows one example of them. The RGB and $L^*a^*b^*$ values of the component colors of the stimuli are listed in Appendix A.

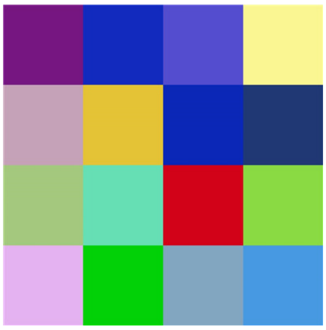

Figure 2.1: One example of the multicolor stimuli used in Experiment 1.

The stimuli were presented using Microsoft PowerPoint. Each stimulus was displayed in the middle of a slide with black background. The index of the stimulus was shown in the top-left corner of the slide to prevent the subjects from filling in a wrong answer sheet. No time limitation was imposed on the subjects' ratings, but the subjects were asked to do the ratings by intuition or, in other words, without deliberation. Reviewing previous ratings was also prohibited. Between every two stimulus-displaying slides, a black slide was presented for 30 s to remove afterimages. An instruction ”Please watch this black screen” was displayed in white in the middle of the black slide throughout the 30-s interval.

The experiment was performed indoors (in Room 213 in Building No.110 at Tokorozawa Campus of Waseda University) with fluorescent lighting (type: National FHF 32EX-N-H, daylight color, color temperature: 5000K). The illuminance was 760lx measured by a HIOKI FT3424 lux meter (JIS C 16091 General Class AA, DIN 5032-7 Class B). The slides were run on a HP EliteBook 8460w PC and projected through a NEC NP64 projector onto a projection screen (refresh rate: 60hz; resolution: 1024 * 768). Each multicolor stimulus was projected onto the screen with size 52.5cm * 52.5cm.

Figure 2.2 shows a sketch of the experiment environment. The projector was placed at the center of the table and the screen was set at a short side of the table. The subjects were sitting at the other three sides of the table. When watching the multicolor stimuli, the subject nearest to the screen had a viewing angle of approximately 12.46°, and the subject farthest from the screen had a viewing angle of approximately 8.32°. The environment was kept quiet throughout the experiment.

2.3.3 Subjects and Adjective Pair Scales

In order to avoid the fatigue effect, this experiment was divided into two sessions which were independently carried out. Eight subjects (six males and two females, 20 to 22 years of age), all of whom were undergraduate students of Waseda University, participated in the first session. Stimuli No. 1-20 were used in this session. In the second session, there were 12 subjects (eight males and four females, 20 to 24 years of age), who were either undergraduate or graduate students of Waseda University. Stimuli No. 21-35 were used in this session. We obtained the informed consents of participation from all the subjects.

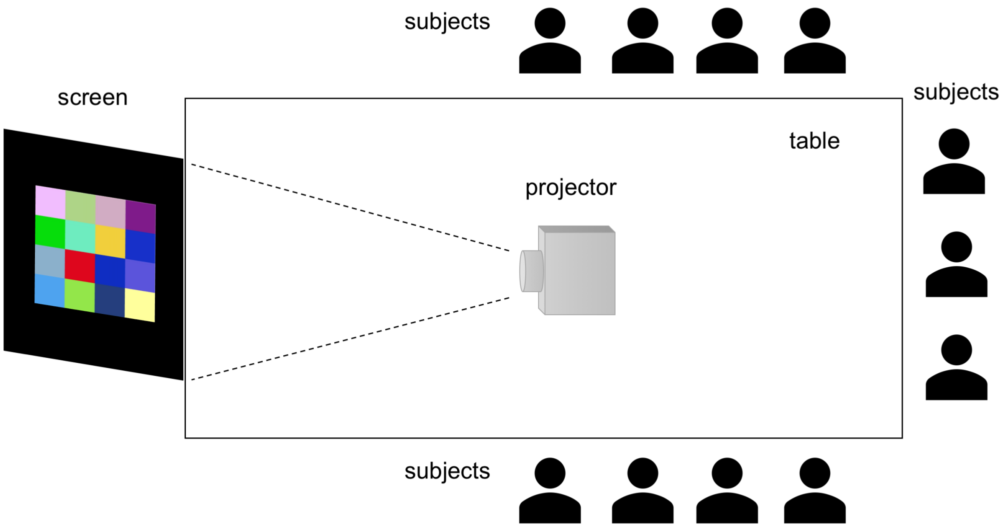

Figure 2.2: A sketch of the environment of Experiment 1.

We gathered 24 adjective pair scales from several past studies pertaining to affective effects of color. These scales were employed in both sessions and were presented to the subjects in form of seven-point rating scales. The psychological experiments in this research were all conducted in Japanese, so we translated the scales into English in writing this thesis. The Japanese version of the scales is displayed in Appendix B.

2.3.4 Data Processing by Factor Analysis

As mentioned in Section 2.3.3, each adjective pair scale was divided into seven degrees. Prior to the factor analysis, we quantified these seven degrees by transforming them into the integers ”-3,” ”-2,” ”-1,” ”0,” ”1,” ”2,” and ”3” from the left end to the right end. The evaluation data for each stimulus were averaged across the subjects on each scale. When there were missing values in a subject’s (or multiple subjects’) evaluation data for a certain stimulus, the evaluation data provided by the other subjects for the stimulus were averaged. The averaged data are imported into IBM SPSS Statistics (version 19) for factor analysis.

The initial factor loadings were calculated using the principal factor method (called ”Princi-

Table 2.1: Communalities of adjective pair scales (Experiment 1).

Adjective Pair Scale	Communality	Adjective Pair Scale	Communality
heavy-light	.780	graceful-awkward	.734
light-dark	.883	clear-dull	.764
warm-cool	.884	static-dynamic	.910
soft-hard	.842	true-false	.658
noisy-quiet	.895	novel-ordinary	.614
like-dislike	.778	beautiful-ugly	.848
ornate-plain	.876	stable-changeable	.714
strong-weak	.828	successful-unsuccessful	.798
pleasant-unpleasant	.847	positive-negative	.799
clean-dirty	.866	relaxed-nervous	.581
harmonious-dissonant	.796	cruel-kind	.848
cheerful-gloomy	.913	passive-active	.867

ple Components" method in SPSS). The communalities of the adjective pair scales were listed in Table 2.1. Because the communality of each adjective pair scale was greater than 0.5, there was no adjective pair scale which should be deleted before entering the next data processing step.

The scree plot (shown in Figure 2.3) exhibits the eigenvalue of each factor. There were three factors whose eigenvalues were greater than 1.0. Thus, these three factors were selected as main factors.

Then, these factors were rotated using the varimax method which preserved the orthogonality among these factors. The varimax method was chosen because it was extensively employed in the literature from which the adjective pair scales used in this experiment were collected. The variances on the original response variables, namely, the adjective pair scales, which were explained by the three factors before and after the rotation operation are summarized in Table 2.2.

Table 2.3 shows the factor loading matrix that resulted from the rotation. All factor loadings whose absolutes are greater than 0.5 are emphasized in bold. For easier reading, the variance

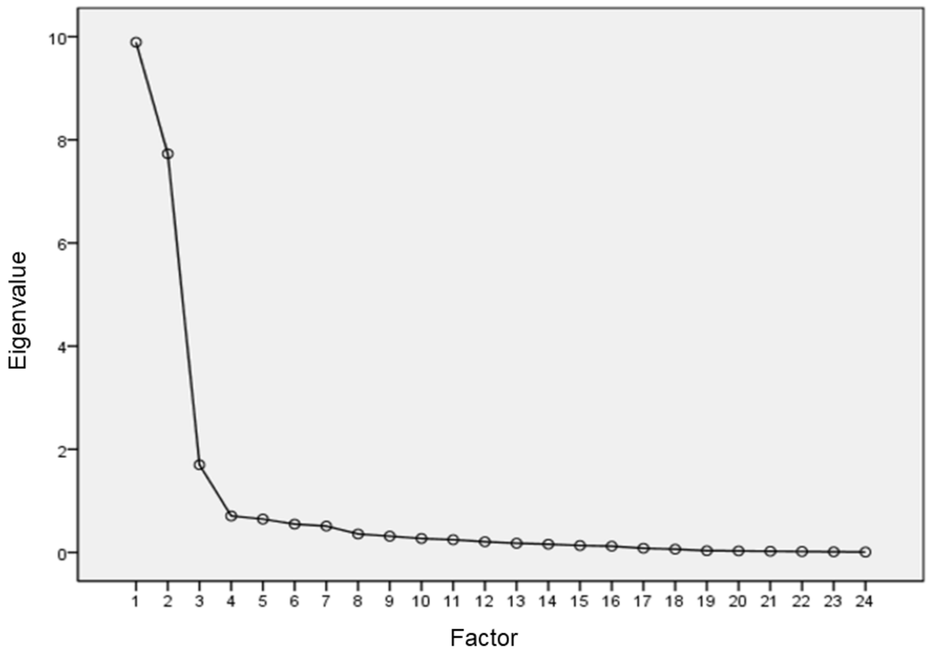

Figure 2.3: Scree plot of Experiment 1.

Table 2.2: Total variance explained (Experiment 1).

Factors	Variance explained before rotation		Variance explained after rotation	
	Individual (%)	Cumulative (%)	Individual (%)	Cumulative (%)
1	41.218	41.218	33.358	33.358
2	32.211	73.430	32.449	65.807
3	7.078	80.508	14.701	80.508

explained by each factor after the rotation, shown in the fourth column in Table 2.2, is duplicated in the final row of Table 2.3.

Finally, the regression method is utilized to calculate the factor scores of each stimulus on the three main factors. The results are listed in Table 2.4.

2.3.5 Results

Each main factor was interpreted based on its factorial structure in the post-rotation factor loading matrix.

The adjective pair scales with the highest loadings on Factor 1, disregarding the plus-minus signs, include ”static-dynamic,” ”passive-active,” ”noisy-quiet,” and ”positive-negative”; thus, we can infer that this factor evaluates the kinetic state of the stimuli. For Factor 2, its representative adjective pair scales include ”beautiful-ugly,” ”clean-dirty,” ”pleasant-unpleasant,” and ”harmonious-dissonant”; thus, this factor evaluates the ”goodness” attribute of the stimuli, or in other words, reveals the evaluator’s preference for the stimuli. The adjective pair scales belonging to Factor 3 are ”soft-hard,” ”strong-weak,” ”relaxed-nervous,” and ”heavy-light.” This implies that the factor measures the degree of strength of the stimuli. Referring to the conventional nomenclature, we can call the three factors ”Activity,” ”Evaluation,” and ”Potency,” respectively. Nonetheless, considering that Osgood[88, 89] interpreted the psychological meaning of the factor Evaluation as gauging the extent of pleasure elicited by a positive judgment of the stimuli in the surroundings, we decided to name Factor 2 ”Pleasure.”

In this experiment, this factor represents, in the fashion of inverse, the extent of pleasure felt by the subjects for the multicolor stimuli. Hence, we define the results of the aesthetic evaluation of the stimuli as the inverses of their factor scores on the factor Pleasure. In this manner, the quantitative concept ”aesthetic evaluation” in Leder’s model is quantified.

In the results of a few other studies on color emotion using the SD method[35, 41, 91],

Table 2.3: Rotated factor loading matrix (Experiment 1).

Factor identity	Adjective pair scale	Factor		
		1	2	3
Factor 1 (Activity)	static-dynamic	**-.948**	.065	-.088
	ornate-plain	**.929**	-.051	.103
	passive-active	**-.918**	-.016	-.157
	noisy-quiet	**.890**	-.285	.148
	positive-negative	**.886**	.058	.098
	cheerful-gloomy	**.834**	.258	.388
	warm-cool	**.745**	.010	**.574**
	novel-ordinary	**.731**	-.278	.038
	light-dark	**.692**	.368	**.518**
	cruel-kind	**-.678**	-.261	**-.566**
Factor 2 (Pleasure)	beautiful-ugly	.103	**.915**	.000
	clean-dirty	.106	**.904**	.191
	pleasant-unpleasant	.135	**.877**	.244
	successful-unsuccessful	.043	**.866**	.213
	harmonious-dissonant	-.230	**.862**	-.010
	clear-dull	.089	**.855**	.156
	graceful-awkward	-.267	**.806**	.113
	true-false	-.104	**.794**	.132
	like-dislike	.175	**.761**	.409
	stable-changeable	-.396	**.723**	.184
Factor 3 (Potency)	soft-hard	.328	.210	**.831**
	relaxed-nervous	.215	.282	**.675**
	strong-weak	**.627**	-.116	**-.649**
	heavy-light	-.425	-.424	**-.648**
Variance explained (%)		33.358	32.449	14.701

Table 2.4: Factor scores (Experiment 1).

Stimulus no.	Factor			Stimulus no.	Factor		
	1 (Ac)	2 (Pl)	3 (Po)		1 (Ac)	2 (Pl)	3 (Po)
1	0.55	-0.23	0.00	19	-0.47	-2.28	-0.97
2	-0.87	0.62	0.71	20	-0.59	1.04	0.87
3	1.80	0.02	-0.32	21	0.94	0.52	0.02
4	0.49	-0.34	-2.42	22	-1.28	-0.41	-0.59
5	-0.60	2.10	0.30	23	0.33	-0.27	0.25
6	-1.71	0.34	0.18	24	-1.58	-1.48	-1.01
7	1.19	0.71	-0.93	25	1.53	0.86	-0.28
8	0.24	0.75	-0.34	26	-0.83	0.18	-0.70
9	0.45	-1.64	2.81	27	0.40	-0.12	-0.87
10	-1.06	1.44	0.46	28	1.51	-1.18	-0.71
11	0.61	-0.47	2.45	29	-0.70	-1.60	-0.98
12	-0.63	1.18	0.93	30	-0.27	-0.46	-0.53
13	1.15	0.92	-1.08	31	0.72	-1.63	1.42
14	-1.04	0.17	0.34	32	-0.65	-0.74	-0.10
15	-0.09	0.36	0.41	33	0.21	1.29	0.13
16	-0.23	1.13	-0.76	34	2.04	-0.18	0.88
17	0.18	0.16	-0.29	35	-1.80	-0.44	0.57
18	0.04	-0.33	0.16				

regarding the common adjective pairs used, most of these adjective pairs are categorized into the same factors as in the present experiment. Hence, it is cogent to argue that the factors extracted in this experiment possess a high degree of psychological reality.

2.4 Experiment 2

2.4.1 objective

[2]Experiment 2 aims to clarify the multicolor-level perceptual features into which the primary color information is transformed during the psychological process of multicolor aesthetic evaluation. To be sure, Leder[55] proposed several perceptual features in the Perceptual Analyses Stage which possibly could influence aesthetic evaluation of multicolor stimuli, such as complexity, contrast, symmetry, color and grouping, and the influence of these perceptual features on multicolor aesthetic evaluation was supported by a few past studies. However, the experimental objectives of these past studies are obviously distinct from the objective of the present research. Thus, their results can hardly be applied to the present research. Therefore, it is necessary to conduct a psychological experiment to specify and quantify the multicolor-level perceptual features engaged in multicolor aesthetic evaluation.

2.4.2 Subjects, Adjective Pair Scales, Stimuli and Environment

Experiment 2 and Experiment 1 were performed in the same room. The experimental environment of Experiment 2, including the lighting condition and the locations of the subjects, the projector and the screen, was the same as that in Experiment 1. The stimuli used in Experiment 2 were those used in Experiment 1, and the stimuli were presented to the subjects in the same manner as in Experiment 1. The environment was kept quiet throughout the experiment.

[2]The content of Section 2.4, except that of the last subsection (Section 2.4.4), is a brief version of Section 5 in Siyuan Fang's master's thesis (pp. 73-109). Please refer to the latter for detailed information. The content of this section was published in Fang, Muramatsu and Matsui's 2017 research paper[18].

Like in Experiment 1, Experiment 2 also consisted of two sessions which were independently carried out. A total of 15 subjects (11 males and four females, 20 to 24 years of age), all of whom were undergraduate students of Waseda University, took part in the first session. Stimuli No.1-20 were used in this session. In the second session, there were 12 subjects (eight males and four females, ages ranging from 20 to 24) who were undergraduate students or graduate students of Waseda University. Stimuli No. 21-35 were used in this session. We obtained the informed consents of participation from all the subjects.

With regard to the adjective pair scales, because the objective of this experiment is to specify the multicolor-level perceptual features, only the adjective pair scales which described the perceptual aspects of color stimuli were chosen. This means that the set of adjective pair sales used in this experiment were much different from that used in Experiment 1, since the set of adjective pair scales used in Experiment 1 included some evaluative ones.

No past research has studied the multicolor-level perceptual features functioning in the psychological mechanism of multicolor aesthetic evaluation, so no list of adjective pair scales exists for this purpose. Hence, we created such a list in the following manner. First of all, we gathered 266 adjective pairs from 60 papers and books of color studies. Then, based on the three standards: 1) whether the adjective pair described a perceptual aspect of multicolor stimuli, 2) how many times the adjective pair has been used, and 3) whether two or more adjective pair scales possessed the same meaning, 45 adjective pair scales were selected. When determining the number of the adjective pair scales used in the experiment, we also considered that the duration of the experiment could not be too long so that the subjects might become fatigued. During the experiment, these scales were presented to the subjects in form of seven-point rating scales.

After the completion of the experiment, for each experimental session, the Cronbach's α coefficient of each adjective pair scale was calculated. Then, the two Cronbach's α coefficients of each adjective pair were averaged. There were 25 adjective pair scales whose average Cronbach's α coefficients were higher than 0.60. This means that these scales possessed relatively high subject-wise consistency. Then, the evaluation data on these scales were imported into IBM SPSS Statistics (Version 19) for factor analysis.

2.4.3 Data Processing by Factor Analysis

As in the data processing procedure in Experiment 1, the seven degrees of each adjective pair scale were transformed into the integers "-3," "-2," "-1," "0," "1," "2," and "3" from the left end to the right end prior to the factor analysis. The evaluation data for each stimulus were averaged across the subjects on each scale. When there were missing values in a subject's (or multiple subjects') evaluation data for a certain stimulus, the evaluation values provided by the other subjects for the stimulus were averaged.

The initial factor loadings were calculated using the principal factor method (called "Principle Components" method in SPSS). The communalities of the original response variables, namely, the adjective pair scales, are shown in Table 2.5. The communality of each original response variable in Table 2.5 was larger than 0.6, so there was no original response variable which should be deleted before entering the following computational steps due to low communality.

The scree plot (shown in Figure 2.4) shows the eigenvalue of each factor. There were three factors whose eigenvalues were larger than 1.0. These three factors were selected as main factors.

Then, the factors were rotated using the promax method which allowed oblique crossing of the rotated factors[3]. The resulted pattern matrix is shown in Table 2.6, and the resulted structure matrix is shown in Table 2.7. All factor loadings whose absolutes are greater than 0.5

[3]Due to the presence of the inter-factor correlations resulting from the promax rotation, the variance explained by the rotated factors could not be computed.

Table 2.5: Communalities of adjective pair scales (Experiment 2).

Adjective pair scale	Communality	Adjective pair scale	Communality
cold-warm	.881	clear-vague	.851
light-heavy	.863	diversified-monotonous	.801
light-dark	.934	strong-weak	.822
noisy-quiet	.875	conspicuous-inconspicuous	.902
soft-hard	.805	gaudy-plain	.736
shallow-deep	.858	neat-disordered	.804
sweet-unsweet	.639	plain-thick	.722
delicious-bad-tasting	.859	leisurely-bustling	.860
dark-pale	.861	calm-restless	.847
wet-dry	.749	plain-ornate	.909
clear-dull	.843	calm-violent	.708
vivid-subdued	.910	distinct-indistinct	.789
dynamic-static	.907		

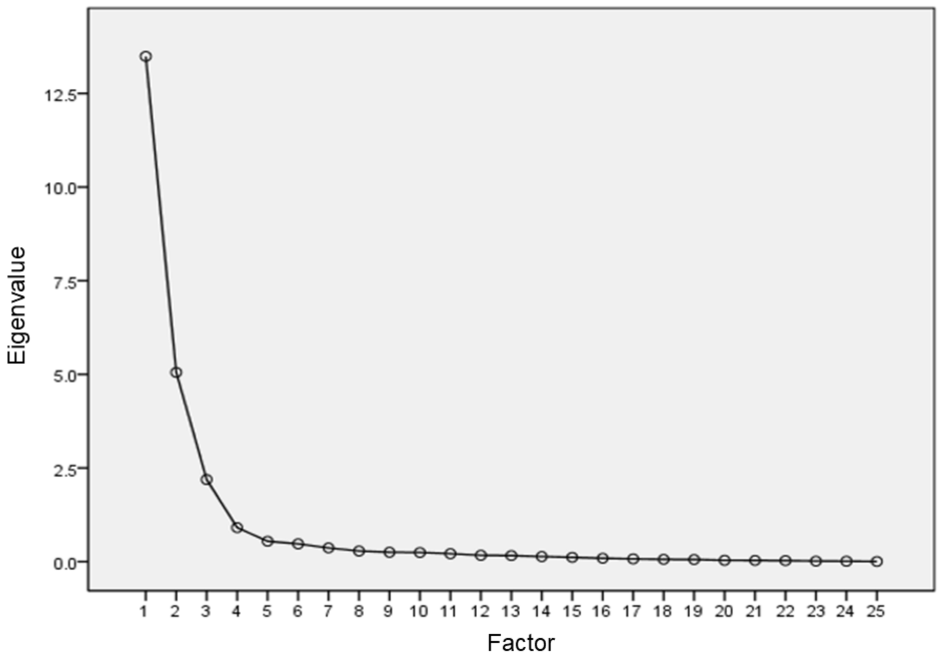

Figure 2.4: Scree plot of Experiment 2.

are emphasized in bold.

The correlation coefficient matrix that demonstrates the pair-wise correlations among the three main factors after the factorial rotation is shown in Table 2.8.

The regression method was utilized to calculate the factor scores of the stimuli on the three main factors. The result is shown in Table 2.9.

2.4.4 Results

As can be seen in the pattern matrix and the structure matrix, three main factors were extracted. We interpreted each main factor by examining its factorial structure.

Adjective pair scales such as ”dynamic-static,” ”leisurely-bustling,” ”noisy-quiet,” ”calm-restless,” ”diversified-monotonous” and ”calm-violent,” disregarding the plus-minus signs, have their highest loadings on Factor 1, which implies that stimuli with high factor scores on this factor bear a stable and static appearance. Hence, we call this factor ”Stability.” Adjective pair scales ”clear-dull,” ”plain-thick,” and ”thick-thin,” irrespective of the plus-minus signs, have their highest loadings on Factor 2, indicating that stimuli with high factor scores on this factor are of dull, turbid, or thick color, which tends to elicit a feeling of heaviness. Thus, we named this factor ”Heaviness.” The adjective pair scales, ”dark-pale,” ”shallow-deep,” ”soft-hard,” and ”light-heavy,” disregarding the plus-minus signs, have their highest loadings on Factor 3. For ”dark-pale,” an object with a dark appearance—e.g., the black bathtub in Jacques Louis David’s painting *La Mort de Marat* and the gloomy oceans in the paintings by Ivan K. Aivazovsky—usually evince a strong sense of matter presence. With regard to ”shallow-deep,” it is naturally used to describe the amount of water presence. With regard to the adjective pair scale ”soft-hard,” soft objects, such as cotton and foam, are normally of low density, whereas hard objects, such as a hammer or rock, are usually of high density. In view of this, we think that the measurement of the sense of presence is the core common character shared by these adjective

Table 2.6: Pattern matrix (Experiment 2).

Factor identity	Adjective pair scale	Factor		
		1	2	3
Factor 1 (Stability)	conspicuous-inconspicuous	**.979**	.063	-.123
	plain-ornate	**-.959**	-.028	.029
	leisurely-bustling	**-.950**	.088	.022
	dynamic-static	**.934**	-.086	.080
	clear-vague	**.925**	.208	-.179
	strong-weak	**.917**	-.124	**-.656**
	diversified-monotonous	**.897**	-.264	.047
	noisy-quiet	**.872**	-.144	.191
	calm-restless	**-.821**	.432	-.240
	calm-violent	**-.773**	**.545**	.076
	vivid-subdued	**.762**	.485	-.067
	distinct-indistinct	**.752**	.483	-.215
	wet-dry	**-.722**	.049	-.290
	cold-warm	**-.678**	.200	-.497
	light-dark	**.616**	.117	.480
	sweet-unsweet	.461	.334	.277
Factor 2 (Heaviness)	neat-disordered	-.170	**.945**	-.103
	plain-thick	-.219	**.802**	.172
	gaudy-plain	.341	**-.795**	-.192
	clear-dull	.385	**.752**	.010
	delicious-bad-tasting	**.506**	**.586**	.136
Factor 3 (Presence)	dark-pale	.286	.028	**-1.014**
	soft-hard	-.001	.156	**.826**
	shallow-deep	.280	.046	**.759**
	light-heavy	.345	.205	**.632**

Table 2.7: Structure matrix (Experiment 2).

Factor identity	Adjective pair scale	Factor		
		1	2	3
Factor1 (Stability)	plain-ornate	**-.953**	-.219	-.364
	dynamic-static	**.948**	.141	.419
	conspicuous-inconspicuous	**.944**	.222	.292
	leisurely-bustling	**-.922**	-.103	-.323
	noisy-quiet	**.917**	.113	.483
	clear-vague	**.897**	.334	.269
	diversified-monotonous	**.860**	-.057	.303
	vivid-subdued	**.837**	**.620**	.423
	cold-warm	**-.834**	-.134	**-.690**
	light-dark	**.832**	.431	**.771**
	wet-dry	**-.828**	-.215	**-.559**
	calm-restless	**-.826**	.167	-.401
	distinct-indistinct	**.768**	**.559**	.270
	sweet-unsweet	**.642**	**.538**	**.589**
	strong-weak	**.630**	-.182	-.337
	calm-violent	**-.628**	.411	-.023
Factor 2 (Heaviness)	neat-disordered	-.012	**.869**	.192
	clear-dull	**.547**	**.836**	.451
	plain-thick	.018	**.822**	.392
	thick-thin	.097	**-.797**	-.361
	delicious-bad-tasting	**.683**	**.745**	**.563**
Factor 3 (Presence)	dark-pale	-.112	-.301	**-.889**
	shallow-deep	**.592**	.396	**.888**
	soft-hard	.361	.473	**.885**
	light-heavy	**.640**	**.520**	**.848**

Table 2.8: Correlation coefficient matrix of rotated factors (Experiment 2).

	Factor 1	Factor 2	Factor 3
Factor 1	1.000		
Factor 2	.211	1.000	
Factor 3	.399	.384	1.000

Table 2.9: Factor scores (Experiment 2).

Stimulus	Factor			Stimulus	Factor		
no.	1 (S)	2 (H)	3 (P)	no.	1 (S)	2 (H)	3 (P)
1	-0.01	-0.50	0.41	19	-0.28	-1.15	-1.60
2	-0.55	0.61	1.06	20	-0.63	1.12	0.49
3	2.29	-0.10	0.15	21	1.44	1.31	2.13
4	-0.63	-2.08	-1.58	22	-1.57	-0.82	-0.33
5	-0.70	1.73	0.03	23	0.63	0.81	0.62
6	-0.25	-0.96	-0.68	24	-2.39	-1.18	-1.42
7	0.94	-0.04	-1.13	25	1.25	1.00	-0.41
8	0.77	1.04	0.47	26	-0.88	-0.20	-0.11
9	1.55	-1.45	2.39	27	0.37	1.00	-0.82
10	0.05	1.87	0.37	28	0.87	-1.17	-1.20
11	0.91	-0.21	2.38	29	-1.17	-1.42	-1.29
12	-0.65	0.01	0.16	30	-0.51	-0.74	-1.09
13	0.19	0.56	-0.70	31	0.24	-1.43	-0.19
14	-0.67	0.94	0.49	32	-0.90	0.23	-0.03
15	0.54	0.73	0.24	33	0.09	0.95	0.28
16	0.74	0.54	0.06	34	1.33	-0.21	0.51
17	-0.40	-0.60	-0.48	35	-1.36	-0.20	0.40
18	-0.66	0.01	0.41				

pair scales, and thus ”Presence” is a suitable name for this factor.

We compared our results with those of Gao’s[22, 23] experiments, which investigated the physical visual features engaged in mono-color affective evaluation. The factorial structures of the adjective pairs used in both Gao’s experiments and the present experiment showed considerable inter-research similarity, suggesting a high degree of psychological reality of the factors extracted in the present experiment.

In summary, totally three main factors, namely ”Stability,” ”Heaviness” and ”Presence,” were extracted in the present experiment. Each of them is regarded as corresponding to a multicolor-level perceptual feature into which the primary color information is transformed during the psychological process of multicolor aesthetic evaluation. In other words, these three main factors form the Perceptual Analyses Stage in Leder’s psychological model described in Section 1.2. The values of each stimulus on the three multicolor-level perceptual features are defined as the factor scores of the stimulus on the three factors. In this manner, the multicolor-level perceptual features located in the Perceptual Analyses Stage in Leder’s psychological model are specified and quantified.

2.5 Construction of Computational Model

2.5.1 Macro-Architecture of Computational Model

[4]As mentioned above, based on the results of the two psychological experiments, we defined the concept ”aesthetic evaluation” as the factor Pleasure extracted in Experiment 1, and we defined the multicolor-level perceptual features located in the Perceptual Analysis Stage as the three factors Stability, Heaviness, and Presence extracted in Experiment 2. This refined version of Leder’s model is now composed of three levels: primary color information as the first level,

[4]Section 2.5 is a brief revised version of Section 6 and Section 7 in Siyuan Fang’s master’s thesis (pp. 112-154). The content of this section was published in Fang, Muramatsu and Matsui’s 2017 research paper[18].

multicolor-level perceptual features as the second level, and aesthetic evaluation as the final level.

It is considered that the mapping from the first to the second levels and that from the second to the third levels are so complicated that they can be approximated only through nonlinear functions. We chose to use three-layer backpropagation neural networks (BPNNs), a ”consistent thread”[24] in methodology of computational aesthetics, to implement the inter-level mappings, because a three-layer BPNN is able to approximate any continuous nonlinear function[31, 40, 68]. Thus, the computational model is a hierarchical BPNN system with the macro-architecture shown in Figure 2.5.

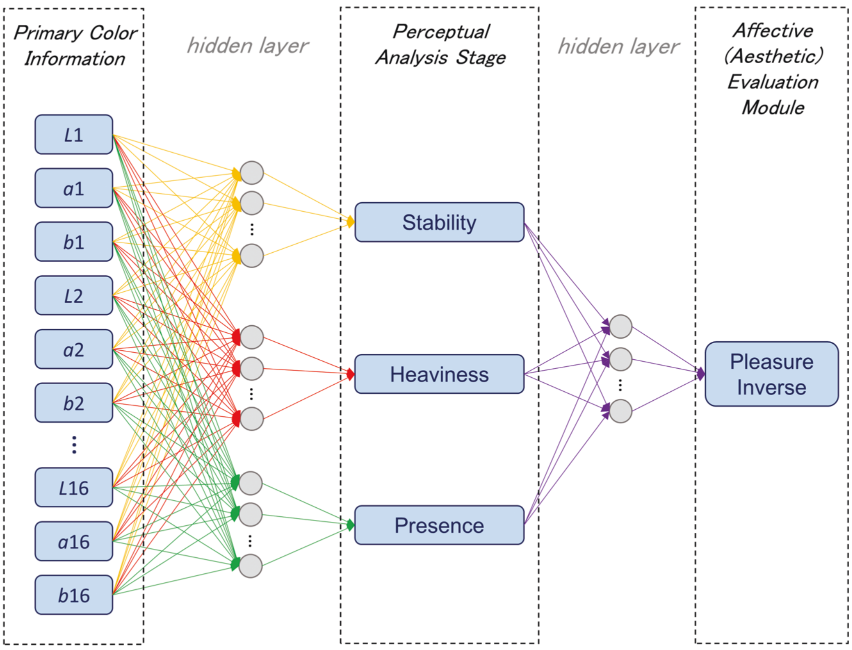

Figure 2.5: The macro-architecture of the computational model. The orange inter-node nexuses form BPNN 1. The red inter-node nexuses form BPNN 2. The green inter-node nexuses form BPNN 3. The purple inter-node nexuses form BPNN 4.

As demonstrated in Figure 2.5, the computational model consists of two levels of BPNNs. The first level contains three BPNNs working in parallel. Given a certain multicolor stimulus, the first BPNN (BPNN 1) transforms the primary color information of the stimulus, defined as the $L^*a^*b^*$ values (the 1976 CIE $L^*a^*b^*$ color space) of the component colors of the stimulus, into the perceptual feature, Stability. BPNN 1 has 48 input nodes corresponding to the $L^*a^*b^*$ values of every component color, and its sole output node represents the factor score of the stimulus on the factor Stability. The second BPNN (BPNN 2) maps the primary color information of the stimulus into the perceptual feature Heaviness. The input nodes of BPNN 2 are the same as those of BPNN 1, and the sole output node of BPNN 2 represents the factor score of the stimulus on the factor Heaviness. The function of the third BPNN (BPNN 3) is to transform the primary color information of the stimulus into the perceptual feature Presence. BPNN 3 shares the same input nodes as BPNN 1 and BPNN 2, but its sole output node represents the factor score of the stimulus on the factor Presence. The fourth BPNN (BPNN 4 forms the second level of the computational model. BPNN 4 maps the three perceptual features, Stability, Heaviness, and Presence, of the stimulus to the variable ”Pleasure Inverse,” which represents the inverse of the aesthetic score of the stimulus. The output nodes of BPNN 1, BPNN 2, and BPNN 3 serve as the input nodes of BPNN 4, and the sole output node of BPNN 4 represents the factor score of the stimulus on the factor Pleasure.

To speed up the convergence of the BPNNs while restraining the instability of their learning processes, we adopt a generalized delta rule which contains a momentum constant in the learning algorithms of the BPNNs[31]. The generalized delta rule is

$$\Delta w_{ji}(n) = \alpha \Delta w_{ji}(n-1) + \eta \delta_j(n) y_i(n) \tag{2.1}$$

where $\Delta w_{ji}(n)$ is the correction to the synaptic weight between the node i to the node j in the nth iteration, $\delta_j(n)$ the local gradient of the node i in the nth iteration, $y_i(n)$ the output value of the node i in the nth iteration, η the learning rate and α the momentum constant. Moreover, the activation function of the nodes in this computational model is the $tansig$ function described below.

$$f(x) = \frac{2}{1 + \exp(-2x)} - 1 \tag{2.2}$$

2.5.2 Hyperparameter Optimization Using Genetic Algorithm

Having determined the macro-architecture of the computational model, the next step is to find for every constituent BPNN the set of hyperparameter values that produce the best performance, because performance of BPNNs are sensitive to their hyperparameter values[105]. The system parameters to be optimized are the number of nodes in the hidden layer, the learning rate, and the momentum constant. The genetic algorithm (GA) is employed to carry out the optimizations. Each BPNN is optimized independently by the procedure shown in Figure 2.6.

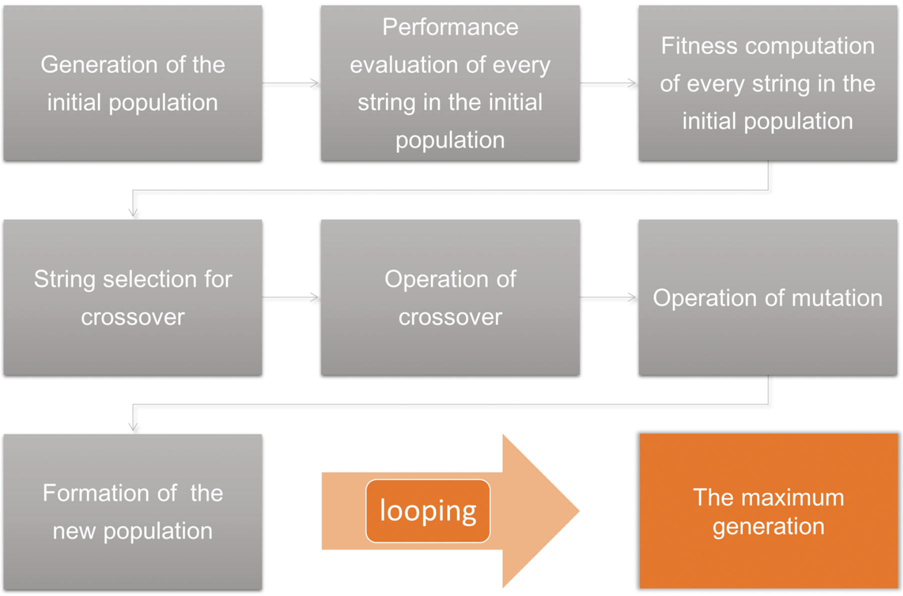

Figure 2.6: The procedure of the GA.

The first step is the random generation of the initial string population. Each string encodes a set of values of the three hyperparameters. The Gray code is adopted because the Hamming distance between two strings encoding adjacent values using this coding method remains the

Table 2.10: Hyperparameters of the GA.

System parameter	Bit length	Range	Step length	BPNN no.
Number of hidden layer nodes	3	{4, 8, 16, 24, 32, 48, 64, 72}	unequal	1, 2, 3
	2	{1, 2, 3, 4}	1	4
Learning rate	3	[1/9, 8/9]	1/9	1, 2, 3, 4
Momentum constant	2	[0.2, 0.8]	0.2	1, 2, 3, 4

same, thus avoiding an inherent coding bias[10]. The values of the coding operation parameters associated with each BPNN are displayed in Table 2.10.

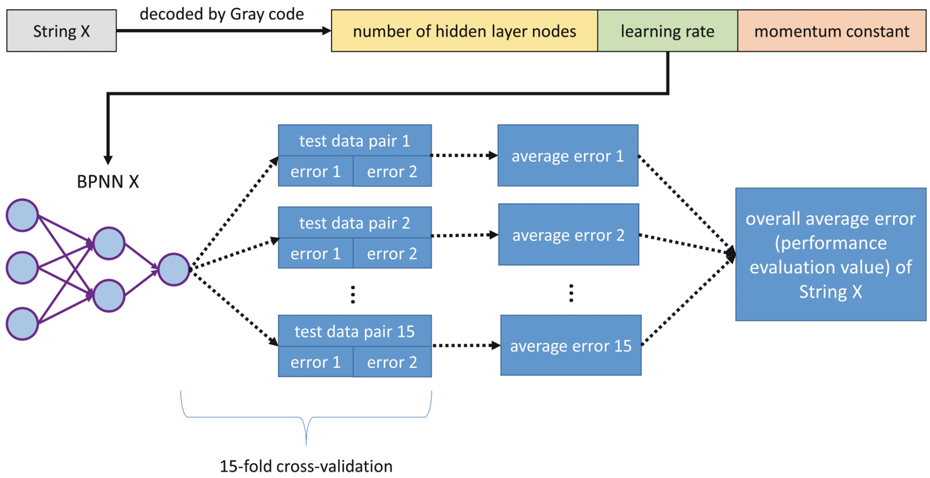

Figure 2.7: The string performance evaluation procedure in the GA.

The next step is to map every string in the initial population into the phenotype space and evaluate its performance by the procedure shown in Fig. 2.7. Given a certain string—say, String X—after being decoded, it is transformed into a hyperparameter value set that specifies a BPNN (BPNN X). The performance of BPNN X, tantamount to the performance of String X, is calculated using 15-fold cross-validation. The $L^*a^*b^*$ values as well as the factor scores on the factors Pleasure, Stability, Heaviness and Presence of the 35 stimuli used in Experiment 1 and Experiment 2 are divided into a training dataset containing 30 stimuli and a validation dataset containing the rest five stimuli. This division is shown in Table 2.11, with the validation-

Table 2.11: Training data and validation data of the GA.

Stimulus no.	Factor score on Pleasure factor	Stimulus no.	Factor score on Pleasure factor
19	-2.2803	17	0.1560
9	-1.6376	14	0.1673
31	-1.6275	26	0.1834
29	-1.6031	6	0.3366
24	-1.4836	15	0.3640
28	-1.1842	21	0.5189
32	-0.7407	2	0.6195
11	-0.4720	7	0.7098
30	-0.4593	8	0.7535
35	-0.4355	25	0.8649
22	-0.4117	13	0.9152
4	-0.3377	20	1.0395
18	-0.3292	16	1.1331
23	-0.2679	12	1.1777
1	-0.2287	33	1.2880
34	-0.1778	10	1.4435
27	-0.1153	5	2.1041
3	0.0170		

set stimuli highlighted in gray. After dividing the training dataset into 15 pairs of stimuli, we use 14 of them to train BPNN X and the remaining one to test it. The two error values for the two testing stimuli are recorded. In the next iteration, another 14 pairs are selected to train the BPNN, and the remaining one is used to obtain the error value pair. This procedure is repeated 15 times, producing a total of 15 pairs of error values. For each error value pair, the absolutes of its two elements are averaged, and the resulting 15 average values are again averaged. This final average value, which we call the ”overall average error (OAE),” is the performance evaluation value of the string (in this case, String X) that produces the BPNN being tested. Obviously, the lower the performance evaluation value—i.e., the OAE—of a string, the better its performance is.

Besides the three parameters to be optimized, the values of the other neural network param-

Table 2.12: Neural network parameters in string performance evaluation procedure.

BPNN no.	Number of input nodes	Number of output nodes	Activation function	Termination criteria of learning algorithm	
				Maximum iteration number	Mean square error
1	48	1		1000	1.00E-05
2	48		*tansig*		
3	48		function		
4	3				

eters are shown in Table 2.12.

The following step is to compute the fitness value of each string in the initial population based on its OAE. The ranking method was chosen for its ability to reduce the risk of premature convergence by restricting the reproduction capacity of the strings[10, 26].

The strings to be imported into the crossover phase are then selected through stochastic universal sampling (SUS), a variant of the conventional roulette wheel method. The SUS method ensures that the number of times a string is sampled is always exactly equal to the fitness value of that string when the fitness value is an integer, or to the nearest integers of the fitness value of that string when the fitness value is a fraction. Thus, the SUS method can avoid the occurrence of extreme sampling results that may appear in a conventional roulette wheel selection45. Next, an operation of single-point crossover with a crossover rate of 0.7 is applied to the selected strings, producing a set of offspring strings. These new strings are further mutated at a mutation rate of 0.7/string length.

When the number of offspring strings is smaller than the population size, we say that a generation gap (GP) exists, and proper means must be taken to determine which strings in the current population to substitute for the offspring strings to form a new generation. A GP is defined as the number of the offspring strings divided by the population size[10, 26]. Through a series of tentative runs of the GA program, we found that a GP tends to cause premature convergence under certain conditions. Therefore, to be conservative, we decided not to introduce a GP into our program; i.e., we set the index of GP to 1.0. In other words, the strings in the new generation are exactly the same strings produced by the preceding operations of crossover and mutation.

Next, the performance of the new generation strings is evaluated individually, and their fitness values are then computed based on their OAEs. Certain strings are then selected and undergo crossover and mutation, producing the next generation. This procedure is repeated until the maximum generation is reached. Considering that a sufficient amount of time must be left to let the evolution converge to a relatively satisfactory extent, we set the maximum generation to 100.

In addition, the performance evaluation procedure described above implies that the OAE of an individual string is not a constant but rather a variable, because the initial synaptic weights and biases of the BPNN are randomly determined during the training processes. In other words, if we evaluate a string multiple times, we will obtain a different result each time because the synaptic weights and biases resulting from the network training vary each time. This phenomenon is called ”approximate evaluation” by Fitzpatrick and Grefenstette[19], who also established a mathematical model for the phenomenon. According to this model, the most efficient way to increase the accuracy of string performance evaluation is to increase the population size to the largest affordable extent. Considering the affordable running time of our program and the maximum generation already determined, we set the population size to 40.

To pinpoint the best solution from the evolution results for each system parameter in each BPNN after the running of our GA program, we employ a modified version of the method put forward by Schaffer, Caruana and Eshelman[104], which chooses the value taken by most strings in the final generation. However, considering that the occurrence frequency distribution of the last generation is quite random, we make use of the occurrence frequency distribution

Table 2.13: Optimal hyperparameter value sets of each BPNN.

BPNN No.	Number of hidden layer nodes	Learning rate	Momentum constant
1	4	1/3	0.2
2	24	5/9	0.2
3	8	7/9	0.2
4	1	1/9	0.4

during the last five generations rather than only the last generation.

The optimal parameter value set of each BPNN is displayed in Table 2.13. With these sets of optimal parameter values, both the typological configuration and the learning algorithms of the BPNNs have been determined. The following step is the conduction of two test simulations intended to adjust the microstructure, namely, the synaptic weights and biases, of the computational model to let the model approximate the actual psychological mechanism of multicolor aesthetic evaluation. (Both the hyperparameter optimization and the model simulations are implemented using MATLAB.)

2.6 Simulation 1 of Computational Model

2.6.1 Procedure

[5]The 35 stimuli used in Experiment 1 and Experiment 2 were divided into two groups: a training dataset and a validation dataset. The training dataset consisted of the 30 stimuli having been used in the string performance evaluation procedure in our GA program. The validation dataset was composed of the remaining five stimuli. Because the validation-set stimuli had not been used in the model construction, they could be deemed as qualified test targets.

The procedure of a single iteration of this simulation, which is shown in Figure 2.8, had two phases: a training phase and a prediction phase. In the training phase, the four BPNNs were trained independently using their respective training information.

[5]The content of Section 2.6 was published in Fang, Muramatsu and Matsui's 2017 research paper[18].

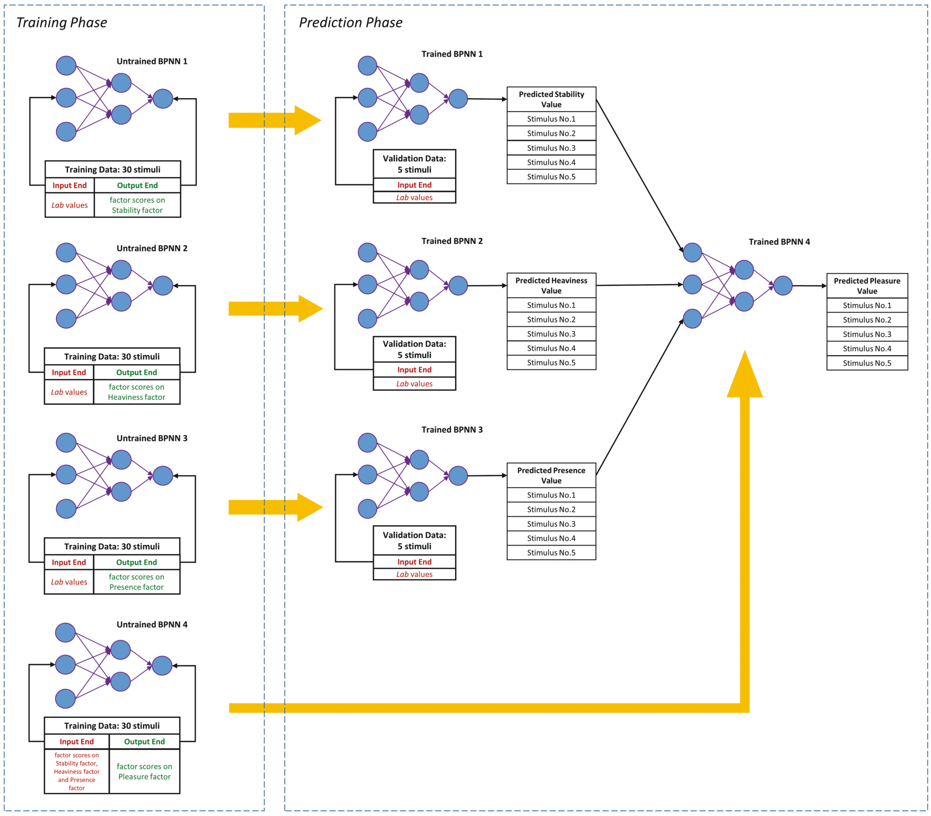

Figure 2.8: The procedure of a single iteration of Simulation 1.

The input-end training information of BPNN 1 comprised the $L^*a^*b^*$ values of the training-set stimuli, and its output-end training information comprised the factor scores of the stimuli on the factor Stability. BPNN 2 and BPNN 3 shared the same input-end training information with BPNN 1 but varied in terms of the output-end. The output-end training information of BPNN 2 comprised the factor scores of the training-set stimuli on the factor Heaviness, and that of BPNN 3 comprised the factor scores of the training-set stimuli on the factor Presence. For BPNN 4, its input-end training information comprised the factor scores of the training-set stimuli on the three factors, Stability, Heaviness and Presence, and its output-end training information comprises the factor scores of the stimuli on the factor Pleasure.

The prediction phase then ensued. In this phase, the $L^*a^*b^*$ values of the five validation-set stimuli were imported, one stimulus after another, into the input layers of BPNN 1, BPNN 2, and BPNN 3. The outputs of BPNN 1, called the ”predicted stability values” for convenience, were the predictions of the factor scores of the stimuli on the factor Stability. The outputs of BPNN 2, called the ”predicted heaviness values” for convenience, were the predictions of the factor scores of the stimuli on the factor Heaviness. In addition, the outputs of BPNN 3, called the ”predicted presence values” for convenience, were the predictions of the factor scores of the stimuli on the factor Presence. Next, the predicted stability values, the predicted heaviness values, and the predicted presence values of the stimuli were imported, one stimulus after another, into the input layer of BPNN 4. The outputs of BPNN 4, called the ”predicted pleasure values” for convenience, were the predictions of the factor scores of the stimuli on the factor Pleasure.

The outputs of BPNN 4—namely, the predicted pleasure values—served as the final outputs of the computational model in the single iteration. It is important to note that because of the existence of the aforementioned approximate evaluation phenomenon, if we run the simulation procedure multiple times, we will obtain a different output result each time. This variation stems from the randomly determined initial synaptic weights and biases of the component BPNNs during the training processes. In other words, each time we run this procedure, we will obtain a different set of initial synaptic weights and biases, which produces different training results during the training phase, eventually leading to different prediction results during the prediction phase. Hence, it is necessary to run this procedure for a number of iterations (50 iterations in this simulation). The predicted pleasure values at every iteration were recorded and then averaged. These average predicted pleasure values were regarded as the eventual model prediction to the aesthetic scores of the validation-set stimuli. The worst-case time complexity of the whole model simulation procedure was $O(n)$, where n was the size of the validation dataset. It suggests that the time consumption of the simulation procedure is generally acceptable.

In addition, we noted that the minimum factor scores on the factors engaged in the simulation process all lay within the range of $(23, 22)$, and the maximum factor scores all lay within the range of $(2, 3)$. The activation function of the BPNNs, however, took the range of $[21, 1]$. We thus contracted all these factor scores into the range of $[21, 1]$ by scaling them by 1/3.

2.6.2 Results

To evaluate the prediction performance of the model, which we regarded as an indicator of the psychological appropriateness of the model, we chose to employ the normalized root mean square error (NRMSE) defined by Eq. (2.3), where N is the number of the predicted objects, $x_i (i = 1, 2, ..., N)$ are the predicted values, $y_i (i = 1, 2, ..., N)$ are the observed values, y_{max} is the maximum of the observed values, and y_{min} is the minimum.

$$NRMSE = \frac{\sqrt{\frac{1}{N}\sum_{i=1}^{N}(x_i - y_i)^2}}{y_{max} - y_{min}} \tag{2.3}$$

Table 2.14: Results of Simulation 1.

Validation-set stimulus no.	1	2	3	4	5
Factor score on the factor Pleasure (1/3)	0.2065	0.0057	0.3926	-0.1372	-0.5344
Average predicted pleasure value	-0.1227	-0.2667	0.0531	-0.1223	0.2362
Error	-0.3292	-0.2724	-0.3394	-0.0149	0.7706

As seen in Eq. (2.3), the NRMSE measures the relative prediction error by scaling the root mean square error by the range of observed values, thus having the merit of not being affected by the unit in use.

The NRMSE of the model in this simulation is 0.4556, and the absolutes of four of the five prediction errors are smaller than 0.34 (shown in Table 2.14). This demonstrates, in our view, that our model may be able to predict the aesthetic scores of most $4 times 4$ grid multicolor stimuli to a certain extent. However, it is difficult to draw statistically meaningful conclusions from a validation dataset of such a small size. This problem will be addressed in the second simulation, where a much larger validation dataset is used.

In addition, the Shapiro-Wilk tests exhibit that the predicted pleasure values across the iterations for each individual validation-set stimulus can be regarded as normally distributed (Stimulus 1, statistic = 50.968, P = 50.183; Stimulus 2, statistic = 50.983, P = 50.676; Stimulus 3, statistic = 50.971, P = 50.266; Stimulus 4, statistic = 50.978, P = 50.475; Stimulus 5, statistic = 50.982, P = 50.627). In view of this, it is reasonable to predict that the model prediction for the aesthetic score of any $4 times 4$ grid multicolor stimulus follows a normal distribution, implying that the prediction by our model for any such stimulus is relatively stable across iterations.

What should also be noticed is that in this simulation the stimuli in both the training data and the validation data have been employed in Experiment 1 and Experiment 2. This suggests the possibility that the predictive power of our model is limited to the stimuli used in the two experiments. We then run a second model simulation to rule out this possibility and to perform a statistically meaningful analysis of the post-training microstructure of the BPNN 4, as introduced in detail in the next section.

2.7 Simulation 2 of Computational Model

2.7.1 Procedure

[6]As mentioned at the end of the previous section, Simulation 2 consists of the following three phases:

Phase 1: Model training

The training dataset in this simulation consisted of all 35 stimuli used in Experiment 1 and Experiment 2.

Phase 2: Model validation

The validation dataset in this simulation was composed of 25 multicolor stimuli newly generated using the method described in the Section 2.3.2. Their aesthetic scores were obtained through a psychological experiment using the SD method with the same procedure as in Experiment 1. This experiment is called ”Experiment 3” in this research. The details of Experiment 3 are introduced in Section 2.7.2. In this manner, we eliminated the possibility that the predictive power of our model was restricted to the stimuli used in Experiment 1 and Experiment 2, and, in the meantime, managed to evaluate the model performance, which is an indicator of the psychological appropriateness of the model, in a statistical sense using

[6]The content of Section 2.7 was published in Fang, Muramatsu and Matsui’s 2017 research paper[18].

a sufficiently large validation dataset. Phase 1 and Phase 2 were carried out using the same procedure as in Simulation 1. The model performance was evaluated in the following two ways:

Error-based Verification of psychological appropriateness

This evaluation, described in Section 2.7.3, was based on the prediction errors of the model. If the microstructure of trained model approximates the actual psychological mechanism of multicolor aesthetic evaluation, the prediction accuracy of the model should be high, or, in other words, the the prediction errors of the model should be small.

Correlation-based Verification of psychological appropriateness

This evaluation, described in Section 2.7.4, was based on the correlation between the overall lightness of the validation-set stimuli and the predicted aesthetic scores of the stimuli. A positive correlation between overall lightness and aesthetic evaluation of color combinations was frequently reported in literature. If the model has successfully learnt the complicated nonlinear mappings from physical color information to the three multicolor-level perceptual features and those from the three perceptual features to the aesthetic evaluation, the model must be able to express simpler psychological rules in multicolor aesthetics, such as the positive linear relationship between overall lightness and aesthetic evaluation of color combinations.

Phase 3: Microstructure analysis of BPNN 4

After the psychological appropriateness of the model was verified at Phase 2, a microstructure analysis of the trained BPNN 4 was performed to figure out how each perceptual feature impacts the aesthetic evaluation (described in Section 2.7.5).

2.7.2 Details of Experiment 3

Eight subjects (three males and five females, 19 to 23 years of age), who were undergraduate students of Waseda University, participated in Experiment 3. The RGB and $L^*a^*b^*$ values of the component colors of each stimulus are listed in Appendix C.

The experiment was performed indoors (in Room 119 in Building No.100 at Tokorozawa Campus of Waseda University) with white fluorescent lighting (type: Panasonic LDL40S · W/22/23, color temperature: 4000K). The illuminance was 464lx measured by a HIOKI FT3424 lux meter (JIS C 16091 General Class AA, DIN 5032-7 Class B). The stimuli were presented to the subjects using Microsoft PowerPoint slides that had the same configurations as those used in Experiment 1 and Experiment 2. The slides were run on a HP EliteBook 8460w PC and projected through two Panasonic PT-F300 projectors onto two projection screens (refresh rate: 60hz; resolution: 1024 * 768). The projectors were fixed on the ceiling, and the screens were set on the front wall. Each stimulus was projected onto the screen with the size 126.5cm * 126.5cm.

Figure 2.9 shows a sketch of the experiment environment. Because we think that the subjects were most likely sitting at the middle area of the room (the middle three columns of desks and approximately from the fourth to eighth row of desks) so that they could look right to the screens, we afterwards measured the largest view angle (measured at the fourth row of desks) and the smallest viewing angle (measured at the eighth row of desks) within this area. The former was about 12.80°, and the latter was about 8.19°. The environment was kept quiet throughout the experiment.

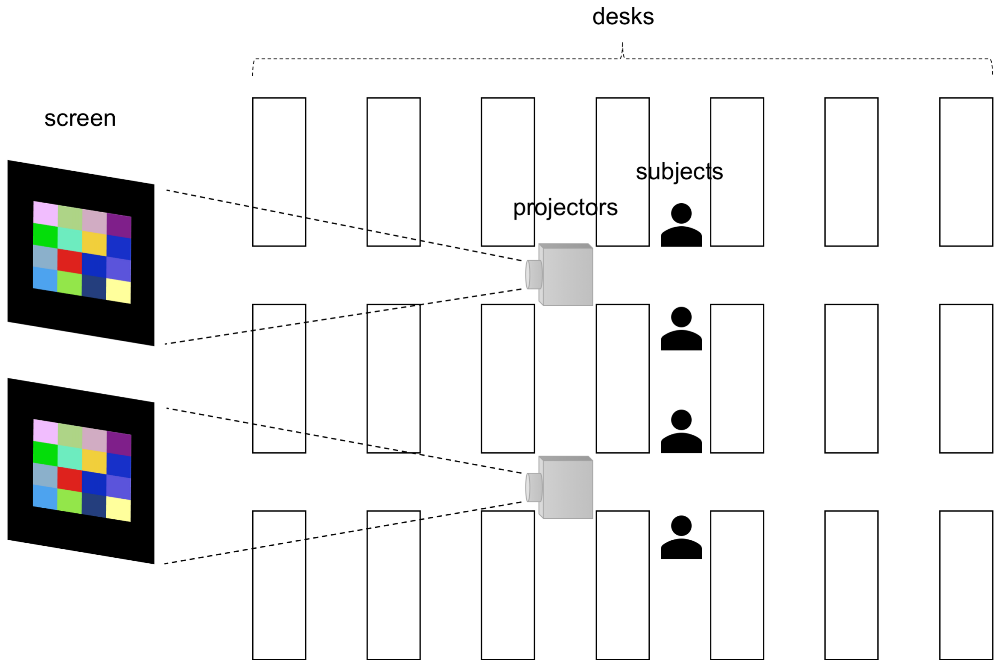

Figure 2.9: A sketch of the environment of Experiment 3 (the middle area of the experiment room).

The subjects were asked to evaluate the validation-set stimuli on the set of adjective pair scales used in Experiment 1. The rating scores on the adjective pair scales with regard to each stimulus were averaged across the subjects and then imported into the factor analysis package in IBM SPSS Statistics (version 19). During the averaging process, the missing values were treated by the same means as in Experiment 1.

The factor extraction was conducted using the principal factor method. The communality of each adjective pair scale is listed in Table 2.15. The communality of every adjective pair scale was larger than 0.6, so there was no adjective pair scale that should be deleted before entering

Table 2.15: Communalities of adjective pair scales (Experiment 3).

Adjective pair scale	Communality	Adjective pair scale	Communality
heavy-light	.916	graceful-awkward	.891
light-dark	.844	clear-dull	.870
warm-cool	.713	static-dynamic	.890
soft-hard	.773	true-false	.881
noisy-quiet	.838	novel-ordinary	.849
like-dislike	.860	beautiful-ugly	.884
ornate-plain	.778	stable-changeable	.895
strong-weak	.956	successful-unsuccessful	.791
pleasant-unpleasant	.980	positive-negative	.836
clean-dirty	.937	relaxed-nervous	.692
harmonious-dissonant	.799	cruel-kind	.654
cheerful-gloomy	.861	passive-active	.724

Table 2.16: Total variance explained (Experiment 3).

Factors	Variance explained before rotation		Variance explained after rotation	
	Individual (%)	Cumulative (%)	Individual (%)	Cumulative (%)
1	44.746	44.746	40.045	40.045
2	31.477	76.224	28.974	69.020
3	7.578	83.802	14.782	83.802

the following computation steps.

The scree plot (Figure 2.10) shows the eigenvalue of each factor. There were three factors whose eigenvalues were larger than 1.0. Therefore, these three factors were selected as main factors.

Then, a varimax rotation was then performed. The total variance of the adjective pair scales which were explained by the three factors before and after the rotation operation were summarized in Table 2.16. The factor loadings resulted from the rotation was shown in Table 2.17. All factor loadings whose absolutes are greater than 0.5 are emphasized in bold. For easier reading, the variance explained by each factor after the rotation, shown in the fourth column in Table 2.16, is duplicated in the final row in Table 2.17.

Table 2.17 shows that nearly all adjective pair scales belong to the factors to which they

Table 2.17: Rotated factor loading matrix (Experiment 3).

Factor identity	Adjective pair scale	Factor		
		1	2	3
Factor 1 (Pleasure)	pleasant-unpleasant	**.964**	.210	.084
	clean-dirty	**.943**	-.128	.178
	stable-changeable	**.921**	-.150	.156
	true-false	**.913**	-.203	.072
	graceful-awkward	**.909**	-.173	.188
	beautiful-ugly	**.906**	-.023	.252
	clear-dull	**.904**	-.102	.207
	like-dislike	**.902**	.176	.128
	successful-unsuccessful	**.846**	-.089	.259
	harmonious-dissonant	**.833**	-.322	.025
Factor 2 (Activity)	positive-negative	.002	**.915**	-.006
	ornate-plain	.086	**.875**	.071
	static-dynamic	.411	**-.848**	-.033
	passive-active	.178	**-.827**	-.092
	noisy-quiet	-.475	**.772**	.122
	cheerful-gloomy	.330	**.756**	.424
	warm-cool	-.320	**.720**	.305
	novel-ordinary	**-.597**	**.702**	.006
	cruel-kind	-.109	**-.684**	-.419
	relaxed-nervous	.181	**.583**	**.566**
Factor 3 (Potency)	soft-hard	.198	.225	**.827**
	heavy-light	-.462	-.298	**-.784**
	light-dark	.111	.482	**.774**
	strong-weak	-.337	**.532**	**-.748**
Variance explained (%)		40.045	28.974	14.782

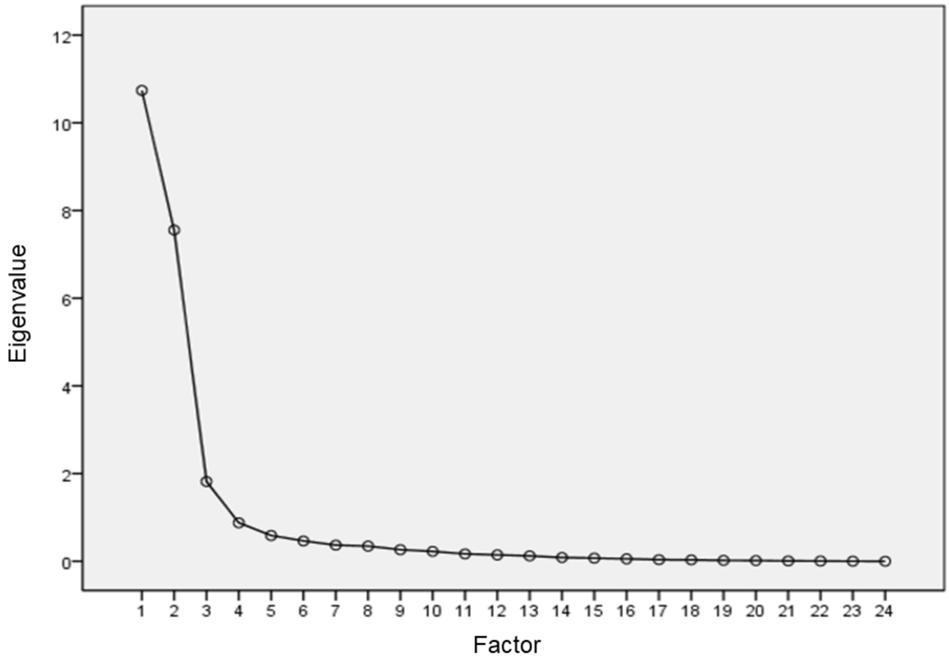

Figure 2.10: Scree plot of Experiment 3.

belonged in Experiment 1. Thus, as in Experiment 1, the factors in the present experiment are named as "Pleasure," "Activity," and "Potency." This agreement between the two experiments in factorial structure suggests that these three factors possess a high degree of psychological reality. As in Experiment 1, the aesthetic score of each stimulus in the present experiment is defined as the inverse of its factor score on the factor Pleasure.

The factor scores of the stimuli on the three factors are listed in Table 2.18 in which the factor scores on the factor Pleasure are emphasized in bold.

2.7.3 Error-Based Verification of Psychological Appropriateness

The simulation results are summarized in Table 2.19, and the NRMSE of the model is 0.2764. To assess the predicting ability of our model reasonably from this numerical result, it is important to recall that, due to the chief concern of our model construction placed on psychological appropriateness, our model, based on Leder's aesthetic psychological model, adopted an ar-

Table 2.18: Factor scores (Experiment 3).

Stimulus	Factor			Stimulus	Factor		
no.	**1 (Pl)**	2 (Ac)	3 (Po)	no.	**1 (Pl)**	2 (Ac)	3 (Po)
1	**0.25**	0.99	-0.79	14	**0.58**	1.29	0.37
2	**-0.58**	0.58	-0.93	15	**-1.53**	1.11	-1.07
3	**0.99**	0.93	-0.91	16	**-1.44**	0.23	0.74
4	**-0.71**	0.93	1.94	17	**1.23**	-0.26	-0.36
5	**-1.27**	-1.13	-1.36	18	**0.33**	-0.08	-0.80
6	**-2.54**	0.93	1.45	19	**0.69**	0.26	-0.29
7	**0.96**	-0.71	1.90	20	**0.91**	0.77	1.24
8	**-0.61**	-2.24	-0.89	21	**0.54**	-0.44	1.14
9	**0.32**	-0.40	-0.79	22	**0.65**	1.12	-0.33
10	**1.53**	0.20	-0.47	23	**-0.04**	-0.69	0.62
11	**0.15**	1.29	-1.13	24	**0.63**	-1.13	0.66
12	**0.52**	-1.67	0.56	25	**-0.97**	-1.03	-0.76
13	**-0.57**	-0.84	0.28				

chitecture composed of two levels of neural networks. During the model simulations, because the outputs of the first-level neural networks served as the inputs to the second-level neural network, the prediction errors occurring at the first level were inherited by the second level, combining with the inherent prediction error of the second level to generate the overall prediction error. In view of the existence of this error pooling effect, we deem the NRMSE of the model in this simulation to be acceptable.

Furthermore, the Shapiro-Wilk test shows that the prediction error values across the validation-set stimuli can be considered as normally distributed (statistic = 50.961, P = 50.432). The distribution has a relatively small standard deviation (0.382) along with a mean (0.016) and a median (20.028) fairly close to zero. This implies that the model prediction has a large likelihood to have a small error and a small likelihood to have a large error. In other words, this suggests that the model prediction deviates little from the actual aesthetic scores of most stimuli.

Combining the above analyses, it is reasonable to argue that our model is able to predict

Table 2.19: Results of Simulation 2.

Validation-set stimulus no.	Factor score on the factor Pleasure (1/3)	Average predicted Pleasure value (inversed)	Error
1	0.0832	-0.1581	-0.2413
2	-0.1940	-0.0192	0.1748
3	0.3297	0.2254	-0.1043
4	-0.2357	0.1750	0.4106
5	-0.4225	-0.0658	0.3567
6	-0.8461	0.2093	1.0554
7	0.3189	0.2719	-0.0470
8	-0.2037	0.2776	0.4812
9	0.1061	0.2081	0.1020
10	0.5092	-0.1320	-0.6412
11	0.0484	-0.3092	-0.3576
12	0.1717	-0.2839	-0.4556
13	-0.1905	-0.0866	0.1038
14	0.1943	-0.1410	-0.3354
15	-0.5114	-0.0905	0.4209
16	-0.4793	-0.0587	0.4206
17	0.4096	-0.0227	-0.4324
18	0.1112	-0.0025	-0.1138
19	0.2311	-0.1884	-0.4194
20	0.3038	0.2769	-0.0269
21	0.1788	0.1506	-0.0282
22	0.2158	0.1561	-0.0597
23	-0.0148	0.2874	0.3022
24	0.2090	-0.0207	-0.2297
25	-0.3231	-0.2465	0.0766

with some degree of accuracy the aesthetic scores of the stimuli whose origins are completely independent of those of the training data. In other words, our model may possess some generalizing power of what it has learned from the training data that originated in Experiment 1 and Experiment 2. Therefore, it is reasonable to regard our model as a relatively successful approximation of the psychological information processing mechanism from which the sense of beauty attributed to multicolor objects emerged.

In addition, as in the prediction performance evaluation procedure of Simulation 1, the inter-iteration variation of the model prediction for each individual validation-set stimulus was also assessed. The Shapiro-Wilk tests show that the P values for 60% of those stimuli exceeded 0.05. This means that for most of the stimuli, the prediction by our model across the iterations can be considered to follow a normal distribution, implying that our model makes prediction with a certain amount of inter-iteration stability.

2.7.4 Correlation-Based Verification of Psychological Appropriateness

To our knowledge, the strongest and most robust physical property reported in the literature that positively influences the aesthetic evaluation of a two-color or three-color combination is the overall lightness of the color combination. In this research, we defined the overall lightness of each multicolor stimulus as the average L^* value of the component colors. We found that in both the training and the validation datasets, there exists statistically significant positive correlation between the overall lightness and the actual aesthetic scores (for the training dataset: Pearson's correlation coefficient = 50.370, P = 50.029, scatter plot shown in Figure 2.11; for the validation dataset: Pearson's correlation coefficient = 50.461, P = 50.024, scatter plot shown in Figure 2.12). Considering that the two datasets were obtained from two independent psychological experiments, we think that the strong positive impact of overall lightness on aesthetic evaluation of color combinations remains to be true in the cases of color combinations consisting of more than three (for example, sixteen in this research) colors.

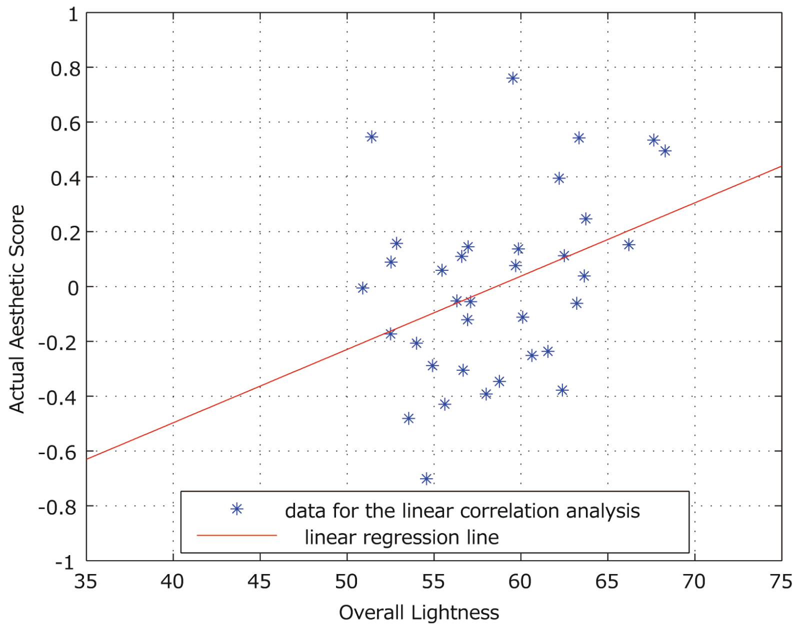

Figure 2.11: The overall lightness is plotted against the actual aesthetic scores over the training dataset. There is statistically significant positive correlation between these two variables (Pearson's correlation coefficient = 0.370, $P = 0.029$).

A further analysis of the simulation results showed a statistically significant positive correlation between the overall lightness and the predicted aesthetic scores (Pearson's correlation coefficient = 50.648, $P = 50.0005$, scatter plot shown in Figure 2.13). This means that, through training, the computational model successfully learned the psychological relationship between the overall lightness and the aesthetic scores of the multicolor stimuli, which implies that a relatively high degree of psychological appropriateness was achieved by the trained computational model.

Yet, we note that the Pearson's correlation coefficient (0.648) between the overall lightness and the predicted aesthetic score is somewhat higher than that (0.461) between the overall lightness and the actual aesthetic score. Additionally, the positive correlation between the actual and the predicted aesthetic scores did not reach statistical significance (Pearson's correlation coef-

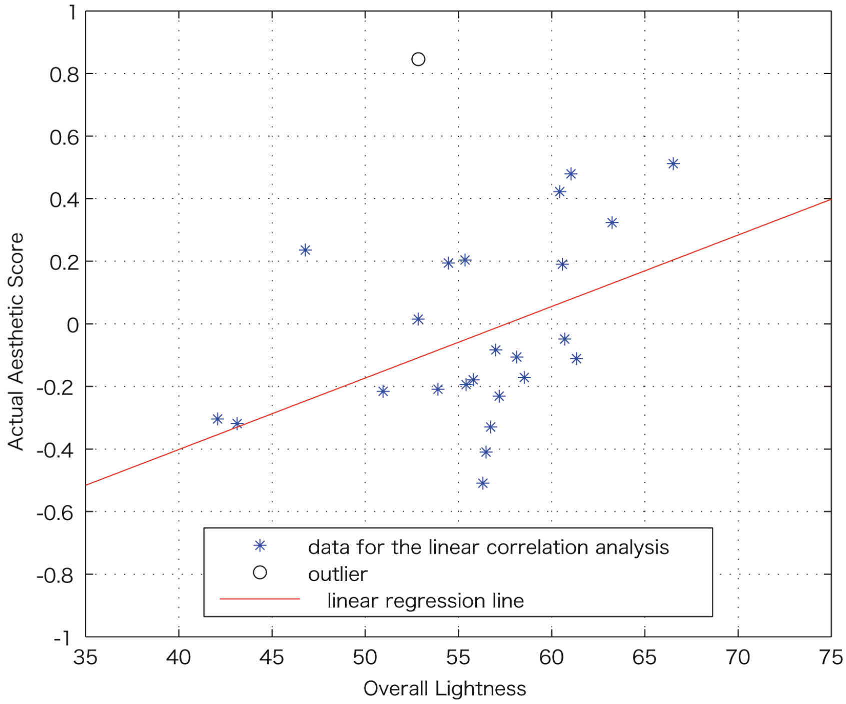

Figure 2.12: The overall lightness is plotted against the actual aesthetic scores over the validation dataset. There is statistically significant positive correlation between these two variables (Pearson's correlation coefficient = 0.461, $P = 0.024$).

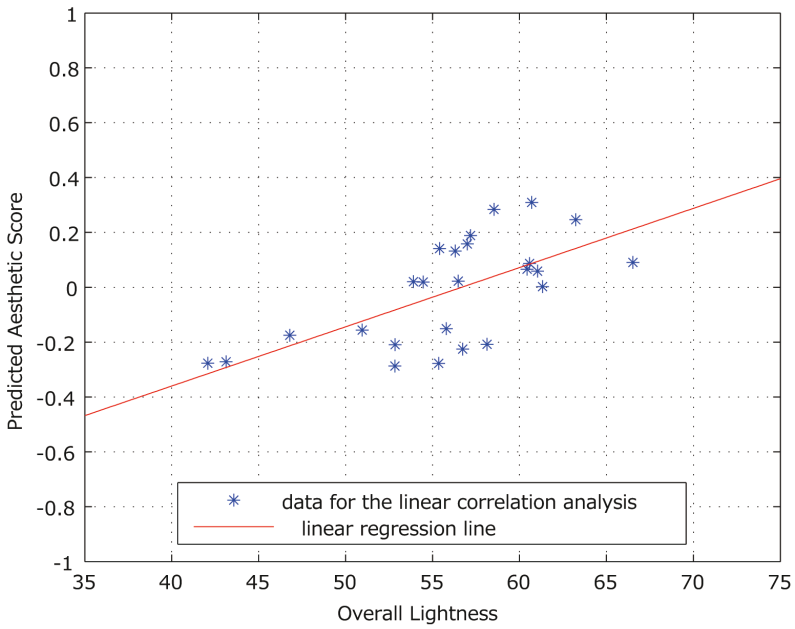

Figure 2.13: The overall lightness is plotted against the predicted aesthetic scores over the validation dataset. There is statistically significant positive correlation between these two variables (Pearson's correlation coefficient = 0.648, P = 0.0005).

ficient = 50.151, P = 50.481, scatter plot shown in Figure 2.14). This perhaps resulted from the computational model's incomplete learning of some minor aesthetically relevant psychological rules. Due to the time limitations of the psychological experiments, the training dataset in this simulation is of a relatively small size. Although the correlation between the overall lightness and the aesthetic evaluation is strong and robust enough to be learned even from a small amount of training data, the relationships between some other physical properties and the aesthetic evaluation may be too weak to be adequately learned from a training dataset of a small size, such as that used in this model simulation. Thus, we think that the problem will be solved by providing the computational model with more training data in future studies.

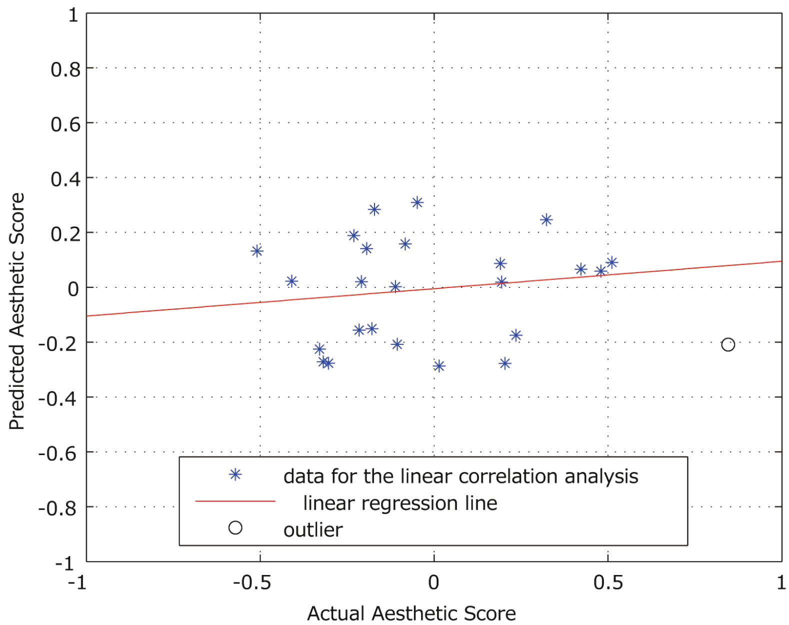

Figure 2.14: The actual aesthetic scores are plotted against the predicted aesthetic scores over the validation dataset. The positive correlation between these two variables has not reached statistical significance (Pearson's correlation coefficient = 0.151, $P = 0.481$).

2.7.5 Microstructure Analysis of BPNN 4

As the psychological appropriateness of the computational model was verified in an error-based way and a correlation-based way, we examined the microstructure—namely, the synaptic weights and biases—of BPNN 4 (Figure 2.15), which mapped the three multicolor-level perceptual features Stability, Heaviness and Presence to the aesthetic score. This analysis aims to clarify the "indirect effect" of each input node on the output node of BPNN 4, which represents how each perceptual feature influences the aesthetic evaluation in a quantitative fashion.

For a BPNN, we divide the influences of an individual input node on another node into two types: "direct effect" and "indirect effect." "Direct effect" is defined qualitatively as the influence of an input node on another node that is directly linked to the input node. The strength of such influence, which we call the "direct effect index (DEI)," is defined as the synaptic weight from the input node to the target node. When the DEI is greater than zero, we call

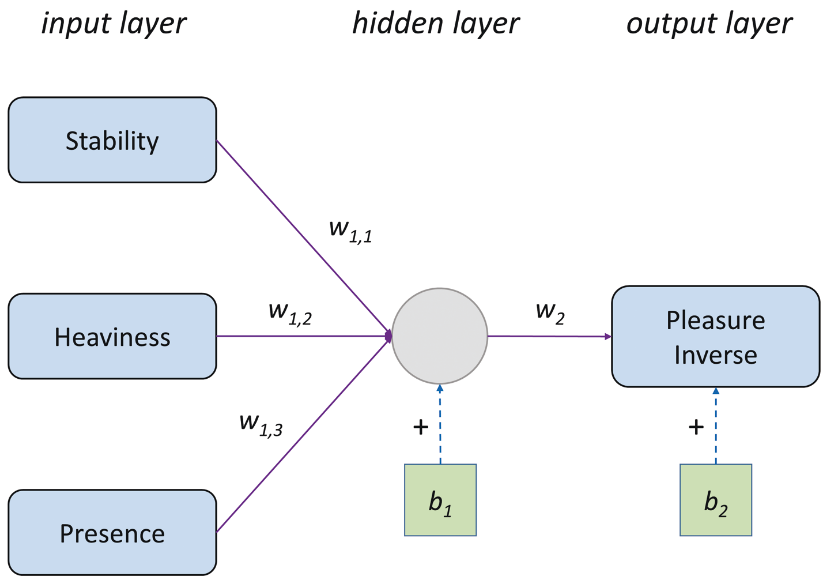

Figure 2.15: The microstructure of BPNN 4.

the direct effect a ”stimulatory direct effect.” Conversely, when the DEI is less than zero, we call the direct effect an ”inhibitory direct effect.” ”Indirect effect” is defined qualitatively as the influence of an input node on a node to which the input node is not directly linked but is bridged by another node(s). For instance, in a three-layer BPNN, the influence of an input node on an output node is relayed by the hidden layer nodes. The quantification of an indirect effect is more complicated than that of a direct effect because there is no direct synaptic link between the two nodes. We quantify an indirect effect as the partial derivative of the output value of the target node with respect to the input value from the input node. The reason for using this definition is that this partial derivative function shows how the output of the target node varies with the input from the input node. When the function value is greater than zero, we call the indirect effect a ”stimulatory indirect effect,” and when the function value is less than zero, we call the indirect effect an ”inhibitory indirect effect.”

Before investigating the indirect effects of the three input nodes on the output node of BPNN 4 using the definitions introduced above, we examined the 50 sets of convergence points of the synaptic weights and biases produced by the 50 iterations during the training process of BPNN 4 to see whether BPNN 4 learned the same mapping at these iterations. This is necessary because if the mappings learned by the BPNN at these iterations differ in a qualitative sense, then there is no significance in further examining the microstructure of the BPNN. Figure 2.16 displays the distribution of the convergence points on three of the microstructure parameters.

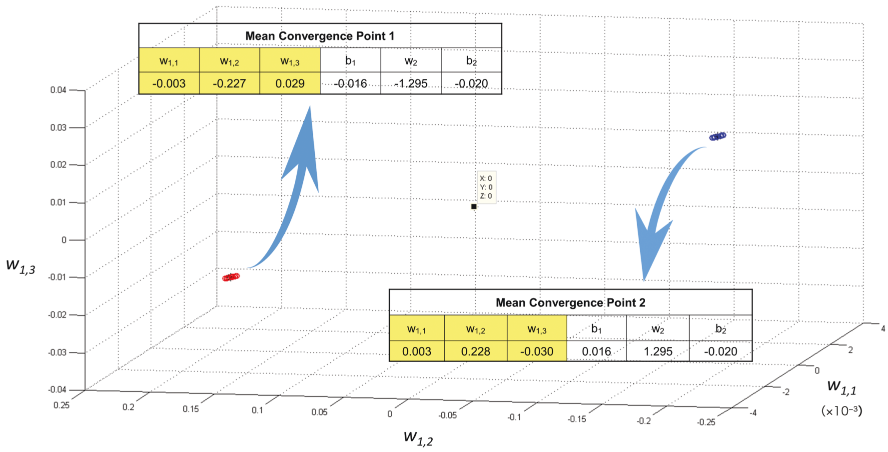

Figure 2.16: The distribution pattern of the convergence points on the microstructure parameters $w_{1,1}$, $w_{1,2}$ and $w_{1,3}$.

As Figure 2.16 shows, the 50 convergence points gathered into two clusters. The average points of the two clusters, which we call "Mean Convergence Point 1 (MCP 1)" and "Mean Convergence Point 2 (MCP 2)," are nearly symmetric with respect to the zero point in the weight-bias space except on the dimension b_2, where the coordinate values of the two average points are nearly equal. Here below is a mathematical explanation of this phenomenon.

Suppose that the input from the node Stability is x, the input from the node Heaviness is y, and the input from the node Presence is z. The activation function of BPNN 4—namely, the $tansig$ function introduced at the end of Section 2.5.1—is labeled as $f(x)$. When the weights and biases are $(w_{1,1}, w_{1,2}, w_{1,3}, b_1, w_2, b_2)$, the output of the output node that represents the aesthetic evaluation—namely, the node Pleasure—is p_{ori}:

$$p_{ori} = f(w_2 f(w_{1,1}x + w_{1,2}y + w_{1,3}z + b_1) + b_2) \tag{2.4}$$

When the values of $w_{1,1}$, $w_{1,2}$, $w_{1,3}$, b_1, and w_2 are inverted against the zero point—that is, when the weights and biases become $(-w_{1,1}, -w_{1,2}, -w_{1,3}, -b_1, -w_2, b_2)$—the output value of the node Pleasure is p_{oppo}:

$$p_{oppo} = f(-w_2 f(-w_{1,1}x - w_{1,2}y - w_{1,3}z - b_1) + b_2) \tag{2.5}$$

Because $f(x)$ is an odd function, which means that

$$f(-x) = -f(x) \tag{2.6}$$

we can know the following:

$$\begin{aligned} p_{oppo} &= f(-w_2 f(-w_{1,1}x - w_{1,2}y - w_{1,3}z - b_1) + b_2) \\ &= f(-w_2 f(-(w_{1,1}x + w_{1,2}y + w_{1,3}z + b_1)) + b_2) \\ &= f(w_2 f(w_{1,1}x + w_{1,2}y + w_{1,3}z + b_1) + b_2) \\ &= p_{oppo} \end{aligned} \tag{2.7}$$

Eq. 2.7 tell us that, when a training process searches for the optimal solution(s) in the weight-bias space, if the point

$$(w_{1,1}, w_{1,2}, w_{1,3}, b_1, w_2, b_2) \tag{2.8}$$

is a solution, then the point

$$(-w_{1,1}, -w_{1,2}, -w_{1,3}, -b_1, -w_2, b_2) \tag{2.9}$$

must also be a solution. The two BPNNs that correspond to the two weight-bias sets always produce the same outputs if the same inputs are given. Namely, the two BPNNs encode the same mappings. This explains why MCP 1 and MCP 2 are nearly symmetric with respect to the zero point on the dimensions $w_{1,1}$, $w_{1,2}$, $w_{1,3}$, b_1, and w_2 and why MCP 1 and MCP 2 have almost the same value on the dimension b_2.

Taking one step further, we can see that

$$\frac{\partial p_{ori}}{\partial x} = \frac{\partial p_{oppo}}{\partial x} = a\frac{4w_2 w_{1,1} G}{(\cos h(b_1 + w_{1,1}x + w_{1,2}y + w_{1,3}z))^2 (G+1)^2} \tag{2.10}$$

$$\frac{\partial p_{ori}}{\partial y} = \frac{\partial p_{oppo}}{\partial y} = \frac{4w_2 w_{1,2} G}{(\cos h(b_1 + w_{1,1}x + w_{1,2}y + w_{1,3}z))^2 (G+1)^2} \tag{2.11}$$

$$\frac{\partial p_{ori}}{\partial z} = \frac{\partial p_{oppo}}{\partial z} = \frac{4w_2 w_{1,3} G}{(\cos h(b_1 + w_{1,1}x + w_{1,2}y + w_{1,3}z))^2 (G+1)^2} \tag{2.12}$$

where

$$G = \exp(-2b_2 - 2w_2(\frac{2}{\exp(-2b_1 - 2w_{1,1}x - 2w_{1,2}y - 2w_{1,3}z) + 1} - 1)) \tag{2.13}$$

Eqs. 2.10-2.12 mean that the indirect effects of the input nodes on the output node when BPNN 4 takes the weight-bias set $(w_{1,1}, w_{1,2}, w_{1,3}, b_1, w_2, b_2)$ are the same as when it takes the weight-bias set $(-w_{1,1}, -w_{1,2}, -w_{1,3}, -b_1, -w_2, b_2)$. In other words, when investigating the indirect effects of the input nodes on the output node, no matter which of the two weight-bias sets we select, we will obtain the same results. Hence, we pool together the group of convergence points centered by MCP 1 and the group of convergence points centered by MCP 2. The average point of the pooled data is $(0.003, 0.227, -0.029, 0.016, 1.295, -0.020)$. With this weight-bias set, the output (p) of the node Pleasure is

$$p = f(1.295f(0.003x + 0.227y - 0.030z - 0.016) - 0.020) \tag{2.14}$$

where

$$f(a) = \frac{2}{1 + \exp(-2a)} - 1 \tag{2.15}$$

In Eq. 2.14 and Eq. 2.15, as mentioned above in this section, x is the input from the node Stability, y is the input from the node Heaviness, and z is the input from the node Presence.

The indirect effect of the node Stability on the output node Pleasure is depicted by the function $FIE_{S,PL}$:

$$FIE_{S,PL} = \frac{\partial p}{\partial x} = \frac{0.016G(x,y,z)}{\cosh(0.003x + 0.227y - 0.029z - 0.016)^2(G(x,y,z) + 1)^2)} \tag{2.16}$$

where

$$G(x,y,z) = \exp(2.631 - \frac{5.180}{\exp(0.059z - 0.455y - 0.006x + 0.032) + 1}) \tag{2.17}$$

The function value of $FIE_{H,PL}$ is always greater than zero. This means that, regardless of the values taken by x, y, and z, the node Heaviness consistently exerts a stimulatory indirect effect on the output node Pleasure.

The indirect effect of the node Heaviness on the output node Pleasure is described by the function $FIE_{H,PL}$:

$$FIE_{H,PL} = \frac{\partial p}{\partial y} = \frac{1.178G(x,y,z)}{\cosh(0.003x + 0.227y - 0.029z - 0.016)^2(G(x,y,z) + 1)^2)} \tag{2.18}$$

The function value of $FIE_{H,PL}$ is always greater than zero. This means that, regardless of the values taken by x, y, and z, the node Heaviness consistently exerts a stimulatory indirect effect on the output node Pleasure.

The indirect effect of the node Presence on the output node Pleasure is depicted by the function $FIE_{PR,PL}$:

$$FIE_{PR,PL} = \frac{\partial p}{\partial z} = \frac{-0.153G(x,y,z)}{\cosh(0.003x + 0.227y - 0.029z - 0.016)^2(G(x,y,z)+1)^2)} \quad (2.19)$$

Contrary to $FIE_{S,PL}$ and $FIE_{H,PL}$, the function value of $FIE_{PR,PL}$ is always less than zero. This tells us that regardless of how x, y, and z change their values, the node Presence consistently has an inhibitory indirect effect on the output node Pleasure.

We can see that the node Heaviness always has the strongest indirect effect, followed by the node Presence and then the node Stability. The strength of the indirect effect of the node Heaviness remains 7.70 times that of the node Presence and 73.63 times that of the node Stability.

In summary, the microstructure analysis of BPNN 4 shows that, in the psychological process of multicolor aesthetic evaluation, the perceptual feature Heaviness bears the principal influence on aesthetic evaluation. The heavier a multicolor stimulus is perceived to be, the less aesthetically pleasing it is (please keep in mind that the output of the node Pleasure represents the inverse of the aesthetic score). This matches well with our daily experience in which objects sensed as heavy tend to evoke psychological stress, thus leading to negative emotions. The other two perceptual features—namely, Presence and Stability—exert minor influence on aesthetic evaluation. The stronger the sense of matter presence a multicolor object can elicit, the higher its aesthetic score. This is perhaps because objects that are felt as possessing a dense or tight inner quality are not easily damaged and thus are able to offer a sense of firmness or security. This association helps relieve mental tension and therefore can bring about mood improvement. Conversely, the feeling of stability aroused by multicolor stimuli negatively affects their aesthetic valence. As a possible explanation, an object that rarely changes often makes people feel bored and becomes associated with weakness or fragility or, in other words, an impression of lacking vigor, thereby rendering itself unappealing.

2.8 Stability of the Model Training

As introduced in Section 2.6.1 and Section 2.7.1, in Simulation 1, the computational model was trained using a dataset consisting of 30 multicolor stimuli, and, in Simulation 2, the computational model was trained using a dataset consisting of 35 multicolor stimuli. Since the same training procedure was used in the two simulations and the computational model was trained 50 times in either simulation, we compared the convergence states of the training procedure of BPNN 4 in the two simulations to evaluate the stability of the model training. The convergence state of the training procedure of BPNN 4 in Simulation 1 is shown in Figure 2.17, and the convergence state of the training procedure of BPNN 4 in Simulation 2 is shown in Figure 2.18 (which is a duplicate of Figure 2.16).

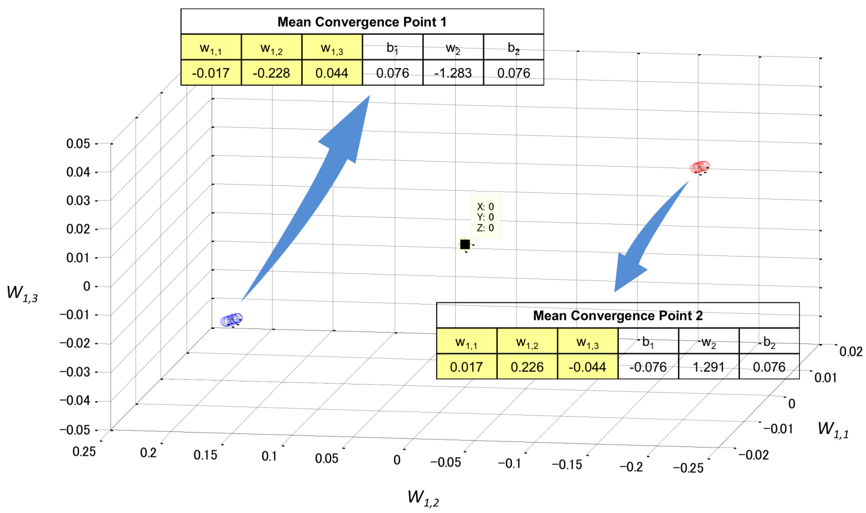

Figure 2.17: Convergence state of the training procedure of BPNN 4 in Simulation 1 (size of training dataset = 30).

Figure 2.17 and Figure 2.18 reveal that, in the two simulations, the training procedure of BPNN 4 converged into almost the same locations in the weight-bias space (Simulation 1:

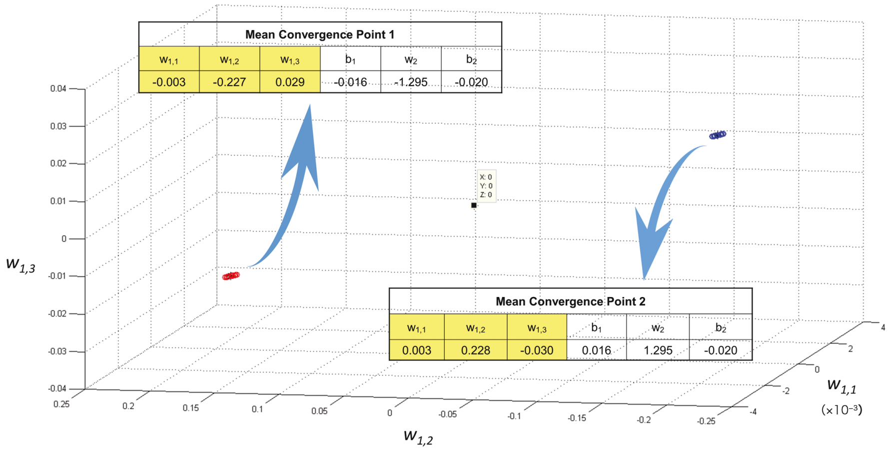

Figure 2.18: Convergence state of the training procedure of BPNN 4 in Simulation 2 (size of training dataset = 35).

Mean Convergence Point 1 = (-0.017, -0.228, 0.044, 0.076, -1.283, 0.076), Mean Convergence Point 2 = (0.017, 0.226, -0.044, -0.076, 1.291, 0.076); Simulation 2: Mean Convergence Point 1 = (-0.003, -0.227, 0.029, -0.016, -1.295, -0.020), Mean Convergence Point 2 = (0.003, 0.228, -0.030, 0.016, 1.295, -0.020)). In view of the fact the training procedure of BPNN 4 was run 50 times in either simulation, this similarity in convergence state can hardly be an accident. This suggests that the training process of the computational model is relatively stable even when the size of training data has changed a little. However, the stability of the model training needs further testing by using training dataset of other sizes. It is also important to test how the prediction accuracy of the trained computational model changes as the amount of training data changes. These testings will be performed in our future studies.

2.9 Summary of Study 1

In Study 1 of the research, we discovered three multicolor-level perceptual features that can influence the aesthetic evaluation of multicolor stimuli—i.e., Stability, Heaviness and Presence —through a psychological experiment using the SD method (Experiment 2). These three perceptual features were extracted as three main factors in this experiment. Another psychological experiment that used the SD method (Experiment 1) was conducted to quantify the aesthetic evaluation of the multicolor stimuli. The aesthetic score of each stimulus is defined as the inverse of its factor score on the factor Pleasure extracted in Experiment 1.

Then, we constructed a computational model that describes the roles of the three multi-color perceptual features in the psychological mechanism of multicolor aesthetic evaluation using BPNNs. Three BPNNs formed the first level of the computational model and described the mappings from the primary color information to each perceptual feature. The fourth BPNN formed the second level of the computational model and depicted the mapping from the three perceptual features to the aesthetic evaluation. For each BPNN, the number of its hidden layer nodes, learning rate, and momentum constant were optimized by the GA. The psychological appropriateness of the computational model was validated through two simulations, especially through Simulation 2.

In Simulation 1, 30 multicolor stimuli used in Experiment 1 and Experiment 2 formed the training data, and the remaining five stimuli constituted the validation data. The model performance evaluated using the NRMSE implied that the computational model may possess certain predicting capacity, but the size of the validation dataset was too small to derive any statistically meaningful conclusion. Moreover, because both the training and validation data were obtained through Experiment 1 and Experiment 2, it is possible that the predicting power of the computational model is limited to the stimuli used in the two experiments.

Simulation 2 addressed these two problems. The training data were composed of all 35 stimuli used in Experiment 1 and Experiment 2, and the validation data consisted of 25 newly generated stimuli. The aesthetic scores of the validation-set stimuli were obtained through a third psychological experiment (Experiment 3). The error-based analysis of the simulation results show that the model has achieved an acceptable degree of prediction accuracy given its hierarchical architecture where error pooling occurs at the second level. Moreover, the prediction error values across the validation-set stimuli follow a normal distribution with an mean value fairly close to zero and a relatively small standard deviation. This implies that our model is able to generalize to some extent the knowledge it learned from the empirical data obtained from the first two experiments to new $4\ times 4$ grid multicolor stimuli. The correlation-based performance evaluation also supports the psychological appropriateness of the model by demonstrating that the model has successfully learned the positive correlation between overall lightness and aesthetic evaluation, which is a psychological effect of considerable importance in the field of multicolor aesthetics. Thus, this computational model can be regarded as a general approximation of the psychological mechanism through which the three multicolor-level perceptual features—i.e., Stability, Heaviness and Presence—influence multicolor aesthetic evaluation.

We then analyzed the post-training synaptic weights and biases of BPNN 4 in Simulation 2. Through this analysis, we clarified in a quantitative fashion how each multicolor-level perceptual feature influences aesthetic evaluation of multicolor stimuli. The analysis shows that, during the psychological process of multicolor aesthetic evaluation, the perceptual feature Heaviness exerts the main influence on the aesthetic evaluation, whereas the other two perceptual features Stability and Presence have only a minor impact. The aesthetic score of a multicolor object increases as the sense of matter presence elicited by the object increases and decreases as the sense of heaviness and (or) stability elicited by the object increases.

Chapter 3 Component-Level Perceptual Features

This chapter introduces Study 2 of the research, that is, the investigation of the component-level perceptual features engaged in the psychological process of multicolor aesthetic evaluation. Section 3.1 introduces the research background and objective of Study 2. Section 3.2 describes the first psychological experiment in Study 2. Section 3.3 describes the second psychological experiment in Study 2. Section 3.4 integrates the results of the two experiments and provides a general discussion of them.

Because the relationship between focality score and memory accuracy score has a relatively large amount of literature as compared to other psychological relationships investigated in Study 2, this relationship is discussed in detail based on the literature at the end of Section 3.3, namely, prior to the general discussion put in Section 3.4. It is also because the discussions about this relationship, together with other content of Section 3.3, were published as a whole in Fang and Matsui's 2017 conference paper[14].

3.1 Background and Objective

[1]Language provides itself as a means for conceptualizing our color sensations. Every language contains a set of basic color terms in its lexicon, such as *black*, *white*, *red*, *green*, *blue*, *yellow*, *brown*, *gray*, *orange*, *pink* and *purple* in the English language[4]. The categories signified by the basic color terms (called "basic color categories" for short), which are natural

[1]Parts of the literature review in Section 3.1 was published in Fang and Matsui's 2017 conference paper[14] and is to be published in a research paper authored by Fang and Matsui[15] .

categories, have their inner structures formed around their prototypes. This means that, within a basic color category, the member colors differ in their focality, namely their closeness to the prototype, or in other words, their goodness as a typical example of the category[102, 45].

Martindale and Moore's[64] study showed that their subjects, most of whom were probably English speakers because they were students of the University of Maine, tended to prefer colors with high focality to colors with low focality. This implies that color focality has the ability to influence color preference. Furthermore, Oyama's[92] and Oyama and Miyata (Ito)'s[94] studies found that, for a two-color combination, its preference, namely its aesthetic evaluation, is greatly influenced by the preference for its component colors. These studies suggest the possibility that the focality of the component colors of a multicolor stimulus can influence the aesthetic preference for the multicolor stimulus by first influencing the preference for the component colors. This process is illustrated in Figure 3.1. This means that color focality possibly serves as a component-level perceptual feature that can affect aesthetic evaluation of multicolor stimuli.

However, we noticed that Martindale and Moore[64] only investigated the English language, so we do not know whether the effect of color focality on color preference that they reported also holds for the Japanese language. Thus, we find it important to investigate whether this effect can also be detected in the Japanese language. What is more important is that, in their study, color focality was defined as a discrete variable which had only five levels. This precluded any investigation of the continuous pattern of the focality-preference relationship. To solve this problem, in Study 2 of the present research, we quantify the concept of color focality in a continuous fashion and delve into the continuous relationship between color focality and color preference, that is, how color preference changes gradually along the continuum of color focality. This is the first aim of Study 2.

The second aim of Study 2 is to clarify the psychological variable(s) that mediate the re-

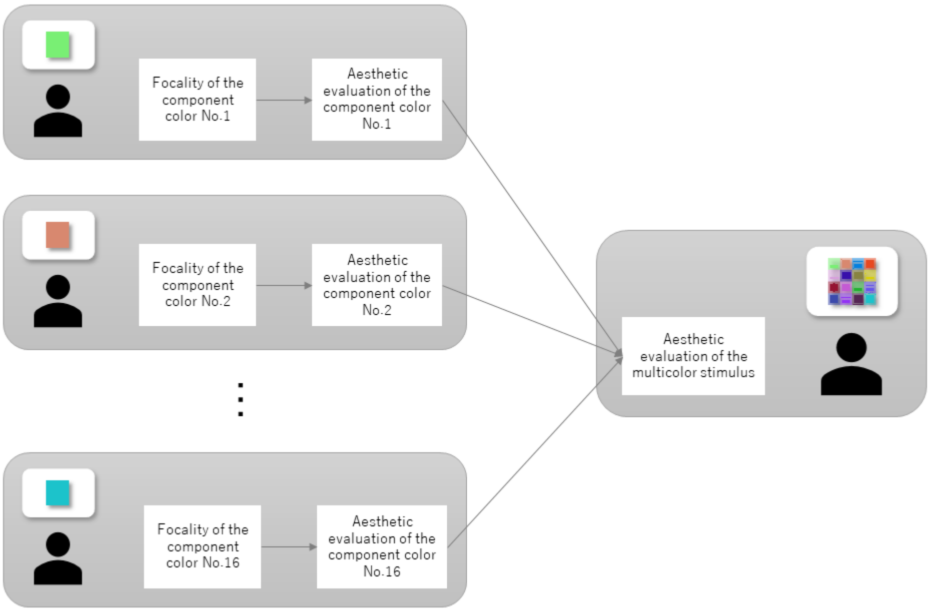

Figure 3.1: Scheme depicting the psychological process that the focality of the component colors of a multicolor stimulus influences the aesthetic preference for the multicolor stimulus by first influencing the preference for the component colors.

lationship between color focality and color preference, or, in other words, how color focality exerts its impact on aesthetic preference of colors. To our knowledge, our research is the first experimental research for this purpose. Two sorts of candidates for such mediating variables were investigated in this research, which are introduced below.

The first is the ”psychological processing fluency (PPF)” of colors. Reber, Schwarz, and Winkielman[98] proposed that the aesthetic preference for a stimulus is largely determined by the fluency of the psychological information processing of the stimulus. Concerning the results of Martindale and Moore’s[64] experiments, Reber et al. argue that the high-focality colors are more likely to be preferred than low-focality colors because the high focality of a color facilitates the information processing of the color. Although this hypothesis, which we call the ”fluency hypothesis,” is theoratically intriguing, it still remains empirically unchecked.

The other candidates for the mediating psychological variables are the color impressions measured by the adjective pair scales that were collected from literature on affective evaluations of colors and used in Experiment 1, except the scale ”harmonious-dissonant.” The scale ”harmonious-dissonant” is excluded because it cannot be applied to single colors.

In brief, the first objective of Study 2 of our research is to investigate the continuous relationship between color focality and color preference, and the second objective is to clarify what psychological vairable(s) mediate this relationship. In Study 2, two psychological experiments are conducted using the Japanese language. The first experiment (Experiment 4) aims to obtain the data of the focality and the PPF of the test colors. The second experiment (Experiment 5) is intended to obtain the data of the aesthetic evaluation and the color impressions of the test colors.

3.2 Experiment 4

3.2.1 Objective

[2]Experiment 4 aims to measure the focality of the color stimuli and quantify the PPF of the color stimuli.

In this research, the short-term memory (STM) performance of the color stimuli is used as the operative indicator of the PPF of the color stimuli for the two reasons: 1) it is generally acknowledged that a stimulus which can be easily processed in our mind is also easy to memorize, or conversely, that a stimulus which is difficult to comprehend is also hard to memorize; 2) The STM of a stimulus reflects multiple presumably fluency-related properties of the stimulus, for example, linguistic codability[7, 59, 60], perceptual discriminability[7], and ecological relevance[66].

[2]The content of Section 3.2 was published in Fang and Matsui’s 2017 conference paper[14] and is to be published in a research paper authored by Fang and Matsui[15].

3.2.2 Experimental Paradigm

Heider[33] developed an experimental paradigm to measure STM performance of color stimuli. She used a simplified version of the color array which was originally developed by Lenneberg and Roberts[57] and later used by Berlin and Kay[4]. The array was composed of 160 Munsell color chips, 24 of which were selected as test chips. Eight of these chips were focal colors, that is, the colors of the highest focality for each of the eight chromatic basic color categories that were shared by numerous languages but generally corresponded to the English categories *Red*, *Green*, *Yellow*, *Blue*, *Brown*, *Purple*, *Pink*, and *Orange, Roberson2000, Roberson2005*. The other 16 chips were of lower focality for these categories, and thus were classified as nonfocal colors. The selection and categorization of the test chips were based on the color-naming data gathered by Berlin and Kay[4]. In each trial in her experiment, a subject was required to watch a color chip for 5 s and then search for it in the color array after a 30-s interval, where the chip was hidden from the subject. For either stimulus type, two indexes of STM performance were measured. The first index was the ”memory accuracy score (MAS),” which was defined as the mean number of correct recognitions for this stimulus type across the subjects. The second index was the ”error distance score (EDS),” which measured the mean error distance across the incorrect trials of this stimulus type. The English-speaking subjects showed superior performance for both measures of the focal colors relative to the nonfocal ones.

Roberson et al.[100] employed the same experimental paradigm and stimuli. Regarding their English-speaking subjects, a focality effect similar to that reported by Heider was detected in terms of MAS. However, no focality effect was found in terms of EDS. Roberson et al[101], which also used this experimental paradigm and these stimuli, found that the mean d' score (a modified version of MAS) of the test chips that were focal only in Himba (a language mainly spoken in Southern Africa) were significantly higher than that of the test chips that were focal only in English. This effect was also detected in the language of Berinmo, which is mainly spoken in Papua New Guinea. The index of EDS was not used in this study. Overall, these studies have provided some evidence for the universal existence of a focality effect across languages in terms of MAS. On the other hand, no robust focality effect has been observed in terms of EDS.

We modified and adopted this experimental paradigm in our Experiment 4. Because MAS is the major measure of STM performance in the literature, we use it to measure the STM performance of the test colors in this research. In other words, MAS is the operative measure of PPF of the test colors in this research.

In addition, we should note that, in the past studies described above, color focality was treated as a categorical variable with only two values: ”focal” and ”nonfocal.” This made any elaborate description of the relationship between color focality and MAS of colors impossible. Therefore, in the present research, we defined the concept of color focality as a continuous variable to delve into the continuous patterns of the focality-MAS relationship, that is, how MAS of colors changes gradually with color focality.

3.2.3 Two Possible Confounding Factors

Mainly two possible confounding factors were examined in the past studies.

The first one is the positive influence of color discriminability on STM performance of colors (called the ”discriminability effect” for short), which was initially reported by Brown and Lenneberg[7]. Heider[33] noticed the possibility that the ”focality effects” that she detected was actually caused by a superiority of the focal colors relative to the nonfocal ones in discriminability. She computed the discriminability scores of the test chips in the array using Brown and Lenneberg’s[7] method. The method calculated the perceptual difference between two colors in the following manner: First, the Munsell coordinate values of the two were transformed into their CIE indexes using O.S.A-developed conversion tables[80]. Then, a color distance value was computed on each of the dimensions of hue, lightness, and saturation. Finally, a weighted sum of the three distance values was given. Since no significant contrast was found between the focal and the nonfocal colors, she argued that this possibility was eliminated.

By contrast, Lucy and Shweder[59] specified the discriminability of the test chips through a psychological experiment. This experiment used the same paradigm that Heider[33] used to explore the relationship between color focality and STM performance of colors, except that the subjects always had the test chip being presented in sight when searching for the chip in the array. The accuracy and latency of the subjects' responses for a color were taken as indicators of the discriminability of the color. The results showed that the focal colors had both a significantly higher response accuracy and a significantly shorter response latency compared to the nonfocal colors, which led them to contend that the discriminability of the focal colors was significantly larger than that of the nonfocal colors. Then, Lucy and Shweder[59] modified the design of the color array to equalize the focal and nonfocal colors in their discriminability measures, namely response accuracy and latency.

Several experiments have been performed to figure out whether the focality effects detected in the past could survive this change in array configuration. The first one was conducted by Lucy and Shweder[59] themselves. They found no significant contrast in MAS between the focal and the nonfocal colors. However, Garro[25] found a significant contrast in all of his experiments using English-speaking participants, and he pointed out that the reason for Lucy and Shweder's[59] failure in detecting this effect might be that they allowed verbal conversations during their experiment. Lucy and Shweder[60] revised their experimental paradigm following Garro's[25] proposal and obtained the same results as in Garro's[25] study. On the other hand, Roberson et al.[100] replicated the results of Lucy and Shweder's[59] experiment for discriminability specification, but could not find a focality effect on STM performance despite using

Lucy and Shweder's[59] revised experimental procedure.

In the face of these contradictory results, Berry et al.[5] pointed out that Lucy and Shweder's[59] modification of the array design, which was based on subjects' response behaviors rather than physical parameters of colors, might have already brought in a bias in subjects' response performance, which could facilitate what subjects were good at and hinder what subjects were poor at, thus making the recognition performance for all the test chips more equal. Poortinga and Van de Vijver[97] agreed with Berry et al.'s[5] view and regarded Lucy and Shweder's[59] experimental results as "not very relevant" to the question whether a focality effect on STM performance exists. It is presumably for this reason that Heider's[33] array design still remains the benchmark one.

The second possible confounding factor, which can distort MAS-related experimental results, is subjects' guessing tendencies during the searching phases when they are unsure about the right answers. Heider[33] raised the concern that if the focal colors were more likely than the nonfocal ones to be selected in the case of guessing, a superiority of the focal colors in MAS relative to the nonfocal ones would appear even if the focal colors were actually no easier than the nonfocal colors to be correctly recognized. She compared the focal and the nonfocal colors in the number of incorrect designations, which indicated the possibility of being selected when the subjects were guessing and found no significant difference. Hence, she concluded that the focal effect in terms of MAS that she detected was not a spurious one that resulted from the biased guessing tendency of the subjects.

Later, Roberson et al.[100], Roberson et al.[101], and Pitchford and Mullen[96] tackled the problem of subjects' guessing biases by replacing MAS with d' score, a measure of recognition performance that derived from the Signal Detection Theory. They adopted the following formula (Eq. 3.1) to calculate the d' scores of each test color, where the hit rate was MAS and the false-alarm rate was the rate of incorrect designations.

$$d' = z(hit\,rate) - z(false - alarm\,rate) \tag{3.1}$$

During the data analyses in our research, both of the possible confounding factors were checked with regard to how they affected the continuous patterns of the relationships between the focality measure and the measures of STM performance.

3.2.4 Subjects

Twenty-two subjects (eleven males and eleven females of ages M = 31.45 and SD = 14.34, native Japanese speakers), who are either undergraduate or graduate students at Waseda University, took part in the experiment. None of them reported having color-related art experience. They all passed the Ishihara Color Vision Test (38 plates, International Edition), and no one reported having color-vision deficiencies. Hence, these subjects were considered to have normal color vision.

We obtained the informed consents of participation from all the subjects. They all participated in both experimental sessions.

3.2.5 Materials and Environment

A color array of Heider's[33] design was used. Its layout is shown in Figure 3.2. This array was made of cardboard (58.5cm * 28.5cm) and had color chips embedded in its white surface. Thirty colors (called "test colors" for short) were tested in the formal trials of Session 1. The locations of these colors are shown in Figure 3.2. These colors were selected as test colors mainly for the following two reasons: (1) These colors formed a large contiguous region on the array, so their degrees of focality could be expected to cover the entire continuum of color focality; and (2) our common sense on the Japanese language showed that these colors could sample a relatively large number of Japanese basic color categories (the specific categories are listed in Section 3.2.6). The test colors were mounted on the white surface of a 5.0cm * 5.0cm piece of cardboard when being presented to the subjects. Chips in the Munsell Book of Color (Glossy Edition) were used, conforming to the convention of past studies.

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
A	5BG 9/2	10BG 9/2	5B 9/2	10B 9/2	5PB 9/2	10PB 9/2	5P 9/2	10P 9/2	5RP 9/2	10RP 9/2	5R 9/2	10R 9/2	5YR 9/2	10YR 9/4	5Y 9/6	10Y 9/6	5GY 9/4	10GY 9/4	5G 9/2	10G 9/2
B	5BG 8/4	10BG 8/4	5B 8/4	10B 8/6	5PB 8/6	10PB 8/4	5P 8/4	10P 8/6	5RP 8/6	10RP 8/6	5R 8/6	10R 8/6	5YR 8/8	10YR 8/14	5Y 8/14	10Y 8/12	5GY 8/10	10GY 8/8	5G 8/6	10G 8/6
C	5BG 7/8	10BG 7/8	5B 7/8	10B 7/8	5PB 7/8	10PB 7/8	5P 7/8	10P 7/8	5RP 7/10	10RP 7/8	5R 7/10	10R 7/10	5YR 7/14	10YR 7/14	5Y 7/12	10Y 7/12	5GY 7/12	10GY 7/10	5G 7/10	10G 7/8
D	5BG 6/10	10BG 6/8	5B 6/10	10B 6/10	5PB 6/10	10PB 6/10	5P 6/8	10P 6/10	5RP 6/12	10RP 6/12	5R 6/12	10R 6/14	5YR 6/14	10YR 6/12	5Y 6/10	10Y 6/10	5GY 6/10	10GY 6/12	5G 6/10	10G 6/10
E	5BG 5/10	10BG 5/10	5B 5/10	10B 5/12	5PB 5/12	10PB 5/10	5P 5/10	10P 5/12	5RP 5/12	10RP 5/14	5R 5/14	10R 5/16	5YR 5/12	10YR 5/10	5Y 5/8	10Y 5/8	5GY 5/10	10GY 5/12	5G 5/10	10G 5/10
F	5BG 4/8	10BG 4/8	5B 4/10	10B 4/10	5PB 4/12	10PB 4/12	5P 4/12	10P 4/12	5RP 4/12	10RP 4/14	5R 4/14	10R 4/12	5YR 4/8	10YR 4/8	5Y 4/6	10Y 4/6	5GY 4/8	10GY 4/8	5G 4/10	10G 4/10
G	5BG 3/8	10BG 3/8	5B 3/8	10B 3/10	5PB 3/10	10PB 3/10	5P 3/10	10P 3/10	5RP 3/10	10RP 3/10	5R 3/10	10R 3/10	5YR 3/6	10YR 3/6	5Y 3/4	10Y 3/4	5GY 3/6	10GY 3/6	5G 3/8	10G 3/8
H	5BG 2/6	10BG 2/6	5B 2/6	10B 2/6	5PB 2/8	10PB 2/10	5P 2/8	10P 2/6	5RP 2/8	10RP 2/8	5R 2/8	10R 2/6	5YR 2/4	10YR 2/2	5Y 2/2	10Y 2/2	5GY 2/2	10GY 2/4	5G 2/6	10G 2/6

Figure 3.2: The layout of the color array (chips marked in Munsell indexes). The test colors are those within the area surrounded by bold lines.

The experiment was performed indoors with fluorescent lighting (type: National FHF 32EX-N-H, daylight color, color temperature: 5000K). The illuminance was 744lx measured by a HIOKI FT3424 lux meter (JIS C 16091 General Class AA, DIN 5032-7 Class B). Because the experiments in relevant previous studies are conducted in natural daylight or fluorescent light that simulates daylight, the results of our experiment can be compared with the results of those previous experiments.

The experimenter and subject being tested sat opposite each other at a table. The distance between the stimuli and the subject's eyes was controlled at 50cm. To separate the two people, a cardboard wall was erected along the middle of the table so the subject would be unable to see the experimenter's face while observing the stimuli, waiting during the 30-s intervals, and filling out the answer sheets.

3.2.6 Procedure

Overall Scheme

The whole experiment consisted of two sessions. Figure 3.3 shows the overall procedure of this experiment. Session 1 aimed to measure the subjects' STM performance for the test colors. Session 2, which immediately followed session 1, was intended to obtain the categorical memberships and the focality of the test colors. One subject was tested at a time.

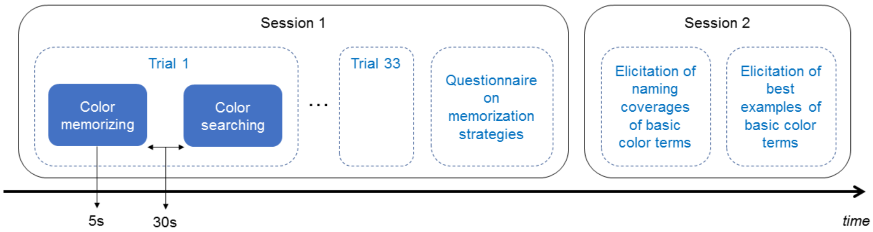

Figure 3.3: Overall procedure of Experiment 4.

Session 1

Session 1 used a procedure similar to that used by Heider[33]. It consisted of 33 trials. In each trial, a test color was presented to the subject for 5 s and then retrieved by the experimenter. After a 30-s interval, the color array was presented to the subject, and the subject was asked to report which color in the array he/she thought was the previously presented one by writing the coordinates of the color on an answer sheet. There was no conversation between the experimenter and the subject during the interval. When a trial was completed, the experimenter retrieved the completed answer sheet to prevent the subject from referencing the previous answers during the following trials. Each test color was tested at least once with each subject, and for each subject, the order of color testing was randomly determined. Thus, for each subject, there were three repeated trials, which were intended to prevent the subject from using a strategy of excluding the already tested colors. Before the formal experiment began, a two-trial training session using a different set of test colors was conducted. For each subject, the colors tested during the training were randomly selected.

After all 33 formal trials were completed, a questionnaire was given to the subject. This questionnaire asked the subject to freely report on the strategies that he/she adopted to memorize the test colors during that session.

Session 2

Session 2 was targeted to elicit the coverages of the six basic color categories corresponding to the six basic color terms *akairo* [*red*], *pinkuiro* [*pink*], *kiiro* [*yellow*], *orenjiiro* [*orange*], *chairo* [*brown*] and *murasakiiro* [*purple*]. Then, the focality of each test color was quantified using a modified version of Berlin and Kay's[4] method. The status of these terms as Japanese basic color terms was substantiated empirically by Uchikawa and Boynton[116]. This session was divided into the following two parts.

First, the subject was required to write on six answer sheets (one for each basic color term) all colors that he/she thought could be named by the term. When a subject completed an answer sheet, the experimenter retrieved the answer sheet before handing him/her the next one to prevent the interference of the past answers that might occur if the subject referred to the already filled-in answer sheets. The answer sheets were provided to each subject in random order.

Next, the subject was asked to report the colors that he/she thought were the best examples of each of the six basic color terms by writing the coordinates of the colors on an answer sheet. Multiple answers were allowed for each basic color term, but the subject was instructed to narrow down his/her choices as much as possible. The detailed instructions of Session 2 were given after the completion of Session 1, which ensured that during Session 1 the subject was ignorant of the involvement of the six basic color terms in this experiment.

3.2.7 Variable Definitions

Focality Score

We used the data obtained from Session 2 to specify the coverages of the six basic color categories over the array and quantified the focality of the test colors.

We first computed the six attributes for each test color: Red Index, Pink Index, Yellow Index, Orange Index, Brown Index, and Purple Index. These attributes measured the intersubject naming consistency of the color in terms of each basic color term. The Red Index of a color was defined by Eq. 3.2, where NS_r was the number of the subjects who named the color as *Akairo* [*Red*], and NS_{total} was the total number of the subjects. The other five indexes were similarly defined.

$$Red\ Index = \frac{NS_r}{NS_{total}} \tag{3.2}$$

Then, we defined the Overall Index (OI) of a color as the largest of the six single-term-based indexes of the color, as Eq. 3.3 shows.

$$Overall\ Index = \max \left\{ \begin{array}{l} Red\ Index,\ Pink\ Index,\ Yellow\ Index, \\ Orange\ Index,\ Brown\ Index,\ Purple\ Index \end{array} \right\} \tag{3.3}$$

We classified a color into the color category *Red* if its OI was its Red Index, the color category *Pink* if its OI was its Pink Index, and so forth. Figure 3.4 shows the distribution of the nonzero OIs and the partition of the six basic color categories.

Table 3.1, which is a summary of the data gathered during the second part of session 2, gives the proportion of the responses in which the colors having large OIs (designated as those $\geqslant 0.80$) were selected as best examples accounts for relative to total responses, in terms of each basic color category. From this table, we can see that, in each category, nearly all the colors that the subjects selected as best examples were the high-OI colors. Thus, it is reasonable to

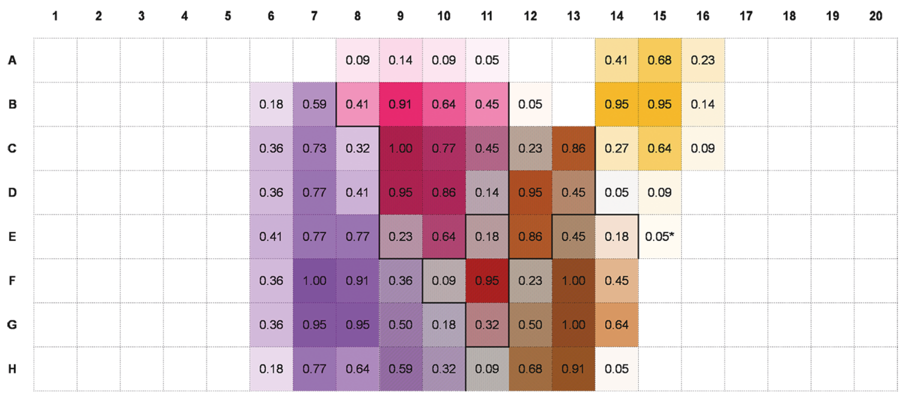

Figure 3.4: Distribution of Overall Indexes of test colors (colors within area covered by thin diagonal stripes) and other relevant colors, and partition of the six basic color categories. Color depth represents Overall Index magnitude. The asterisk (*) denotes that, for that color, Orange Index = Brown Index.

Table 3.1: The percentage of the responses in which the high-OI colors, e.g., the colors whose OIs $\geqslant 0.80$, were selected as best examples (computed for each basic color category).

Basic color category	Percentage of high-OI responses
Akairo [*Red*]	100%
Pinkuiro [*Pink*]	83%
Kiiro [*Yellow*]	100%
Orenjiiro [*Orange*]	100%
Chairo [*Brown*]	81%
Murasakiiro [*Purple*]	75%

deem the OI of a color as reflecting the appropriateness of the color as a typical example of the category to which the color belongs. In this manner, we defined the focality score (FS) of a color as its OI value[3]

Discriminability Score

We defined the discriminability score (DS) of a test color on the color array as the average of the color differences between the test color and its eight adjacent colors. Eq. 3.4 expresses this

[3]The OIs of the chips on the left border of the category *Murasakiiro* do not necessarily signify the categorical membership and the focality of the chips, because the subjects possibly prefer to name them as *aoiro* [*blue*]. So are the chips on the right and the lower borders of the category *Kiiro*, which adjoins the category *Midoriiro* [*Green*].

definition, where $\Delta C_i (i = 1, 2, \cdots, 8)$ is the color difference between the test color and one of its eight neighbor colors.

$$DS = \frac{\sum_{i-1}^{8} \Delta C_i}{8} \tag{3.4}$$

In this research, a color difference was defined as a Euclidean distance in the CIE $L^*a^*b^*$ color space. Thus, before calculating color differences, we transformed the Munsell coordinate values of all relevant colors into the CIE xyY indexes using the O.S.A.-developed conversion tables[80], then the XYZ indexes, and finally the $L^*a^*b^*$ indexes. At the XYZ-to-$L^*a^*b^*$ transformation step, the parameter values of the CIE D50 standard illuminant, which resembled the light source used in this experiment, were used. DS measures the perceptual distinctiveness of a color on the color array. The DSs of the test colors are displayed in Figure 3.5. The DSs of the test colors have a SD of 4.61, which is not a small quantity when compared to the mean (22.437). This confirms the heterogeneity of the test colors in discriminability.

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
A																				
B																				
C									18.46	18.60	22.79	29.59	32.54							
D									19.23	19.44	25.52	29.82	24.23							
E									18.82	20.35	25.18	34.89	22.81							
F									18.80	21.64	25.08	25.50	23.91							
G									18.90	19.81	22.19	24.27	20.99							
H									16.07	16.80	18.28	19.08	19.50							

Figure 3.5: Distribution of discriminability scores of test colors.

Our definition of DS, like Brown and Lenneberg's[7] definition, uses the physical attributes of colors, so the possible interference with subjects' recognition performance that Berry et al.[5] points out (described in Section 1.3) can be avoided. In addition, our definition employs the CIE $L^*a^*b^*$ color space. Perceptual color differences represented by Euclidean distances are largely uniform across this color space, which enables relatively precise and easy computation of perceptual color differences[106]. This color space, which was issued in 1975[106], was unavailable to Brown and Lenneberg at the time of their study.

STM Performance Index (Main): Memory Accuracy Score

We adopted MAS as the major index of STM performance, which we used to measure the PPF of the test colors. MAS measures the probability for which a color can be accurately recognized. Since the variable FS is continuous in our research, rather than categorical as in the past studies, it is necessary to take the variable MAS also as continuous.

We defined the MAS of a test color as the percentage of the trials in which the subjects correctly recognized the color.

$$MAS = \frac{NT_{correct}}{NT_{total}} \tag{3.5}$$

Eq. 3.5 expresses this definition, where $NT_{correct}$ is the number of trials in which the test color is correctly recognized, and NT_{total} is the total number of the trials in which the test color is presented to the subjects as the probe. The MASs of the test colors are displayed in Figure 3.6.

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
A																				
B																				
C									0.72	0.28	0.54	0.91	0.96							
D									0.65	0.50	0.58	0.73	0.39							
E									0.42	0.79	0.67	0.96	0.40							
F									0.56	0.88	0.87	0.63	0.65							
G									0.54	0.56	0.52	0.50	0.62							
H									0.43	0.25	0.64	0.67	0.83							

Figure 3.6: Distribution of memory accuracy scores of test colors.

STM Performance Index (Minor): Error Distance Score

The EDS is used as another index of STM performance. EDS measures the expected error extent in the case of misrecognition. As in Heider's[33] and Roberson et al.'s[100] studies, the EDS for a test color is defined as the mean of the color differences between the test color and the colors mistaken for the test color in the incorrect recognition trials. Eq. 3.6 expresses this definition, where m is the number of the trials in which the test color is the probe but is not selected at the searching phase, and $\Delta C_j (j = 1, 2, \cdots, m)$ is the color difference between the test color and the selected color in the jth incorrect designation trial of the test color.

$$EDS = \frac{\sum_{j=1}^{m} \Delta C_j}{m} \tag{3.6}$$

The EDSs of the test colors are displayed in Figure 3.7.

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
A																				
B																				
C									16.64	15.18	20.84	22.74	27.96							
D									16.07	18.52	19.97	25.45	15.00							
E									18.03	14.04	30.02	30.20	16.17							
F									21.26	24.40	18.08	28.72	27.23							
G									21.15	17.61	18.86	18.57	15.64							
H									16.97	16.40	20.88	14.99	12.21							

Figure 3.7: Distribution of error distance scores of test colors.

3.2.8 Statistical Analyses on Relationship Between Color Focality and PPF of Colors

Synopsis

To know the continuous pattern of the relationship between color focality and PPT of colors, which is measured by MASs of colors, we first conducted regression analyses on the FS data and the original MAS data of the test colors to obtain a general impression of the relationship pattern. Then, we examined whether the subjects' guessing behavior distorted this pattern, and finally we performed refined regressions on the FS data and the MAS data with the influence of the discriminability effect, which was regarded as a possible confounding factor, controlled.

Initial Linear and Quadratic Regressions

The initial regressions, which were carried out on the FS data and the original MAS data, show no statistically significant linear relationship between FS and MAS (R^2 = 0.066, P = 0.171 [B_{FS} = 0.152, P = 0.171]), but they do show a significant quadratic relationship between the two variables (R^2 = 0.237, P = 0.026 [B_{FS} = -1.064, P = 0.045; B_{FS*FS} = 1.073, P = 0.021]). The results are plotted in Figure 3.8A. We chose to use quadratic regression besides linear regression because we tried polynomial regressions from second to sixth order and found that the quadratic one had the smallest Bayesian information criterion (BIC) value.

Confounding Factor Check 1: Test of the Guessing Hypothesis

There exists the possibility that the quadratic pattern came from a tendency of the subjects to select the color with the highest or lowest degree of focality when they felt unsure about which color was the right answer within a group of candidates. We call this possibility the "guessing hypothesis." To check this hypothesis, not only the hits, which were measured by MAS, should

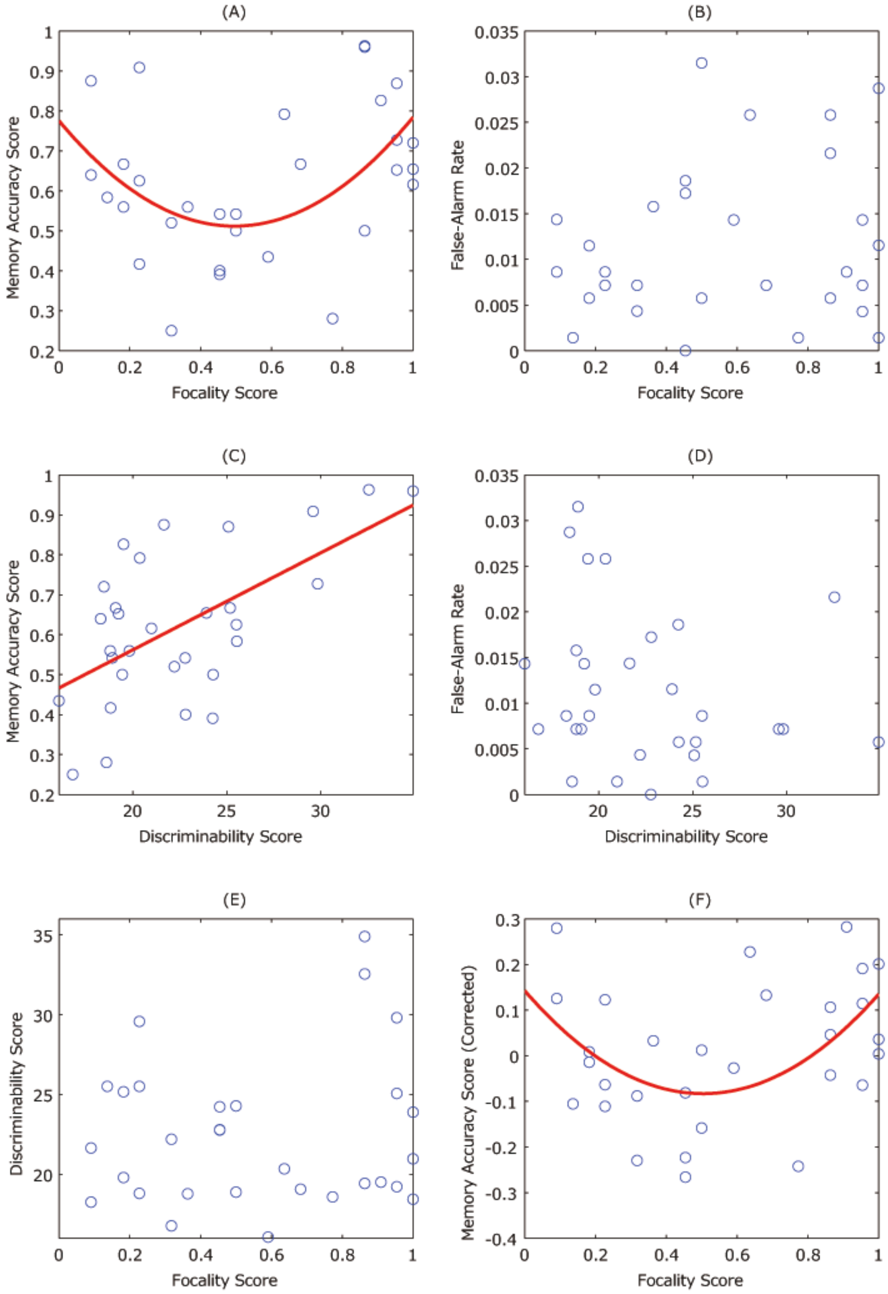

Figure 3.8: Results of the data analyses for investigating the relationship between color focality and correct recognition possibilities of colors: Plots of the data points and the statistical significant regression models in terms of (A) focality score and memory accuracy score, (B) focality score and false-alarm rate, (C) discriminability score and memory accuracy score, (D) discriminability score and false-alarm rate, (E) focality score and discriminability score, and (F) focality score and corrected memory accuracy score.

be considered but also the false alarms. We defined the false-alarm rate (FAR) for a test color as the number of the trials in which the test color was erroneously selected divided by the number of the trials in which the test color was not the actual presented one. Similar to Heider's[33] method, we looked into the relationship between FS and FAR for the test colors. The result of the quadratic regression analysis, which is plotted in Figure 3.8B, failed to achieve statistical significance (R^2 = 0.068, P = 0.386 [B_{FS} = 0.030, P = 0.245; B_{FS*FS} = -0.022, P = 0.318]). This implies that the colors with the highest or lowest focality were no more likely than the medium-focality colors to be selected when the subjects were guessing. Thus, the guessing hypothesis was ruled out.

As described in Section 3.2.3, Roberson et al.[100], Roberson et al.[101], and Pitchford and Mullen[96], and approached the problem of guessing bias by calculating the d' score of each test color using Eq. 3.1. However, according to the Signal Detection Theory (SDT), the application of this equation is restricted in the Yes-No experimental paradigm[27, 62]. This means that the equation is unsuitable for the m-alternative forced choice (m-AFC) experimental paradigm with oblique coordinates that these three studies and our research have adopted. Moreover, with regards to the m-AFC paradigm with oblique coordinates, no response behavior model and, therefore, discrimination measure have been established in SDT so far. Thus, in the present research, we did not choose the SDT approach.

Confounding Factor Check 2: Elimination of the Discriminability Effect

The next step is to check whether the influence of FS on MAS is just a spurious correlation caused by the discriminability effect.

First, we looked into the relationship between DS and MAS. A significant positive linear regression model could be established between these two variables (R^2 = 0.353, P = 0.001 [B_{DS} = 0.024, P = 0.001], which is plotted in Figure 3.8C). The guessing hypothesis was tested using the same method as in the case of FS. No significant linear relationship was found between DS and FAR (R^2 = 0.042, P = 0.275 [$B_{DS} < 0.001$, P = 0.276]; plotted in Figure 3.8D), which indicated that the linear model established between DS and MAS was not an artifact caused by the subjects' biases in guessing.

Then, the regression analyses studying the relationship between FS and DS were conducted, producing neither a significant linear model (R^2 = 0.014, P = 0.534; [B_{FS} = 1.710, P = 0.534]) nor a significant quadratic model (R^2 = 0.029, P = 0.675 [B_{FS} = -7.032, P = 0.618; B_{FS*FS} = 7.712, P = 0.527]). Nevertheless, we noticed that a slight U-shaped relationship could be recognized when we scrutinized the scatter plot (Figure 3.8E). This means that the possibility that DS mediated the FS-to-MAS relationship could not be ruled out.

Hence, we conducted partial linear and quadratic regressions on FS and MAS while partialing out the influence of DS on MAS. No significant linear relationship was found (R^2 = 0.054, P = 0.217 [B_{FS} = 0.111, P = 0.217]), but a significant quadratic one was detected (R^2 = 0.233, P = 0.028 [B_{FS} = -0.893, P = 0.037; B_{FS*FS} = 0.885, P = 0.018]), which was similar to the results of the initial regressions. This means that a significant quadratic relationship exists between FS and MAS even if DS has been treated as a control variable. The results are plotted in Figure 3.8F.

Summary

A significant U-shaped quadratic relationship between FS, which measures color focality, and MAS, which measure PPF of colors, showed up even when the impact of DS was controlled. In addition, no significant linear or quadratic relationship could be found between FS and FAR, which falsified the guessing hypothesis.

3.2.9 Statistical Analyses on Relationship Between Color Focality and Misrecognition Error Distances of Colors (Minor Measure of STM Performance of Colors)

Synopsis

The continuous pattern of the relationship between color focality and misrecognition error distances of colors, which is a minor measure of STM performance of colors, was also investigated. The first step was to obtain a preliminary understanding of what the relationship pattern looks like by carrying out linear and quadratic regressions on the FS data and the original EDS data. Then, these regressions were repeated but with the impact of DS controlled, which was intended to remove the possible distorting influence of the discriminability effect.

Initial Linear and Quadratic Regressions

The initial regressions, which were run on the FS data and the original EDS, produced no significant linear model (R^2 = 0.110, P = 0.084 [B_{FS} = -4.694, P = 0.084]), but did produce a significant quadratic one (R^2 = 0.250, P = 0.027 [B_{FS} = -30.113, P = 0.019; B_{FS*FS} = 22.376, P = 0.041]). The results are plotted in Figure 3.9A.

We chose to use quadratic regression besides linear regression because we tried polynomial regressions from second to sixth order and found that the quadratic one had the smallest BIC value.

Confounding Factor Check: Elimination of the Discriminability Effect

Then, a positive linear relationship was found between DS and EDS through a regression analysis (R^2 = 0.299, P = 0.003 [B_{DS} = 0.692, P = 0.003], which is plotted in Figure 3.9B). To remove the possible distorting influence of the discriminability effect (as in the case of

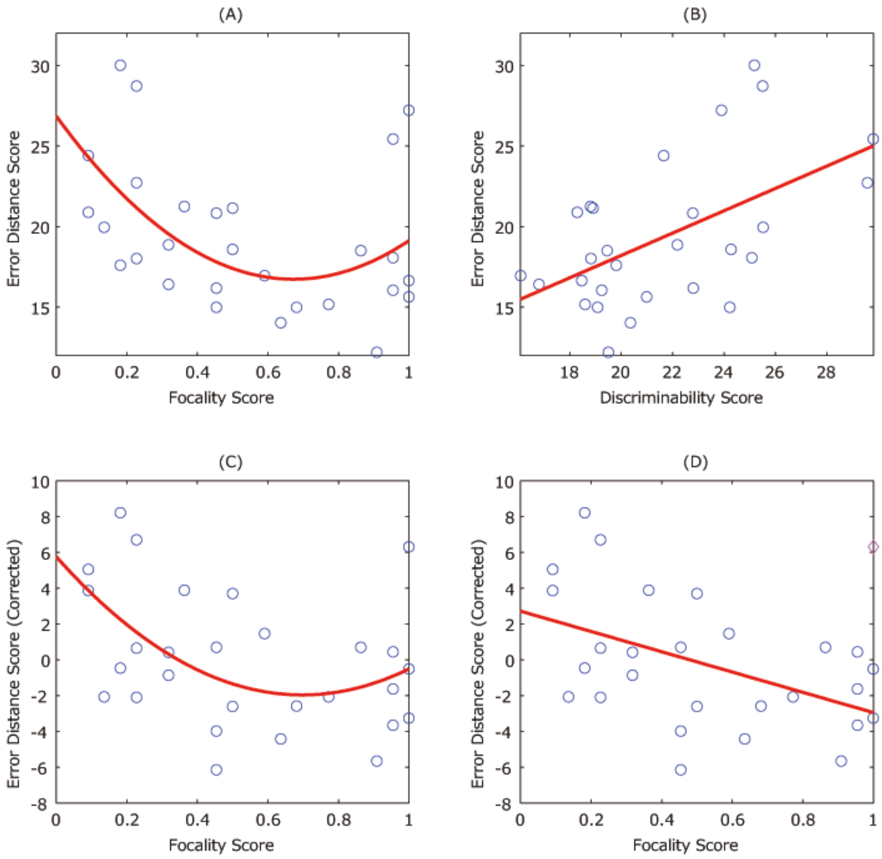

Figure 3.9: Results of the data analyses for investigating the relationship between color focality and misrecognition error distances of colors: Plots of the data points and the statistical significant regression models in terms of (A) focality score and error distance score, (B) discriminability score and error distance score, (C) focality score and corrected error distance score, and (D) focality score and corrected error distance score (the general-case pattern, with the data point unexplained by the regression model represented by a magenta-colored rhombus).

MAS), we partialed out the influence of DS on EDS and ran linear and quadratic regressions on the FS data and the refined EDS data. A significant quadratic model showed up (R^2 = 0.221, P = 0.044 [B_{FS} = -22.166, P = 0.041; B_{FS*FS} = 15.895, P = 0.085]), but not a significant linear one (R^2 = 0.121, P = 0.070 [B_{FS} = -4.110, P = 0.070]), which resembles the results of the initial regressions. The results are plotted in Figure 3.9C.

Nevertheless, there exists a test color that appears to be isolated from the cluster of other high-FS test colors at the EDS coordinates. Owing to the employment of the least squares method, this data point could have exerted a disproportionately strong influence on the relationship pattern. To determine what pattern the relationship actually takes in the general case, we reran the regressions on the corrected dataset but did not include this data point. This time we obtained a linear relationship (R^2 = 0.236, P = 0.010 [B_{FS} = -5.669, P = 0.010], which is plotted in Figure 3.9D) instead of a quadratic one. (When adding $FS * FS$ to the regression as a predictive variable, neither B_{FS} nor B_{FS*FS} achieved significance, although the model remained significant). Because this linear regression model is free from the effects of extreme data and therefore represents the general-case relationship between the FS and the corrected EDS, the discussions on FS-EDS relationship in Section 3.2.11 are based on this model.

Summary

With the EDS data corrected by eliminating the confounding influence of DS, a statistically significant negative linear tendency appears between FS and EDS under the general circumstance.

3.2.10 Discussion 1: The General Pattern of the Relationship Between Color Focality and PPF of Colors

Our experiment results demonstrated that, in the Japanese language, color focality is able to affect PPF of colors, which is measured by the MASs of colors in a statistically significant way. Specifically speaking, a significant U-shaped quadratic regression function can be established between FS and MAS, which implies that the PPF is best for colors at the two terminals of the focality continuum and decreases as the focality moves toward the intermediate level.

This pattern was computed with the influence of the discriminability effect controlled, since this experiment verified the existence of the discriminability effect. Specifically, we confirmed the positive impact of color discriminability on the MASs of colors. This tallies well with the experimental result of Brown and Lenneberg's[7] study, which, like our research, used the physical attributes of color to define color discriminability.

In addition, we investigated the continuous pattern of the relationship between the color focality and the EDS of colors, although EDS is unfrequently used in this area. We found that a significant negative linear regression function can be established between FS and EDS under the general circumstance. This suggests that the average error extent in the case of misrecognitions for a color decreases as its focality increases. We also found that color discriminability positively influences EDS. No past study has ever probed the existence of the discriminability effect in terms of EDS.

3.2.11 Discussion 2: Memorization Strategy of Linguistic Categorical Coding as One Cause of the Relationship Between Color Focality and PPF of Colors

How Linguistic Categorical Coding is Applied

To determine what caused the continuous pattern of the relationship between color focality and PPF of colors, we examined the memorization strategies reported by the subjects (Table 3.2). We noticed that the strategy of encoding colors using linguistic color categories (called ”linguistic categorical coding” for short), which has the highest number of reports, might have played an important role in the formation of the detected U-shaped pattern.

A detailed description of the procedure of this strategy in a single trial is as follows: The subjects consciously encoded the test color using the basic color terms while observing the test color. The basic color terms were used as reference points, which means that the subject anchored the test color to the central points of the basic color categories, namely, the most typical colors of these categories. The role of prototype colors as reference points has also been demonstrated in the task of hedge-sentence fulfilling and the task of spatial distance judgement[103]. The subject then retained this linguistic code in his/her STM during the waiting period. Finally, during the phase of color searching, the subject decoded the code to recover the test color.

For convenience of discussion, color focality is generally divided into the levels of ”high,” ”medium,” and ”low,” and their respective ways of being coded are described as follows.

High-focality colors

A high-focality color can be encoded using only one basic color term since it is, or is substantially close to, the central point of the basic color category.

Medium-focality colors

Coding a medium-focality color needs some modifiers in hue, saturation,

Table 3.2: The memorization strategies reported by the subjects in the questionnaire.

Name	Number of reports	Brief description
Linguistic categorical coding	16	Use basic color concepts as reference points, and then fine-tune along the dimensions of hue, lightness, and/or saturation
Absent object association	14	Associate the test color with the color of a familiar object in memory, e.g., the banner of Waseda University, a Bordeaux wine, or lipstick, and then fine-tune along the dimensions of hue, lightness, and/or saturation
Direct retention of visual image	3	Directly memorize the visual image of the test color
Preference evaluation	2	Use the degree of preference for the test color as a cue
Present object association	1	Use the color of an object located in the experimental environment, e.g., an answer sheet, as a reference point, and then fine-tune along the dimensions of hue, lightness, and/or saturation
Reference on past test colors	1	Use a previously presented test color as a reference point, and then fine-tune along the dimensions of hue, lightness, and/or saturation

and/or lightness besides a basic color term. For example, in the questionnaire, the subjects reported having used the codes ”*azayakasugiru* [*very saturate* (a modifier in saturation)] *orenji* [*orange* (a basic color term)],” ”*usui* [*light* (a modifier in lightness)] *chairo* [*brown* (a basic color term)],” and ”*sukoshi kiroi* [*a bit yellowish* (a modifier in hue)] *pinku* [pink (a basic color term)].”

Low-focality colors

Because a low-focality color is situated at the border region between two basic color categories, the two basic color terms corresponding to the two categories are used to constitute the code for this color. In other words, a subject only needs to memorize these two terms when encoding the color and focus on the border of these two categories when looking for the color. As an example, a subject reported having used the basic color term *orenji* [*orange*] and the basic color term *pinku* [*pink*] to encode a test color.

How Color Focality Affects MAS-Measured PPF of colors Through Linguistic Categorical Coding

With the employment of this strategy, it is obvious that the MAS of a color, which indicates the PPF of the color, is mainly determined by 1) how easily the code for the color can be retained in STM during the waiting period and 2) the semantic ambiguity of the code for the color, or, in other words, how accurately the encoded color can be recovered from the code. Since codes for colors of all three types can be formed by just a few words, they will not cause a memory burden. This implies that the rate of successful retention should be high for each color type. On the other hand, the variable semantic ambiguity bears a much larger intertype variance, which indicates its chief role in mediating the impact of color focality on PPF of colors. The semantic ambiguity of color codes in the three situations of high-focality colors, medium-focality colors, and low-focality colors are described, respectively, as follows.

High-focality colors
The code of a high-focality color generally consists of a sole basic color term, which possesses a fairly plain meaning since any Japanese speaker is able to understand what a basic color term means. Thus, during the searching phase, the decoding procedure can be done in high fluency, or in other words, the signifier of the code, namely, the coded color, can be pinpointed in high precision.

Medium-focality colors
For a medium-focality color, the modifiers in its code have much vaguer meanings. Even if the subjects have carried the code into the decoding phase without mistakes, they will find themselves lost in numerous possible answers, all of which more or less match the description. This will surely lower their chance of finding the color that they had actually coded. In other words, the vague linguistic meanings of the codes reduced the fluency of the mental processing of the codes.

Low-focality colors The code of a low-focality color is unequivocal in its meaning because the code contains basic color terms but no modifiers. The central points of the basic color terms, as in the case of a high-focality color, can serve as reliable reference points for the localization of the encoded color. This implies that codes of low-focality colors can be used in high fluency, such as in decoding during the color searching phase in our Experiment 4.

In brief, the semantic ambiguity of color codes, which negatively influences the fluency of the mental processing of the codes and thus the likelihood of correct recognition (i.e., the MASs) of colors, is low for high- and low-focality colors and high for medium-focality colors. Thus, high- and low-focality colors tend to have higher MASs than those of medium-focality ones. This is exactly what our experimental results showed.

In addition, because for any color, the semantic ambiguity of its code is a language-inherent and thus subject-independent attribute, this continuous pattern can be expected to have a high degree of intralanguage consistency, or, in other words, a high likelihood to be replicated if the experiment is repeated using the same language.

The findings of the past studies that explored the relationship between language usage and the effect of color focality on MASs of colors support our argument that subjects' conscious encoding of colors using linguistic terms is one vital cause of this effect. Brown and Lenneberg[7] showed that the linguistic codability of colors possesses a strong positive influence on the MASs of colors in the English language. Also using English-speaking subjects, Garro[25] and Lucy and Shweder[60] demonstrated that the superiority of the focal colors relative to the nonfocal colors in MAS vanished when verbal communication was allowed during the waiting intervals. One explanation that Lucy and Shweder[60] proposed is that the conversations interrupted the subjects' ability to retain the color codes in their STMs, which prevented the recognition of high-focality colors from benefiting from their advantageous verbal encoding. While these studies place their emphases on the relationship between the strategy of linguistic color coding and the MASs of colors, our hypothesis mainly concerns how color focality exerts its impact through this strategy on PPF of colors.

How Color Focality Affects EDSs of Colors Through Linguistic Categorical Coding

It is probable that the continuous pattern of the relationship between the color focality and the EDSs of colors also derived from the memorization strategy of linguistic categorical coding.

We can know that the continuous relationship pattern itself is unstable for the reason explained later in this subsection. Considering within the framework of linguistic categorical coding, the EDS of a color mainly depends on which parts of the code the subjects have forgotten, and how many times each of these parts have been forgotten. How these factors vary across the three situations of high-focality colors, medium-focality colors, and low-focality colors is described in the following paragraphs.

High-focality colors For a high-focality color, once a subject has forgotten the sole basic color term during the waiting period, in the searching phase he/she is unable to tell the basic color category to which the test color belongs. His/her selection will thus be random, although other memory clues, such as the visual image of the test color, can be of help. It is easy to imagine that, under this circumstance, a large error will occur.

Medium-focality colors

For a medium-focality color, when only the modifiers have been forgotten, given that the basic color term has become the only guide, the central point of this basic color category may pull the subjects' selections toward it. In this case, a misrecognition is expected to occur, but within a moderate error range that is approximately half the "category radius." On the other hand, the loss of the basic color term may lead to a much larger error distance, as in the case of high-focality colors.

Low-focality colors With regard to the code of a low-focality color, when one of its two basic color terms has been forgotten, the remaining one will tend to drag the subjects' selections toward the central point of the category it represents. On this occasion, because the low-focality color is situated at the border region of the category, a selection with an error distance of approximately one category-radius long might take place. On the other hand, the loss of both basic color terms may result in a much larger error distance, as what will happen when the sole basic color term is forgotten for a high-focality color.

Note that owing to the small total number of memory losses suggested by the small memory burden imposed by the color codes, it is possible that some of these ”forgetting types” did not occur in our experiment. Thus, one explanation for the continuous pattern of the focality-EDS relationship that we detected is that our subjects have never forgotten the basic color terms in the codes for the high-[4] and medium-focality colors. In addition, the small sample size of memory losses means that the distribution of occurrence frequency across the forgetting types can hardly be consistent across experiments even when using the same language. In other words, if the experiment is repeated, a substantially different frequency distribution across the forgetting types will occur, which will lead to a very different pattern of the focality-EDS relationship. Perhaps this susceptibility to the unpredictability in subject behavior is one reason why EDS was employed much less frequently than MAS in past studies.

3.2.12 Discussion 3: Expected Universality of the Relationship Between Color Focality and PPF of Colors

Several past studies on STM performance of colors, which used English-speaking subjects, also recorded their subjects' memorization methods.

Lucy and Shweder[60] recorded the subjects' incidental remarks on memorization strategies during the course of their experiments, and they carried out a questionnaire on memorization

[4] As mentioned in Section 3.2.9, a data point representing a high-focality test color (5YR 4/8, depicted by a magenta-colored rhombus in Figure 3.9D) appears separated from the regression model. The test color was mistaken as the color one-unit above it (5YR 5/12) in all its misrecognition cases. This misrecognition pattern is difficult to explain by the strategy of linguistic categorical color coding. We thus conjecture that it might result from other memorization strategies, which needs further exploration.

Table 3.3: Comparison of the memorization strategies used by subjects in Lucy and Shweder's[60] study with the memorization strategies recorded in the present research.

	Lucy and Shweder [Incidental remarks]	Lucy and Shweder [Follow-up questionnaire]	The present research [Follow-up questionnaire]
Linguistic categorical coding	1 (50%)	1 (70%)	1 (73%)
Absent object association	3 (10%)	4 (20%)	2 (64%)
Direct retention of visual image	3 (10%)	2 (60%)	3 (14%)
Present object association	2 (40%)	3 (50%)	5 (5%)

strategies when the experiments were finished. They provided a quantitative report of their findings, which is summarized in Table 3.3 with comparison to the findings of our research. In Table 3.3, each cell shows the rank of the memorization strategy in terms of proportion of the subjects who reported having used it. The proportion is shown in brackets beside the rank. This table shows that the strategy of linguistic categorical coding was the most frequently adopted, followed by the strategies of direct retention of visual image, present object association, and absent object association. This coincides well with the results of our questionnaire.

Brown and Lenneberg[7], Lucy and Shweder[59], and Garro[25] also reported the memorization strategies used by their subjects, although they did not provide detailed statistics. Brown and Lenneberg[7] mentioned that their subjects transformed the colors that they had to remember into their ”names” and stored the names in their memory. In our view, this method generally encompasses the strategies of linguistic categorical coding and object association. Lucy and Shweder[59], by examining their subjects' incidental comments, found that linguistic categorical coding, absent object association, and present object association were the three most frequently used memorization strategies. The direct retention of visual image was used as a method of supplementation for linguistic categorical coding. The recordings of Garro's[25] post-experiment questionnaire reveal that their subjects mainly employed the strategies of linguistic categorical coding, direct retention of visual image, and object association. We can see that all three studies reported the usage of linguistic categorical coding.

The fact that linguistic categorical coding is employed as a chief memorization strategy by both Japanese speakers and English speakers suggests that its applicability is possibly universal across languages. Moreover, considering the likely close ties of this strategy to the formation of the continuous pattern of the focality-PPF relationship, this further implies that all languages may share a common language-based mechanism for the generation of the relationship between color focality and PPF of colors[5].

Regarding the continuous pattern of the relationship between color focality and PPF of colors, given the likely intralanguage consistency (explained in the subsection "How Color Focality Affects MAS-Measured PPF of colors Through Linguistic Categorical Coding" in Section 3.2.11) of this pattern, we can expect to also observe this pattern in other languages because of the likely universal applicability of linguistic categorical coding.

This speculation is empirically supported by the agreement between the continuous FS-to-MAS (PPF) relationship detected in our experiment and the superiority of focal colors to nonfocal colors in correct recognition possibility reported by Heider[33], Roberson et al.[100], and Roberson et al.[101]. Specifically, their definition of focal colors generally corresponds to the colors rated high on our focality continuum, which have high MASs according to our experimental results, and their definition of nonfocal colors covers the low region of our focality continuum, which have high MASs in our research, along with the medium region, which have low MASs in our research. Thus, if we bisect our focality continuum using these defini-

[5]Reports show that there exit languages that possibly lack basic color terms, e.g., Piraha[13] and Warlpiri[118] (but see Regier, Kay, and Khetarpal's 2009 study[99]). This means that we can hardly apply the linguistic definition of color focality and the memorization strategy of linguistic categorical coding to such languages. Thus, the discussions in this section are probably unsuitable for these languages.

tions and compare the two categories of colors in MAS based on our experimental results, a focal-color superiority will show up, just as in these previous studies.

3.2.13 Discussion 4: Expected Universality of the Relationship Between Color Focality and EDSs of Colors

Since the continuous pattern of the relationship between the color focality and the EDSs of colors lacks consistency inside the language being tested (explained in the subsection ”How Color Focality Affects EDSs of Colors Through Linguistic Categorical Coding” in Section 3.2.11), finding a continuous pattern that is consistent across languages for this relationship must be even harder.

The comparison of the results of Heider’s[33] and Roberson et al.’s[100] experiments and the present experiment supports this speculation. The results of Heider’s[33] experiment show that colors that have higher degrees of focality tend to have lower EDSs. This agrees with the negative linear FS-to-EDS relationship detected in our experiment. In contrast, Roberson et al.[100] found no significant difference between the focal colors and the nonfocal colors in EDS.

3.2.14 Summary of the Findings of Experiment 4

Through Experiment 4, we quantified the PPF, which was measured by MAS, and focality of the test colors. Regression analyses show that a U-shaped quadratic relationship exists between focality and PPF, even when the influence of the subjects’ guessing behaviors and the Discriminability Effect were removed. Specifically speaking, PPF is highest at both ends of the continuum of color focality and decreases as color focality moves toward the medium region from either end. From the results of a questionnaire that recorded the memorization strategies used by the subjects, we infer that the subjects’ frequent and conscious employment of the memorization strategy of using linguistic color categories to code the test colors may be one crucial cause of the detected focality-PPF relationship. Because the continuous pattern of this relationship tallies well with experimental findings of the past studies testing the English language, this continuous pattern is likely universal across languages.

We also tested EDS-a less important index of STM performance which was used in some of the previous studies. Regression analyses show that, in general, EDS decreases as focality increases. This relationship pattern can also be explained by the use of the memorization strategy of linguistic categorical coding. However, due to the lack of robustness in this experimental paradigm, it is difficult to expect this relationship pattern to show universality across languages, which is also suggested by the inconsistency among the experimental results of our research and those of previous studies testing the English language. Probably for this reason, EDS remains a minor index of STM performance compared to MAS.

3.3 Experiment 5

3.3.1 Objective

[6]This experiment aims to quantify the aesthetic evaluation of test colors and obtain their evaluation data on the 22 color impressions.

3.3.2 Subjects, Materials, and Environment

Totally, 32 subjects (15 males and 17 females of ages M = 30.34 and SD = 14.23, native Japanese speakers), who were either undergraduate or graduate students at Waseda University, participated in the experiment. They all passed the Ishihara Color Vision Test (38 plates, International Edition), and none reported having color-vision deficiencies. Hence, these subjects

[6]The content of Section 3.3 is to be presented as a conference paper authored by Fang and Matsui[16].

were considered to have normal color vision. We obtained informed consent for participation from all the subjects. They all participated in both experimental sessions.

The evaluation data of three subjects were excluded from the data processing because they reported having color-related art experience. In this manner, the conditions of the subjects in this experiment match those in Experiment 4 in which no subject reported having color-related art experience. The evaluation data of the remaining 29 subjects (14 males and 15 females of ages M = 28.41 and SD = 12.24) were used in the data processing.

The stimuli and the environment, including the lighting condition, were the same as in Experiment 4.

3.3.3 Procedure

The experiment comprised 30 trials. In each trial, a test color was presented to the subject, who was asked to rate it on 22 adjective pair scales, which represented the 22 color impressions, and another adjective pair scale ”like-dislike,” which measured the degree of preference. Although no time limitation was imposed on the subjects’ ratings, they were asked to provide ratings without deliberation. To remove the afterimages, we presented a gray chip (N 5.5/) for 30 s between every two trials and asked the subjects to look at it when it was presented. There was no conversation between the experimenter and the subject during the ratings. For each subject, each test color was tested once, and the test colors were tested in a random order. The procedure is shown in Figure 3.10.

Before the formal experiment began, a training trial was conducted. A gray chip (N 5.5/) was used as the test color in this trial because it had the middlemost value (i.e., degree of lightness) among the achromatic chips in the Munsell Book of Colors (Glossy Edition).

The online survey tool SurveyMonkey (https://jp.surveymonkey.com/) was used as the platform for the ratings. The interface was set to be achromatic.

Figure 3.10: Overall procedure of Experiment 5.

The reason we used this procedure is because we adopted a Likert scale-based definition of color preference, which will be described in the next section (Section 3.3.4). Because we wanted to investigate the relationships between each color impression and color preference, color impressions had to be measured using the same method that was used to measure color preference.

3.3.4 Variable Definitions

Preference Score

The continuum of the adjective pair scale ”like-dislike” was divided into seven degrees on the SurveyMonkey answer sheets. We quantified these seven degrees by transforming them into the integers ”1,” ”2,” ”3,” ”4,” ”5,” ”6,” and ”7” from the left end to the right end. The preference score (PS) of each test color is defined as the average of the evaluation scores of the test color across the subjects. This definition is expressed in Eq. 3.7, where N is the total number of the subjects, $k(k = 1, 2, ..., N)$ is the subject index, and PE_k is the evaluation score rated by the kth subject on the preference scale. We chose to use this Likert scale-based definition because it was used by most previous studies on color preference, including Martindale and Moore’s

1988 study[63].

$$PS = \frac{\sum_{k=1}^{N} PE_k}{N} \tag{3.7}$$

The preference scores of the test colors are shown in Figure 3.11.

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
A																				
B																				
C									4.52	4.93	4.48	4.86	4.52							
D									4.34	4.72	4.97	4.69	4.93							
E									4.45	4.24	4.97	4.59	3.38							
F									4.48	4.55	4.90	4.10	3.14							
G									4.76	4.59	4.66	3.93	3.66							
H									4.62	4.34	4.72	3.72	3.97							

Figure 3.11: Distribution of preference scores of test colors.

Color Impression Scores

The 22 adjective pair scales were quantified using the same method that was used to quantify the "like-dislike" scale. Then, like the definition of PS, each color impression score of the test colors is defined as the average of the evaluation scores of the test colors on the corresponding adjective pair scale across the subjects. As an example, the score on the scale "graceful-awkward," namely the scale *gracefulness*, is named "gracefulness score (GS)" and is defined by Eq. 3.8, where N is the total number of the subjects, $k(k = 1, 2, ..., N)$ is the subject index, and GE_k is the evaluation score rated by the kth subject on the scale *gracefulness*.

$$GS = \frac{\sum_{k=1}^{N} GE_k}{N} \tag{3.8}$$

The GSs of the test colors are listed in Appendix D. The other 22 color impression scores of the test colors are listed in Appendix E.

3.3.5 Statistical Analyses on Relationships Between Color Focality and Color Impressions

We examined the plots of the FSs and the 22 color impression scores of the test colors and ran some regression analyses to explore whether there were color impression scores that had a significant relationship with FS.

We found a clear descending tendency between FS and GS, namely the score of the color impression *gracefulness*. The linear regression analysis showed that there was a significant negative linear relationship between FS and GS (Pearson's correlation coefficient = -0.584, $P < 0.001$, plotted in Figure 3.12). No other color impression score showed any relationship with FS.

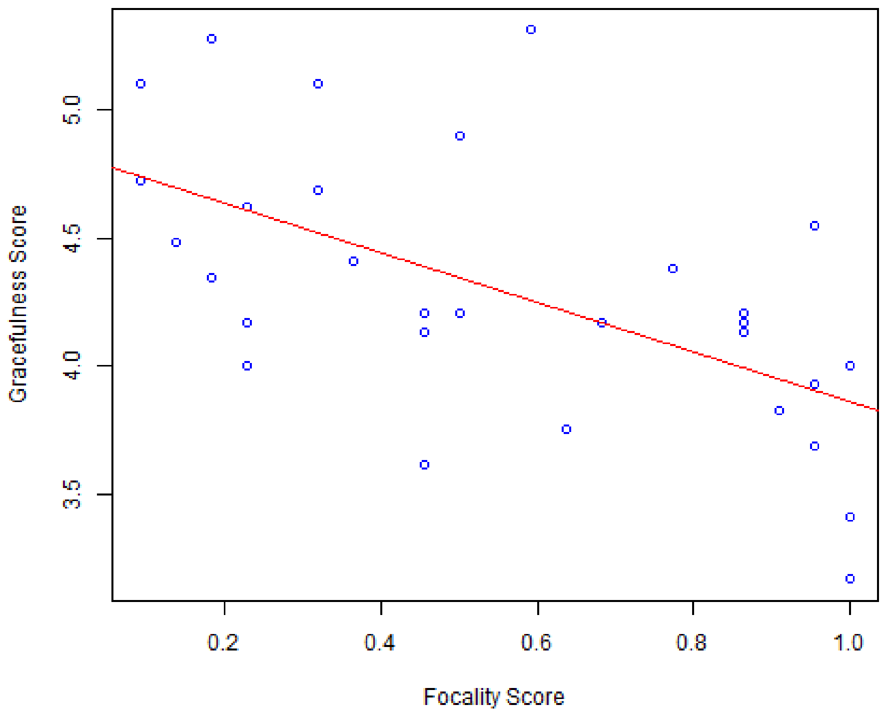

Figure 3.12: Plot of the data points and the linear regression model in terms of focality score and gracefulness score.

Regarding the discriminability of the test colors in the color array examined in Experiment

4, which was a possible confounding factor for the FS-PPF relationship, because the color array was not used in Experiment 5, it was impossible for discriminability to interfere with the statistical analyses concerning color impressions or color preference.

3.3.6 Statistical Analyses on Relationships Between Color Impressions and Color Preference

Since GS is the only color impression that has a significant relationship with FS, we plotted the PSs and the GSs of the test colors as shown in Figure 3.13. We found a clear ascending trend in this figure, and the regression analysis showed that there was a significant positive linear relationship between GS and PS (Pearson's correlation coefficient = 0.623, $P < 0.001$). Because GS has a significant relationship with both FS and PS, it is reasonable to argue that GS bridges FS and PS.

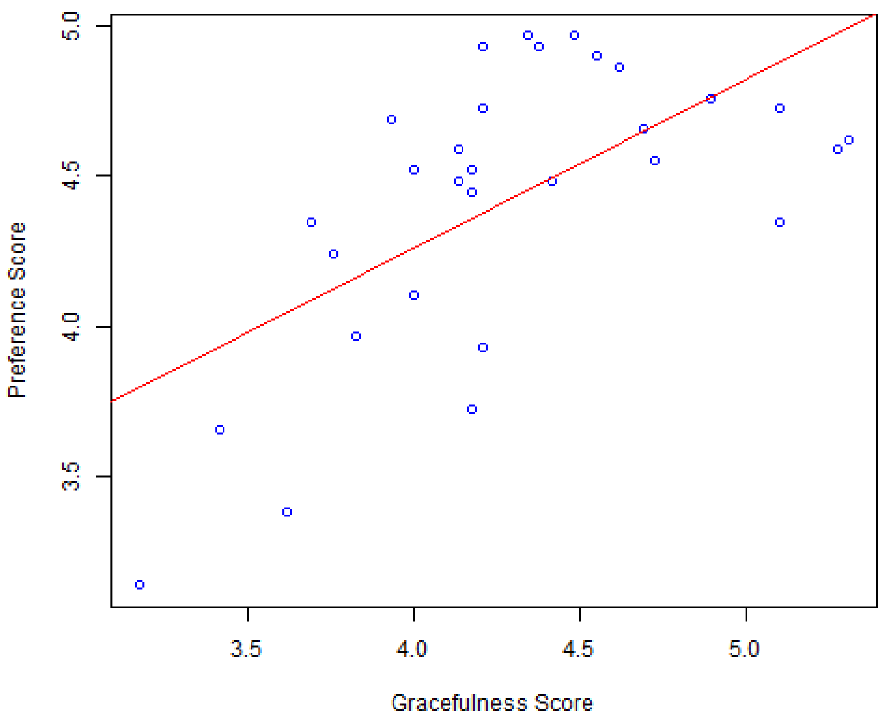

Figure 3.13: Plot of the data points and the linear regression model in terms of the gracefulness score and the preference score.

With regard to other color impression scores, heaviness score (Pearson's correlation coefficient = -0.439, P = 0.015), lightness score (Pearson's correlation coefficient = 0.555, P = 0.001), noisiness score (Pearson's correlation coefficient = 0.509, P = 0.004), ornateness score (Pearson's correlation coefficient = 0.668, $P < 0.001$), pleasantness score (Pearson's correlation coefficient = 0.910, $P < 0.001$), cleanness score (Pearson's correlation coefficient = 0.673, $P < 0.001$), cheerfulness score (Pearson's correlation coefficient = 0.630, $P < 0.001$), clearness score (Pearson's correlation coefficient = 0.673, $P < 0.001$), dynamicness score (Pearson's correlation coefficient = 0.507, P = 0.004), trueness score (Pearson's correlation coefficient = 0.793, $P < 0.001$), novelty score (Pearson's correlation coefficient = 0.581, $P < 0.001$), beauty score (Pearson's correlation coefficient = 0.922, $P < 0.001$), successfulness score (Pearson's correlation coefficient = 0.818, $P < 0.001$), positivity score (Pearson's correlation coefficient = 0.657, $P < 0.001$), and activity score (Pearson's correlation coefficient = 0.549, P = 0.002) also had significant linear relationships with PS. However, because these color impression scores do not have significant relationships with FS, it is impossible for them to bridge FS and PS.

3.3.7 Statistical Analyses on Relationship Between PPF of Colors and Color Preference

We plotted the MASs of the test colors, which measure the degrees of PPF of the test colors, and the PSs of the test colors as shown in Figure 3.14, but we could not find any trend in this plot. In view of the quadratic relationship found between MAS and FS, we ran a linear and a quadratic regression analysis on the MAS and the PS data. No significant linear or quadratic regression model was found between MAS and PS (linear regression model: R^2 = 0.001, P = 0.845; quadratic regression model: R^2 = 0.042, P = 0.562).

There also exists the possibility that the PPF of colors can influence color preference by

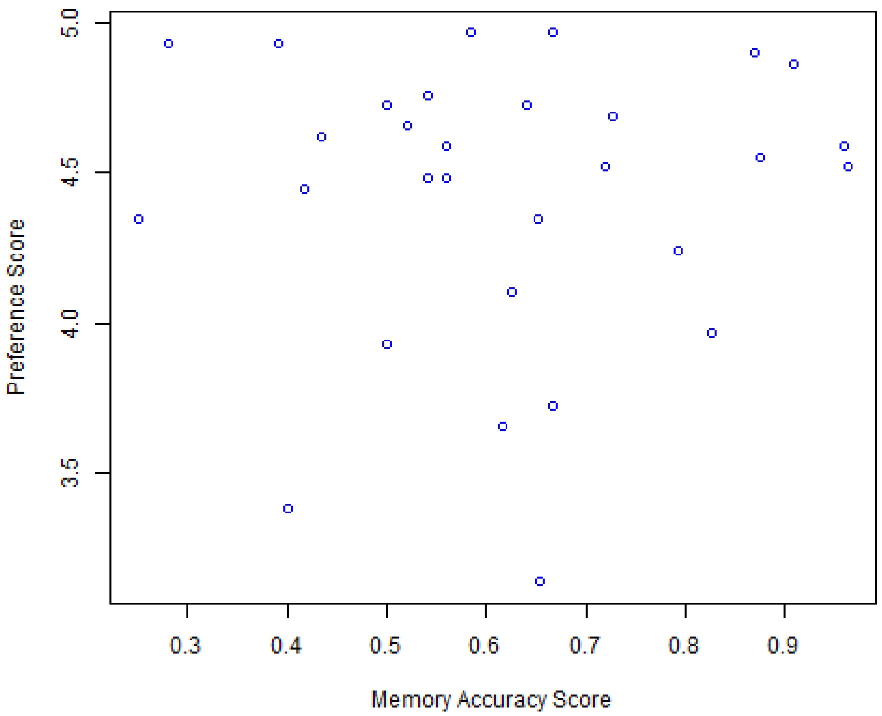

Figure 3.14: Plot of the data points in terms of memory accuracy score, which measures PPF, and preference score.

first influencing the evaluation of the color impression *gracefulness*, or in other words, that the evaluation of gracefulness bridges the impact of PPF on preference. However, no trend can be discerned in the plot of the MAS and GS data, which is shown in Figure 3.15. We ran a linear and a quadratic regression analysis on the MAS and GS data and found no significant model (linear regression model: R^2 = 0.040, P = 0.290; quadratic regression model: R^2 = 0.082, P = 0.317). These results imply that there is no relationship between PPF and the evaluation of gracefulness, which means that the possibility that the evaluation of gracefulness relays an influence of PPF on preference is ruled out.

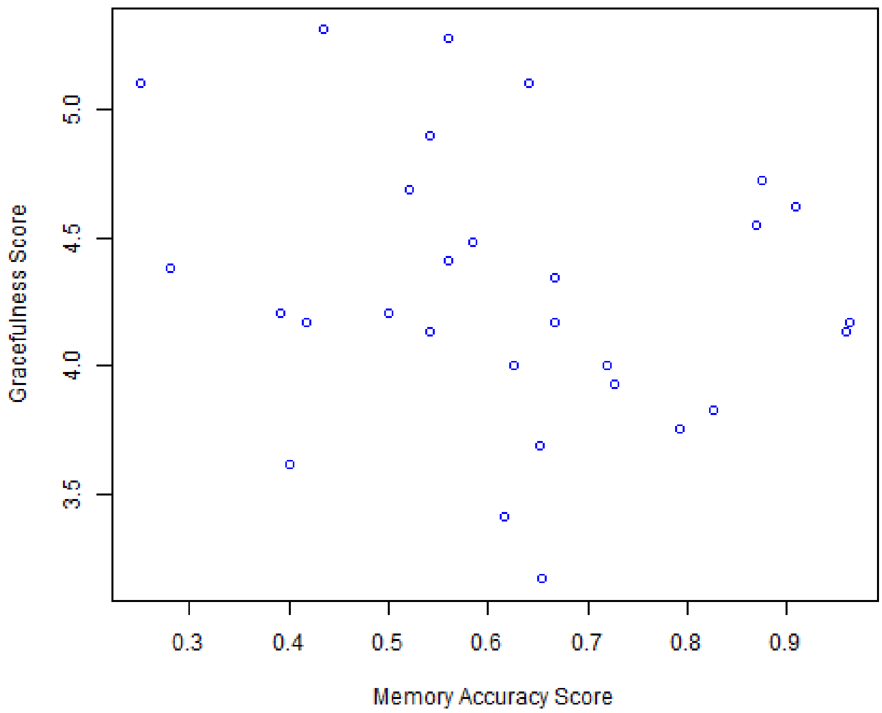

Figure 3.15: Plot of the data points in terms of memory accuracy score, which measures PPF, and gracefulness score.

3.3.8 Statistical Analyses on Relationship Between Color Focality and Color Preference

We plotted the FSs and the PSs of the test colors as shown in Figure 3.16, and we discerned a descending trend in the plot. Yet, the result of a linear regression analysis run on the FS and PS data did not achieve statistical significance (Pearson's correlation coefficient = -0.300, R^2 = 0.090, P = 0.108). This result seems counterintuitive since both FS and PS have a significant linear relationship with GS. A possible cause of this result is discussed in Section 3.4.

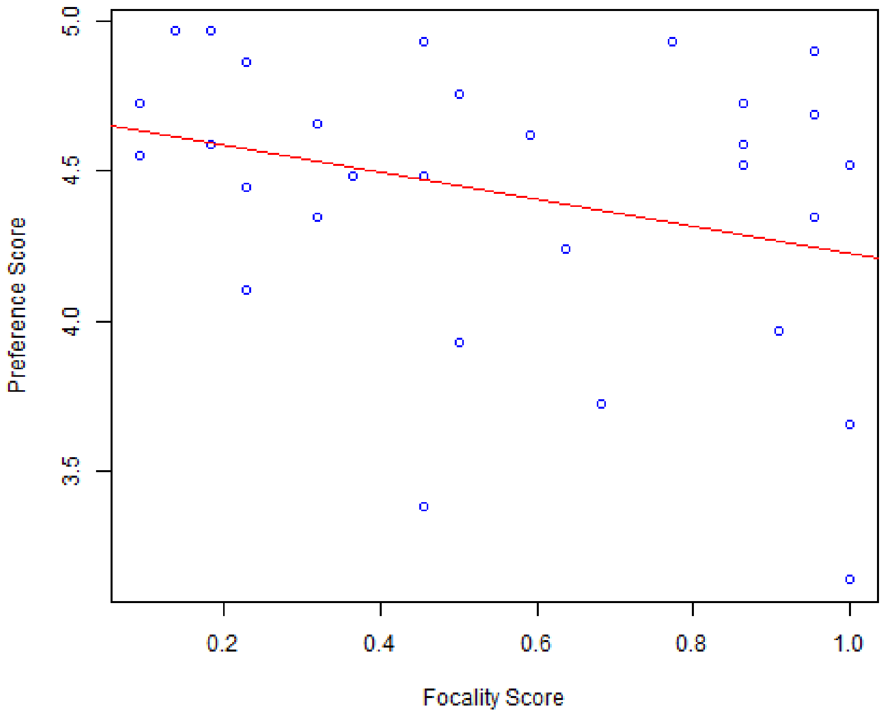

Figure 3.16: Plot of the data points in terms of focality score and preference score.

Besides, in view of the quadratic relationship detected between FS and MAS, we also ran a quadratic regression analysis on the FS and the PS data, although we could not discern any quadratic relationship between FS and MAS. The regression result shows no significant quadratic relationship between FS and PS (R^2 = 0.094, P = 0.266).

3.4 General Discussion

3.4.1 Main Findings

[7]The results of the two experiments (Experiment 4 and Experiment 5) in Study 2 of this research show that there is a negative linear relationship between FS, which measures color focality, and GS, which measures the color impression *gracefulness*, and a positive linear relationship between GS and PS, which measures color preference. Within the framework of Leder's model, because *gracefulness* is a culture-derived affective evaluation, it is located at the Implicit Memory Integration Stage, which is the second stage of the visual information processing module. This means that the evaluation of gracefulness is one stage after the processing of color focality, which is a perceptual feature, and, thus, is located at the Perceptual Analyses Stage. Also considering that the aesthetic preference is the final output, the experiment results can be interpreted into such a causal chain: the color focality negatively influences the evaluation of the color impression *gracefulness*, and the evaluation of gracefulness positively influences color preference. In other words, color focality has a negative influence on color preference which is mediated by the evaluation of gracefulness.

Because both the FS-GS relationship and the GS-PS relationship have a significant correlation coefficient larger than 0.50, it is reasonable to argue that a linear relationship really exists between FS and PS. However, probably because the FS-PS relationship contains both the noises in the FS-GS relationship and the noises in the GS-PS relationship, the FS-PS relationship did not reach statistical significance in our experiments.

On the other hand, although the experimental indicator of PPF, namely MAS, has a significant U-shaped quadratic relationship with FS, no significant relationship was found between MAS and PS. This means that the experiment' s results do not support PPF's role as a mediating variable between color focality and color preference. In addition, no color impression

[7]The content of Section 3.4 is to be presented as a conference paper authored by Fang and Matsui[16].

score other than GS was found to have significant relationships with FS. Therefore, in our experiments, the evaluation of color impression *gracefulness* was the only psychological variable that mediated color focality and color preference.

3.4.2 Noisy Colors Phenomenon: A Possible Cultural Cause of the Focality-Gracefulness-Preference Relationship

Based on a literature review, we noticed that there exists a ”Noisy Colors Phenomenon (NCP)” in many color-related areas in present-day Japan, for example, city planning, apparel designing, and cosmetics. We speculate that this phenomenon may be the cause of the focality-gracefulness-preference relationship detected in our experiments.

The term NCP originated in the 1960s in the field of city planning and referred to the phenomenon that high-focality colors, especially those in the range of warm colors, tend to be considered ”noisy,” namely flashy, gaudy, and thus unpleasant and dissonant with other colors in the surroundings. This phenomenon began to receive public attention after the Setagaya McDonald’s Incident in 1985 at Setagaya District in Tokyo in which McDonald’s set up a huge neon sign on the roof of a tall building, with the three symbol colors of the enterprise—red, white, and yellow–flashing repeatedly. Severe protests broke out from nearby residents who claimed that this neon sign had brought unbearable color pollution. This incident, together with the Tokyo Bus Incident, in which Tokyo’s city buses were painted pure yellow and red in 1981, and the Takasaki BigCamera Incident, in which the exterior of the BigCamera shopping mall near the Takasaki Railway Station was painted an extremely gaudy orange in 1986, are referred to as ”the three major incidents of NCP in Japan” by Muneo Mitsuboshi. To prevent such incidents from occurring or, in other words, to avoid noisy colors affecting the harmony of the color coordination of street landscapes, Mitsuboshi suggested refraining from using high-focality warm colors in designing public transport carriers, buildings, and advertising boards.

In fact, nowadays, a number of enterprises have already adopted this method. For example, McDonald's and Seven Eleven substituted the original high-focality red color for darkish brown on the logo plates of some of their branches, and the vending machines in some districts were placed in wooden cases to cover their pure red exterior[70, 71]. Like in city planning, with regard to clothing, high-focality colors are considered gaudy, distasteful, and therefore unsuitable for many everyday occasions such as dating[121], working in offices[42, 114, 119], attending parties[38], and attending children's school entrance ceremonies as their guardians[49]. This convention also holds for facial[109, 110] and nail makeup[39]. We speculate that our Japanese subjects unconsciously acquired these negative impressions on high-focality colors, which were detected in our experiments, from their everyday color-using experience of which NCP has long become a routine. This means that the focality-gracefulness-preference relationship detected in our experiments is presumably an unconsciously acquired one that is culture-dependent, rather than universal.

The relationship between color focality and color preference likely varies across cultures as does the relationship between color focality and the evaluation of gracefulness, since there are cultures in which high-focality colors are reported to be positively evaluated or even popular. Some examples are introduced below:

- Vivid, bright, and high-focality colors are frequently used in apparel design in Scandinavian countries probably because they can bring an element of vigor and liveliness to indoor life during harsh winters[81].

- London is famous for its vivid red buses and telephone booths. This vivid red, in Muneo Mitsuboshi's view, acts as an accent color against the dominating grey colors of the streetscapes. In addition, many patrol cars, taxies, and bridges are painted in high-focality colors. Thus, it is hard to imagine the existence of NCP in London[69, 70].

- In China's traditional culture, high-focality colors are highly preferred, especially pure red[77].

- In Korea, high-focality colors are frequently used in general design[95, 107], apparel design[52, 58], architecture[52, 117], craftworks[51], and traditional cuisine[52, 67, 117]. It is probably because pure black, white, red, blue, and yellow are the most important colors in Korea's traditional culture. These five colors are called Obangsaek, which means "five cardinal colors." This tradition originated from the Yin-Yang philosophy in ancient China. Each cardinal color is associated with a direction, a type of substance, and an animal, and each except yellow is associated with a season. Pure blue is associated with the concepts *east*, *wood*, *Chinese dragon*, and *spring*; pure red is associated with the concepts *south*, *fire*, *Chinese phoenix*, and *summer*; the focal white is associated with the concepts *west*, *gold*, *tiger*, and *autumn*; the focal black is associated with the concepts *north*, *water*, *turtle*, and *winter*; the focal yellow is associated with the concepts *center*, *earth*, and *Chinese dragon*[52, 53, 34, 95, 117].

In view of these reports, NCP likely does not exist in the abovementioned cultures, or in American modern culture since Martindale and Moore's[63] experiments conducted in the United States revealed that color focality positively influences color preference. Hence, it is difficult to expect that the focality-gracefulness-preference relationship detected in our research will also be detected in these cultures.

With regard to the cause of NCP in Japanese culture, Yoshimura[123] pointed to the Japanese tradition of valuing modesty as a vital virtue. Yoshimura wrote that[8]

> There was a tradition in Japan that "bitter colors" are considered beautiful. "Bitter colors" refer to subtle and composite colors that are composed of various hues such as red, yellow, green, blue, and purple and contain a bit grey···. In the

[8]We translated this quoted portion of Yoshimura's 2007 paper[123], which was originally written in Japanese, into English for easy reading. Please refer to Yoshimura's 2007 paper[123] for the original Japanese version.

> Japanese language, there are proverbs that consider an inconspicuous existence as daunting behavior, such as ”unexpected piles [or nails] are to be struck” and ”hawks that have brains hide nails,” as well as proverbs derived from provisions legislated during the feudal periods, such as ”there is no winner in a quarrel.” In traditional Japanese culture in which humility is respected as a virtue, ”bitter colors” contain a sense of beauty of modest composite colors.

This is an interesting hypothesis but needs further verification.

Another point to note is that while the existing reports on NCP are all records of external observations, our research provided the first experimental evidence of the existence of NCP in modern Japanese culture.

Chapter 4 General Discussion

4.1 General Conclusions of Study 1 and Study 2

This research aimed to clarify what perceptual features influence multicolor aesthetic evaluation and how these features influence the aesthetic evaluation. Based on the framework of Helmut Leder's model, the objective of this research can be rephrased as aiming to figuring out the perceptual features located at the Perceptual Analyses Stage in the visual information processing module. The entire research consists of two studies. Study 1 to clarify what multicolor-level perceptual features are engaged in multicolor aesthetic evaluation and how these multicolor-level perceptual features affect the aesthetic evaluation. Multicolor-level perceptual features are perceptual features that are attributes of multicolor stimuli themselves. Study 2 investigated the role of color focality, which is a component-level perceptual feature, in multicolor aesthetic evaluation. Component-level perceptual features are perceptual features that are properties of the component colors of multicolor stimuli, in other words, single colors that constitute multicolor stimuli.

In Study 1, we first conducted two psychological experiments (Experiment 1 and Experiment 2) using the SD method. Thirty-five 4×4 color grids were used as multicolor stimuli in both experiments. Experiment 1 quantified the aesthetic evaluation of the multicolor stimuli by defining the aesthetic score of each multicolor stimulus as the inverse of its factor score on the factor "Pleasure" extracted in this experiment. In Experiment 2, three factors, "Stability," "Heaviness," and "Presence," were extracted and each of them was regarded as a multicolor-level perceptual feature. The feature values of each stimulus on the three perceptual features were defined as the factor scores of each stimulus on the three factors. Following this, we built a computational model composed of two levels of BPNNs. The first level represents the mappings from the primary color information (the $L^*a^*b^*$ values of the component colors) to the three multicolor-level perceptual features, and the second level represents the mappings from the three multi-color level perceptual features to the aesthetic evaluation. The learning rate, number of hidden-layer nodes, and momentum constant of each BPNN were optimized through the GA. We then performed two simulations to figure out how each of the three perceptual features influences the aesthetic evaluation. Simulation 2 verified the psychological appropriateness of the model by evaluating the model performance in a statistically meaningful sense. In this simulation, the data obtained in Experiments 1 and 2 were used as the training data, and the data obtained in another experiment, namely, Experiment 3, were used as the validation data. The prediction of the model had a small NRMSE, which is the indicator of the prediction error, and the prediction errors across the validation-data stimuli were normally distributed around a value which was very close to zero. This means that the prediction performance of the model achieved a certain degree of accuracy and robustness, which backs the psychological appropriateness of the trained model. In addition, a correlation-based evaluation of the model performance indicated that the model automatically learned the relationship between overall lightness and aesthetic evaluation of color combinations, which is a critical psychological rule in multicolor aesthetics. This serves as another piece of evidence for the psychological appropriateness of the trained model. Next, we analyzed the microstructure, namely, the synaptic weights and biases, of the BPNN 4 to investigate how each multicolor-level perceptual feature influenced the aesthetic evaluation. The analysis showed that the perceptual feature Heaviness had the principal impact on the aesthetic evaluation of multicolor objects, whereas the other two perceptual features Stability and Presence had a minor influence. The heavier and/or more stable a multicolor object is perceived to be, the less aesthetically pleasing it is. Conversely, the stronger the sense of matter presence a multicolor object elicits, the more aesthetically appealing it is.

In Study 2, we considered color focality as a candidate for component-level perceptual features engaged in multicolor aesthetic evaluation because previous studies have suggested that the focalities of the component colors of a multicolor stimulus can influence the aesthetic evaluation of the multicolor stimulus by first influencing the aesthetic preference of the component colors. Through two psychological experiments (Experiment 4 and Experiment 5), we investigated the continuous relationship between color focality and color preference as well as the psychological variable(s) that mediate this relationship. The candidates for the mediating variables included the fluency of the psychological information processing (PPF) of the colors and 22 color impressions. Thirty Munsell chips were used as test colors in these experiments. Experiment 4 consisted of two sessions. In each trial of Session 1, the subject was asked to watch a test color for 5 s, and after a 30-s interval, find it in a color array. The PPF of each test color was quantified as its mean memory accuracy because short-term memory accuracy reflects multiple PPF-related perceptual properties. After the completion of Session 1, the subjects were asked to report the methods they used to memorize the test colors, on a questionnaire. In Session 2, to quantify the focality of each test color, we asked the subjects to report the colors in the color array that belonged to each of the six Japanese basic color categories that covered the test colors in the color array. In Experiment 2, the subjects rated the test colors on 22 Likert scales that represented the 22 color impressions and another Likert scale measuring the preference degree. We found a U-shaped quadratic relationship between color focality and PPF of colors, which were measured by memory accuracy scores of colors. In other words, the degree of PPF was highest at both ends of the continuum of color focality and decreased as the color focality moved towards the medium region from either end. This relationship was still maintained after the false-alarm rate and the Discriminability Effect—two possible confounding factors—were ruled out. The results of the questionnaire revealed that the strategy of encoding colors using linguistic color categories was the most-frequently used memorization strategy by the subjects. We speculate that the subjects' frequent and conscious use of this strategy was an important cause of the detected focality-PPF relationship. Despite the quadratic relationship between color focality and PPF of colors, no significant relationship was found between the PPF of colors and color preference. Thus, the role of PPF of colors as a mediating variable between color focality and color preference was not supported in this research. However, with regard to the other type of potential mediating variables, namely, the 22 color impressions, we found a significant negative linear relationship between color focality and color impression *gracefulness* and a significant positive linear relationship between *gracefulness* and color preference. According to the framework of Leder's model, this implies that high-focality colors were evaluated as less graceful, and thus, were less preferred than were low-focality colors. We speculate that this focality-gracefulness-preference relationship is derived from the prevailing Noisy Colors Phenomenon (NCP) in various color-related areas, e.g., city planning, apparel design, and cosmetics, in the modern Japanese culture. NCP refers to the phenomenon, wherein high-focality colors are generally perceived as flashy, gaudy, and therefore, unpleasant. No other color impression was found to have a significant relationship with color focality. In conclusion, Study 2 experimentally verified that color focality is a component-level perceptual feature engaged in the psychological process of multicolor aesthetic evaluation. Specifically, color focality has a negative influence on color preference, and the evaluation of the color impression *gracefulness* bridges this influence.

In terms of the methodology of data analysis, in Study 1, we used the BPNN to clarify how the three multicolor-level perceptual features, Stability, Heaviness, and Presence, influence the aesthetic evaluation of multicolor stimuli, and in Study 2, we performed regression analyses to investigate the impact of color focality on the aesthetic preference for colors. In other words, the BPNN and the regression analyses were used in this research for the same purpose, that is, to clarify the role of the perceptual feature(s) in question in the psychological mechanism of multicolor aesthetic evaluation. The reason for adopting different methods for data analysis in Study 1 and Study 2 is that Study 1 and Study 2 differ in terms of the complexity of the intervariable relationships being extracted. Specifically, since Study 1 deals with the mappings from primary color information to three perceptual features and the mapping from these perceptual features to the aesthetic evaluation, the BPNN was chosen because of its great capacity to process complicated multivariate nonlinear mappings. On the other hand, Study 2 examined the relationship between one perceptual feature and the aesthetic evaluation, and therefore, one-variable regression analyses were employed.

In summary, three multicolor-level perceptual features that engaged in multicolor aesthetic evaluation—Stability, Heaviness, and Presence—were detected in Study 1 of the research, and one component-level perceptual feature involved in multicolor aesthetic evaluation—color focality—was detected in Study 2 of the research. This research also clarified how each perceptual feature exerts its influence color aesthetic evaluation in a quantitative manner. Figure 4.1 demonstrates Leder's model refined by the findings of Study 1 and Study 2.

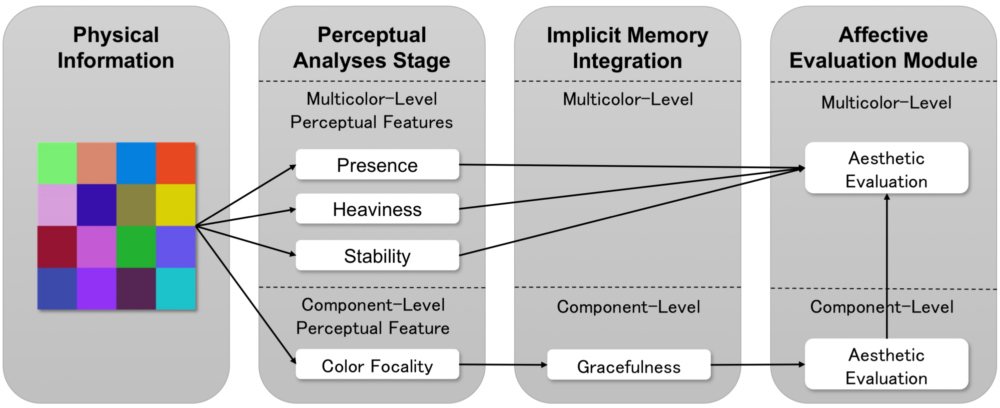

Figure 4.1: Leder's model refined by the findings of Study 1 and Study 2 of this research.

4.2 Integration of Study 1 and Study 2

4.2.1 Synopsis

As described above, Study 1 of the research investigated the multicolor-level information processing of multicolor stimuli, and Study 2 looked into the component-level information processing of multicolor stimuli. In this section, we discuss three possible links between these two levels on the basis of a literature review and a reexamination of our experimental data.

4.2.2 Link 1: Focality of Component Colors → Multicolor-Level Feature "Stability" → Aesthetic Evaluation of the Multicolor Stimulus

Hard, Sivik, and Tonnquist[28, 29] developed the Natural Color System (NCS), which represents colors using six elementary attributes. The six elementary attributes of a color denote the degrees of resemblance of the color to the six basic color categories *black*, *white*, *red*, *yellow*, *blue*, and *green*. Because the concepts of the elementary attributes accord with Berlin and Kay's[4] definition of basic color categories, the elementary attributes in the NCS resemble the concept of "color focality" in our research, or in other words, the elementary attributes in the NCS and the color focality in our research encode similar information. Later, Hard and Sivik, in their 2001 study[30], proposed an equation that computed the degree of complexity, which they termed "complexity factor," of a color combination from the elementary attributes of the component colors. This equation (Eq. 4.1) is shown below.

$$CF = MA + 0.1SA(6 - MA) \tag{4.1}$$

In Eq. 4.1, CF is the complexity factor, MA is the total number of the main attributes that the component colors have, and SA is the total number of the secondary attribute(s) that the component colors have. The main attribute of a color refers to the dominating elementary at-

Table 4.1: The rating scales that were used in the original factor analysis in Experiment 2 and had strong correlations with the complexity scale ”simple-complex.”

Rating scale	Pearson’s correlation coefficient	P value
noisy-quiet	-0.511	0.002
dynamic-static	-0.575	<0.001
diversified-monotonous	-0.699	<0.001
leisurely-bustling	0.598	<0.001
calm-restless	0.579	<0.001
calm-violent	0.59	<0.001

tribute of the color, namely, the elementary attribute of the color that has the greatest value. The secondary attribute(s) of the color refers to the elementary attributes other than the main attribute[29]. Through a psychological experiment, Hard and Sivik[30] found that the complexity factors computed in this manner positively correlated with the rating scores on the semantic rating scale ”complicated.” This implies that the complexity factors match well with humans’ subjective impression of the complexity of color combinations; in other words, complexity factor can be deemed a reliable psychological index of the degree of the complexity of color combinations.

In this research, we measured the degrees of the complexity of multicolor stimuli in Experiment 2 using the semantic rating scale ”simple-complex.” Because the Cronbach’s α coefficient of the scale was 0.433, which is lower than 0.60, this scale was not used for the factor analysis. However, a reexamination of the rating data collected in Experiment 2 showed that within the rating scales used in the factor analysis, those having strong correlations with the scale ”simple-complex” (listed in Table 4.1) all belonged to the factor Stability. Thus, we conjecture that the complexity scale also represents the sense of stability. We tested this by running a factor analysis that included the complexity scale. The factor analysis was performed on R (version 3.3.0).

The parameters of the new factor analysis were the same as those of the original factor analysis in Experiment 2. There were three factors whose eigenvalues were greater than 1.0, and therefore, the three factors were selected as main factors. The principal factor method was used to compute the initial factor loadings, and the promax method was used to rotate the factors. The factor loadings after the rotation (i.e., the pattern matrix) are shown in Table 4.2. All factor loadings whose absolutes are greater than 0.5 are indicated in bold. The proportion of variance explained by each factor is shown in the last row of the table.

Table 4.2 shows that as in the original factor analysis, three factors—Stability, Heaviness, and Presence—were extracted, and the complexity scale belonged to the factor ”Stability.” Because the rating scales that strongly correlated with the scale ”simple-complex” in previous studies on color affective effects[1, 78, 115] also strongly correlated with the complexity scale in our new factor analysis, it is reasonable to contend that the factorial structure of the complexity scale in our new factor analysis is a robust one. Thus, we verified the conjecture that the subjective evaluation of the complexity of multicolor stimuli represents the sense of stability of multicolor stimuli. Combining this finding with Hard and Sivik's[30] finding described above, we can see that the elementary attributes, namely the focality information, of component colors have a large impact on the multicolor-level perceptual feature Stability. Further, because the multicolor-level perceptual feature Stability influences the aesthetic evaluation of multicolor stimuli, which is a result of Study 1 of this research, it is cogent to argue that focality information of the component colors of a multicolor stimulus has an influence on the aesthetic evaluation of the multicolor stimulus by first influencing the multicolor-level perceptual feature Stability. This influence chain is shown in Figure 4.2.

4.2.3 Link 2: Focality of Component Colors → Gracefulness Evaluation of the Multicolor Stimulus → Aesthetic Evaluation of the Multicolor Stimulus

Another finding of Hard and Sivik's[30] aforementioned psychological experiment was that the complexity factors computed from the elementary attributes of the component colors neg-

Table 4.2: Rotated factor loading matrix of the factor analysis that included the complexity scale ”simple-complexity.”

Factor identity	Adjective pair scale (rating scale)	Factor		
		1	2	3
Factor 1 (Stability)	conspicuous-inconspicuous	**0.947**	0.146	-0.097
	leisurely-bustling	**-0.929**	0.022	-0.003
	plain-ornate	**-0.929**	-0.106	0.003
	dynamic-ststic	**0.917**	-0.026	0.106
	clear-vague	**0.886**	0.287	-0.151
	diversified-monotonous	**0.885**	-0.209	0.072
	strong-weak	**0.874**	-0.032	**-0.604**
	noisy-quiet	**0.848**	-0.090	0.221
	calm-restless	**-0.807**	0.382	-0.269
	calm-violent	**-0.748**	0.466	0.048
	vivid-subdued	**0.739**	**0.552**	-0.066
	distinct-indistinct	**0.707**	**0.536**	-0.184
	wet-dry	**-0.684**	-0.006	-0.310
	simple-complex	**-0.674**	0.305	0.146
	cold-warm	**-0.661**	0.175	**-0.525**
	light-dark	**0.589**	0.142	**0.512**
	sweet-unsweet	0.423	0.352	0.286
Factor 2 (Heaviness)	neat-disordered	-0.194	**0.924**	-0.106
	clear-dull	0.356	**0.777**	0.009
	plain-thick	-0.209	**0.726**	0.159
	gaudy-plain	0.326	**-0.715**	-0.172
	delicious-bad-tasting	0.476	**0.619**	0.142
Factor 3 (Presence)	dark-pale	0.280	0.086	**-1.001**
	soft-hard	-0.008	0.131	**0.809**
	shallow-deep	0.270	0.034	**0.765**
	light-heavy	0.319	0.209	**0.647**
Variance explained		44.0%	16.7%	15.6%

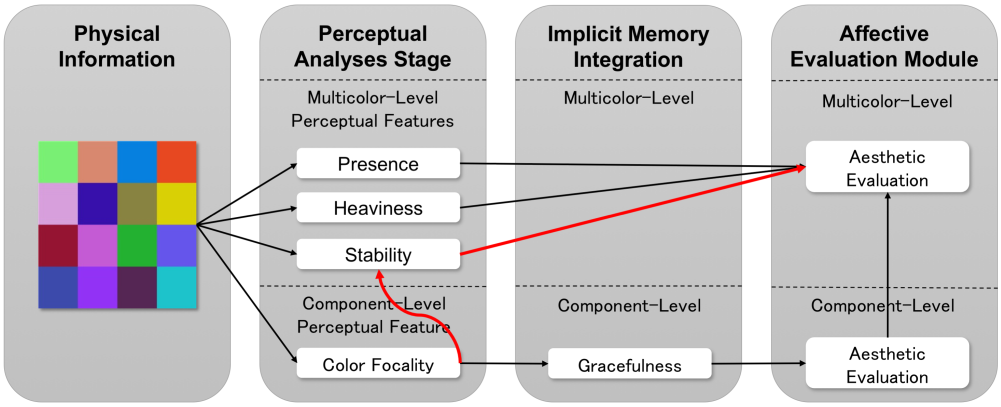

Figure 4.2: Link 1 between component-level information processing and multicolor-level information processing.

atively correlated with the rating scores on the semantic rating scale ”cultured.” This suggests that it is possible to predict the evaluation of culturedness of a multicolor stimulus using the focality information of its component colors. Considering that, in the Japanese language, the concept of *culturedness* has a similar meaning as the concept of *gracefulness*, Hard and Sivik's[30] finding implies that the focality information of component colors has the ability to influence the evaluation of gracefulness of multicolor stimuli.

In our research, the gracefulness of multicolor stimuli was measured in Experiment 1 using the semantic rating scale ”graceful-awkward.” The results of Experiment 1 show that the scale ”graceful-awkward” belongs to the factor Pleasure. Because the aesthetic scores of the multicolor stimuli were defined using the factor scores of the stimuli on the factor Pleasure, the gracefulness degrees of the multicolor stimuli positively correlated with the aesthetic scores of the stimuli. In other words, the higher the gracefulness degree of a multicolor stimulus, the higher the aesthetic score of the stimulus. Combining this with Hard and Sivik's[30] finding that focality of component colors influences gracefulness evaluation of multicolor stimuli, it is reasonable to argue that the focality information of the component colors of a multicolor stimulus has the ability to influence the aesthetic evaluation of the multicolor stimulus by first influencing the gracefulness evaluation of the multicolor stimulus. Figure 4.3 shows the link between the component level and multicolor level. As explained in Section 3.4.1, because the evaluation of gracefulness—despite being an unconscious subjective evaluation—is influenced by memory factors, such as personal experience and cultural background, it is likely situated in the Implicit Memory Integration Stage in the visual information processing module.

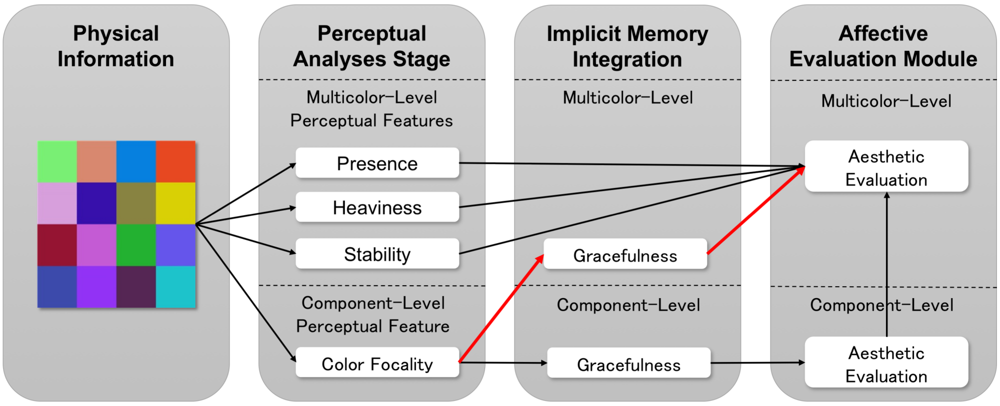

Figure 4.3: Link 2 between component-level information processing and multicolor-level information processing.

4.2.4 Link 3: Focality of Component Colors → Gracefulness Evaluation of Component Colors → Gracefulness Evaluation of the Multicolor Stimulus → Aesthetic Evaluation of the Multicolor Stimulus

Oyama and Miyata[94] found that the rating score of a two-color combination on the semantic rating scale "like-dislike" could be largely predicted using an "averaging model," or in other words, by computing the mean of the rating scores of the two component colors on the scale (R = 0.658, $P < 0.001$). The averaging model is expressed as Eq. 4.2.

$$Y = \frac{X_i + X_j}{2} \tag{4.2}$$

In Eq. 4.2, Y is the rating score of a two-color combination on the scale "like-dislike," and X_i and X_j are the rating scores of the two component colors on the scale.

In our research, the results of Experiment 1 showed that the scale "graceful-awkward" positively correlated with the scale "like-dislike" (Pearson's correlation coefficient = 0.660, $P < 0.001$), and both scales belonged to the factor Pleasure, which suggests the psychological similarity between the two scales. Thus, we conjecture that, similar to the evaluation on the scale "like-dislike," the gracefulness evaluation of a multicolor stimulus can be predicted to some extent using the gracefulness evaluation of its component colors. In other words, it is possible that the gracefulness evaluation of the component colors of a multicolor stimulus can influence the gracefulness evaluation of the multicolor stimulus. Considering that the gracefulness evaluation of a multicolor stimulus can influence the aesthetic evaluation of the multicolor stimulus (explained in Section 4.2.3), as well as the focality of the component colors influences the gracefulness evaluation of the component colors (which is a result of Study 2 of our research), this conjecture suggests the possibility that the focality information of the component colors of a multicolor stimulus first influences the gracefulness evaluation of the component colors. Following this, the gracefulness evaluation of the component colors influences the gracefulness evaluation of the multicolor stimulus, and finally the gracefulness evaluation of the multicolor stimulus influences the aesthetic evaluation of the multicolor stimulus. This relationship chain is illustrated in Figure 4.4. Because the influence of the component-level gracefulness evaluation on the multicolor-level gracefulness evaluation is just a conjecture at present, it is marked by a dotted curve.

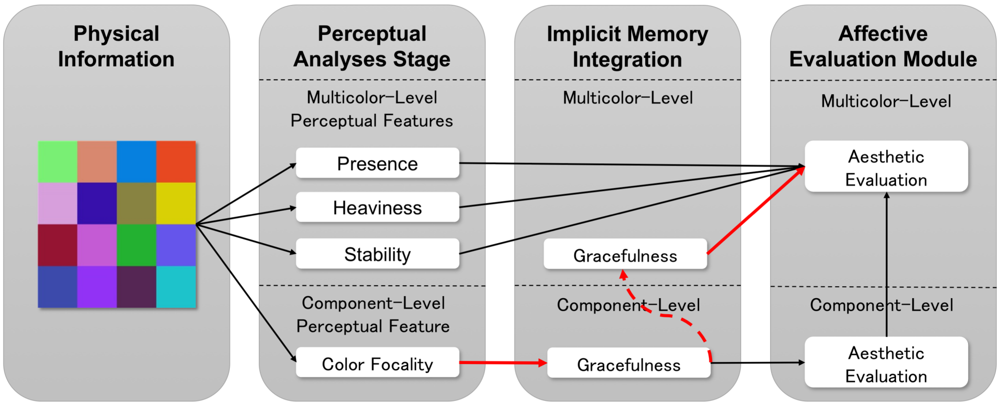

Figure 4.4: Link 3 between component-level information processing and multicolor-level information processing.

4.2.5 Summary of the Three Links Between Component Level and Multicolor Level

Figure 4.5 integrates the three links between the component-level visual information processing and the multicolor-level visual information processing within the psychological mechanism of multicolor aesthetic evaluation.

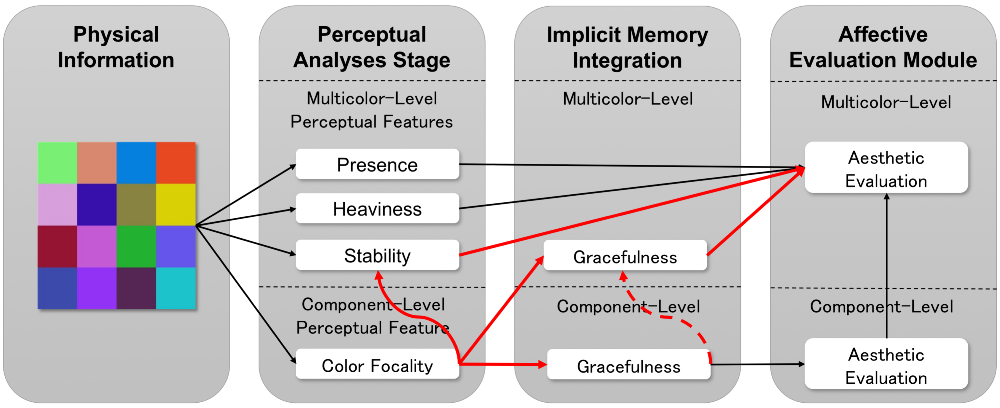

Figure 4.5: Integration of the three links between component-level information processing and multicolor-level information processing.

Chapter 5 Implications for Future Work

In Study 1 of the research, the experimental stimuli were composed of uniform-color patches. However, natural scenes are filled with colors that gradually transition to adjacent colors. According to the hypothesis of natural harmony, initially put forward by Rood and later supported by von Bezold and von Brucke, an assembly of colors that mirrors a pattern of hue gradation that often occurs in nature, namely, a natural sequence of hues, tends to be sensed as harmonious[11, 87] (introduced in Section 1.3.2). In our future studies, we will take spatial patterns of chromatic changes into consideration as an aesthetically relevant factor.

In addition, as described in Section 2.8, the stability of the training procedure of the computational model constructed in Study 1 of this research needs to be further tested using training datasets of various sizes. Besides the convergence state of the training procedure, the prediction accuracy of the trained model should also be examined. In our future studies, we intend to test how the prediction accuracy of the trained model changes as the size of the training dataset changes.

In Study 2 of this research, we experimentally confirmed the role of color focality as a component-level perceptual feature involved in the psychological process of multicolor aesthetic evaluation. Apart from color focality, there might be other component-level perceptual features that affect multicolor aesthetic evaluation. We will continue to search for such component-level perceptual features in our future studies.

Further, through Experiment 4, we found that a U-shaped quadratic relationship exists between color focality and the PPF of colors. Although not directly related to the topic of multicolor aesthetics, it is interesting to see whether this pattern of focality-PPF relationship can also be found in the categories of other domains. These domains could be simple perceptual categories such as shapes and phonemes, complicated multimodal concepts such as animals and tools, or even emotionally or socially meaningful signals such as human facial expressions.

In addition, in a series of psychological experiments, Rosch[32, 102, 103] and Kay and McDaniel[45] reported that the inner structure, namely, the membership functions of basic color categories take unimodal shapes. However, it is possible that the shapes of membership functions of basic color categories vary across people, and basic color categories that have multimodal membership functions also exist. We will explore these possibilities in our future studies.

We will also consider the color appearance modes of the experimental stimuli. Modes of color appearance can be generally divided into two types: object color and light source color. Object color refers to color sensation generated by light reflected from or penetrating through surfaces of objects. Thus, object color is determined by the spectral distribution of the illuminant, the spectral reflectance, or penetration and the sensibility of the human eyes. On the other hand, light source color refers to color sensation caused by light directly emitted from light sources. Thus, light source color is determined by the spectral distribution of the illuminant and the sensibility of the human eyes[6, 50, 120, 111, 112]. Since the experimental stimuli in Study 1 were images projected on screens, and the experimental stimuli in Study 2 were Munsell color chips, all findings of this research pertain to object color[1]. In future studies, we will explore whether the experimental results of this research can also be found in experiments that use light source color as stimuli.

It should also be noted that all experiments in this research were conducted in the Japanese language. It will be interesting to test other languages using the same experimental paradigms and compare the results across languages. From the perspective of neuroaesthetics, it is also interesting to delve into the neural correlates of the psychological relationships between the

[1] Yoshida, Horii, and Sato[122] also explains why the color appearance mode of projected images is object color.

perceptual features and the aesthetic evaluation of multicolor stimuli detected in this research.

With regard to practical applications, as described in Section 1.5.2, the results of this research can be used to develop systems of image aesthetic evaluation, which could help improve automation in fields such as industrial design, interior design in architecture, and human-computer interface construction.

ACKNOWLEDGEMENTS

My heart-felt gratitude is primarily given to Professor Tatsunori Matsui for his inspiring, generous and patient guidance and encouragement. Thanks to his warm-hearted, intelligent and enlightening advice, I have been successful in overcoming multitudes of difficulties and obstacles in the process of my research and the fulfillment of this thesis. In addition, he has provided me with a lot of sincere and timely help in my study and life in Japan.

I owe my gratitude to Professor Keiichi Watanuki, Professor Hideaki Kikuchi and Professor Miho Saito for their kindly being examiners of this thesis and for their insightful and valuable advice on this thesis.

I would also like to thank Dr. Keiichi Muramatsu, former research associate in Matsui Laboratory, now assistant professor in the Graduate School of Science and Engineering of Saitama University. He gave me many strict and precious suggestions during my experiment designing, data analyses and revisions of journal manuscripts.

My thanks also go to Mr. Yoshimasa Tawatsuji, now research associate in Matsui Laboratory. During the past four years, he helped me with my study and life by offering me valuable advice, support and aid.

I sincerely thank my fellow doctoral and master students in Matsui Laboratory. Thanks to their kindness, friendliness and enthusiasm in research, I have been imbued in a warm atmosphere of teamship.

I would also like to express my gratitude to Japanese Cognitive Science Society for offering me passage allowance for my presentation at the 39th annual meeting of the Cognitive Science Society held in London on July 26-29, 2017.

The accomplishment of this thesis is greatly dependent upon the above people's kindness and generosity, without which all these achievements are unconceivable.

Bibliography

[1] Aizaki, T., Sekine, S., & Ohkawa, M. (2003). Evaluation of comfort of picture by semantic differential method. *Journal of the Color Science Association of Japan*, 27, 102-103.

[2] Arnheim, R. (1998). *Art and visual perception* (S. Teng & J. Zhu, Trans.). Sichuan People's Publishing House, Chengdu, Sichuan. (Original work published 1974).

[3] Bar, M., & Neta, M. (2007). Visual elements of subjective preference modulate amygdala activation. *Neuropsychologia*, 45(10), 2191-2200.

[4] Berlin, B., & Kay, P. (1991). *Basic color terms: Their universality and evolution*. University of California Press. (Original work published 1969).

[5] Berry, J. W., Poortinga, Y. H., Segall, M. H., & Dasen, P. R. (1992) *Cross-cultural psychology: Research and applications*. Cambridge University Press, Cambridge, England.

[6] Buttaishoku to kogenshoku no chigai ni tsuite [On the differences between object colors and light source colors]. *Official Website of KONIKA MINOLTA*. Retrieved from https://www.konicaminolta.jp/instruments/knowledge/color/section5/17.html

[7] Brown, R. W., & Lenneberg, E. H. (1954). A study in language and cognition. *The Journal of Abnormal and Social Psychology*, 49(3), 454.

[8] Chatterjee, A. (2004). Prospects for a cognitive neuroscience of visual aesthetics. *Bulletin of Psychology and the Arts*, 4(2), 56-60.

[9] Chijiiwa, H. (2001). *An outline of color science*. University of Tokyo Press, Tokyo.

[10] Chipperfield, A. J., Fleming, P., & Pohlheim, H. (1994). *Genetic algorithm toolbox: For use with MATLAB; User's guide (version 1.2)*. University of Sheffield, Department of Automatic Control and Systems Engineering.

[11] Color Science Association of Japan (Eds.). (2007) *Encyclopedia of color science*. University of Tokyo Press, Tokyo.

[12] Cupchik, G. C. (1986). A decade after Berlyne: New directions in experimental aesthetics. *Poetics*, 15(4-6), 345-369.

[13] Everett, D. (2005). Cultural constraints on grammar and cognition in Piraha: Another look at the design features of human language. *Current Anthropology*, 46(4), 621-646.

[14] Fang, S., & Matsui, T. (2017). Experimental investigation into the continuous pattern of the relationship between color focality and short-term memory performance for colors. In G. Gunzelmann, A. Howes, T. Tenbrink, & E. J. Davelaar (Eds.), *Proceedings of the 39th Annual Conference of the Cognitive Science Society* (pp. 1985-1990). Austin, TX: Cognitive Science Society.

[15] Fang. S., & Matsui, T. (in press). An experimental study on the continuous patterns of the influence of color focality on short-term memory performance of colors. *International Journal of Affective Engineering*.

[16] Fang. S., & Matsui, T. (in press). Experimental investigation into the mediating variables of the relationship between color focality and color preference. Oral presentation at International Colour Association Interim Meeting 2018 (AIC 2018), Sep. 25-29, 2018, Lisbon, Portugal [full paper accepted].

[17] Fang, S., Muramatsu, K., & Matsui, T. (2015). Experimental study of aesthetic evaluation to multi-color stimuli using semantic differential method. *Transactions of Japan Society of Kansei Engineering*, 14(1), 37-47.

[18] Fang, S., Muramatsu, K., & Matsui, T. (2017). A computational model simulating the mental function of multicolor aesthetic evaluation. *Color Research & Application*, 42(2), 216-235.

[19] Fitzpatrick, J. M., & Grefenstette, J. J. (1988). Genetic algorithms in noisy environments. *Machine Learning*, 3(2-3), 101-120.

[20] Fukuda, T., & Fukuda, R. (Eds.). (2009). *Ningen kogaku gaido* [*A guide to ergonomics*]. SAIENSU-SHA Co., Ltd., Tokyo.

[21] Fujimoto, H. From the dean. *Official Website of School of Human Sciences of Waseda University*. Retrieved from https://www.waseda.jp/fhum/hum/about/message/

[22] Gao, X. P., & Xin, J. H. (2006). Investigation of human's emotional responses on colors. *Color Research & Application*, 31(5), 411-417.

[23] Gao, X. P., Xin, J. H., Sato, T., Hansuebsai, A., Scalzo, M., Kajiwara, K., Guan, S., Valldeperas, J., Lis, M. J., & Billger, M. (2007). Analysis of cross - cultural color emotion. *Color Research & Application*, 32(3), 223-229.

[24] Galanter, P. (2012). Computational aesthetic evaluation: Past and future. In J. McCormack & M. d'Inverno (Eds.), *Computers and Creativity* (pp. 255-293). Springer, Berlin, Heidelberg.

[25] Garro, L. C. (1986). Language, memory, and focality: A reexamination. *American Anthropologist*, 88(1), 128-136.

[26] Goldberg, D. (1989). *Genetic algorithms in search, optimization and machine learning*. Massachusetts: Addison-Wesley.

[27] Green, D. M., & Swets, J. A. (1988). *Signal detection theory and psychophysics*. Peninsula Publishing, Los Altos, California.

[28] Hard, A., Sivik, L., & Tonnquist, G. (1996). NCS, natural color system—from concept to research and applications. Part I. *Color Research & Application*, 21(3), 180-205.

[29] Hard, A., Sivik, L., & Tonnquist, G. (1996). NCS, natural color system—From concept to research and applications. Part II. *Color Research & Application*, 21(3), 206-220.

[30] Hard, A., & Sivik, L. (2001). A theory of colors in combination—A descriptive model related to the NCS color - order system. *Color Research & Application*, 26(1), 4-28.

[31] Haykin, S. (2010). *Neural networks and learning machines (3rd edition)*. China Machine Press ,Beijing. (Original work published 2008).

[32] Heider, E. R. (1971). "Focal" color areas and the development of color names. *Developmental psychology*, 4(3), 447-455.

[33] Heider, E. R. (1972). Universals in color naming and memory. *Journal of Experimental Psychology*, 93(1), 10.

[34] Hogarth, H. K. K. (2012). Changing traditions in colour symbolism. *Proceedings of the 6th World Congress of Korean Studies*.

[35] Horita, Y., Kanda, A., Murai, T., & Nakajima, Y. (1997). Estimation of the sensitivity factor of the 2 color combination with monochromatic color sensitivity factor. *Technical Report of the Institute of Image Information and Television Engineers*, 21 (28), 1-6.

[36] Ito, K. (2004). Affective effects of two-color combinations within the same hue, *Journal of the Color Science Association of Japan*, 28(1), 3-15.

[37] Ito, K., & Oyama, Y. (2005). Affective effects of two-color combinations between different hues, *Journal of the Color Science Association of Japan*, 29(4), 91-302.

[38] Johin de ochitsuita yosooi no pathi doresu [Graceful and relaxing party dresses]. (2014, May 30). Retrieved from http://www.crsp-bordeaux.org/archives/78

[39] Johin kawai ga kanau! Nyuansu neiru de kire na yubisaki e [Dreams of becoming graceful and lovely come true! Making fingertips beautiful using nuance nails]. (2017, August 30). Retrieved from https://cyuncore.com/beauty/2904

[40] Kaastra, I., & Boyd, M. (1996). Designing a neural network for forecasting financial and economic time series. *Neurocomputing*, 10(3), 215-236.

[41] Kansaku, J. (1963). The analytic study of affective values of color combinations: a study of color pairs. *Japanese Journal of Psychology*, 34, 11-12.

[42] Kanzen joshi rosen! Utsukushiku sakihokore! OL no tame no tsukin fasshon [Perfect lady style! Beautifully blooming! Office fashion for OL]. (2017, February 15). Retrieved from https://matome.naver.jp/odai/2141119873760703601

[43] Kawabata, H., & Zeki, S. (2004). Neural correlates of beauty. *Journal of Neurophysiology*, 91(4), 1699-1705.

[44] Kawachi, Y., Suzuki, M., Shibata, M., Kawabata, H., & Gyoba, J. (2008). Distributions of neural activities corresponding to psychological structures underlying the impressions of pictures. *Technical Report of the Institute of Electronics, Information and Communication Engineers*, 108(264), 25-30.

[45] Kay, P., & McDaniel, C. K. (1978). The linguistic significance of the meanings of basic color terms. *Language*, 610-646.

[46] Kay, P., Berlin, B., Maffi, L., Merrifield, W. R., & Cook, R. (2009). *The world color survey*. Center for the Study of Language and Information (CSLI) Publications.

[47] Kirk, U., Skov, M., Christensen, M. S., & Nygaard, N. (2009). Brain correlates of aesthetic expertise: A parametric fMRI study. *Brain and Cognition*, 69(2), 306-315.

[48] Kirk, U., Skov, M., Hulme, O., Christensen, M. S., & Zeki, S. (2009). Modulation of aesthetic value by semantic context: An fMRI study. *Neuroimage*, 44(3), 1125-1132.

[49] Kiwado wa ”johin”. Nyugakushiki no mama no fukuso wa koko ni kiwotsukete! [The keyword is ”to be graceful”. Here are some notes for mothers to choose clothes for attending children’s school entrance ceremonies]. (2004, April 10). Retrieved from https://kosodate-march.jp/jyouhin5566/

[50] ”Kogenshoku” to ”buttaishoku” [”Light source colors” and ”object colors”]. *Official Website of CCS Inc.*. Retrieved from https://www.ccs-inc.co.jp/guide/column/light_color/vol12.html

[51] Korean Cultural Center. Kankoku bunka no aru hi – Oyako de tsukuru shippo koge [One day of Korean culture – Seven-treasures crafts made by parents and children]. (2017, May 22). Retrieved from http://www.koreanculture.jp/info_news_view.php?cate=5&page=1&number=5401&keyfield=&keyfield1=&key=

[52] Korean Cultural Center Osaka. Sikisai de yomu Kankoku [Reading Korea through colors]. Retrieved from http://www.k-culture.jp/korea_story_01.php

[53] Kwon, Y. G. (2003). A study on the signification and symbol of RED in Korea, China, Japan. *Journal of Korean Society of Color Studies*, 113-117.

[54] Lacey, S., Hagtvedt, H., Patrick, V. M., Anderson, A., Stilla, R., Deshpande, G., Hu, X., Sato, J. R., Reddy, S. & Sathian, K. (2011). Art for reward's sake: Visual art recruits the ventral striatum. *Neuroimage*, 55(1), 420-433.

[55] Leder, H., Belke, B., Oeberst, A., & Augustin, D. (2004). A model of aesthetic appreciation and aesthetic judgments. *British Journal of Psychology*, 95(4), 489-508.

[56] Lee, Y., Kobayashi, M., & Sobagaki, H. (2010). Relation between "beauty" and metric perceived quantities of two-color combinations, *Journal of the Color Science Association of Japan*, 34(2), 131-142.

[57] Lenneberg, E. H., & Roberts, J. M. (1956). The language of experience: A study in methodology. *Memoir 13 of International Journal of American Linguistics*.

[58] Liu, X., & Xin, C. (2007). Korean traditional costume's colors and modern design. *Journal of Zhengzhou University of Light Industry (Social Science)*, 2007 February, 47-48, 55.

[59] Lucy, J. A., & Shweder, R. A. (1979). Whorf and his critics: Linguistic and nonlinguistic influences on color memory. *American Anthropologist*, 81(3), 581-615.

[60] Lucy, J. A., & Shweder, R. A. (1988). The effect of incidental conversation on memory for focal colors. *American Anthropologist*, 90(4), 923-931.

[61] Machado, P., & Cardoso, A. (1998, November). Computing aesthetics. In *Brazilian Symposium on Artificial Intelligence* (pp. 219-228). Springer, Berlin, Heidelberg.

[62] Macmillan, N. A., & Creelman, C.D. (2005). *Detection theory: A user's guide. 2nd ed.*. Lawrence Erlbaum Associates, IncMahwah, New Jersey.

[63] Martindale, C., & Moore, K. (1988). Priming, prototypicality, and preference. *Journal of Experimental Psychology: Human Perception and Performance*, 14(4), 661.

[64] Martindale, C., Moore, K., & Borkum, J. (1990). Aesthetic preference: Anomalous findings for Berlyne's psychobiological theory. *The American Journal of Psychology*, 53-80.

[65] Matsuura, A., & Matsui, T. (2008). Rule extraction of impression for harmony among multiple colors using neural network. *Journal of the Color Science Association of Japan*, 32, 48-49.

[66] McCarthy, C. R. J. (1990). Ecologically relevant stimuli and color memory. *The Journal of General Psychology*, 117(4), 369-377.

[67] Minari, Y. (2013). Kankoku kyute ryori to yakuzen [Korean court cuisine and medicinal cooking]. *Journal of Cookery Science of Japan*, 46(2), 139-141.

[68] Mitchell, T. M. (1997). *Machine learning*. McGraw Hill, Burr Ridge, Illinois.

[69] Mitsuboshi, M. (2010). The colors of London. *Studies in Humanities*, 170, 1-21.

[70] Mitsuboshi, M. (2011). A line of noisy color pollutions in Japan and their solutions. *Bulletin of the Institute for Humanities Research of Kanagawa University*, 46, 35-51.

[71] Mitsuboshi, M. (2013). Public color pollutions and Issues of landscape: Reality and solutions. *Bulletin of the Institute for Humanities Research of Kanagawa University*, 50, 41-67.

[72] Moon, P., & Spencer, D. E. (1943). A metric based on the composite color stimulus. *Journal of the Optical Society of America*, 33(5), 270-277.

[73] Moon, P., & Spencer, D. E. (1944). Geometric formulation of classical color harmony. *Journal of the Optical Society of America*, 34(1), 46-59.

[74] Moon, P., & Spencer, D. E. (1944). Area in color harmony. *Journal of the Optical Society of America*, 34(2), 93-103.

[75] Moon, P., & Spencer, D. E. (1944). Aesthetic measure applied to color harmony. *Journal of the Optical Society of America*, 34(4), 234-242.

[76] Mori, N., Nayatani, Y., Tsujimoto, A., Ikeda, T., & Nanba, S. (1966). An appraisement of two-color harmony by paired comparison technique. *The Japanese Journal of Ergonomics*, 2(4), 2-14_1.

[77] Murayama, K. (2016, May 15). Nihonjin ga konomu ”jimi de johinna iro”, chugokujin ga keensuru riyu ha nani? [Why do Chinese people not like the ” restrained and graceful colors” that Japanese people like?]. *Searchina*. Retrieved from http://news.searchina.net/id/1609736?page=1

[78] Nakayama, K., Sato, M., Shiraga, K., Yamamoto, S., & Inui, M. (2001). Study on color appearance of architectural surface consisted in multiple colors. *Journal of Architecture, Planning & Environmental Engineering*, 544, 1-7.

[79] Nayatani, Y., Tujimoto, A., Sobagaki, H., Asano, C., Machihara, S., Ikeda, J.,Nanba, S., Tsujimoto, A., Sobagaki, H., & Hirata, M. (1968). Analysis of three-color harmony affection by Semantic Differential. *Television*, 22, 441-450.

[80] Newhall, S. M., Nickerson, D., & Judd, D. B. (1943). Final report of the OSA subcommittee on the spacing of the Munsell colors. *The Journal of General Psychology*, 33(7), 385-418.

[81] Noda, T. (2012). Study on the survey research of the apparel and fashion industry in Northern Europe. *Bulletin of Gifu City Women's College*, 62, 101-106

[82] Ogawa, K.. (2011). *Kansei evaluation by using multidimensional neural networks based on affective dimensional model* (Unpublished master thesis). Waseda University, Tokyo, Japan.

[83] Ogawa, K., Muramatsu, K., & Matsui, T. (2010). Construction of simulation for impression evaluation of colors using neural network. *Journal of Color Science Association of Japan*, 34, 46-47.

[84] Ogawa, K., Muramatsu, K., & Matsui, T. (2011). Color kansei evaluation using multiple linked neural network. *Journal of Color Science Association of Japan*, 35, 70-71.

[85] Ogawa, K., Muramatsu, K. & Matsui, T. (2011). Construction of affective dimensions for colors by using neural networks. *Technical Report of the Institute of Electronics, Information and Communication Engineers*, 111(241), 89-94.

[86] Ogawa, K., Muramatsu, K. & Matsui, T. (2012). Kansei evaluation by using multidimensional neural networks based on affective dimensional model. *Journal of Color Science Association of Japan*, 36, 78-79.

[87] Ohmi, G., et al. (2011). Shikisai Chowa Ron [Theories on Color Harmony]; In the Color Science Association of Japan (Eds.), *Handbook of Color Science (3rd Edition)* (pp. 293-335). University of Tokyo Press, Tokyo.

[88] Osgood, C. E. (1962). Studies on the generality of affective meaning systems. *American Psychologist*, 17(1), 10.

[89] Osgood, C. E. (1969). On the whys and wherefores of E, P, and A. *Journal of Personality and Social Psychology*, 12(3), 194.

[90] Ou, L. C., & Luo, M. R. (2006). A colour harmony model for two - colour combinations. *Color Research & Application*, 31(3), 191-204.

[91] Oyama, T. (2001). Affections induced by two-color combinations: Comparisons with those of single colors. *Journal of Color Science Association of Japan*, 25, 98-99.

[92] Oyama, T. (2001). Color harmony or color-combination effects : From a psychological point of view. *Journal of the Color Science Association of Japan*, 25(4), 283-287.

[93] Oyama, T. (2004). *Shikisai shinrigaku nyumon* [*A guide to psychology of color*]; Chuokoron-Shinsha, Inc., Tokyo.

[94] Oyama, T., & Miyata (Ito), K., (2012). Effects of hue-, value-, chroma-differences on affections induced by color-combinations, *Journal of the Color Science Association of Japan*, 36(4), 277-282.

[95] Piao, M. (2002), Hanguo de chuantong secai he shenmei yishi [Korean traditonal colors and aesthetics]. *Contemporary Korea*, 2002 April, 86-89.

[96] Pitchford, N. J., & Mullen, K. T. (2002). Is the acquisition of basic-colour terms in young children constrained?. *Perception*, 31(11), 1349-1370.

[97] Poortinga, Y. H., & Van de Vijver, F. J. (1997). Is there no cross-cultural evidence in colour categories of psychological laws, only of cultural rules?. *Behavioral and Brain Sciences*, 20(2), 205-206.

[98] Reber, R., Schwarz, N., & Winkielman, P. (2004). Processing fluency and aesthetic pleasure: Is beauty in the perceiver's processing experience?. *Personality and Social Psychology Review*, 8(4), 364-382.

[99] Regier, T., Kay, P., & Khetarpal, N. (2009). Color naming and the shape of color space. *Language*, 85(4), 884-892.

[100] Roberson, D., Davies, I., & Davidoff, J. (2000). Color categories are not universal: Replications and new evidence from a stone-age culture. *Journal of Experimental Psychology: General*, 129(3), 369.

[101] Roberson, D., Davidoff, J., Davies, I. R., & Shapiro, L. R. (2005). Color categories: Evidence for the cultural relativity hypothesis. *Cognitive Psychology*, 50(4), 378-411.

[102] Rosch, E. H. (1973). Natural categories. *Cognitive Psychology*, 4(3), 328-350.

[103] Rosch, E. (1975). Cognitive reference points. *Cognitive Psychology*, 7(4), 532-547.

[104] Schaffer, J. D., Caruana, R. A., & Eshelman, L. J. (1990). Using genetic search to exploit the emergent behavior of neural networks. *Physica D: Nonlinear Phenomena*, 42(1-3), 244-248.

[105] Schaffer, J. D., Whitley, D., & Eshelman, L. J. (1992, June). Combinations of genetic algorithms and neural networks: A survey of the state of the art. In *Combinations of Genetic Algorithms and Neural Networks, 1992., COGANN-92. International Workshop on* (pp. 1-37). IEEE.

[106] Schanda, J. (2007). CIE colorimetry. In: J. Schanada (Ed.), *Colorimetry: Understanding the CIE System* (pp.25-79). John Wiley & Sons, Hoboken, New Jersey.

[107] Schelwach, S., & Jeon, C. (2017). Contemporary traditions: South Korea's emerging design identity–An exploration among professional and academic design experts aged 30-50. *Archives of Design Research*, 30(3), 43-54.

[108] Schloss, K. B., & Palmer, S. E. (2011). Aesthetic response to color combinations: Preference, harmony, and similarity. *Attention, Perception, & Psychophysics*, 73(2), 551-571.

[109] Shin shakaijin wa hikken! Ofisu de shippaishinai meiku no pointo [New working adults must read this! Points of make-up never failing in office]. (2017, January 20). Retrieved from http://web.mbkr.jp/plusdays/3772/

[110] 2017 haru natsu meiku wa kono waza de kimari! Enkinkan to kireime kara de otona bijin no yoyu o enshutsu [The key of make-up in spring and summer in 2017 is this technique!

Show the leisure of adult beauty using beautiful skin feelings and beautiful colors]. (2017, April 13). Retrieved from https://andb.life/8004

[111] Suzuki, T. (2003). Kogenshoku [Light source color]. In: The Color Science Association of Japan (Ed.), *Encyclopedia of Color Science* (pp.176-177). University of Tokyo Press, Tokyo.

[112] Suzuki, T. (2003). Buttaishoku [Object color]. In: The Color Science Association of Japan (Ed.), *Encyclopedia of Color Science* (p.408). University of Tokyo Press, Tokyo.

[113] Szabo, F., Bodrogi, P., & Schanda, J. (2010). Experimental modeling of colour harmony. *Color Research & Application*, 35(1), 34-49.

[114] Tsukin kode, kyo wa dore ni suru? [What color of clothing to wear to go to work today?]. (2017, January 24). Retrieved from https://www.rcawaii.com/blog/fashion/1685

[115] Tsutsui, A., & Ohmi, G. (2009). Pleasantness, interestingness and color affective scales in color combinations, *Bulletin of Japanese Society for Science of Design*, 56(3), 81-88.

[116] Uchikawa, K., & Boynton, R. M. (1987). Categorical color perception of Japanese observers: Comparison with that of Americans. *Vision Research*, 27(10), 1825-1833.

[117] Utsukushiki Kankoku no dentoshoku [Beautiful traditional colors in Korea]. Retrieved from http://s.plus-han.com/?mode=f2

[118] Wierzbicka, A. (2008). Why there are no 'colour universals' in language and thought. *Journal of the Royal Anthropological Institute*, 14(2), 407-425.

[119] Yamasaki, T. (2005). *Johinna hito, Gehinna hito* [Men of good taste, Men of bad taste]. PHP Institute, Tokyo.

[120] Yamashita, L. (2016). Iro o hyogen suru tame ni: I. Iro no shurui ni tsuite [For color representation: I. On types of color]. *Technical Sheet of Technology Research Institute of Osaka Prefecture*, No.16001.

[121] Yochui! Danse ga kirau jose fasshon [Watch out! Female fashion disliked by men]. (2014, June 16). Retrieved from http://magazine.campus-web.jp/archives/2590

[122] Yoshida, T., Horii, C., & Sato, K. (2004). Projective texture for mixed reality museum based on color reflection measurement. *Transactions of the Virtual Reality Society of Japan*, 9(1), 97-104.

[123] Yoshimura, K. (2007). Cultural Differences between English and Japanese Synesthetic Expressions with Color Terms : The Real Meaning of Synesthesia and Japanese Situation-Centered Expressions for Onomatopoeia. *Journal of Inquiry and Research of Kansai Gaidai College*, 86, 19-37.

[124] Zeki, S. (2001). Artistic creativity and the brain. *Science*, 293(5527), 51-52.

[125] Zeki, S. (2013) Statement on Neuroesthetics, *Online Journal of Neuroesthetics*, 2013.

Appendix A. RGB and $L^*a^*b^*$ Values of Multicolor Stimuli Used in Experiment 1 and Experiment 2

The indexes (1-16) of the component colors of each multicolor stimulus follows the configuration shown in figure 5.1.

13	14	15	16
9	10	11	12
5	6	7	8
1	2	3	4

Figure 5.1: Arrangement of the indexes of the component colors of each multicolor stimulus.

The RGB and $L^*a^*b^*$ values of the component colors of each multicolor stimulus employed in Experiment 1 and Experiment 2 are listed in the following tables (see next page), and the images of the stimuli are displayed after the tables in Figure 5.2.

Table 5.1: RGB and $L^*a^*b^*$ values of component colors of Stimulus 1.

Component color no.	R	G	B	L*	a*	b*
1	227.83	178.80	239.01	78.87	25.97	-23.10
2	2.78	209.03	8.74	73.39	-68.00	68.90
3	131.62	166.60	192.30	66.42	-7.73	-17.37
4	71.40	153.78	226.64	60.91	-7.94	-45.09
5	164.02	200.45	125.24	76.75	-22.30	33.46
6	102.88	223.75	180.59	81.34	-43.25	10.42
7	210.76	2.72	24.37	45.04	69.94	51.74
8	138.54	218.34	67.41	79.89	-43.26	62.40
9	197.76	162.67	184.89	70.60	16.10	-6.38
10	228.14	195.84	52.75	80.18	1.62	70.45
11	11.04	39.16	183.17	25.00	38.79	-78.27
12	32.80	56.30	115.66	24.48	7.86	-37.67
13	118.64	22.90	129.22	29.47	49.92	-37.53
14	19.51	42.46	190.67	26.70	39.58	-79.87
15	85.31	77.79	208.81	40.12	32.93	-68.21
16	250.42	247.09	147.04	95.94	-9.00	48.34

Table 5.2: RGB and $L^*a^*b^*$ values of component colors of Stimulus 2.

Component color no.	R	G	B	L*	a*	b*
1	59.54	74.49	171.81	34.92	19.29	-55.15
2	146.30	50.92	244.34	44.65	65.07	-80.65
3	85.72	38.16	83.74	23.64	27.75	-17.87
4	28.84	196.84	202.10	72.10	-39.23	-14.66
5	149.70	20.19	49.19	32.82	52.12	20.19
6	197.73	90.11	213.93	56.03	56.45	-44.78
7	34.14	178.03	48.54	63.73	-56.38	51.72
8	101.19	86.43	235.29	45.21	38.09	-74.89
9	213.79	160.53	220.70	72.68	27.87	-22.70
10	55.15	16.55	172.89	22.90	49.05	-75.53
11	136.46	131.92	65.21	54.41	-5.29	36.18
12	217.47	209.92	4.23	82.74	-10.38	81.08
13	121.05	240.75	116.95	86.13	-53.27	48.39
14	216.74	136.57	110.70	65.09	29.50	27.38
15	6.23	128.61	223.50	52.03	-1.81	-57.50
16	231.60	72.73	33.84	54.78	60.92	56.58

Table 5.3: RGB and $L^{*}a^{*}b^{*}$ values of component colors of Stimulus 3.

Component color no.	R	G	B	L*	a*	b*
1	0.03	81.08	135.18	32.66	-4.70	-36.83
2	211.44	171.76	232.82	75.44	23.09	-25.14
3	192.82	42.97	245.82	50.86	76.13	-71.02
4	88.81	152.61	186.50	59.81	-14.00	-24.40
5	222.06	85.02	118.08	55.96	56.30	11.14
6	42.56	118.93	116.37	45.36	-24.62	-5.92
7	32.59	163.18	228.60	62.83	-16.94	-43.25
8	108.99	44.84	75.20	28.64	31.70	-4.21
9	14.98	15.34	6.80	4.25	-1.10	3.52
10	140.37	143.58	77.67	58.22	-8.67	34.39
11	192.40	171.13	239.66	73.73	18.19	-31.62
12	123.71	132.11	201.25	56.67	9.92	-37.32
13	145.65	195.46	234.64	76.50	-9.61	-25.08
14	53.63	155.07	118.76	57.50	-37.28	10.13
15	157.26	51.79	172.79	42.00	55.69	-43.64
16	4.42	57.78	233.19	33.96	43.35	-92.41

Table 5.4: RGB and $L^{*}a^{*}b^{*}$ values of component colors of Stimulus 4.

Component color no.	R	G	B	L*	a*	b*
1	195.95	135.99	17.44	61.62	17.37	63.60
2	50.89	3.06	246.48	30.42	67.10	-105.76
3	212.21	113.55	252.58	63.81	55.52	-53.87
4	202.54	124.60	33.30	59.81	26.39	58.36
5	54.12	211.93	82.57	75.14	-61.05	50.38
6	188.39	40.77	186.53	46.72	68.24	-43.83
7	13.69	182.17	31.04	64.79	-60.17	58.01
8	250.93	151.24	68.73	72.28	33.79	58.18
9	110.16	237.28	129.51	84.84	-53.56	40.85
10	16.34	168.07	114.54	61.00	-47.40	17.12
11	37.53	197.68	135.47	71.04	-52.16	19.56
12	22.41	229.62	221.38	82.43	-47.96	-9.64
13	14.36	163.21	226.76	62.57	-18.55	-42.66
14	88.67	170.01	115.89	63.41	-35.37	19.99
15	113.95	4.34	219.66	34.31	66.70	-83.82
16	239.14	122.09	145.31	65.69	47.61	9.43

Table 5.5: RGB and $L^*a^*b^*$ values of component colors of Stimulus 5.

Component color no.	R	G	B	L*	a*	b*
1	253.16	103.05	56.06	63.23	56.67	55.43
2	220.96	175.70	43.72	74.34	8.33	67.95
3	2.30	43.11	52.63	15.32	-11.08	-10.32
4	229.91	180.83	33.64	76.52	9.39	72.92
5	108.43	168.67	82.21	63.51	-32.50	37.77
6	152.35	53.72	48.35	37.52	41.75	26.51
7	53.12	181.90	64.44	65.44	-53.59	47.41
8	68.85	122.85	56.27	46.61	-29.11	30.18
9	90.44	162.62	191.19	63.00	-17.45	-22.17
10	125.57	157.43	81.46	61.16	-21.21	35.65
11	77.14	36.16	215.31	31.95	54.13	-85.39
12	73.48	231.75	28.87	81.45	-66.36	73.24
13	156.53	72.40	129.57	43.50	41.00	-15.38
14	137.41	13.30	111.46	31.51	54.42	-22.80
15	55.78	234.18	176.32	83.23	-55.77	15.05
16	29.56	91.70	78.04	34.73	-23.46	1.75

Table 5.6: RGB and $L^*a^*b^*$ values of component colors of Stimulus 6.

Component color no.	R	G	B	L*	a*	b*
1	77.07	143.17	203.01	57.09	-7.71	-37.90
2	119.84	201.75	106.79	74.37	-39.95	39.01
3	188.58	91.17	242.20	56.27	57.78	-60.34
4	191.29	4.99	192.57	45.57	75.07	-49.14
5	232.51	88.68	18.18	57.26	54.86	63.72
6	218.05	24.06	140.14	49.32	74.10	-11.73
7	233.46	133.11	13.25	65.81	34.63	69.91
8	112.43	176.25	66.65	65.90	-35.49	47.74
9	228.46	117.23	59.16	62.01	41.25	51.21
10	49.97	216.67	189.11	78.09	-47.84	1.10
11	168.86	27.31	213.22	43.73	70.62	-64.30
12	116.00	165.91	230.63	66.53	-3.62	-38.35
13	28.23	153.98	4.18	55.51	-52.02	55.29
14	23.26	230.95	139.64	81.31	-63.19	31.00
15	164.27	60.15	158.25	43.69	52.38	-32.23
16	20.14	164.67	44.96	59.13	-54.27	48.36

Table 5.7: RGB and $L^*a^*b^*$ values of component colors of Stimulus 7.

Component color no.	R	G	B	L*	a*	b*
1	157.83	247.43	170.64	90.32	-39.51	28.20
2	84.61	116.76	6.36	45.24	-23.12	47.87
3	133.39	252.24	184.46	90.83	-46.40	21.74
4	101.22	150.64	146.16	58.72	-18.02	-3.22
5	129.46	37.11	74.19	31.05	42.23	0.50
6	127.79	36.18	103.63	31.56	44.83	-17.89
7	151.24	206.39	120.00	77.57	-30.09	36.91
8	171.90	193.08	46.44	74.57	-20.86	65.28
9	158.01	196.33	107.38	74.97	-24.33	40.04
10	111.94	124.64	55.29	50.03	-12.99	35.50
11	212.39	127.89	26.20	61.84	28.72	62.79
12	92.91	209.30	31.33	75.06	-55.19	67.64
13	165.59	187.84	181.22	74.33	-8.86	0.64
14	26.86	244.44	108.04	85.14	-70.60	50.61
15	142.83	25.19	123.20	33.99	54.43	-26.26
16	39.17	78.66	33.38	29.68	-22.38	21.97

Table 5.8: RGB and $L^*a^*b^*$ values of component colors of Stimulus 8.

Component color no.	R	G	B	L*	a*	b*
1	87.60	169.08	131.63	63.35	-32.88	11.35
2	203.15	147.03	17.05	65.18	14.78	66.62
3	63.36	231.71	226.82	83.68	-44.26	-10.59
4	252.40	82.53	251.79	64.77	78.25	-51.48
5	156.54	166.28	153.34	67.15	-5.41	5.47
6	203.44	155.66	96.20	67.86	13.27	38.15
7	44.53	15.28	38.85	9.36	18.19	-9.12
8	97.77	3.34	246.17	34.74	69.89	-98.26
9	81.14	243.68	7.38	85.26	-68.67	78.91
10	101.28	70.84	92.27	34.11	16.26	-7.29
11	199.02	113.64	14.61	56.80	30.25	61.17
12	12.24	135.36	195.53	52.82	-14.06	-40.50
13	156.73	237.77	199.30	88.02	-31.93	10.47
14	192.50	83.19	133.11	51.21	48.46	-5.32
15	164.62	130.37	168.84	58.82	18.71	-15.28
16	58.98	250.56	21.91	86.95	-73.88	78.66

Table 5.9: RGB and $L^*a^*b^*$ values of component colors of Stimulus 9.

Component color no.	R	G	B	L*	a*	b*
1	105.23	251.25	123.74	88.77	-60.07	48.29
2	156.03	2.86	195.98	39.25	69.35	-61.69
3	189.10	55.42	101.32	45.51	56.17	5.64
4	66.95	20.05	56.13	15.44	26.53	-11.92
5	123.67	131.40	213.95	56.99	13.13	-43.95
6	105.35	102.55	242.94	49.77	31.86	-71.75
7	85.45	34.11	177.34	29.09	47.72	-67.83
8	235.31	124.32	128.95	65.28	43.95	18.00
9	1.18	48.06	198.81	28.34	38.25	-81.94
10	155.76	202.08	213.66	78.35	-13.74	-11.02
11	175.28	20.77	205.57	43.98	71.91	-59.43
12	34.37	145.33	146.55	54.57	-30.49	-9.87
13	4.11	238.38	156.77	83.75	-63.64	25.62
14	170.49	107.55	179.21	54.44	34.53	-27.78
15	120.03	104.27	252.10	51.91	36.38	-73.31
16	123.74	29.90	218.81	37.25	63.25	-78.43

Table 5.10: RGB and $L^*a^*b^*$ values of component colors of Stimulus 10.

Component color no.	R	G	B	L*	a*	b*
1	87.20	80.68	96.27	35.39	5.33	-8.10
2	221.82	135.52	6.41	64.76	28.59	69.33
3	63.33	184.54	110.26	67.02	-47.67	27.51
4	66.67	48.30	54.69	22.34	9.51	-0.65
5	126.20	8.42	95.02	28.21	50.95	-17.50
6	90.95	70.05	67.42	31.96	8.90	5.20
7	189.85	33.58	204.79	47.30	72.35	-53.49
8	77.62	87.43	168.13	39.51	14.53	-45.52
9	9.94	196.90	175.97	71.33	-46.05	-1.69
10	236.45	15.81	175.30	53.73	81.73	-25.56
11	185.05	215.86	141.22	82.69	-19.95	33.33
12	214.81	12.47	193.50	50.23	79.14	-41.97
13	172.38	222.98	84.39	83.26	-32.19	60.07
14	132.22	115.76	45.77	49.23	0.21	39.88
15	216.09	11.84	133.52	48.24	74.58	-9.37
16	59.34	230.33	138.02	81.53	-59.88	32.19

Table 5.11: RGB and $L^{*}a^{*}b^{*}$ values of component colors of Stimulus 11.

Component color no.	R	G	B	L*	a*	b*
1	25.16	218.56	10.90	76.55	-69.26	71.29
2	36.22	55.46	198.25	30.91	35.62	-77.34
3	91.19	6.46	212.61	30.63	62.67	-85.97
4	92.62	201.08	224.31	75.36	-27.39	-21.55
5	172.51	139.04	229.22	63.77	26.86	-41.47
6	25.37	23.19	8.73	7.76	-0.91	7.95
7	82.66	42.35	254.51	37.16	61.23	-98.96
8	105.95	192.36	176.79	72.10	-30.21	-0.71
9	196.51	232.63	79.22	87.50	-27.45	67.02
10	52.20	18.58	149.74	20.30	42.75	-65.76
11	138.33	164.24	152.49	65.20	-11.16	2.95
12	114.76	41.45	195.05	35.45	53.69	-67.68
13	250.74	36.50	211.06	58.66	85.16	-38.31
14	224.69	17.96	121.86	49.87	75.25	0.36
15	188.38	187.10	208.14	76.52	3.53	-10.45
16	136.04	143.01	88.36	57.82	-9.31	28.17

Table 5.12: RGB and $L^{*}a^{*}b^{*}$ values of component colors of Stimulus 12.

Component color no.	R	G	B	L*	a*	b*
1	140.44	5.38	187.88	36.05	65.37	-62.28
2	212.66	213.26	8.45	83.23	-13.94	80.83
3	232.60	24.21	216.98	55.00	83.75	-47.69
4	133.16	4.17	32.94	27.89	49.77	24.00
5	151.26	184.06	9.44	70.40	-25.99	68.89
6	44.92	59.27	182.33	30.86	29.43	-68.04
7	196.50	150.29	222.60	68.58	27.59	-30.17
8	115.77	250.90	66.29	88.59	-61.56	70.87
9	176.12	152.17	71.23	63.85	1.95	45.08
10	71.20	174.36	206.88	66.27	-23.28	-25.94
11	19.44	94.19	205.59	41.11	12.92	-64.92
12	244.37	253.40	108.52	96.73	-16.62	66.32
13	110.45	86.30	115.94	40.04	14.59	-12.99
14	196.34	192.70	253.06	79.83	10.68	-29.46
15	51.78	95.73	180.71	41.03	7.57	-50.58
16	41.69	82.61	206.30	38.55	22.24	-69.44

Table 5.13: RGB and $L^*a^*b^*$ values of component colors of Stimulus 13.

Component color no.	R	G	B	L*	a*	b*
1	55.83	83.74	117.81	34.57	-2.74	-23.01
2	89.84	18.75	128.82	23.84	44.70	-46.67
3	69.92	144.75	7.02	53.59	-39.96	53.91
4	156.50	108.02	164.55	52.37	27.22	-22.67
5	78.06	74.15	109.78	33.07	8.36	-20.17
6	140.32	31.23	33.26	31.70	45.75	28.49
7	58.34	27.27	23.72	14.42	15.32	9.65
8	48.09	174.29	138.08	63.88	-42.08	8.35
9	132.96	240.16	162.56	86.99	-44.80	27.58
10	183.23	108.23	63.33	53.67	27.72	38.02
11	238.68	204.54	194.22	84.98	11.23	9.84
12	15.75	231.20	167.19	81.76	-59.35	17.65
13	241.73	139.04	77.86	68.52	36.46	49.61
14	22.57	107.49	93.72	40.42	-28.10	0.32
15	217.39	250.90	13.73	93.72	-28.82	88.15
16	91.07	254.76	105.83	89.27	-65.74	56.61

Table 5.14: RGB and $L^*a^*b^*$ values of component colors of Stimulus 14.

Component color no.	R	G	B	L*	a*	b*
1	97.07	26.67	150.35	27.55	47.35	-53.95
2	37.31	96.03	238.77	44.27	24.08	-78.56
3	178.59	61.95	189.24	47.65	59.53	-44.07
4	83.29	94.34	125.66	40.05	1.72	-19.19
5	51.16	73.33	74.61	29.37	-8.63	-3.70
6	109.61	121.92	237.64	54.59	20.48	-61.03
7	213.24	73.94	244.31	57.31	70.98	-59.61
8	184.38	220.53	133.43	83.82	-23.05	38.64
9	13.46	244.80	171.41	85.96	-63.07	21.23
10	16.69	249.68	75.19	86.45	-75.11	65.09
11	242.23	44.22	0.23	53.62	72.14	67.57
12	224.01	216.89	85.78	85.42	-9.03	63.15
13	25.07	160.17	102.11	58.31	-46.08	20.25
14	236.65	224.55	157.87	89.20	-2.86	34.29
15	227.95	53.31	65.87	52.02	66.77	37.16
16	1.05	23.07	151.93	18.03	37.66	-70.99

Table 5.15: RGB and $L^*a^*b^*$ values of component colors of Stimulus 15.

Component color no.	R	G	B	L*	a*	b*
1	76.32	197.74	163.81	72.38	-41.54	6.52
2	45.78	72.25	34.64	27.79	-16.68	19.10
3	143.69	36.79	218.29	40.82	64.60	-72.16
4	7.75	216.83	59.36	76.05	-68.00	60.11
5	192.51	14.38	180.45	45.50	73.19	-42.14
6	46.95	253.62	131.76	88.37	-69.14	43.78
7	137.23	69.24	96.46	38.89	31.98	-2.20
8	130.60	131.43	119.58	54.59	-1.99	6.32
9	39.31	54.86	164.52	27.98	25.85	-62.16
10	154.57	46.50	221.26	43.71	64.19	-69.06
11	134.57	248.51	72.80	88.62	-55.32	69.07
12	242.02	154.94	198.00	73.97	37.68	-7.97
13	16.19	150.75	229.44	58.96	-11.42	-49.79
14	130.00	135.87	203.64	58.25	10.27	-36.16
15	167.05	118.08	142.12	55.33	22.56	-5.05
16	189.47	131.15	30.80	59.69	17.29	58.25

Table 5.16: RGB and $L^*a^*b^*$ values of component colors of Stimulus 16.

Component color no.	R	G	B	L*	a*	b*
1	68.36	235.84	150.30	83.54	-58.71	28.75
2	137.50	239.00	65.19	85.91	-51.43	69.02
3	232.53	142.06	188.36	69.84	39.66	-8.89
4	174.50	35.63	248.24	47.52	74.88	-77.93
5	108.26	19.43	27.52	23.35	38.90	20.68
6	247.92	145.44	123.38	71.22	38.04	29.30
7	164.47	55.64	177.20	43.92	56.68	-43.10
8	62.67	161.57	244.83	63.65	-8.22	-50.79
9	185.24	247.26	194.35	92.11	-27.85	18.97
10	133.90	113.93	162.78	51.33	15.66	-23.48
11	78.32	113.07	169.64	46.96	0.16	-34.65
12	209.68	81.35	167.55	54.66	58.04	-20.12
13	128.88	40.15	77.79	31.50	41.11	-1.23
14	26.80	222.01	173.13	79.13	-54.73	10.93
15	210.85	241.25	160.80	91.61	-20.07	35.48
16	130.92	141.78	251.79	61.88	18.15	-57.13

Table 5.17: RGB and $L^*a^*b^*$ values of component colors of Stimulus 17.

Component color no.	R	G	B	L*	a*	b*
1	159.86	217.26	125.01	81.25	-31.26	39.16
2	5.09	62.18	140.27	27.03	9.32	-49.00
3	25.90	120.95	44.51	44.35	-40.47	32.96
4	82.97	28.15	109.20	22.63	36.12	-36.17
5	183.34	195.74	236.43	79.12	2.09	-21.79
6	193.02	84.79	224.54	54.96	59.00	-52.52
7	134.03	68.17	42.90	37.15	27.19	28.35
8	11.48	43.55	236.07	31.66	51.67	-97.86
9	27.70	212.71	38.71	74.78	-66.89	65.07
10	218.25	152.63	50.43	68.43	19.05	60.40
11	203.37	200.78	24.37	79.14	-11.82	75.49
12	43.72	153.72	82.90	56.24	-44.16	27.86
13	68.16	201.27	157.24	73.12	-45.55	11.02
14	175.05	177.85	237.67	73.91	8.76	-30.42
15	30.65	154.76	124.92	57.03	-39.48	5.94
16	135.07	56.58	190.79	40.31	51.11	-57.15

Table 5.18: RGB and $L^*a^*b^*$ values of component colors of Stimulus 18.

Component color no.	R	G	B	L*	a*	b*
1	65.83	50.77	222.24	33.98	48.02	-86.05
2	6.93	9.27	111.69	10.73	32.83	-57.70
3	16.75	249.74	160.37	87.31	-66.18	28.57
4	167.36	83.62	121.25	47.10	38.37	-4.77
5	4.14	51.01	218.92	31.09	42.80	-89.02
6	17.22	79.24	210.79	37.31	23.79	-74.10
7	156.03	233.22	225.45	87.25	-25.98	-4.19
8	62.55	244.34	115.34	85.60	-66.40	48.04
9	226.58	118.54	149.23	63.31	45.17	3.54
10	2.98	185.75	105.90	66.51	-55.04	28.89
11	30.25	152.05	1.20	54.89	-51.27	55.27
12	239.56	176.31	69.31	76.63	16.63	61.64
13	192.20	167.87	72.35	69.60	0.94	51.31
14	136.58	15.36	172.74	34.81	61.26	-55.35
15	235.34	118.99	94.05	63.71	44.41	35.41
16	233.94	2.95	236.95	55.57	87.77	-58.04

Table 5.19: RGB and $L^{*}a^{*}b^{*}$ values of component colors of Stimulus 19.

Component color no.	R	G	B	L*	a*	b*
1	176.82	209.42	218.72	81.81	-9.72	-8.44
2	28.20	179.77	140.67	65.35	-45.91	8.94
3	132.76	80.00	41.16	39.72	20.07	32.12
4	81.53	171.12	133.61	63.78	-34.98	10.82
5	154.99	193.88	41.78	73.58	-28.19	65.23
6	126.06	195.00	76.71	72.37	-37.72	50.67
7	213.91	173.35	83.76	73.27	7.96	50.81
8	137.26	213.64	61.96	78.41	-42.13	62.85
9	4.97	154.91	137.39	57.10	-38.77	-1.02
10	57.72	234.58	238.68	84.72	-43.44	-15.26
11	124.60	177.90	131.73	67.87	-25.39	17.85
12	72.73	2.41	74.70	15.21	37.28	-25.18
13	169.95	39.55	57.52	39.04	53.29	23.62
14	223.79	154.46	164.94	70.97	27.74	5.58
15	17.64	55.73	198.90	30.39	34.37	-78.59
16	160.41	18.01	167.59	39.00	64.80	-45.27

Table 5.20: RGB and $L^{*}a^{*}b^{*}$ values of component colors of Stimulus 20.

Component color no.	R	G	B	L*	a*	b*
1	132.51	241.62	5.32	86.31	-55.33	80.74
2	185.82	101.62	175.25	55.03	42.15	-24.40
3	18.64	193.30	12.93	68.41	-63.26	64.49
4	163.81	127.70	135.17	57.18	15.26	1.58
5	188.25	70.01	60.70	46.84	48.06	31.95
6	193.79	244.82	42.38	90.54	-34.92	80.65
7	96.54	24.68	238.31	35.03	64.71	-93.35
8	204.74	186.20	202.65	77.54	8.99	-5.80
9	146.26	92.23	41.14	44.61	19.35	37.97
10	226.02	230.01	193.44	90.28	-5.91	17.43
11	92.26	126.54	167.96	51.68	-3.85	-26.28
12	73.69	36.48	78.57	21.11	23.51	-18.53
13	88.60	35.58	55.14	22.56	26.60	-0.11
14	141.73	13.05	106.42	32.18	55.02	-18.46
15	181.45	224.21	153.38	84.91	-23.80	30.03
16	118.57	207.93	82.14	75.93	-44.84	52.11

Table 5.21: RGB and $L^*a^*b^*$ values of component colors of Stimulus 21.

Component color no.	R	G	B	L*	a*	b*
1	216.92	142.43	229.95	69.28	39.48	-32.89
2	106.98	91.32	124.69	41.31	11.86	-16.45
3	65.27	236.94	119.02	83.43	-63.41	43.75
4	64.77	109.96	179.15	45.74	1.31	-42.15
5	102.59	46.37	218.34	36.45	54.28	-79.60
6	148.97	95.26	56.53	45.88	19.62	31.34
7	55.84	133.17	110.52	50.32	-29.31	4.73
8	189.03	17.96	216.07	46.87	76.11	-60.66
9	173.37	34.85	218.89	45.39	70.65	-64.82
10	50.96	154.87	138.48	57.79	-33.87	-0.53
11	41.39	1.44	196.73	23.36	56.84	-88.97
12	195.02	107.37	14.49	54.86	31.78	59.62
13	149.37	44.41	185.80	40.45	58.30	-53.86
14	136.24	64.53	233.85	44.54	55.91	-74.99
15	193.34	226.19	17.54	85.29	-27.57	80.66
16	46.80	187.95	177.66	69.00	-39.55	-5.94

Table 5.22: RGB and $L^*a^*b^*$ values of component colors of Stimulus 22.

Component color no.	R	G	B	L*	a*	b*
1	198.13	127.99	108.50	60.76	26.30	22.42
2	155.87	218.22	171.05	81.96	-27.73	16.67
3	133.52	76.20	179.51	43.09	39.29	-46.11
4	97.31	144.76	226.40	59.46	2.00	-47.14
5	214.95	229.19	239.45	90.16	-3.42	-6.72
6	207.94	0.35	0.79	44.28	69.27	59.65
7	22.30	66.49	5.81	24.18	-24.98	28.97
8	108.14	86.97	138.05	40.76	17.29	-25.36
9	236.17	76.12	86.21	56.54	62.60	31.13
10	219.17	86.82	35.22	54.76	51.15	54.85
11	129.49	218.45	98.00	79.70	-44.24	49.56
12	177.40	160.12	114.85	66.60	1.56	26.08
13	120.77	242.18	21.29	86.03	-58.75	78.80
14	71.36	113.99	149.83	46.07	-7.22	-24.47
15	223.80	119.62	111.54	62.42	40.77	23.66
16	190.28	119.32	219.51	60.88	41.10	-40.42

Table 5.23: RGB and $L^*a^*b^*$ values of component colors of Stimulus 23.

Component color no.	R	G	B	L*	a*	b*
1	118.96	127.02	124.29	52.45	-3.51	0.40
2	58.51	21.82	17.18	12.99	18.22	12.15
3	226.54	59.46	219.71	56.60	75.64	-46.78
4	181.49	222.57	239.19	86.13	-11.14	-12.82
5	35.62	100.44	250.04	46.18	24.91	-81.73
6	164.42	228.58	122.97	84.71	-34.65	44.47
7	3.59	158.83	58.93	57.16	-52.58	40.51
8	134.50	184.87	154.89	70.91	-22.00	9.67
9	150.03	110.53	62.26	49.95	12.05	32.99
10	109.38	2.60	155.25	28.06	55.90	-56.02
11	244.28	24.34	9.08	52.80	76.20	65.47
12	225.99	62.97	2.27	52.38	62.22	64.28
13	207.80	35.83	224.37	51.60	77.83	-57.59
14	24.32	89.90	151.32	36.75	-2.41	-40.08
15	149.22	170.26	165.25	67.95	-8.34	-0.02
16	110.51	35.64	191.74	33.84	54.39	-68.40

Table 5.24: RGB and $L^*a^*b^*$ values of component colors of Stimulus 24.

Component color no.	R	G	B	L*	a*	b*
1	61.66	165.87	218.63	63.86	-17.77	-36.13
2	21.51	247.88	8.02	85.71	-76.74	78.79
3	213.03	213.11	12.71	83.22	-13.68	80.46
4	139.20	240.51	81.98	86.52	-50.59	63.76
5	205.65	153.36	201.35	69.60	26.01	-16.82
6	203.79	12.64	72.22	44.34	68.43	22.80
7	166.63	124.86	248.08	60.71	36.23	-56.72
8	190.86	144.80	76.24	63.55	12.41	42.82
9	65.31	226.07	113.93	80.05	-60.60	42.02
10	208.08	25.08	219.20	50.74	79.09	-56.02
11	7.05	229.28	229.48	82.46	-46.52	-13.86
12	133.65	30.65	45.34	30.40	44.22	18.84
13	180.06	212.00	8.88	80.39	-26.73	77.40
14	193.25	244.06	87.43	90.49	-33.19	66.87
15	162.75	87.47	55.20	45.99	29.65	32.82
16	200.48	184.39	71.10	74.72	-3.82	57.44

Table 5.25: RGB and $L^*a^*b^*$ values of component colors of Stimulus 25.

Component color no.	R	G	B	L*	a*	b*
1	148.52	107.36	23.48	48.65	11.51	49.25
2	6.13	125.24	70.96	45.89	-40.68	20.98
3	86.64	73.27	43.58	31.95	2.48	19.91
4	101.81	177.90	51.94	65.77	-40.63	53.32
5	169.91	112.98	110.49	53.73	22.96	11.10
6	44.69	49.27	157.19	26.30	27.83	-60.45
7	68.60	142.72	240.92	58.26	1.55	-57.14
8	182.19	173.20	244.64	73.53	14.71	-34.70
9	197.71	154.97	241.74	70.44	28.67	-37.75
10	15.21	68.52	251.60	38.19	43.16	-95.76
11	196.91	121.22	173.63	60.54	35.86	-14.98
12	106.32	96.94	54.38	41.23	-0.88	25.57
13	97.65	7.57	120.44	23.70	47.70	-41.50
14	85.01	248.84	141.64	87.71	-61.37	38.45
15	215.81	104.06	117.81	58.30	45.70	14.32
16	210.71	252.76	133.61	94.37	-27.54	51.91

Table 5.26: RGB and $L^*a^*b^*$ values of component colors of Stimulus 26.

Component color no.	R	G	B	L*	a*	b*
1	235.99	188.45	144.69	79.96	13.43	29.00
2	247.04	210.25	244.70	88.23	17.59	-11.93
3	164.82	96.79	121.53	49.56	30.68	-1.41
4	232.53	3.79	39.96	49.81	75.45	49.87
5	120.25	138.46	15.22	54.64	-17.88	54.84
6	167.80	226.86	27.96	83.83	-37.54	77.22
7	111.63	71.46	251.24	44.25	51.73	-85.34
8	155.23	64.71	33.82	39.86	37.28	37.50
9	138.98	211.09	213.44	79.92	-22.95	-8.72
10	212.50	51.95	138.83	50.39	66.63	-9.49
11	223.11	30.86	218.37	53.68	81.08	-50.69
12	229.44	55.56	19.63	52.22	65.69	60.13
13	120.92	212.93	119.70	77.90	-42.70	37.08
14	105.51	128.20	31.99	50.44	-18.83	45.94
15	33.73	221.97	153.75	78.84	-57.57	20.61
16	67.65	220.52	14.82	77.81	-64.49	72.21

Table 5.27: RGB and $L^{*}a^{*}b^{*}$ values of component colors of Stimulus 27.

Component color no.	R	G	B	L*	a*	b*
1	127.74	84.66	44.34	40.24	15.34	30.72
2	159.54	146.66	191.50	63.14	11.80	-21.61
3	39.15	90.98	36.71	34.11	-27.33	25.35
4	216.90	86.16	70.18	54.55	51.60	36.93
5	1.53	204.49	126.84	72.77	-57.82	26.29
6	137.15	222.08	184.33	82.33	-33.20	10.08
7	170.36	45.60	140.38	42.02	57.39	-23.96
8	244.77	151.99	206.19	73.88	41.02	-12.54
9	251.06	225.91	54.53	89.94	-3.24	79.25
10	8.83	115.04	3.52	41.79	-43.13	44.33
11	120.80	242.56	63.48	86.27	-57.61	69.62
12	98.54	110.02	211.88	49.39	18.01	-54.95
13	210.28	115.51	97.04	59.45	36.71	27.53
14	236.10	188.92	188.10	80.84	17.38	7.20
15	241.46	130.08	201.93	68.94	50.01	-17.62
16	115.30	216.55	99.56	78.52	-47.65	47.31

Table 5.28: RGB and $L^{*}a^{*}b^{*}$ values of component colors of Stimulus 28.

Component color no.	R	G	B	L*	a*	b*
1	178.31	145.70	160.32	63.83	14.46	-2.87
2	223.83	168.90	223.23	75.73	27.41	-19.37
3	119.22	36.04	17.37	27.91	36.20	32.33
4	182.13	78.54	171.15	50.02	51.37	-29.77
5	166.37	135.42	182.35	60.66	18.71	-20.15
6	128.73	124.44	126.95	52.63	2.06	-0.76
7	238.67	99.27	29.87	60.09	52.69	62.54
8	61.31	174.65	214.01	66.31	-23.21	-29.82
9	247.39	54.87	193.89	58.75	79.17	-28.42
10	148.95	102.75	130.06	49.14	22.36	-7.38
11	126.39	166.10	189.65	65.87	-9.61	-16.76
12	77.00	22.85	210.62	29.89	57.55	-86.10
13	99.34	197.70	45.74	71.76	-48.98	61.11
14	27.89	230.82	223.47	82.88	-47.68	-10.07
15	254.95	220.38	9.40	88.81	0.38	86.70
16	138.89	254.39	130.31	90.96	-52.08	48.25

Table 5.29: RGB and $L^*a^*b^*$ values of component colors of Stimulus 29.

Component color no.	R	G	B	L*	a*	b*
1	38.20	177.64	32.90	63.59	-56.56	56.57
2	241.22	226.04	131.33	89.64	-3.84	47.98
3	173.25	249.08	31.99	90.49	-44.35	82.01
4	191.82	210.90	199.26	82.82	-8.29	3.47
5	48.68	109.30	3.69	40.75	-33.44	43.95
6	82.95	34.35	114.88	24.08	34.23	-37.49
7	145.93	201.97	107.03	75.83	-31.27	41.04
8	135.80	236.05	229.27	87.18	-31.95	-6.42
9	138.93	229.79	13.22	83.11	-49.02	77.80
10	206.19	85.40	58.31	52.50	47.68	40.59
11	209.71	88.80	42.20	53.56	46.99	50.00
12	7.17	243.62	173.47	85.61	-62.53	19.70
13	219.44	239.47	173.45	91.92	-14.87	29.72
14	233.94	65.46	225.83	58.73	76.30	-46.79
15	234.61	76.52	18.71	55.79	60.34	62.98
16	195.69	21.66	185.83	46.67	73.49	-43.41

Table 5.30: RGB and $L^*a^*b^*$ values of component colors of Stimulus 30.

Component color no.	R	G	B	L*	a*	b*
1	242.79	248.84	7.88	95.05	-17.84	90.55
2	125.94	219.99	61.93	79.75	-48.24	63.94
3	212.74	207.47	160.39	82.84	-3.61	24.19
4	0.57	96.83	230.62	43.31	19.01	-75.58
5	173.49	96.60	161.15	51.86	38.67	-21.23
6	62.03	145.71	250.34	59.34	1.72	-60.61
7	216.67	72.28	174.03	54.69	64.20	-23.71
8	91.32	251.67	21.42	87.95	-68.79	79.86
9	63.83	206.89	21.53	73.45	-61.15	67.73
10	135.47	204.16	188.40	77.19	-25.19	0.62
11	36.12	111.66	89.35	42.07	-28.52	5.35
12	122.02	149.79	37.18	58.27	-21.84	51.89
13	230.86	163.25	41.55	72.39	18.78	67.55
14	144.31	237.56	199.69	87.37	-35.34	9.25
15	174.85	118.89	66.38	55.13	18.66	37.70
16	145.16	63.44	81.42	38.61	36.55	7.02

Table 5.31: RGB and $L^*a^*b^*$ values of component colors of Stimulus 31.

Component color no.	R	G	B	L*	a*	b*
1	232.25	225.73	202.62	89.85	-0.54	12.18
2	236.08	45.60	131.97	54.03	73.96	0.57
3	159.89	232.86	169.31	86.29	-32.99	23.50
4	99.24	188.70	208.50	71.54	-23.86	-18.74
5	153.09	21.67	235.20	42.31	72.16	-79.27
6	13.67	134.39	30.31	48.73	-47.13	43.36
7	96.94	207.27	62.24	74.71	-52.15	58.39
8	225.53	181.72	96.43	76.67	9.27	49.02
9	63.47	64.48	195.65	34.36	33.02	-70.07
10	12.71	174.75	158.17	64.00	-41.46	-2.61
11	190.40	249.20	97.90	91.81	-35.81	64.08
12	66.35	223.75	205.55	80.88	-45.21	-3.51
13	117.59	23.20	143.89	30.18	51.67	-45.53
14	47.78	135.58	90.53	50.46	-35.62	16.27
15	80.27	185.32	131.52	68.09	-41.29	17.86
16	201.61	52.15	172.92	49.60	66.78	-31.15

Table 5.32: RGB and $L^*a^*b^*$ values of component colors of Stimulus 32.

Component color no.	R	G	B	L*	a*	b*
1	234.33	153.65	179.05	72.33	33.52	-0.07
2	189.64	98.20	64.13	52.41	35.51	36.47
3	9.37	120.39	164.49	46.83	-16.04	-32.05
4	71.14	132.05	62.64	49.81	-31.58	30.73
5	75.86	165.87	227.31	64.61	-13.15	-39.67
6	219.58	53.53	101.77	50.99	65.76	13.99
7	226.41	65.41	246.53	58.45	77.03	-58.88
8	157.89	42.16	210.68	43.06	64.54	-64.02
9	167.20	139.35	64.09	59.47	4.33	43.39
10	10.24	59.51	92.08	23.36	-5.99	-24.37
11	161.53	251.46	52.82	90.62	-48.77	77.40
12	193.06	226.01	120.42	85.78	-23.35	47.46
13	40.52	206.78	121.51	73.73	-56.90	30.31
14	29.65	223.31	161.97	79.35	-56.94	17.05
15	24.81	231.65	8.93	80.67	-72.57	74.74
16	10.14	252.08	174.98	88.21	-64.86	22.47

Table 5.33: RGB and $L^*a^*b^*$ values of component colors of Stimulus 33.

Component color no.	R	G	B	L*	a*	b*
1	121.35	93.92	167.18	45.13	23.49	-35.76
2	239.24	158.21	72.12	72.38	25.69	56.20
3	52.32	111.98	6.95	41.84	-33.22	44.14
4	223.43	155.57	51.92	69.74	19.79	61.30
5	132.58	13.73	219.86	37.52	67.95	-78.53
6	112.95	139.74	144.55	56.19	-9.00	-5.90
7	173.50	94.70	19.95	48.95	29.34	52.34
8	116.37	12.20	188.26	32.34	60.70	-68.76
9	9.69	243.33	189.30	85.81	-59.64	11.89
10	239.05	130.91	61.43	66.33	38.83	54.99
11	66.29	193.54	253.30	73.32	-21.18	-40.26
12	90.96	191.98	28.06	69.59	-50.01	63.67
13	152.25	109.80	186.33	53.10	28.38	-34.16
14	66.60	24.18	115.00	19.78	35.00	-44.67
15	163.22	33.67	115.47	38.46	56.71	-14.16
16	166.31	210.88	78.56	79.40	-29.62	58.24

Table 5.34: RGB and $L^*a^*b^*$ values of component colors of Stimulus 34.

Component color no.	R	G	B	L*	a*	b*
1	102.60	225.48	178.65	81.81	-44.31	12.08
2	61.68	193.76	74.19	69.50	-54.96	47.77
3	70.75	1.56	95.55	16.15	40.16	-37.83
4	111.42	77.60	74.17	36.72	14.35	8.03
5	61.84	238.85	219.35	85.59	-48.73	-3.90
6	101.29	122.25	144.07	50.24	-3.87	-14.56
7	124.85	68.80	252.38	45.24	55.60	-84.30
8	46.84	219.72	8.32	77.16	-67.51	72.13
9	84.65	190.93	164.31	70.58	-37.00	3.73
10	43.16	242.81	138.53	85.12	-65.48	36.48
11	64.11	147.54	233.45	59.15	-3.91	-51.64
12	228.38	123.04	112.90	63.78	40.91	24.82
13	79.50	14.11	192.22	27.52	56.23	-79.26
14	33.65	90.76	100.95	35.19	-16.63	-11.74
15	225.81	5.42	215.24	53.10	83.91	-49.80
16	73.46	63.84	124.54	30.38	16.49	-33.63

Table 5.35: RGB and $L^*a^*b^*$ values of component colors of Stimulus 35.

Component color no.	R	G	B	L*	a*	b*
1	185.90	51.67	55.15	43.62	54.20	31.40
2	248.97	151.28	77.62	72.10	33.29	53.84
3	246.76	228.47	48.46	90.19	-6.36	80.89
4	0.46	181.50	221.27	67.82	-29.06	-31.55
5	30.17	9.95	152.54	17.58	44.34	-72.08
6	154.10	131.69	1.91	55.75	0.67	59.57
7	175.68	241.24	222.75	90.45	-24.19	1.98
8	28.89	90.42	61.69	33.90	-26.30	10.61
9	142.88	156.25	76.71	62.05	-13.98	39.44
10	203.53	202.89	199.18	81.62	-0.26	1.94
11	89.53	13.85	180.72	27.71	55.28	-72.08
12	253.20	41.43	28.96	55.67	75.88	61.53
13	232.78	122.82	217.21	66.87	52.41	-29.36
14	206.53	47.62	63.04	47.24	62.06	32.08
15	13.82	155.29	198.19	59.08	-23.00	-32.26
16	130.32	7.08	252.55	39.71	73.80	-93.46

(a) Stimulus 1 (b) Stimulus 2 (c) Stimulus 3 (d) Stimulus 4 (e) Stimulus 5

(f) Stimulus 6 (g) Stimulus 7 (h) Stimulus 8 (i) Stimulus 9 (j) Stimulus 10

(k) Stimulus 11 (l) Stimulus 12 (m) Stimulus 13 (n) Stimulus 14 (o) Stimulus 15

(p) Stimulus 16 (q) Stimulus 17 (r) Stimulus 18 (s) Stimulus 19 (t) Stimulus 20

(u) Stimulus 21 (v) Stimulus 22 (w) Stimulus 23 (x) Stimulus 24 (y) Stimulus 25

(z) Stimulus 26 (aa) Stimulus 27 (ab) Stimulus 28 (ac) Stimulus 29 (ad) Stimulus 30

(ae) Stimulus 31 (af) Stimulus 32 (ag) Stimulus 33 (ah) Stimulus 34 (ai) Stimulus 35

Figure 5.2: Multicolor stimuli used in Experiment 1 and Experiment 2.

Appendix B. Original Japanese Version and English Translation of Adjective Pair Scales

In Experiments 1-3 and 5 , the adjective pair scales were used in Japanese. In writing this manuscript, we translated them into English. Parts of the translation were based on the book *Ningen Kogaku Gaido* [*A guide to ergonomics*] (pp.153-155) edited by T. Fukuda and R. Fukuda[20]. Table 5.36 shows the original Japanese version and the English translation of the adjective pair scales used in Experiment 1, Experiment 3 and Experiment 5. Table 5.37 shows the original Japanese version and the English translation of the adjective pair scales used in Experiment 2.

Table 5.36: The original Japanese version and the English translation of the adjective pair scales used in Experiment 1, Experiment 3 and Experiment 5.

Original Japanese version	English translation
重い-軽い	heavy-light
明るい-暗い	light-dark
暖かい-涼しい	warm-cool
柔らかい-硬い	soft-hard
うるさい-静かな	noisy-quiet
好きな-嫌いな	like-dislike
派手な-地味な	ornate-plain
強い-弱い	strong-weak
快い-不快な	pleasant-unpleasant
きれいな-汚い	clean-dirty
調和した-不調和な	harmonious-dissonant
陽気な-陰気な	cheerful-gloomy
上品な-下品な	graceful-awkward
澄んだ-濁った	clear-dull
静的な-動的な	static-dynamic
正しい-間違った	true-false
奇抜な-無難な	novel-ordinary
美しい-醜い	beautiful-ugly
安定した-不安定な	stable-changeable
成功した-失敗した	successful-unsuccessful
積極的な-消極的な	positive-negative
緩んだ-緊張した	relaxed-nervous
冷淡な-親切な	cruel-kind
受動的な-能動的な	passive-active

Table 5.37: The original Japanese version and the English translation of the adjective pair scales used in Experiment 2.

Original Japanese version	English translation
つめたい-あたたかい	cold-warm
明るい-暗い	light-dark
騒がしい-静かな	noisy-quiet
軽い-重い	light-heavy
湿った-乾いた	wet-dry
動的な-静的な	dynamic-static
鮮やかな-くすんだ	vivid-subdued
強い-弱い	strong-weak
閑散とした-にぎやかな	leisurely-bustling
目立つ-目立たない	conspicuous-inconspicuous
変化のある-変化のない	diversified-monotonous
濃い-浅い	dark-pale
落ち着きのある-落ち着きのない	calm-restless
浅い-深い	shallow-deep
くどい-あっさりした	gaudy-plain
地味な-派手な	plain-ornate
やわらかい-かたい	soft-hard
はっきりした-ぼんやりした	clear-vague
澄んだ-濁った	clear-dull
鮮明な-不鮮明な	distinct-indistinct
甘い-甘くない	sweet-unsweet
あっさりした-こってりした	plain-thick
おいしい-まずい	delicious-bad-tasting
穏やかな-激しい	calm-violent
すっきりした-ごちゃごちゃした	neat-disordered

Appendix C. RGB and $L^*a^*b^*$ Values of Multicolor Stimuli Used in Experiment 3

The RGB and $L^*a^*b^*$ values of the component colors of each multicolor stimulus employed in Experiment 3 are listed in the following tables (see next page), and the images of the stimuli are displayed after the tables in Figure 5.3.

Table 5.38: RGB and $L^*a^*b^*$ values of component colors of Stimulus 1.

Component color no.	R	G	B	L*	a*	b*
1	196.06	42.65	219.81	49.89	73.43	-57.88
2	252.42	131.18	225.49	71.31	55.68	-26.88
3	149.95	39.46	50.97	34.94	46.71	21.66
4	103.77	190.92	210.52	72.43	-23.40	-18.48
5	201.44	81.22	136.19	52.38	52.55	-5.23
6	22.94	28.48	34.75	10.15	-1.22	-5.23
7	173.06	126.27	48.38	56.60	13.14	47.90
8	126.23	37.64	14.02	29.55	38.05	36.04
9	216.93	142.94	237.05	69.64	40.27	-36.24
10	177.65	148.61	207.93	65.83	19.74	-26.48
11	224.15	252.17	0.13	94.53	-26.80	89.94
12	220.69	156.20	252.44	73.45	37.10	-38.66
13	134.56	122.28	204.34	55.12	19.32	-41.41
14	58.10	127.01	229.72	52.89	5.39	-59.53
15	146.54	215.52	188.35	81.01	-26.96	6.19
16	149.43	62.92	169.94	42.33	48.64	-41.51

Table 5.39: RGB and $L^*a^*b^*$ values of component colors of Stimulus 2.

Component color no.	R	G	B	L*	a*	b*
1	21.29	159.62	168.54	59.55	-32.45	-14.93
2	186.09	227.14	250.49	87.92	-9.97	-16.03
3	196.10	148.27	236.72	68.55	30.80	-37.95
4	147.92	4.33	30.82	31.25	53.67	30.15
5	219.99	123.50	215.44	65.26	47.82	-31.01
6	53.40	140.83	160.62	54.02	-22.13	-18.68
7	8.16	156.75	92.41	56.88	-47.65	23.54
8	12.63	124.84	49.09	45.54	-42.86	32.04
9	31.39	52.40	37.36	19.73	-11.75	6.76
10	48.21	10.88	161.98	20.40	47.58	-73.11
11	71.88	137.34	177.27	54.18	-12.48	-27.88
12	127.27	136.63	113.52	55.67	-6.83	11.23
13	31.60	125.04	217.51	50.98	-0.34	-55.79
14	222.85	68.93	53.16	52.80	59.87	44.75
15	144.07	163.28	106.34	64.52	-14.35	27.27
16	52.52	241.72	20.93	84.13	-72.40	76.45

Table 5.40: RGB and $L^*a^*b^*$ values of component colors of Stimulus 3.

Component color no.	R	G	B	L*	a*	b*
1	26.96	36.22	42.45	13.62	-2.81	-5.57
2	158.34	146.30	13.28	60.22	-4.83	61.16
3	237.46	185.81	188.15	80.23	19.45	6.34
4	16.17	219.41	238.27	79.79	-40.32	-22.47
5	251.02	219.03	200.32	89.76	9.36	13.53
6	130.91	45.29	101.64	33.45	41.88	-13.69
7	34.15	7.88	239.48	28.86	64.30	-104.44
8	76.83	75.36	84.90	32.45	2.25	-5.43
9	119.10	165.29	6.43	62.68	-30.68	62.34
10	214.76	142.55	217.80	68.66	36.97	-27.19
11	88.71	113.74	13.83	44.57	-19.78	45.86
12	45.16	169.02	84.36	61.29	-48.89	33.38
13	229.11	30.13	252.05	56.38	86.05	-65.22
14	137.70	180.26	254.87	72.57	0.12	-42.00
15	73.40	105.70	118.53	42.58	-9.46	-11.15
16	194.81	208.64	25.56	80.58	-19.03	76.02

Table 5.41: RGB and $L^*a^*b^*$ values of component colors of Stimulus 4.

Component color no.	R	G	B	L*	a*	b*
1	45.42	91.71	14.46	34.48	-27.21	35.71
2	133.08	85.64	44.80	41.21	17.24	31.79
3	53.28	230.81	172.22	82.10	-55.60	15.59
4	119.46	232.59	26.52	83.12	-55.81	75.69
5	190.11	187.75	143.27	75.51	-4.46	22.94
6	46.97	152.29	76.48	55.76	-43.87	30.61
7	34.20	54.21	228.21	33.42	45.02	-90.44
8	18.22	61.83	13.71	22.30	-23.79	23.74
9	112.64	3.39	228.78	34.92	68.23	-88.03
10	50.15	23.81	78.38	14.55	22.85	-29.03
11	116.29	25.93	253.82	38.97	69.55	-95.46
12	84.68	75.82	15.82	32.24	-1.62	34.51
13	76.05	11.82	128.88	20.82	44.06	-51.75
14	194.16	160.92	22.92	67.72	3.78	66.77
15	20.62	198.20	230.81	73.26	-33.17	-28.37
16	136.11	27.83	210.58	38.27	64.36	-71.94

Table 5.42: RGB and $L^*a^*b^*$ values of component colors of Stimulus 5.

Component color no.	R	G	B	L*	a*	b*
1	86.21	74.96	190.31	38.29	29.90	-60.46
2	2.64	12.35	170.32	18.95	46.97	-80.65
3	153.88	134.16	186.08	59.24	15.63	-24.54
4	180.35	199.25	73.43	77.08	-19.16	58.27
5	176.60	141.95	101.11	61.57	9.48	26.98
6	15.71	198.94	86.08	70.59	-60.74	43.52
7	155.01	189.02	26.73	72.13	-26.43	67.69
8	32.61	140.13	123.73	52.27	-33.74	-0.18
9	227.07	203.73	187.26	83.59	6.52	11.38
10	13.09	18.59	22.57	5.36	-1.33	-3.25
11	203.58	240.47	174.35	91.09	-20.96	27.94
12	33.68	184.29	28.14	65.68	-59.15	59.59
13	29.96	163.38	83.85	59.17	-49.21	30.97
14	166.72	191.03	148.71	74.66	-14.35	18.55
15	188.71	59.88	187.41	48.93	62.64	-40.87
16	247.50	221.07	21.99	88.31	-2.96	84.65

Table 5.43: RGB and $L^*a^*b^*$ values of component colors of Stimulus 6.

Component color no.	R	G	B	L*	a*	b*
1	93.44	94.15	174.68	42.91	16.50	-43.88
2	152.48	201.29	93.75	75.89	-29.67	47.51
3	52.54	22.10	196.84	25.97	52.53	-84.66
4	52.45	99.01	140.70	40.13	-6.08	-28.32
5	58.38	163.69	123.54	60.49	-38.89	11.64
6	38.72	199.39	25.65	70.62	-62.74	64.40
7	74.99	60.53	135.37	30.25	21.64	-40.55
8	23.33	103.36	26.74	37.94	-36.61	34.08
9	28.63	200.03	74.35	70.94	-61.25	49.11
10	153.90	245.93	110.28	88.99	-46.13	55.12
11	177.16	193.32	110.32	75.53	-15.28	39.65
12	167.15	27.99	238.11	45.13	73.72	-76.21
13	47.80	67.88	203.45	34.84	31.07	-73.86
14	124.34	196.08	100.98	72.83	-36.78	40.02
15	69.60	9.49	171.69	23.81	52.03	-73.23
16	109.54	115.19	155.51	49.22	5.11	-22.70

Table 5.44: RGB and $L^*a^*b^*$ values of component colors of Stimulus 7.

Component color no.	R	G	B	L*	a*	b*
1	15.15	80.53	197.04	36.60	18.21	-67.32
2	177.59	31.96	33.19	39.55	56.99	39.35
3	23.55	1.99	107.89	10.17	36.19	-56.13
4	167.17	184.35	135.46	72.63	-12.50	22.84
5	27.75	161.10	32.26	57.97	-53.22	51.54
6	34.25	25.14	36.22	10.34	6.62	-5.86
7	42.90	50.04	80.96	21.30	4.09	-20.15
8	80.69	55.48	64.02	26.48	12.68	-0.60
9	227.70	179.32	141.71	76.85	14.48	26.30
10	47.03	54.07	19.72	21.30	-7.77	20.14
11	233.02	180.21	142.24	77.63	16.19	27.16
12	79.92	42.38	158.74	28.28	38.76	-58.00
13	251.92	43.46	65.74	55.82	75.78	42.97
14	101.18	18.87	174.44	29.28	54.69	-65.68
15	102.61	250.62	102.56	88.32	-62.33	57.02
16	158.27	39.36	97.24	37.56	52.12	-4.12

Table 5.45: RGB and $L^*a^*b^*$ values of component colors of Stimulus 8.

Component color no.	R	G	B	L*	a*	b*
1	41.09	193.32	222.13	71.76	-31.91	-25.97
2	89.45	174.81	75.01	64.43	-41.69	42.00
3	135.31	212.27	152.36	78.85	-34.17	21.92
4	85.50	76.30	115.41	34.63	10.96	-21.20
5	107.77	91.70	142.37	42.14	15.27	-25.83
6	189.35	108.21	109.49	55.17	33.12	14.15
7	31.84	6.23	74.00	7.85	26.74	-37.05
8	80.97	166.69	244.02	65.67	-8.22	-47.12
9	238.61	116.76	61.32	63.44	45.42	52.19
10	194.79	193.63	188.87	78.29	-0.21	2.55
11	189.64	27.01	173.80	45.29	70.11	-38.59
12	118.13	54.10	25.12	31.24	27.20	31.16
13	210.01	44.63	41.71	47.41	63.34	44.72
14	169.83	228.07	131.72	84.98	-31.75	40.68
15	179.19	39.17	243.13	48.21	74.15	-73.93
16	137.93	173.33	9.32	66.33	-26.64	65.38

Table 5.46: RGB and $L^*a^*b^*$ values of component colors of Stimulus 9.

Component color no.	R	G	B	L*	a*	b*
1	206.35	190.90	30.65	76.78	-5.72	72.62
2	133.89	83.09	139.34	42.89	28.88	-22.56
3	101.71	105.85	46.09	43.37	-8.45	32.36
4	65.12	5.24	235.54	30.29	65.25	-99.76
5	166.69	237.82	41.70	86.94	-42.05	77.04
6	234.88	202.64	147.24	83.38	6.17	32.11
7	112.21	65.69	191.75	39.03	41.48	-59.96
8	58.31	16.37	195.67	25.76	54.40	-84.30
9	171.16	182.38	163.73	72.90	-6.63	8.12
10	106.86	99.64	208.12	47.18	25.12	-56.31
11	80.94	207.71	201.21	76.35	-37.42	-7.77
12	217.33	128.94	162.09	64.49	37.97	-2.26
13	242.48	113.21	15.30	63.14	47.71	68.76
14	221.02	160.95	90.54	71.04	17.58	45.32
15	254.24	57.16	166.38	59.32	78.08	-11.53
16	154.27	98.75	36.26	47.27	19.19	43.33

Table 5.47: RGB and $L^*a^*b^*$ values of component colors of Stimulus 10.

Component color no.	R	G	B	L*	a*	b*
1	6.41	107.38	46.95	39.30	-38.09	25.82
2	185.07	94.44	214.60	55.08	51.00	-46.80
3	187.23	145.61	45.10	63.12	9.10	56.07
4	244.13	67.66	235.77	61.07	79.12	-48.54
5	57.06	95.26	22.31	36.39	-24.03	34.87
6	163.23	46.06	11.49	38.13	47.76	46.97
7	184.41	88.60	168.46	52.06	47.14	-25.02
8	97.89	159.97	5.52	59.81	-36.46	59.65
9	232.20	204.14	190.19	84.21	8.56	10.78
10	207.34	97.74	157.41	56.78	49.11	-11.08
11	146.75	135.16	70.14	56.34	-2.12	36.21
12	63.40	115.17	58.07	43.81	-27.46	25.80
13	205.13	251.46	7.65	93.04	-33.84	87.66
14	136.59	22.20	204.53	37.47	64.71	-69.75
15	252.23	17.07	239.55	59.31	90.47	-53.31
16	4.64	174.38	199.85	64.98	-31.52	-24.16

Table 5.48: RGB and $L^*a^*b^*$ values of component colors of Stimulus 11.

Component color no.	R	G	B	L*	a*	b*
1	136.21	225.77	229.25	84.30	-27.37	-10.63
2	159.61	35.16	55.54	36.42	51.39	21.10
3	46.45	10.66	27.27	8.29	20.03	-1.89
4	157.19	239.61	90.39	87.19	-43.98	61.51
5	104.71	251.01	241.12	90.61	-42.65	-7.71
6	172.54	252.02	195.54	92.74	-33.74	19.09
7	85.86	168.91	62.26	62.31	-41.55	45.31
8	75.35	173.45	134.60	64.30	-37.24	10.96
9	104.96	153.67	191.38	61.11	-9.21	-25.07
10	148.80	140.71	148.81	59.43	4.28	-3.19
11	130.51	21.06	183.49	34.87	60.65	-61.70
12	254.02	90.41	247.67	65.85	75.48	-47.50
13	88.34	226.07	115.95	80.63	-56.14	41.93
14	105.42	55.52	32.04	29.53	21.09	24.66
15	78.77	185.16	199.63	69.51	-28.51	-17.05
16	176.92	2.50	215.02	44.15	75.34	-64.57

Table 5.49: RGB and $L^*a^*b^*$ values of component colors of Stimulus 12.

Component color no.	R	G	B	L*	a*	b*
1	235.19	196.59	10.88	80.95	3.58	80.16
2	96.44	179.61	186.03	68.14	-25.12	-11.58
3	57.19	68.61	171.62	33.18	22.41	-57.86
4	121.76	159.05	60.29	61.18	-25.02	45.64
5	45.17	211.56	195.57	76.61	-45.10	-4.47
6	238.29	27.51	46.47	51.90	74.63	48.74
7	25.27	124.89	49.28	45.75	-41.38	32.24
8	228.45	25.27	11.26	49.58	72.00	61.37
9	142.11	196.99	79.54	73.89	-32.64	51.41
10	45.64	86.43	53.59	33.05	-21.12	14.51
11	130.09	231.12	160.38	84.21	-42.33	24.96
12	25.89	99.67	13.93	36.61	-35.64	37.48
13	127.83	110.09	254.38	54.04	35.50	-71.09
14	206.96	123.84	228.08	64.15	45.33	-39.90
15	35.07	99.45	236.48	44.96	21.09	-76.15
16	233.96	181.96	157.68	78.39	17.01	20.02

Table 5.50: RGB and $L^*a^*b^*$ values of component colors of Stimulus 13.

Component color no.	R	G	B	L*	a*	b*
1	87.54	238.69	31.82	83.93	-65.56	74.96
2	186.30	164.85	212.45	70.80	14.90	-21.35
3	101.56	191.20	212.98	72.49	-23.51	-19.72
4	82.23	140.83	249.68	58.70	7.39	-61.22
5	140.07	84.26	157.97	44.67	33.34	-30.85
6	91.96	192.91	105.54	70.48	-44.73	34.51
7	125.55	177.16	248.05	70.79	-2.96	-41.15
8	83.58	213.64	188.46	77.90	-41.85	1.28
9	243.31	8.14	91.00	52.74	79.05	23.87
10	168.98	71.78	58.75	43.81	40.14	28.56
11	181.34	159.27	150.61	67.38	7.32	7.66
12	168.41	12.13	88.94	37.17	60.43	0.93
13	115.09	61.43	182.34	37.95	42.39	-56.16
14	218.33	71.78	186.42	55.34	66.08	-29.87
15	35.13	213.36	35.34	75.05	-66.54	66.16
16	149.99	93.37	205.72	50.01	39.95	-50.12

Table 5.51: RGB and $L^*a^*b^*$ values of component colors of Stimulus 14.

Component color no.	R	G	B	L*	a*	b*
1	128.46	124.85	223.65	56.09	20.32	-50.76
2	90.05	114.61	245.70	51.89	22.61	-69.93
3	10.79	248.10	48.25	85.81	-76.29	72.58
4	170.12	149.54	172.15	64.31	10.99	-8.92
5	92.06	158.17	206.84	62.28	-11.16	-31.95
6	4.91	21.39	248.57	30.16	62.93	-107.42
7	166.09	58.97	102.89	41.92	47.65	-1.12
8	31.12	68.45	65.75	26.25	-14.54	-2.71
9	84.57	38.82	88.74	23.82	27.77	-20.91
10	31.02	225.46	24.04	78.80	-70.34	71.51
11	237.16	101.75	12.09	60.22	50.88	67.22
12	87.31	187.67	202.64	70.63	-27.27	-16.98
13	138.95	174.99	227.88	70.29	-2.66	-30.93
14	13.97	77.43	11.78	28.12	-30.59	30.24
15	49.85	183.64	184.05	67.89	-36.01	-11.05
16	223.84	148.52	18.02	68.01	22.92	69.77

Table 5.52: RGB and $L^*a^*b^*$ values of component colors of Stimulus 15.

Component color no.	R	G	B	L*	a*	b*
1	235.30	204.09	72.92	83.06	1.23	66.28
2	138.63	251.12	182.50	90.69	-44.88	22.57
3	213.94	110.48	120.01	59.32	42.11	14.39
4	142.98	68.62	191.00	43.49	47.70	-52.10
5	128.49	164.94	78.48	63.56	-23.77	40.00
6	35.37	121.27	92.43	45.37	-32.05	8.16
7	200.97	198.98	170.47	79.83	-2.85	14.61
8	34.04	5.50	142.76	16.18	43.66	-68.34
9	76.71	239.55	250.13	86.82	-40.63	-18.06
10	73.09	204.21	228.51	75.82	-31.21	-23.17
11	152.37	225.42	240.65	85.27	-20.77	-15.07
12	140.04	185.74	147.07	71.35	-21.63	14.55
13	6.59	113.87	164.81	44.73	-12.70	-35.50
14	132.91	94.94	238.97	50.25	41.36	-68.62
15	211.53	216.52	95.00	84.36	-13.46	57.79
16	151.26	222.50	238.04	84.34	-20.22	-15.09

Table 5.53: RGB and $L^*a^*b^*$ values of component colors of Stimulus 16.

Component color no.	R	G	B	L*	a*	b*
1	170.46	52.73	166.73	44.03	57.95	-36.67
2	18.37	103.72	170.07	41.91	-4.74	-43.02
3	238.10	206.79	123.56	84.42	3.66	45.36
4	192.97	106.35	247.81	59.84	52.86	-57.73
5	251.93	220.36	99.17	88.87	1.33	62.07
6	115.96	62.91	200.03	39.47	45.50	-64.02
7	225.12	233.00	142.36	90.25	-12.57	43.05
8	152.71	37.96	229.43	43.08	67.67	-74.74
9	114.85	52.45	229.41	39.70	56.13	-80.51
10	194.46	225.03	72.66	85.23	-25.21	67.13
11	171.67	169.39	31.32	67.88	-10.15	63.83
12	103.87	70.20	182.75	38.46	35.04	-55.66
13	72.26	228.53	210.78	82.54	-45.17	-3.82
14	99.46	126.97	177.18	52.60	0.14	-30.12
15	212.76	155.46	146.56	69.59	21.18	13.56
16	83.14	116.39	182.02	48.73	2.74	-39.01

Table 5.54: RGB and $L^*a^*b^*$ values of component colors of Stimulus 17.

Component color no.	R	G	B	L*	a*	b*
1	225.52	183.82	4.75	76.76	5.71	77.40
2	172.07	111.82	111.64	53.75	24.56	10.51
3	29.84	207.74	82.84	73.48	-62.58	48.30
4	62.79	87.39	95.80	35.40	-7.89	-8.05
5	139.37	143.29	100.93	58.36	-6.84	22.11
6	101.52	131.42	167.67	53.68	-3.79	-22.99
7	242.48	184.20	102.02	79.10	15.32	49.67
8	212.13	34.26	15.42	46.79	66.06	56.53
9	21.48	41.79	82.68	17.17	4.02	-27.94
10	76.94	2.98	137.68	21.00	47.93	-56.98
11	24.32	37.36	160.94	22.70	33.75	-68.71
12	219.13	248.43	145.56	93.91	-21.59	45.92
13	254.20	141.15	131.44	71.29	43.00	25.18
14	84.32	109.65	125.41	44.72	-6.55	-11.97
15	18.11	226.37	16.48	78.98	-71.55	72.63
16	111.23	210.79	100.61	76.64	-46.59	44.56

Table 5.55: RGB and $L^*a^*b^*$ values of component colors of Stimulus 18.

Component color no.	R	G	B	L*	a*	b*
1	156.44	208.75	225.99	80.54	-14.52	-14.34
2	237.43	48.65	65.94	53.41	70.69	39.27
3	228.96	151.31	128.48	70.20	28.12	24.65
4	156.27	208.95	135.63	78.81	-27.83	30.71
5	51.53	115.74	109.12	44.43	-22.79	-2.93
6	246.34	158.11	177.32	74.70	35.52	4.49
7	183.64	88.46	131.83	50.79	42.85	-5.37
8	141.96	39.91	143.32	36.37	52.14	-35.01
9	177.17	108.75	213.25	56.84	41.42	-43.36
10	186.50	91.81	115.82	51.48	40.78	5.22
11	98.53	197.77	187.24	73.64	-32.41	-4.15
12	109.72	176.91	241.03	69.66	-8.30	-39.19
13	199.98	179.92	27.88	73.29	-3.13	70.22
14	99.43	150.68	117.14	58.00	-23.46	12.07
15	12.84	58.32	212.72	32.20	37.00	-83.62
16	3.99	220.25	19.91	76.98	-70.53	70.54

Table 5.56: RGB and $L^*a^*b^*$ values of component colors of Stimulus 19.

Component color no.	R	G	B	L*	a*	b*
1	170.61	127.55	55.59	56.69	11.66	44.60
2	145.76	31.16	171.15	37.55	59.43	-49.90
3	152.89	14.27	14.37	32.76	53.31	41.25
4	38.89	5.00	110.97	12.74	36.99	-53.81
5	212.22	157.43	132.63	69.76	18.68	21.40
6	220.29	24.91	231.55	53.53	82.81	-58.43
7	27.54	131.83	36.50	48.08	-44.34	40.47
8	142.64	1.17	195.50	36.80	67.29	-65.53
9	216.42	233.79	251.68	91.73	-3.31	-10.75
10	128.81	69.21	25.69	36.41	23.69	36.52
11	129.50	149.33	194.54	61.54	0.98	-26.01
12	21.16	168.71	131.83	61.56	-44.17	8.46
13	43.62	239.33	150.57	84.22	-62.74	29.41
14	112.36	240.19	167.26	86.28	-49.39	24.12
15	115.25	214.12	135.82	78.22	-42.93	29.37
16	141.24	173.42	93.63	67.08	-21.46	36.97

Table 5.57: RGB and $L^*a^*b^*$ values of component colors of Stimulus 20.

Component color no.	R	G	B	L*	a*	b*
1	61.02	147.63	221.06	58.54	-7.97	-45.75
2	103.73	28.72	113.18	26.55	42.07	-32.23
3	76.55	102.35	212.51	45.96	17.75	-60.91
4	102.93	99.49	91.91	42.31	0.31	4.83
5	35.77	66.33	22.14	24.90	-19.08	22.57
6	109.50	65.61	75.88	33.42	20.63	2.28
7	108.34	30.40	126.24	28.33	44.57	-37.62
8	180.13	62.11	200.19	48.44	61.33	-49.14
9	18.89	100.44	0.87	36.68	-37.68	40.99
10	56.27	0.33	48.24	9.61	30.80	-15.48
11	36.33	68.36	44.60	25.99	-17.02	10.59
12	35.36	152.72	229.77	59.81	-11.10	-48.63
13	239.54	56.40	123.08	55.46	71.65	8.03
14	95.88	133.56	67.54	51.74	-22.93	30.90
15	17.43	111.26	44.33	40.79	-38.52	28.96
16	6.66	243.44	109.80	84.74	-70.98	49.34

Table 5.58: RGB and $L^*a^*b^*$ values of component colors of Stimulus 21.

Component color no.	R	G	B	L*	a*	b*
1	245.20	194.42	1.87	81.41	8.84	81.88
2	173.41	180.02	164.51	72.45	-4.52	7.10
3	140.84	55.62	196.95	41.36	53.93	-58.99
4	58.15	94.57	227.19	43.85	23.63	-72.69
5	218.38	102.62	81.09	57.85	45.19	34.79
6	155.20	232.10	231.82	87.05	-24.61	-7.82
7	150.86	84.81	217.53	48.98	46.73	-58.47
8	112.81	230.61	8.46	82.24	-57.27	76.96
9	135.77	182.71	45.72	69.07	-30.71	59.42
10	85.82	47.87	82.09	25.81	22.06	-13.45
11	102.98	139.88	12.43	53.82	-25.90	54.10
12	140.95	70.08	61.58	38.92	30.10	19.63
13	62.00	39.31	243.89	34.03	58.07	-98.26
14	238.59	208.77	185.71	85.85	8.19	15.52
15	44.83	91.89	48.14	34.83	-24.62	19.98
16	0.31	80.69	178.40	35.13	10.92	-58.81

Table 5.59: RGB and $L^*a^*b^*$ values of component colors of Stimulus 22.

Component color no.	R	G	B	L*	a*	b*
1	159.44	138.48	111.95	59.07	5.12	17.32
2	73.29	127.92	194.19	52.09	-2.71	-40.75
3	194.41	146.89	190.65	66.59	23.88	-15.55
4	164.61	31.42	128.62	38.99	58.85	-21.51
5	88.55	23.50	37.70	19.84	31.03	8.19
6	50.53	171.43	110.04	62.53	-45.38	21.69
7	177.07	65.48	2.49	43.71	44.72	54.31
8	135.73	71.24	241.29	46.11	54.29	-76.62
9	231.14	100.13	6.34	58.95	49.35	66.95
10	171.22	213.48	247.73	83.29	-8.42	-21.63
11	14.52	114.83	148.53	44.47	-17.75	-26.36
12	175.09	183.46	165.76	73.54	-5.44	7.97
13	185.36	95.33	148.30	52.85	41.84	-12.02
14	29.61	14.70	249.84	30.40	65.18	-107.71
15	72.63	151.72	245.35	61.16	-1.41	-54.97
16	47.37	49.23	87.12	21.73	7.78	-23.51

Table 5.60: RGB and $L^*a^*b^*$ values of component colors of Stimulus 23.

Component color no.	R	G	B	L*	a*	b*
1	237.89	99.62	69.67	60.28	53.17	44.70
2	38.75	101.26	95.55	38.78	-21.99	-3.09
3	33.43	110.94	23.34	40.86	-37.34	38.40
4	156.73	2.80	146.18	36.52	63.57	-36.31
5	201.38	60.02	114.25	48.66	59.13	2.56
6	145.19	15.66	126.55	33.89	57.34	-28.49
7	163.79	56.42	213.45	45.86	61.50	-61.05
8	247.62	215.83	129.03	87.62	3.45	46.81
9	71.11	190.39	60.42	68.58	-52.98	52.52
10	244.12	158.17	153.07	73.97	32.48	16.69
11	44.01	23.04	65.09	12.59	18.31	-23.03
12	218.94	232.32	178.41	90.09	-11.24	24.76
13	184.92	58.62	146.89	46.51	57.56	-20.75
14	206.71	102.98	252.05	61.19	58.78	-57.80
15	22.95	81.84	130.41	33.08	-4.75	-33.21
16	15.45	185.05	141.92	66.94	-48.39	10.46

Table 5.61: RGB and $L^*a^*b^*$ values of component colors of Stimulus 24.

Component color no.	R	G	B	L*	a*	b*
1	134.99	211.65	218.98	80.01	-23.27	-11.55
2	201.20	81.05	115.31	51.81	50.87	6.38
3	191.82	28.01	27.98	42.16	61.84	45.75
4	68.82	133.78	248.03	56.12	8.63	-64.47
5	181.15	79.52	74.32	47.50	41.83	24.42
6	216.84	232.47	163.02	89.74	-13.64	32.05
7	65.12	22.61	213.76	29.01	57.19	-89.40
8	149.10	241.77	15.56	87.12	-50.06	80.91
9	149.08	72.70	211.07	46.15	50.89	-59.34
10	48.70	112.85	100.32	43.15	-24.00	0.42
11	210.78	172.60	52.94	72.55	5.77	62.83
12	81.12	34.12	171.22	28.02	45.65	-65.95
13	145.60	43.29	37.65	34.45	43.30	29.13
14	121.40	231.57	140.80	83.70	-47.32	34.03
15	8.40	13.74	205.29	23.92	55.04	-93.08
16	115.10	97.57	201.36	47.04	26.88	-52.61

Table 5.62: RGB and $L^*a^*b^*$ values of component colors of Stimulus 25.

Component color no.	R	G	B	L*	a*	b*
1	92.89	135.75	181.47	54.89	-5.43	-29.07
2	222.23	83.82	165.78	57.00	60.62	-15.33
3	248.58	19.37	149.69	55.40	82.25	-7.66
4	105.54	78.83	67.28	36.18	10.29	11.43
5	193.49	253.78	47.58	93.02	-38.59	81.56
6	199.19	49.93	253.05	52.92	76.60	-71.66
7	204.58	108.18	185.86	59.23	46.47	-23.80
8	127.08	206.29	90.91	75.91	-40.84	48.38
9	18.68	150.70	232.10	59.09	-10.47	-51.06
10	49.41	110.25	191.04	45.75	1.85	-49.04
11	9.99	241.31	194.74	85.31	-57.96	8.38
12	142.50	46.88	126.98	36.70	47.65	-24.42
13	132.05	253.53	217.99	91.75	-42.08	5.96
14	245.41	173.13	102.89	76.81	22.09	46.72
15	238.42	122.27	59.11	64.47	42.68	54.22
16	101.05	179.79	142.43	67.42	-31.99	11.22

(a) Stimulus 1 (b) Stimulus 2 (c) Stimulus 3 (d) Stimulus 4 (e) Stimulus 5

(f) Stimulus 6 (g) Stimulus 7 (h) Stimulus 8 (i) Stimulus 9 (j) Stimulus 10

(k) Stimulus 11 (l) Stimulus 12 (m) Stimulus 13 (n) Stimulus 14 (o) Stimulus 15

(p) Stimulus 16 (q) Stimulus 17 (r) Stimulus 18 (s) Stimulus 19 (t) Stimulus 20

(u) Stimulus 21 (v) Stimulus 22 (w) Stimulus 23 (x) Stimulus 24 (y) Stimulus 25

Figure 5.3: Multicolor stimuli used in Experiment 3.

Appendix D. Main Variable Values of the Test Colors used in Experiment 4 and Experiment 5

Table 5.63 summarizes the values of the variables that were used in data analysis, namely, FS, MAS, EDS, PS, GS, and DS, of the 30 test colors used in Experiment 4 and Experiment 5.

Table 5.63: Main variable values of the test colors used in Experiment 4 and Experiment 5.

Test color no.	Munsell index	FS	MAS	EDS	PS	GS	DS
1	5RP 7/10	1.00	0.72	16.64	4.52	4.00	18.46
2	5RP 6/12	0.95	0.65	16.07	4.34	3.69	19.23
3	5RP 5/12	0.23	0.42	18.03	4.45	4.17	18.82
4	5RP 4/12	0.36	0.56	21.26	4.48	4.41	18.80
5	5RP 3/10	0.50	0.54	21.15	4.76	4.90	18.90
6	5RP 2/8	0.59	0.43	16.97	4.62	5.31	16.07
7	10RP 7/8	0.77	0.28	15.18	4.93	4.38	18.60
8	10RP 6/12	0.86	0.50	18.52	4.72	4.21	19.44
9	10RP 5/14	0.64	0.79	14.04	4.24	3.76	20.35
10	10RP 4/14	0.09	0.88	24.40	4.55	4.72	21.64
11	10RP 3/10	0.18	0.56	17.61	4.59	5.28	19.81
12	10RP 2/8	0.32	0.25	16.40	4.34	5.10	16.80
13	5R 7/10	0.45	0.54	20.84	4.48	4.14	22.79
14	5R 6/12	0.14	0.58	19.97	4.97	4.48	25.52
15	5R 5/14	0.18	0.67	30.02	4.97	4.34	25.18
16	5R 4/14	0.95	0.87	18.08	4.90	4.55	25.08
17	5R 3/10	0.32	0.52	18.86	4.66	4.69	22.19
18	5R 2/8	0.09	0.64	20.88	4.72	5.10	18.28
19	10R 7/10	0.23	0.91	22.74	4.86	4.62	29.59
20	10R 6/14	0.95	0.73	25.45	4.69	3.93	29.82
21	10R 5/16	0.86	0.96	30.20	4.59	4.14	34.89
22	10R 4/12	0.23	0.63	28.72	4.10	4.00	25.50
23	10R 3/10	0.50	0.50	18.57	3.93	4.21	24.27
24	10R 2/6	0.68	0.67	14.99	3.72	4.17	19.08
25	5YR 7/14	0.86	0.96	27.96	4.52	4.17	32.54
26	5YR 6/14	0.45	0.39	15.00	4.93	4.21	24.23
27	5YR 5/12	0.45	0.40	16.17	3.38	3.62	22.81
28	5YR 4/8	1.00	0.65	27.23	3.14	3.17	23.91
29	5YR 3/6	1.00	0.62	15.64	3.66	3.41	20.99
30	5YR 2/4	0.91	0.83	12.21	3.97	3.83	19.50

Appendix E. Color Impression Scores of the Test Colors used in Experiment 4 and Experiment 5

The following tables (see next page) summarize the color impression scores, except gracefulness score, of the 30 test colors used in Experiment 4 and Experiment 5.

Table 5.64: Heaviness scores, lightness scores, and warmness scores of the test colors.

Test color no.	Munsell index	Heaviness score	Lightness score	Warmness score
1	5RP 7/10	2.48	5.97	4.66
2	5RP 6/12	3.17	5.48	4.90
3	5RP 5/12	4.03	4.69	4.62
4	5RP 4/12	4.62	4.31	4.48
5	5RP 3/10	4.76	3.62	4.45
6	5RP 2/8	5.14	2.90	3.76
7	10RP 7/8	2.76	5.52	4.86
8	10RP 6/12	3.14	5.69	4.97
9	10RP 5/14	3.48	5.72	4.66
10	10RP 4/14	4.10	4.62	4.72
11	10RP 3/10	5.03	4.00	4.45
12	10RP 2/8	5.07	3.03	4.41
13	5R 7/10	2.90	5.17	4.69
14	5R 6/12	2.90	5.45	4.93
15	5R 5/14	3.62	5.66	5.07
16	5R 4/14	4.34	5.62	5.59
17	5R 3/10	5.03	4.00	5.03
18	5R 2/8	5.31	3.21	4.59
19	10R 7/10	2.66	5.69	4.90
20	10R 6/14	3.79	5.86	5.45
21	10R 5/16	4.31	5.59	5.38
22	10R 4/12	4.72	4.00	5.14
23	10R 3/10	4.72	3.97	4.79
24	10R 2/6	5.62	2.86	4.62
25	5YR 7/14	2.76	5.90	5.00
26	5YR 6/14	3.66	5.24	5.34
27	5YR 5/12	4.21	3.59	4.86
28	5YR 4/8	4.79	3.62	4.59
29	5YR 3/6	5.14	3.14	4.66
30	5YR 2/4	5.66	2.86	4.28

Table 5.65: Hardness scores, noisiness scores and ornateness scores of the test colors.

Test color no.	Munsell index	Hardness score	Noisiness score	Ornateness score
1	5RP 7/10	2.72	5.03	5.21
2	5RP 6/12	3.17	4.72	5.00
3	5RP 5/12	3.48	4.55	4.76
4	5RP 4/12	3.90	4.45	4.62
5	5RP 3/10	4.00	3.55	4.03
6	5RP 2/8	4.41	2.86	3.17
7	10RP 7/8	2.76	4.17	4.66
8	10RP 6/12	2.83	5.21	5.41
9	10RP 5/14	3.62	5.41	5.79
10	10RP 4/14	3.76	4.21	5.31
11	10RP 3/10	4.17	3.38	4.00
12	10RP 2/8	4.34	3.10	3.55
13	5R 7/10	3.17	4.38	4.76
14	5R 6/12	3.14	4.72	4.97
15	5R 5/14	3.24	4.83	5.45
16	5R 4/14	3.93	5.21	5.66
17	5R 3/10	4.24	4.34	4.79
18	5R 2/8	4.38	3.34	3.48
19	10R 7/10	2.66	4.55	4.90
20	10R 6/14	3.52	5.59	5.66
21	10R 5/16	4.17	5.38	5.48
22	10R 4/12	3.83	3.79	3.79
23	10R 3/10	4.34	3.69	3.72
24	10R 2/6	4.66	3.00	2.55
25	5YR 7/14	3.07	5.41	5.41
26	5YR 6/14	3.14	4.62	4.90
27	5YR 5/12	3.72	3.79	3.34
28	5YR 4/8	3.41	3.28	3.00
29	5YR 3/6	3.69	2.79	2.55
30	5YR 2/4	4.90	2.62	2.41

Table 5.66: Strength scores, pleasantness scores and clearness scores of the test colors.

Test color no.	Munsell index	Strength score	Pleasantness score	Cleaness score
1	5RP 7/10	4.00	4.59	5.52
2	5RP 6/12	4.28	4.52	4.62
3	5RP 5/12	4.34	4.21	4.55
4	5RP 4/12	4.76	4.28	4.62
5	5RP 3/10	4.86	4.38	4.59
6	5RP 2/8	5.00	4.31	4.45
7	10RP 7/8	3.66	4.90	5.21
8	10RP 6/12	4.48	4.79	5.21
9	10RP 5/14	5.28	4.31	4.52
10	10RP 4/14	4.69	4.14	4.69
11	10RP 3/10	5.07	4.45	4.66
12	10RP 2/8	4.93	4.00	4.38
13	5R 7/10	3.76	4.66	4.97
14	5R 6/12	4.07	4.93	4.97
15	5R 5/14	5.28	4.83	5.28
16	5R 4/14	5.72	4.93	5.07
17	5R 3/10	5.28	4.24	4.45
18	5R 2/8	4.72	4.59	4.34
19	10R 7/10	3.66	5.07	5.07
20	10R 6/14	5.48	4.72	5.03
21	10R 5/16	5.66	4.52	4.69
22	10R 4/12	4.34	4.34	3.59
23	10R 3/10	4.45	3.79	3.52
24	10R 2/6	4.69	3.62	3.31
25	5YR 7/14	4.69	4.93	5.07
26	5YR 6/14	4.38	4.83	4.72
27	5YR 5/12	3.90	3.24	3.28
28	5YR 4/8	3.52	3.34	2.86
29	5YR 3/6	3.90	3.55	3.17
30	5YR 2/4		3.72	3.31

Table 5.67: Cheerfulness scores, clearness scores and dynamicness scores of the test colors.

Test color no.	Munsell index	Cheerfulness score	Clearness score	Dynamicness score
1	5RP 7/10	5.59	5.24	4.10
2	5RP 6/12	4.93	4.45	4.59
3	5RP 5/12	4.76	4.00	4.24
4	5RP 4/12	4.38	3.86	4.45
5	5RP 3/10	3.55	3.66	3.52
6	5RP 2/8	3.45	3.79	3.00
7	10RP 7/8	5.03	4.83	3.83
8	10RP 6/12	5.41	4.83	4.83
9	10RP 5/14	5.41	4.66	5.45
10	10RP 4/14	4.69	3.62	4.59
11	10RP 3/10	3.83	4.31	3.69
12	10RP 2/8	3.45	3.45	3.03
13	5R 7/10	5.03	4.66	4.24
14	5R 6/12	5.31	4.28	4.83
15	5R 5/14	5.59	4.83	5.17
16	5R 4/14	5.34	4.59	5.21
17	5R 3/10	4.07	3.52	4.66
18	5R 2/8	3.55	3.14	3.28
19	10R 7/10	5.38	5.14	4.41
20	10R 6/14	5.97	5.07	5.34
21	10R 5/16	5.59	4.52	5.59
22	10R 4/12	4.00	3.48	3.83
23	10R 3/10	3.59	3.21	3.72
24	10R 2/6	3.00	3.07	3.21
25	5YR 7/14	5.52	5.24	4.97
26	5YR 6/14	5.31	4.38	4.45
27	5YR 5/12	3.41	3.00	3.41
28	5YR 4/8	3.62	3.00	3.55
29	5YR 3/6	2.86	2.86	2.93
30	5YR 2/4	2.93	3.10	2.59

Table 5.68: Trueness scores, novelty scores and beauty scores of the test colors.

Test color no.	Munsell index	Trueness score	Novelty score	Beauty score
1	5RP 7/10	4.28	4.66	5.03
2	5RP 6/12	4.21	4.72	4.66
3	5RP 5/12	4.03	4.76	4.52
4	5RP 4/12	3.93	4.72	4.52
5	5RP 3/10	4.31	3.97	4.86
6	5RP 2/8	4.07	3.97	4.93
7	10RP 7/8	4.59	3.83	4.97
8	10RP 6/12	4.48	4.90	4.83
9	10RP 5/14	3.90	5.66	4.52
10	10RP 4/14	4.21	5.14	4.93
11	10RP 3/10	4.62	3.83	4.93
12	10RP 2/8	4.24	3.62	4.83
13	5R 7/10	4.21	4.45	4.83
14	5R 6/12	4.38	4.69	4.97
15	5R 5/14	4.72	4.86	5.14
16	5R 4/14	4.93	5.31	5.28
17	5R 3/10	4.66	4.17	4.62
18	5R 2/8	4.62	3.07	4.45
19	10R 7/10	4.59	4.28	5.03
20	10R 6/14	4.86	5.14	5.00
21	10R 5/16	4.48	5.03	4.76
22	10R 4/12	4.24	3.48	3.72
23	10R 3/10	3.79	3.55	3.52
24	10R 2/6	4.14	2.86	3.72
25	5YR 7/14	4.69	4.97	5.00
26	5YR 6/14	4.72	4.34	4.79
27	5YR 5/12	3.69	3.72	3.14
28	5YR 4/8	3.66	2.93	2.97
29	5YR 3/6	3.86	2.79	3.28
30	5YR 2/4	4.17	2.62	3.45

Table 5.69: Stability scores, successfulness scores and positivity scores of the test colors.

Test color no.	Munsell index	Stability score	Successfulness score	Positivity score
1	5RP 7/10	3.86	4.48	4.76
2	5RP 6/12	4.14	4.31	5.00
3	5RP 5/12	4.00	4.34	4.52
4	5RP 4/12	4.00	4.28	4.14
5	5RP 3/10	4.59	4.41	3.93
6	5RP 2/8	4.59	4.17	4.00
7	10RP 7/8	4.28	4.69	4.28
8	10RP 6/12	4.45	4.66	5.03
9	10RP 5/14	3.52	4.52	5.45
10	10RP 4/14	4.24	4.03	4.72
11	10RP 3/10	4.41	4.66	4.41
12	10RP 2/8	4.52	4.21	3.66
13	5R 7/10	4.07	4.59	4.41
14	5R 6/12	4.45	4.45	5.03
15	5R 5/14	4.45	4.97	5.31
16	5R 4/14	4.31	4.97	5.72
17	5R 3/10	4.45	4.83	4.83
18	5R 2/8	4.83	4.45	3.72
19	10R 7/10	4.34	4.86	4.41
20	10R 6/14	4.66	5.07	5.55
21	10R 5/16	4.24	4.69	5.52
22	10R 4/12	4.52	4.24	4.10
23	10R 3/10	4.07	3.76	3.86
24	10R 2/6	4.86	4.24	3.48
25	5YR 7/14	4.14	5.03	5.34
26	5YR 6/14	4.45	4.45	4.86
27	5YR 5/12	3.72	3.48	3.45
28	5YR 4/8	4.21	3.34	3.28
29	5YR 3/6	4.52	3.72	3.31
30	5YR 2/4	4.83	4.07	3.14

Table 5.70: Nervousness scores, kindness scores and activeness scores of the test colors.

Test color no.	Munsell index	Nervousness score	Kindness score	Activenss score
1	5RP 7/10	2.79	5.00	4.86
2	5RP 6/12	3.41	4.55	4.59
3	5RP 5/12	3.66	4.21	4.28
4	5RP 4/12	3.93	4.00	4.48
5	5RP 3/10	4.41	3.62	3.72
6	5RP 2/8	4.86	3.28	3.76
7	10RP 7/8	2.93	5.24	3.86
8	10RP 6/12	3.52	4.86	4.86
9	10RP 5/14	4.03	4.10	5.52
10	10RP 4/14	4.03	4.34	4.69
11	10RP 3/10	4.17	3.79	3.90
12	10RP 2/8	4.31	3.72	3.83
13	5R 7/10	3.28	4.90	3.97
14	5R 6/12	3.52	4.97	4.62
15	5R 5/14	3.76	4.86	5.10
16	5R 4/14	4.83	4.62	5.45
17	5R 3/10	4.38	4.14	4.66
18	5R 2/8	4.66	4.00	3.83
19	10R 7/10	2.90	4.97	4.34
20	10R 6/14	3.62	5.03	5.52
21	10R 5/16	4.28	4.79	5.34
22	10R 4/12	4.17	4.62	3.86
23	10R 3/10	4.03	4.28	4.07
24	10R 2/6	4.21	3.72	3.24
25	5YR 7/14	3.72	5.17	5.38
26	5YR 6/14	3.28	5.28	4.76
27	5YR 5/12	3.52	4.34	3.69
28	5YR 4/8	3.10	4.07	3.48
29	5YR 3/6	3.83	4.48	3.28
30	5YR 2/4	4.41	4.14	3.41

Publications

Please refer to the next two pages for all publications relating to this doctoral research, which are summarized in the format specified by Graduated School of Human Sciences, Waseda University.

早稲田大学 博士（人間科学） 学位申請 研究業績書

［学位論文・学術論文・著書・その他（学会発表等）の順に記入してください］

氏　名　**方 思源**　印

（2018年05月14日現在）

学位論文

1． Extraction of Aesthetic Rules to Multi-Color Stimuli Using Artificial Intelligence Technology: Towards the Construction of an Artificial KANSEI System　2015　早稲田大学大学院人間科学研究科修士論文

学術論文

○**1．** Siyuan Fang, Keiichi Muramatsu, Tatsunori Matsui： 2015　Experimental study of aesthetic evaluation to multi-color stimuli using semantic differential method．日本感性工学会論文誌，14巻1号，37-47頁．

○**2．** Siyuan Fang, Keiichi Muramatsu, Tatsunori Matsui： 2017　A computational model simulating the mental function of multicolor aesthetic evaluation．Color Research & Application, Volume 42, Issue 2, pp.216-235．

○**3．** Siyuan Fang, Tatsunori Matsui： 2018　An experimental study on the continuous patterns of the influence of color focality on short-term memory performance of colors．International Journal of Affective Engineering（採録決定）．

その他

《国際学会発表（査読あり）》

1． Siyuan Fang, Tatsunori Matsui：　2017　Experimental Investigation into the Continuous Pattern of the Relationship Between Color Focality and Short-Term Memory Performance for Colors．Proceedings of the 39th Annual Meeting of the Cognitive Science Society, 1985-1990頁．

2． Yoshimasa Tawatsuji, Tatsuro Uno, Keita Okazaki, Siyuan Fang, Tatsunori Matsui：　2017　Extraction of Relationships between Learners' Physiological Information and Learners' Mental States by Machine Learning．Proceedings of the 25th International Conference on Computers in Education, 56-61頁．

《国際学会発表（査読なし）》

1． Yoshimasa Tawatsuji, Tatsuro Uno, Keita Okazaki, Siyuan Fang, Tatsunori Matsui： 2017　Formalization of Relationships Between Learners' Physiological Information and Learners' Mental States by Deep Neural Network．The 1st International Conference on LASI (Learning Analytics Summer Institute) - Asia．

２．Siyuan Fang，Tatsunori Matsui：　2018　Experimental Evidence for the Existence of Noisy Colors Effect in Modern Japanese Culture．International Association of Empirical Aesthetics (IAEA) Congress, Aug. 30-Sep. 2, 2018, Toronto, Canada（採録決定）．
３．Siyuan Fang，Tatsunori Matsui：　2018　Experimental Investigation into the Mediating Variables of the Relationship Between Color Focality and Color Preference． International Colour Association (AIC) Interim Meeting, Sep. 25-29, 2018, Lisbon, Portugal (採録決定)．

《国内学会発表（査読なし）》

１．方思源，村松慶一，松居辰則： 2015　Construction of an Artificial KANSEI System Simulating the Mental Function of Multi-Color Aesthetic Evaluation．　第29回日本人工知能学会大会，4J1-3．
２．村松慶一，方思源，田中英一郎，綿貫啓一，松居辰則：　2015　多色配色に対する潜在的態度と顕在的態度の比較．第17回日本感性工学会大会，B25（※優秀発表賞）．
３．村松慶一，方思源，田中英一郎，綿貫啓一，松居辰則： 2015　多色配色に対する潜在的態度としての嗜好に関する実験的検討．第46回日本色彩学会大会，日本色彩学会誌，39 (5_SUPPLEMENT)，128-129頁．
４．吉川花里，方思源，松居辰則： 2015　人工共感覚(数字に対する色イメージ)の記憶力に対する影響の定量的分析．第46回日本色彩学会大会，日本色彩学会誌，39 (5_SUPPLEMENT)，178-181頁．
５．方思源，松居辰則： 2016　An Experimental Study of the Relationship Between Focality and Short-Term Memory of Colors in the Japanese Language．第33回日本認知科学会大会予稿集，484-490頁．
６．近藤将，方思源，田和辻可昌，松居辰則：　2017　ポジティブな気分が生起する過程の心理モデルの提案．2016年度教育システム情報学会学生研究発表会，A04．
７．Siyuan Fang，Tatsunori Matsui：　2017　Experimental Probe into the Continuous Pattern of the Relationship Between Color Focality and Short-Term Memory Performance for Colors: A Case Study Testing the Japanese Language．第12回日本感性工学会春季大会，2B-07．
８．奥村和生，方思源，田和辻可昌，松居辰則：　2017　自閉スペクトラム症の物体認識特性を表現する定性的脳機能モデルの構築：ナボン図形（Navon figure）の認識を題材として．第12回日本感性工学会春季大会，2C-16．
９．Siyuan Fang，Keiichi Muramatsu，Tatsunori Matsui：　2017　Construction of a Computational Model of Multicolor Aesthetic Evaluation Based on Psychological Experiments．第19回日本感性工学会大会，E21．
１０．志村友，方思源，田和辻可昌，松居辰則：　2018　エージェントとのインタラクションによる安心感の付与とそのメカニズムの心理モデルの構築．第13回日本感性工学会春季大会，WP2-11．
１１．保戸田健人，方思源，田和辻可昌，松居辰則：　2018　Deep Learningによる色と形の組み合わせに対する審美的評価過程の解析の試み．第13回日本感性工学会春季大会，WP2-13．

《会議報告》

1．日高昇平，方思源，福岡未紗，白砂大，藤崎樹，鳥居拓馬： 2017 The 39th Annual Meeting of the Cognitive Science Society (CogSci 2017）参加報告．認知科学，24巻4号，576-581頁.