

Being red with anger or blue:
Can the recognition of emotional
expressions be enhanced by manipulating
facial skin color?

Monika Žak



Master of Philosophy in Psychology

Cognitive Neuroscience

Department of Psychology

UNIVERSITY OF OSLO

May 2015

Being red with anger or blue: Can the recognition of emotional expressions be enhanced by manipulating facial skin color?

© Monika Žak

2015

Being red with anger or blue: can the recognition of emotional expressions be enhanced by manipulating facial skin color?

Monika Žak

<http://www.duo.uio.no/>

Trykk: Reprosentralen, Universitetet i Oslo

IV

Acknowledgements

I would like to express my sincere gratitude to my supervisor Professor Bruno Laeng for his guidance, support and engagement in this project. Thank you, Bruno for all your ideas, insightful suggestions and inspiring discussions that greatly contributed to this work.

Abstract

A recent revolutionary theory of color vision suggests that seeing facial skin color improves emotion recognition. Depending on an individual's emotional state, the quantity of blood and its oxygenation under the skin may vary, which causes subtle changes in the skin color and especially in the face. According to Mark Changizi and colleagues, specific color information in the skin is thought to be picked up by other co-specifics as signals of specific emotional states. To investigate whether color information affects emotion recognition, we enhanced facial images representing basic emotional expressions like happiness, sadness, anger and fear with colors that were either congruent or incongruent with the presented emotional expression (according to Changizi's model). Specifically, we measured behavioral and pupillary responses in thirty five participants while they performed an emotional expression recognition task while viewing faces of strangers on a computer screen and while their eye pupils were monitored by use of an eye-tracker. We predicted that in the congruent condition, participants' performance would be improved when exposed to color-enhanced images than to their original versions. Respectively, their ability to recognize emotions should worsen while detecting emotions in the incongruent facial images. The findings did not fully support the hypotheses; however, there were some indications that congruent and incongruent conditions differed in both accuracy and response times, as well as in pupillary changes, that was consistent with the idea that facial skin plays a role in the recognition of emotional expressions. Finally, a standard color vision test (FM100) revealed that participants with better color vision also had higher accuracy in the task, indicating that seeing in color can be beneficial for the successful recognition of emotions.

Keywords: emotion recognition, facial skin color, color vision

Contents

- Introduction** 1
 - Why do we see the world in Technicolor? 1
 - A novel insight into color vision 1
 - Detecting skin color changes 3
- The present study** 4
 - The aim of the study 4
 - Hypotheses and predictions 5
- Methods** 5
 - Participants 5
 - Stimuli 6
 - Apparatus 8
 - Procedure 9
- Results** 12
 - Behavioral data 12
 - Accuracy 12
 - Response Times 15
 - Pupillary data 17
- Discussion** 19
- Limitations and recommendations for future studies** 22
- Conclusions** 24
- References** 25
- Appendices** 27

Introduction

Why do we see the world in Technicolor?

The ability to see the world in millions of different colors is something most people take for granted and also they believe to be useful. In fact, the majority of primates and all mammals greatly lack this ability (Kainz, Neitz & Neitz, 1998; Pichaud, Briscoe & Desplan, 1999; Lucas et al., 2003; SurrIDGE, Osorio & Mundy, 2003; Osorio & Vorobyev, 2008), being all dichromats (i.e. missing one of the three types of cones in the retina). This puzzling condition should raise the question: why did apes and some monkeys “regain” the ability to see a chromatic world as most birds and fish do? Most scientists believe that color vision reappeared to help our primate ancestors to spot ripe fruit and young leaves (Osorio & Vorobyev, 1999; Sumner & Mollon, 2003). Therefore they passed the color vision to their descendends and eventually to us, a specific group of hunter-gatherers hominids who would have supposedly benefited from feeding on fruits and vegetables.

A novel insight into color vision

The above dominant, evolutionary account has been strongly criticized by Mark Changizi and colleagues (2006) who claimed that human color vision may have actually evolved for a completely different reason; that is to give individuals greater insight into the emotional and physical states of other individuals. According to Changizi et al. (2006), we use color vision to sense what another person is subjectively experiencing by detecting subtle color changes in their skin. This special ability of skin to achieve a wide variety of colors finds its source in rapid changes in blood circulation under its surface. Specifically, two features of blood would seem to matter: the quantity of blood in the skin and the oxygenation level of that blood. Skin may appear yellowish or blueish depending on how much blood is under the skin, whereas other color coordinates, like redness and greenness of the skin may depend on how much oxygen is in the blood underneath (see Figure 1).

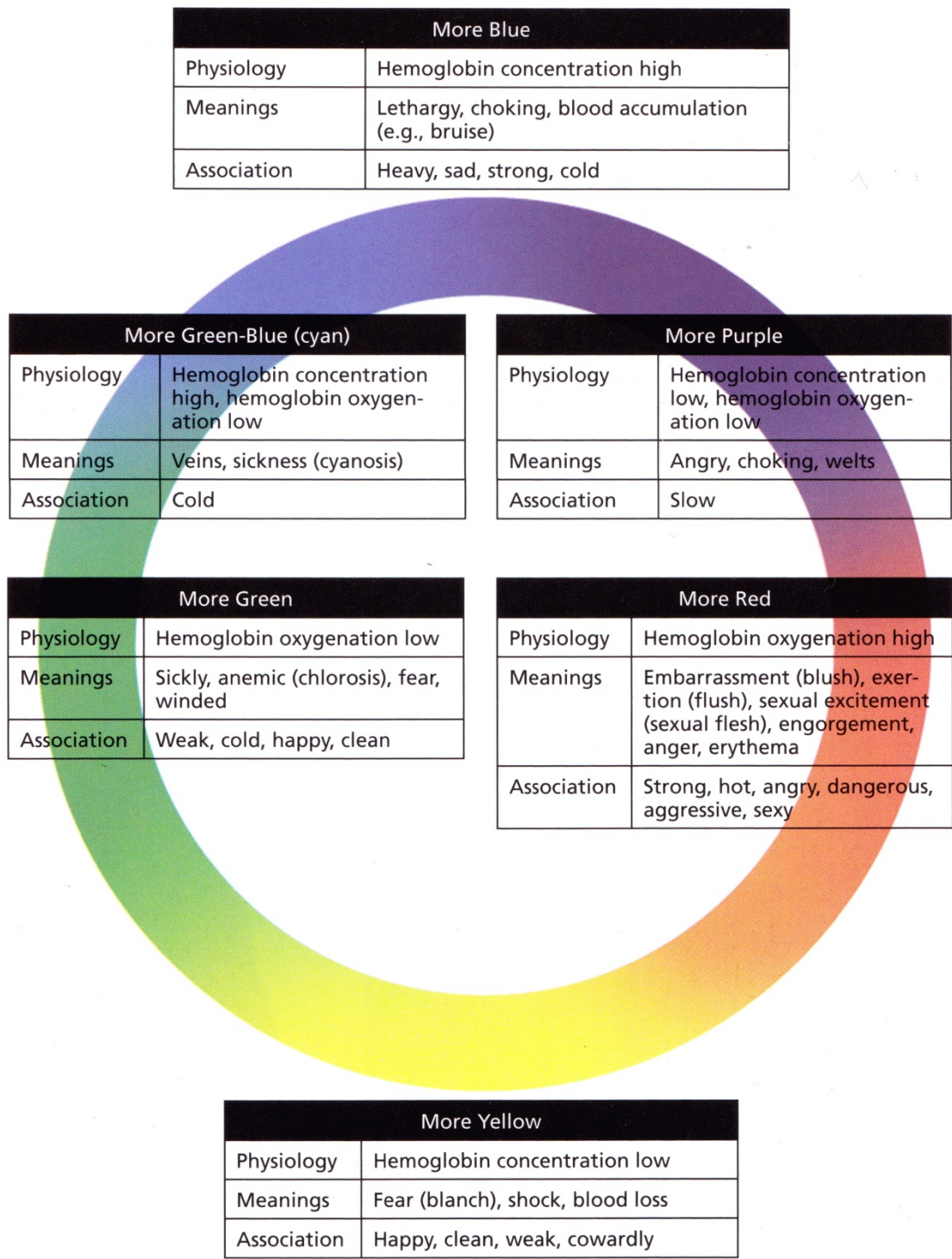


Figure 1. Summary of the manner in which skin color changes with blood, what those colors may sometimes mean to observers, and common associations with these colors. Adapted from “*The vision revolution : How the latest research overturns everything we thought we knew about human vision*” by M. A. Changizi, 2009, Benbella Books, Inc.

Changizi et al. (2006) also revealed connection between bare skin and color vision using photographs of different species of primates. Specifically, they found a positive correlation between amount of facial skin exposed (percentage of bare face) in a variety of primate species and the degree of richness of their color vision. Accordingly, monochromats and dichromats tend to have furry faces, whereas trichromats that see a rich colored world tend to have large bare skin regions in the face. This finding supports the theory that the purpose of color vision may be detecting color changes and recognizing emotional state. Hence, Changizi et al.'s account provides an appealing and logically convincing, alternative view on primates' color abilities. However, it lacks direct empirical evidence that seeing the colors in the face would actually assist detecting or correctly categorizing the emotion that is expressed in the face.

Detecting skin color changes

Various techniques have been used to investigate the relationship between facial skin color changes and emotional states. Jimenez et al. (2010) developed a color appearance model (a reconstruction of the non-contact SIA scope™ system) allowing to capture hemoglobin and melanin distribution within the facial skin. Data were collected from four Caucasian participants who showed different facial emotional expressions, including the six basic ones: anger, disgust, fear, happiness, sadness and surprise (by Ekman, 1972). Highly detailed hemoglobin and melanin maps showed expression-induced blood variations across measurements of an individual, while melanin stayed constant as its concentration typically changes within a longer timeframe, e.g., from hours to weeks. The results revealed that facial expressions may cause drainage of blood in compressed regions of the face, as well as perfusion into other facial areas; this can be observed when one is smiling and the blood concentration increases over the cheek bones.

Yamada and Watanabe (2007) used a digital video camera and thermograph to simultaneously measure changes in both facial skin color and temperature for different emotional expressions. In “communication experiment”, during which subjects could converse freely, expressions of anger, but also laughter and dislike occurred spontaneously. It was demonstrated that immediately after the facial expression of anger was produced, the skin temperature of the nose decreased, whereas the skin temperature of the forehead increased,

resulting in a large temperature difference between these two facial areas, and a visible red tinge on the cheek. Based on the observed skin color changes, Yamada and Watanabe manipulated facial images by increasing hue and saturation in relevant face areas in order to express anger more distinctly. They were showed to 10 participants who found presented faces more angry-looking than the same faces in their original colors, which indicated that the color enhancement made the facial expressions richer and more evident.

The present study

The aim of the study

The aim of the present study was to directly investigate whether color information indeed affects emotion recognition and, more specifically, if the relationship between emotions and skin colors proposed by Changizi et al. (2006) is supported by behavioral decisions. To achieve this goal, an experiment was designed in which participants were asked to look at faces displayed on the screen and choose – as fast and as accurately as possible – the correct emotional expression by pressing one of 5 labeled keys on the keyboard. The stimuli were both original and color-enhanced images of facial expressions that were manipulated in accordance with Changizi et al's circle of emotions-colors (as presented in Figure 1). We would expect that participants' performance can be improved or disrupted by the colors that were specifically enhanced. Thus, two versions or conditions of the experiment were designed, one in which the skin color in the images was congruent (e.g., anger = enhanced red) and one in which it was incongruent (e.g., sad = enhanced red) with the displayed emotional state. We used a normed and standard face database (the Karolinska Directed Emotional Faces) for this purpose, where the recognition of specific emotions from each expression is typically high. We collected behavioral data (accuracy rates and response times, RTs) as well as pupillary data, baseline-corrected to control for luminance, that is pupillary changes.

Hypotheses and predictions

In line with the “theory” of Changizi et al. (2006), we hypothesized that participants will find it easier to recognize emotional expressions when seeing enhanced facial skin color that were congruent with the displayed emotion (e.g., blue-colored sad faces). Therefore, we predicted that accuracy will be higher for congruent enhanced images than for original ones, whereas RTs will be longer for original pictures than for color-enhanced images. We expected the opposite effects for the incongruent condition of the experiment, e.g., lower accuracy for incorrectly enhanced images than for original images, and longer RTs for enhanced images than for incorrectly enhanced colored images than for their original versions. We also expected to observe systematic patterns in pupillary dilations, presuming that pupils will dilate more to enhanced images in the incongruent condition than in congruent condition due to more difficulty in recognizing the facial expressions and therefore, more effortful processing (Hess & Polt, 1964; Kahneman & Beatty, 1966).

Methods

Participants

Fourty participants volunteered to take part in the experiment. Data from 5 participants were excluded due to poor performance (i.e. providing a response to fewer than 50% of trials) or technical problems with the eye-tracking system. The remaining 35 participants were all individuals with normal or corrected-to-normal vision. According to the results of the Farnsworth-Munsell 100-Hue Test (computed using the FM100-Hue scoring program: <http://www.torok.info/fm100>), 10 participants had color discrimination that signaled probable pathologic performance or at least borderline. Ages ranged from 19 to 45 ($M = 25.67$, $SD = 5.88$). Seventeen subjects (8 males and 9 females) were tested with the color-congruent version of the experiment, whereas the remaining 18 (8 males and 10 females) took part in the incongruent version of the experiment. Since there were 2 more conditions (Keyboard: HaNeSaAnAf, AfAnSaNeHa; Version: A, B), participants were distributed across 8 different subgroups in a counter-balanced manner (see Table 1).

	Congruent				Incongruent			
	HaNeSaAnAf		AfAnSaNeHa		HaNeSaAnAf		AfAnSaNeHa	
	A	B	A	B	A	B	A	B
females	2	2	2	3	4	2	2	2
males	3	2	1	2	1	3	2	2

Table 1. A total number and distribution of female and male participants across 8 different subgroups.

Recruitment for the experiment took place at the campus of the University of Oslo or via the Internet by means of social media. Those who expressed interest in participation were provided with a detailed description of experimental procedure and its requirements, without revealing, until the debriefing stage, what was the purpose of the study. All the participants were compensated with 100 NOK at the end of the session.

Stimuli

Participants judged which of five basic facial expressions (neutrality, happiness, sadness, anger, and fear) were displayed by female and male models of the Karolinska Directed Emotional Faces database (KDEF). For the purpose of this study 50 images were selected from the total set of 490 pictures of human facial expressions. Thus, each expression was represented by the 10 best (i.e. with highest emotional labels' hit rates) according to a validation study (by Goeleven, De Raedt, Leyman & Verschuere, 2008). All stimuli, except for the ones representing neutrality, were edited in *Adobe Photoshop CS5 Extended* in order to enhance the color of the skin, either congruently or incongruently to the color-emotion circle presented earlier in Figure 1. The pictures were manipulated by changing RGB values (*Image > Adjustments > Color Balance*) only for the facial area and the visible part of the neck (to avoid visible contrast between the colors), leaving untouched the rest of the image. Adjustments of color levels for the congruent version of images were the following: +30

Yellow for happy facial expressions, +30 Blue for sad facial expressions, +30 Red for angry facial expressions, and +30 Green for fearful facial expressions (see Figure 2).

An incongruent version of the same images was also created by modulating color information in the pictures in such way that color enhancement was always incongruent (according to the color-emotion circle of Changizi and colleagues) with the emotion expressed in a facial image. In this case, adjustments of color levels were the following: +30 Blue for happy facial expressions, +30 Yellow for sad facial expressions, +30 Green for angry facial expressions, and +30 Red for fearful facial expressions. Neutral facial images were not manipulated with color and were always used in their original color version.



Figure 2. Examples of original and color-enhanced stimuli (shown side-by-side) used in the experiment.

The basic idea was to enhance the images in a subtle way so that participants could not notice that skin color was manipulated or would find the faces odd-looking. This aim was successfully achieved; during debriefing none of the participants claimed to notice any unusual changes in the facial skin color. Two sets of stimuli were used in the experiment. Version A consisted of 30 original pictures (10 for neutral facial expressions and 5 for each expression out of the 4 remaining ones) and 20 color-enhanced images, whereas version B included 20 enhanced pictures of those faces that were presented as original ones in version A. Consequently, an experimental session consisted of 50 trials and included 10 images of facial expressions for each emotion.

Apparatus

Throughout the entire experiment, pupillary response data were collected using a Remote Eyetracking Device (SensoMotoric Instruments GmbH). The participant's pupil diameter was measured at a sampling rate of 60 Hz and was recorded with iView X Hi-Speed Software. The software compensates head-motion automatically, allowing the subject free head movements across a wide range (40 x 20 cm at 70 cm distance). Nevertheless, a chinrest was also used to minimize head movements as much as possible. Participants therefore viewed the screen at a distance of 60 cm. All the stimuli were presented with Experiment Center 3.2 by SMI – and were shown on a Dell P2213 VGA LCD monitor. The monitor's size was 18.5", measuring a diagonal length of 47 cm. The display resolution was set to 1680 x 1050 pixels, and held constant throughout the procedure. The experiment was run on a Dell Latitude E6530 powered by an Intel i7-3520M CPU at 2.9 GHz. and 4 GB of RAM, running Windows 7 at 32 bit. All key press data were recorded with a standard HP full-size keyboard. In order to control precisely color, contrast and brightness, the monitor used in this experiment was calibrated with ColorVision Spyder2PRO™. All the participants performed the task under constant light conditions.

Procedure

The experiment was conducted in the Cognitive Laboratory at the Institute of Psychology, University of Oslo. The participants were tested in a quiet, windowless room, with moderate artificial illumination. The chair and the chinrest were adjusted to match the height of each participant before the testing session started. Subjects were instructed to sit comfortably but keeping stable as possible during the entire duration of the experiment. They were also informed that the purpose of the study was to investigate eye movements in response to emotional expressions and the actual purpose of the experiment was only revealed at the final, debriefing stage.

Prior to the actual experiment, a calibration procedure was performed. Participants were asked to follow with their eyes a moving white dot onto 4 fixed positions, while the eye tracker monitored these movements. Then, the SMI software repeated the procedure to validate the precision of gaze tracking and provide an estimation of eye position errors; if these were found to be minimal (i.e. degrees of visual angle were lower or equal to 0.5), the calibration was deemed successful. The experimenter ensured that participant understood how to perform the task and at that point the stimuli presentation began, and detailed task instructions were displayed on the screen at the start of the test

After responding to a single practice trial, each participant performed the actual experiment. Both congruent and incongruent version of the experiment consisted of 60 trials in total. The order of the trials was fixed according to a pseudo-random sequence, so as to ensure that it was unpredictable what will be seen next, and also to prevent that the same expression (despite shown by different individuals) happened to be displayed back-to-back. Each trial started with a 1000 ms presentation of a mosaic, which was a pixelated version of a facial photography (edited in *Adobe Photoshop CS5 Extended* using *Filter > Pixelate > Mosaic*, choosing a cell size of 142 square pixels). The intention was for the mosaic to provide a baseline for each picture as it was composed of the colors of the same intensity. Next, a fixation cross appeared at one of four corners and remained on the screen until participant's gaze dwelled on the cross for at least 400 ms. Then, an image of facial expression was displayed for 2000 ms. At that point participants had to decide as quickly as possible what was the emotional expression and give an answer by pressing the corresponding

keypress on the keyboard. The participant decided when to start a new trial by pressing a spacebar on the keyboard (i.e. in a self-paced fashion).

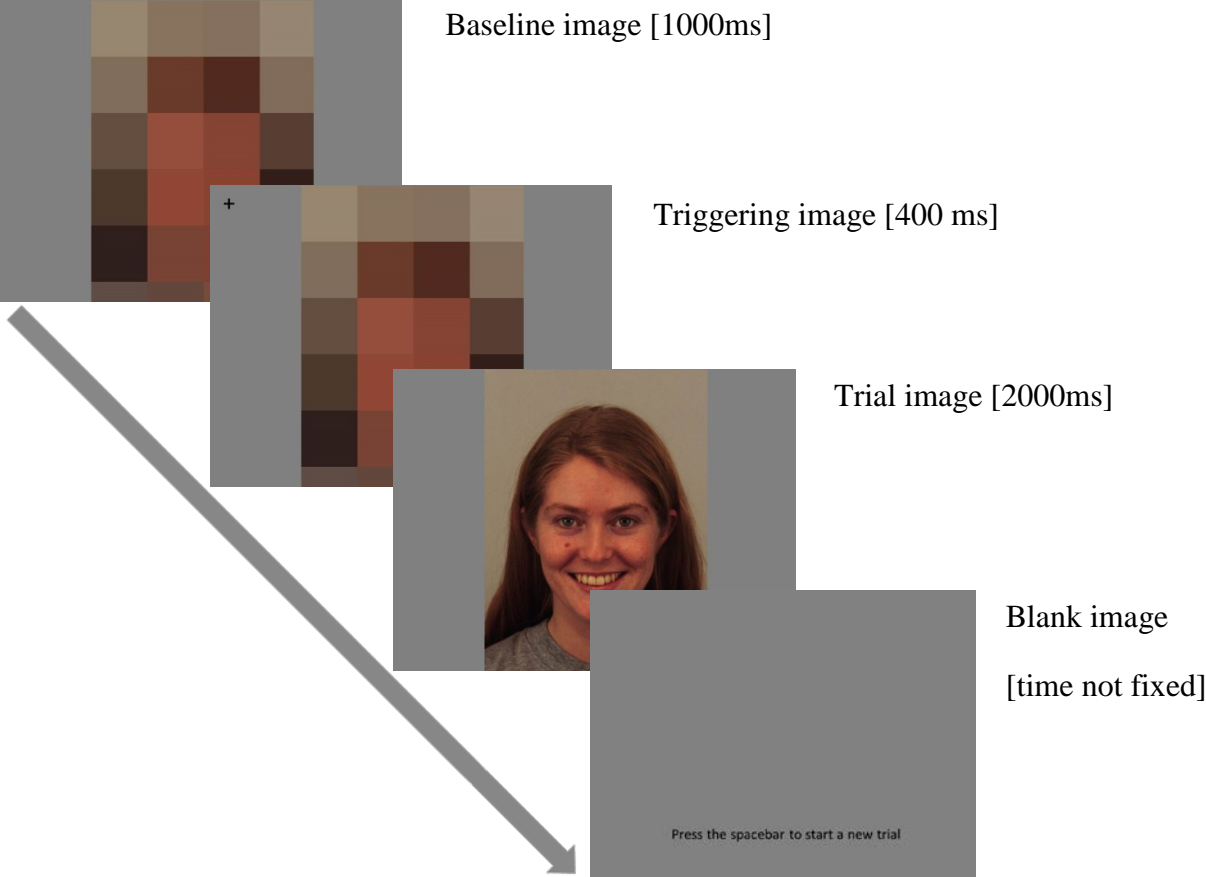


Figure 3. An illustration of the sequence of images within one trial and their display durations.

Stickers representing sketchy faces for each emotion were designed and pasted on the keyboard buttons in the following order: happy – G button, neutral – H button, sad – J button, angry – B button, afraid – N button (see Figure 4). Since three upper buttons (G, H, J) were labeled with emotional expressions of the same dimension (from happy to sad, with neutral in between), we hypothesized that participants could respond quicker to neutral, happy and sad faces than to fearful or angry ones. To avoid prioritizing any of the emotional expressions, participants were divided into either the above keyboard condition or its reversal. In the latter

case, the keys were labeled in a reversed order with the emotional sketchy faces: happy – N button, neutral – B button, sad – J button, angry – H button, afraid – G button.



Figure 4. The response stickers' placement in one keyboard condition (HaNeSaAnAf).

The total duration of the experiment was of approximately 15 min. After performing the task, subjects were given a questionnaire to gather information about age, gender, nationality, educational level, handedness and quality of vision (see Appendix 1, Appendix 2). Additionally, participants were asked to write down whether they had noticed something particular or odd in the presented faces, and what they thought the idea behind the study was.

In the last phase of the testing session, we used the Farnsworth-Munsell 100 Hue Color Vision Test (FM100) as a standardized and normed color vision test. The test consisted of four trays containing a total 85 removable color reference caps spanning the visible spectrum (Figure 5). Participants were asked to order shown color caps in their correct order according to color similarity. Time of performing this task was not measured. Color vision

abnormalities and deficiencies could be detected by the ability of the participant to place the color caps in order of hue. The results of the Farnsworth- Munsell 100 Hue Color Vision Test were calculated using an FM100-Hue scoring program provided by Béla Török on his web-based platform (<http://www.torok.info/colorvision/fm100.htm>).



Figure 5. The Farnsworth- Munsell 100 Hue Test.

Results

Behavioral data

Accuracy. Participants' key presses were extracted using BeGaze™ analysis software provided by SMI. Then, the acquired data was organized in Microsoft Excel and each response was assigned an accuracy value (100 for a correct answer and 0 for an incorrect answer); consequently, when computing the means of these accuracies we obtained values corresponding to the mean percentage of accuracy. Mean accuracy was calculated for each participant for 5 different emotional expressions (neutral, happy, sad, angry, afraid) and the 2 different Manipulation conditions (original, enhanced).

Two-way ANOVA tests analyzed the effect of two independent variables (Manipulation, Expression) on accuracy. Data from congruent and incongruent version of the experiment were analyzed in separate ANOVAs for each group of participants. Because neutral facial expressions were never color-manipulated, responses to these faces were

analyzed separately. Therefore, the factor Expression in the above 2-way ANOVA had a total of four levels (afraid, angry, happy, sad). The keyboard sequence was not used in the ANOVAs since this was meant as a counterbalancing factor in the experiment design.

The ANOVA for the congruent version of the experiment revealed no main effect of either Expression, $F(3,14) = 1.907, p = .139$ or Manipulation, $F(1, 16) = .374, p = .548$. Similarly, there was no significant interaction of Manipulation*Expression, $F(3,14) = 1.652, p = .188$. Thus, the congruent condition failed to support our original hypothesis and predictions of a behavioral advantage when appropriate color information is made more salient.

However, when the same ANOVA was performed for the incongruent version of the experiment, there was a significant effect of Expression on accuracy, $F(3, 15) = 6.362, p = .001$. Yet, there was no main effect of Manipulation, $F(1, 17) = 1.999, p = .178$ and no significant interaction of Manipulation*Expression, $F(3, 15) = 1.011, p = .397$. Further analysis with the paired samples t-tests for incongruent data revealed significant difference between mean accuracy for both afraid and happy emotional expressions, $t(1,17) = -3.718, p = .002$. Participants responded to happy facial images with the highest accuracy ($M = 98.750, SD = 5$), whereas afraid facial images had the lowest hit rate ($M = 79.063, SD = 20.954$). Moreover, subjects tended to recognize each facial expressions more accurately, when color manipulation of the images was congruent with displayed emotion (see Figure 6), which is at least consistent with our hypothesis. Therefore, we decided to perform a one-way ANOVA for data from both congruent and incongruent version of the experiment. The results revealed that effect of Color (congruent, incongruent) on accuracy approached significance, $F(1, 34) = 3.482, p = .07$, and that this trend was in the predicted direction.

A separate one-way ANOVA analysis was performed for only neutral trials from both congruent and incongruent version of the experiment. The effect of Color on accuracy was found insignificant, $F(1, 34) = 1.210, p = .278$. Based on this, we assumed that previous results for accuracy were not affected or confounded by individual differences between participants belonging to either the congruent or incongruent condition group.

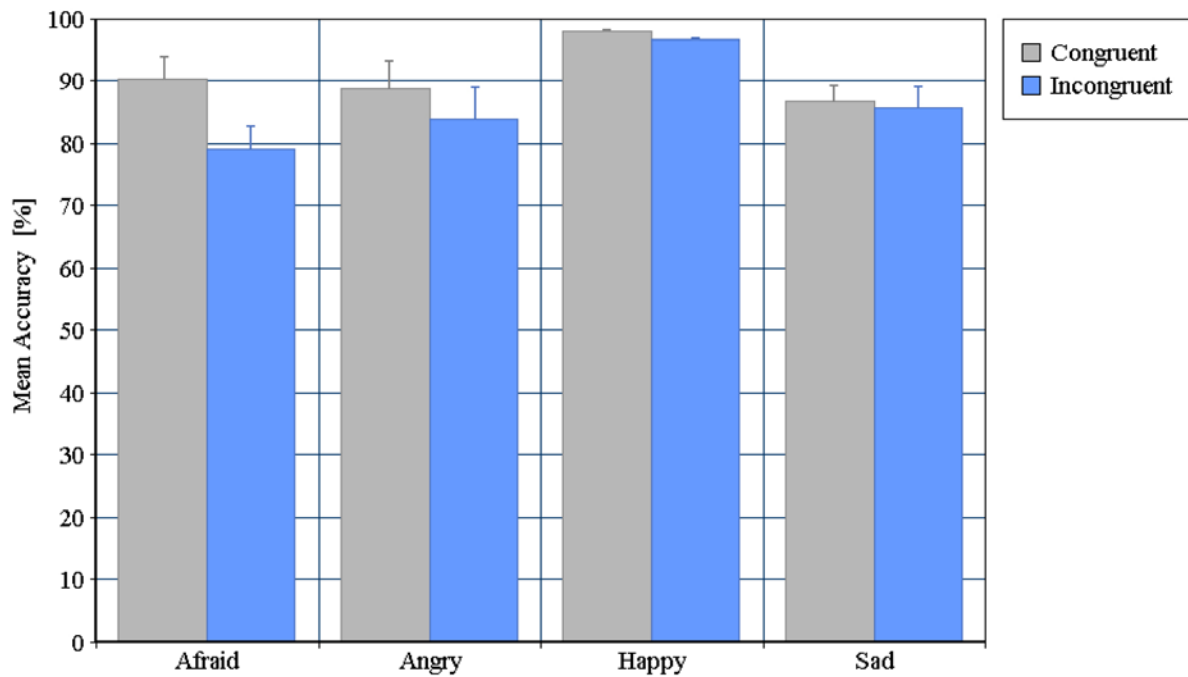


Figure 6. Mean accuracies and Standard Errors for emotional expressions in congruent and incongruent version of the experiment.

Finally, we investigated whether an individual's color vision sensitivity may affect the accuracy of emotion recognition. A simple linear regression analysis was performed on Total Error Scores (FM100) and mean accuracy percentages. The results showed that changes in the Total Error Scores value were significantly and negatively related to changes in accuracy, $t(1,34) = -2.052, p = .048$ (see Figure 7). The coefficient for Total Error Scores was $-.053$, indicating that for every additional score in Total Error Score value, accuracy can be expected to decrease by an average of $.053$.

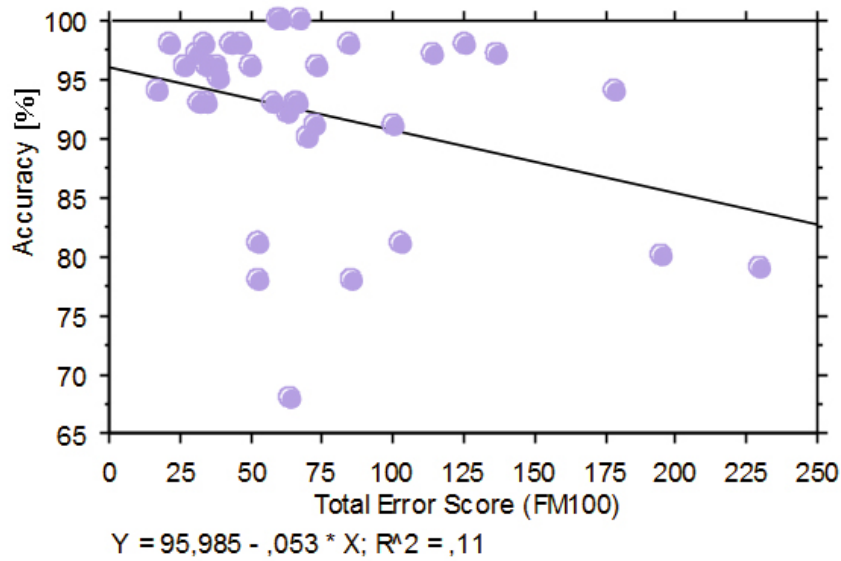


Figure 7. Regression plot presenting the relationship between Total Error Score (FM100) and Accuracy.

Response Times. The keypress events were extracted using BeGaze™ analysis software and organized in Microsoft Excel. RTs of incorrect answers were not included in the analysis. We also excluded outliers by identifying RTs that exceeded 3 standard deviations from the mean. The average values of the remaining response latencies were calculated for each participant for each of the 5 different emotional expressions (afraid, angry, happy, neutral, sad) and in both versions of the experiment (congruent, incongruent).

A two-way ANOVA test examined the influence of color manipulation and emotional expression on RTs. Since neutral facial expressions were never color-manipulated, responses to these images were analyzed separately. Therefore, the factor Expression had four levels (afraid, angry, happy, sad) as seen also in the previous analyses. For the congruent version of the experiment, the analysis revealed no main effect of Manipulation, $F(1, 16) = .682, p = .421$. There was also no significant interaction of Manipulation*Expression, $F(3, 14) = 1.712, p = .177$. However, we observed a main effect of Expression in the task, $F(3, 14) = 20.004, p < .001$. Paired t-tests revealed that participants responded significantly faster to happy facial images than to other emotional expressions (see Figure 8). Results of a similar ANOVA for incongruent version showed no main effect of either Manipulation, $F(1, 17) =$

1.167, $p = .295$ or Expression, $F(3, 15) = .074$, $p = .974$. Similarly, no significant interaction of Manipulation*Expression was observed, $F(3, 15) = .949$, $p = .424$.

As shown in Figure 8, RTs for each emotional expression seemed longer in the incongruent version than in the congruent version of the experiment. These differences, however, resulted to be insignificant when tested with an ANOVA with Color (congruent, incongruent) as independent variable, $F(1, 34) = .002$, $p = .965$.

A separate one-way ANOVA analysis was performed for only neutral trials from both congruent and incongruent version of the experiment. The effect of Color was found insignificant, $F(1, 34) = 2.011$, $p = .166$. Thus, we assumed that previous results for RTs were not affected by individual differences in speed of response between participants in the congruent and incongruent groups of the experiment.

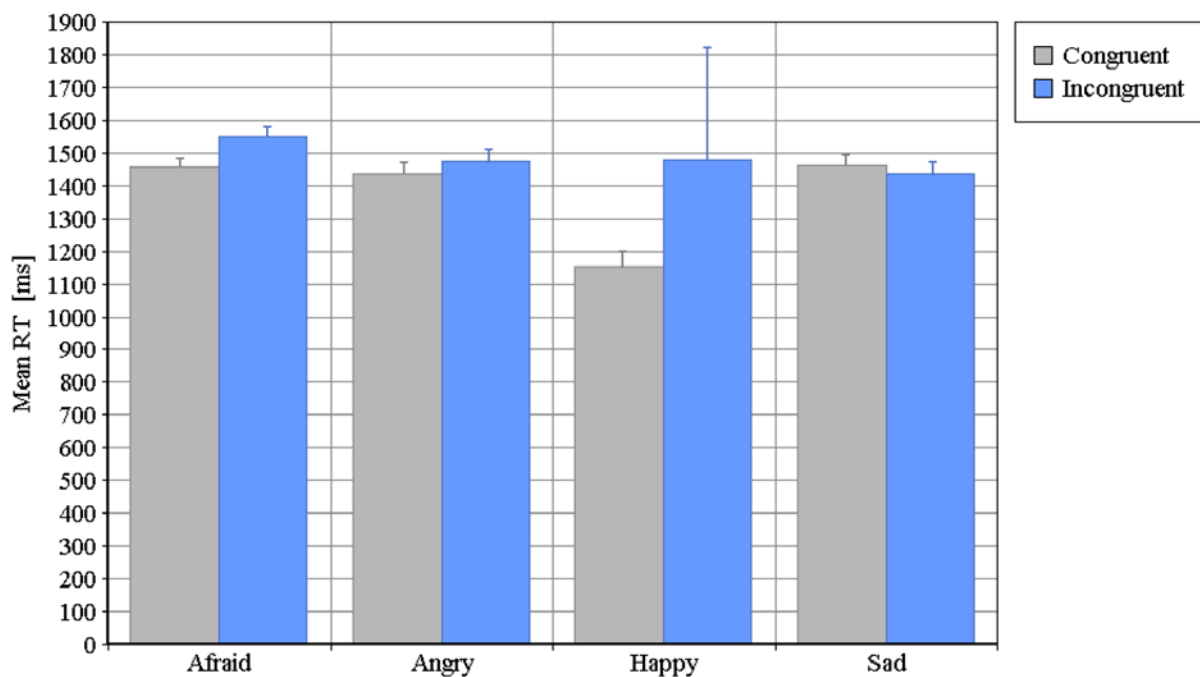


Figure 8. Mean response times and Standard Errors for emotional expressions in the congruent and incongruent conditions.

Pupillary data

The raw pupillometry data were extracted using BeGaze™ analysis software provided by SMI and preprocessed in Microsoft Excel. We calculated average values of pupil size for baseline trials and actual trials (facial images) and subtracted the former from the latter in order to measure changes in pupillary diameter for each stimulus that were independent from changes in luminance across stimuli. Outliers were excluded (i.e. trials showing a pupil size more than three standard deviations away from the mean). The average pupillary changes were calculated for each participant for each emotional expression in both congruent and incongruent version of the experiment.

We conducted a 2x4 ANOVA for the mean pupillary change in a congruent version of the experiment with Manipulation (original, enhanced) and Expression (afraid, angry, happy, sad) as within-subjects factors. The analysis revealed no main effect of Manipulation, $F(1, 16) = .563, p = .461$, but a main effect of Expression, $F(1, 16) = 4.97, p = .004$ and, importantly, a significant interaction of Manipulation*Expression, $F(3, 14) = 4.584, p = .005$. Further analysis with the paired samples t-tests showed significant difference in pupil diameter between the original sad expressions and the enhanced sad expressions, $t(1,16) = -2.421, p = .024$, as well as original angry expressions and enhanced angry expressions, $t(1,16) = 2.062, p = .05$. The difference in pupillary change between afraid original expressions and afraid enhanced expressions barely missed significance, $t(1,16) = -2.008, p = .057$.

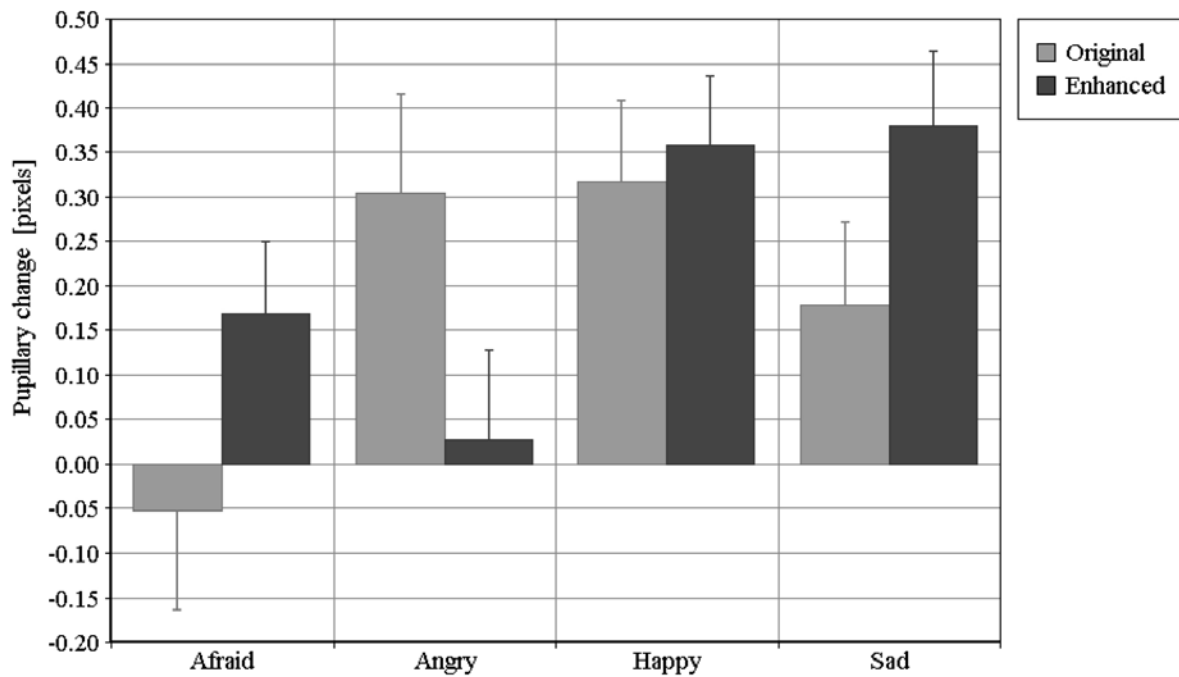


Figure 9. Mean pupillary changes and Standard Errors for original and enhanced emotional expressions in the congruent condition.

A similar separate two-way ANOVA was performed for data from the incongruent version of the experiment. The only significant effect revealed by this analysis was the influence of emotional expression on pupillary change, $F(1, 17) = 3.387, p = .025$. More specifically, the pupil diameter was significantly larger in response to enhanced happy emotional expressions than to other enhanced trials (see Figure 10) as verified by paired t-tests.

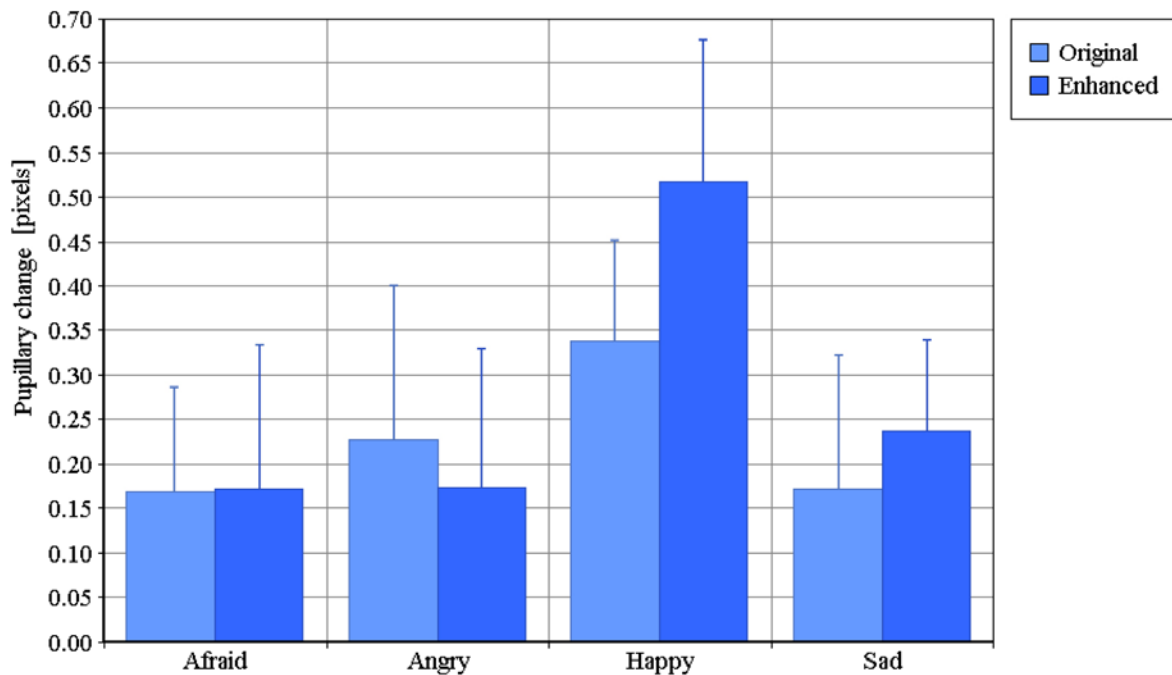


Figure 10. Mean pupillary changes and Standard Errors for the original and enhanced emotional expressions in the incongruent condition.

Lastly, a separate one-way ANOVA analysis was performed for only neutral trials from both congruent and incongruent version of the experiment. The effect of Color on pupillary change was found to be non-significant, $F(1, 34) = .776, p = .384$. Based on this, we assumed that previous results for pupillary changes were not affected by individual differences between groups in their effort employed in either the congruent or incongruent conditions.

Discussion

The purpose of this study was to investigate whether making facial skin color more salient can modulate emotion recognition, as one would predict on the basis of Changizi and colleagues' account (2006). For this aim, we color-enhanced the skin on the face of

individuals and for 4 different but basic emotional expressions, and used these in a task in which participants had to identify the correct label for each emotional expressions. In the incongruent version of the experiment, colors were enhanced incongruently (to Changizi and colleagues' color-emotion circle) and we predicted that participants' performance should worsen due to an incorrect enhancement of a color not associated with the displayed emotion.

The results revealed no significant difference in accuracy between original and enhanced trials in the congruent version of the experiment. Therefore, the results do not bring support to the notion that a more salient facial skin color provides more evidence about emotional state of another person. One possible reason for such result is that the KDEF stimuli already comprise high intensity exaggerated emotions, which are relatively easy to recognize. Using more subtle facial expressions (e.g., morphs of facial expressions) could have increased task difficulty and perhaps, also the effect of color manipulation on emotion recognition ability. Similar results were showed in other experiments using KDEF images, where accuracy for different emotional expressions did not significantly differ (e.g., Sucksmith, Allison, Baron-Cohen & Hoekstra, 2013).

However, we did observe that participants in the incongruent version of the experiment tended to have lower accuracy for the "afraid" facial expression. We did find a significant decrease in performance to "fearful" versus "happy" expressions. A possible explanation is that "fearful" is a more ambiguous facial expression than "happy". Consequently, incongruent skin color might have made these manipulated afraid expressions more difficult to recognize, whereas the accuracy for happy images was not affected by color.

Moreover, participants were generally more accurate when they responded to trials from the congruent version of the experiment compared to the incongruent one, although this difference only approached significance. This trend was observed not only for all trials, but also for each emotional expression. The findings seem at least consistent with our predictions that subjects' performance should worsen when exposed to incongruent images. Presumably, this effect may be more clearly revealed by increasing the power in our test (i.e., by increasing the sample size).

Including the Farnsworth-Munsell 100 Hue Color Vision Test in this study was based on the assumption that ability to detect subtle changes in skin color provides information about what another person is feeling (Changizi, 2009). Therefore, we hypothesized that an

accurate color vision should improve emotion recognition or, conversely, poor color perception should reduce the ability to recognize emotions. This analysis indeed revealed that participants' scores in the FM100 test were negatively correlated with accuracy values.

Furthermore, we found no evidence for differences between RTs to original and to color-enhanced conditions, regardless of whether these were congruent or incongruent. However, participants responded significantly faster to happy facial images than to other emotional expressions in the congruent condition. This effect can be again explained by the relative ease of detecting happy facial expressions. Interestingly, this effect disappeared in the incongruent version of the experiment. It is possible that the incongruent color of the skin made the decision more difficult and latencies to happy images increased. However, these latency changes failed to reach significance in the statistical analysis.

Finally, pupillary results for the congruent version of the experiment revealed a significant effect of emotional expression, as well as a significant interaction of emotional expression and color-manipulation. Pupillary changes were found larger in response to enhanced images than to original images for both sad and fearful emotional expressions. Surprisingly, we observed the opposite relationship for angry faces as the pupil diameter changes were significantly larger in response to the original images than to the color-manipulated images. Moreover, pupillary size changes for happy images were similar for both original and enhanced stimuli, probably due to the relative ease of recognizing this emotional expression.

We observed substantially different results in the incongruent part of the experiment. Importantly, there was no significant interaction of emotional expression and color manipulation. The only significant difference in pupil size was observed for the "happy" expression; mean pupil diameter changes were larger for enhanced images than for original ones. As for the congruent version of the experiment, pupillary changes in response to angry expressions were larger for original images than for enhanced ones. This finding suggests that such relationship is not affected by color manipulation.

It seems difficult to interpret the above pattern of results on pupillary changes. On one hand, one could interpret increases in pupillary dilations as signs of effortful processing (Hess & Polt, 1964; Kahneman & Beatty, 1966), as we had hypothesized at the outset. In this case, one would have expected that pupils should have dilated more to enhanced images in the

incongruent condition, but this was not the case since the pupils dilated more often and differently to various expressions in the congruent condition. On the other hand, a pupillary dilation may reflect instead of effort, the occurrence of a successful recognition of a target. Specifically, Papesh, Goldinger & Hout (2011) found that – in participants that had studied a set of items and later discriminated old from new items, and gave confidence judgments about their discriminations – high-confidence decisions were characterized by large pupils, relative to lower-confidence decisions. In our case, the more frequent pupillary dilations in the congruent than incongruent condition could indicate that enhanced expressions were often accompanied by a “confident acknowledgment” of the presence of a target emotion in the absence of differences in cognitive workload between the two conditions.

Limitations and recommendations for future studies

There are specific limitations of this study that should be taken under consideration. Firstly, the stimuli used in our experiment were images of exaggerated emotional expressions that were relatively easy to recognize. This may have decreased the differences in accuracy and RTs between original and enhanced images. As discussed earlier, a suggested solution is to use more subtle facial expressions in order to increase task difficulty. This would possibly increase the effect of color manipulation on emotion recognition ability.

Another important limitation of our experiment is that the stimuli presented fake emotional expressions played by actors. As professionals, these actors could convincingly obviously pretend to be happy and fake a smile, but we note that pupil size cannot be consciously controlled by most people (Laeng, Sirois, & Gredebäck, 2012) and possibly also by professional actors (especially when posing for static photographs). It is known however that the pupil size is affected by one’s emotion; for instance, experiencing anger or other negative emotions may cause pupil constriction and also that these differences can be (unconsciously) detected by observers and result in activation of the amygdala (Demos, Kelley, Ryan, Davis & Whalen, 2008). Given that monitoring changes of pupil diameter can provide a clue about the emotional state of an individual, we can conclude that the facial images used in the experiment lacked relevant information, that is pupil size, which may improve detecting emotional expressions.

Similarly, no changes in the blood flow were visible in the original pictures since the emotions were just pretended, but not actually experienced by the actors. It is possible that an artificial color enhancement over the whole skin surface is inaccurate as the obtained effect was far from naturally occurring changes of skin color which is more graded and directly reflect the degree of vascularization of different regions of the face (e.g., considering the flush of embarrassment, which is most typically visible in the cheeks). Perhaps not the whole oval of the face, but only some parts should have been colored; therefore one would need to find the optimal and natural spots in which enhancement could facilitate recognition. The level of color manipulation should be also questioned. Ideally, it should be measured to what extent the facial skin color changes in response to real emotions being felt so that this could be simulated and enhanced more accurately.

Moreover, the effect of expression itself (muscular pattern and deformation of the facial surface) is probably more powerful than the effect of color manipulation. If so, the enhanced and congruent color information may improve emotion recognition, but it may not be as effective in the opposite direction – i.e., incongruent coloring may be not enough to worsen accuracy or RTs to a large extent.

There are also limitations related to the credibility of the results of the Farnsworth-Munsell 100 Hue Color Vision Test. Even though participants were instructed to be as accurate as possible, it was difficult to control whether they were truly engaged in the task. Some participants seemed more precise (and often enthusiastic about the color test), whereas others showed lower motivation. Therefore, the unusually high prevalence of individuals with a probable pathologic vision should be treated with caution.

Lastly, it is possible that data were collected with an insufficient sample size to achieve a respectable level of statistical power. Therefore, increasing the sample size would be recommended for further research.

Conclusions

Findings from the present experiment do not provide solid evidence for or against the notion that seeing more saliently the facial skin color of another improves or disrupts emotion recognition. Although our hypotheses were not explicitly confirmed, we observed suggestive differences in participants' performance between congruent and incongruent conditions. Firstly, it was showed that participants had significantly lower accuracy in response to fearful emotional expressions in the incongruent condition, but not in the congruent one. Similarly, participants responded significantly faster to happy images in the congruent condition, but the same effect was not found in the incongruent version of the experiment. There was also a significant interaction of emotional expression and color manipulation on pupillary change in congruent condition only. These findings indicate that using incongruent facial skin color could worsen performance in the task and also influence pupil dilation. Likewise, the results of color vision test revealed that color vision ability is related to accuracy in the task, suggesting that seeing in color may improve emotion recognition. Possibly, by creating a more thorough experimental design, we could gain a better understanding of the evolution of trichromatic color vision in primates, and the relationship between facial skin color and recognition of emotional expressions.

References

- Changizi, M. A., Zhang, Q., & Shimojo, S. (2006). Bare skin, blood and the evolution of primate colour vision. *Biology Letters*, *2*, 217-221.
- Changizi, M. A. (2009). *The vision revolution: How the latest research overturns everything we thought we knew about human vision*. Dallas, Tx: Bella Books.
- Demos, K. E., Kelley, W. M., Ryan, S. L., Davis, F. C., & Whalen, P. J. (2008). Human Amygdala Sensitivity to the Pupil Size of Others. *Cerebral Cortex*, *18*, 2729—2734.
- Ekman, P. (1972). Universal and cultural differences in facial expression of emotion. In Cole, J. (Ed.) *Nebraska Symposium on Motivation, Vol. 19*, 207–282.
- Goeleven, E., De Raedt, R., Leyman, L., & Verschuere, B. (2008). The Karolinska Directed Emotional Faces: A validation study. *Cognition & Emotion*, *22*(6), 1094-1118.
- Hess, E. H., & Polt, J. H. (1964). Pupil Size in Relation to Mental Activity During Simple Problem Solving. *Science*, *143*, 1190-1192.
- Jimenez, J., Scully, T., Barbosa, N., Donner, C., Alvarez, X., Vieira, T., Matts, P., Orvalho, V., Gutierrez, D., & Weyrich, T. (2010). A practical appearance model for dynamic facial color. *ACM Transactions on Graphics*, *29*(6), 1-10.
- Kainz, P. M., Neitz, J., & Neitz, M. (1998). Recent evolution of uniform trichromacy in a New World monkey. *Vision Research*, *38*, 3315-3320.
- Kahneman, D., & Beatty, J. (1966). Pupil diameter and load on memory. *Science*, *154*, 1583-1585.
- Laeng, B., Sirois, S., & Gredebäck, G. (2012). Pupillometry: A window to the preconscious? *Perspectives on Psychological Science*, *7*(1), 18-27.
- Lucas, P.W., Dominy, N. J., Riba-Hernandez, P., Stoner, K. E., Yamashita, N., Loria-Calderon, E., Petersen-Pereira, W., Rojas-Duran, Y, Salas-Pena, R., Solis-Madriral, S., Osorio, D., & Darvell, B. W. (2003). Evolution and function of routine trichromatic vision in primates. *Evolution*, *57*(11), 2636-2643.

- Osorio, D., & Vorobyev, M. (1996). Colour vision as an adaptation to frugivory in primates. *Proceedings of the Royal Society of London B*, 263, 593-99.
- Osorio, D., & Vorobyev, M. (2008). A review of the evolution of animal colour vision and visual communication signals. *Vision Research*, 48, 2042-2051.
- Papesh, M. H., Goldinger, S. D., & Hout, M. C. (2011). Memory strength and specificity revealed by pupillometry. *International Journal of Psychophysiology*, 83, 56–64.
- Pichaud, F., Briscoe, A., & Desplan, C. (1999). Evolution of color vision, *Current Opinion in Neurobiology*, 9, 622-7.
- Sucksmith, E., Allison, C., Baron-Cohen, S., Chakrabarti, B., & Hoekstra, R. A. (2013). Empathy and emotion recognition in people with autism, first-degree relatives, and controls. *Neuropsychologia*, 51(1), 98–105.
- Sumner, P., & Mollon, J. D. (2003). Did primate trichromacy evolve for frugivory or folivory? In Mollon, J. D., Pokorny, J., & Knoblauch, K. (Eds) *Normal and Defective Colour Vision*, Oxford University Press, 21-30.
- SurrIDGE, A. K., Osorio, D., & Mundy, N. I. (2003). Evolution and selection of trichromatic vision in primates. *Trends in Ecology and Evolution*, 18(4), 198-205.
- Yamada, T., & Watanabe, T. (2007). Virtual Facial Image Synthesis with Facial Color Enhancement and Expression under Emotional Change of Anger. In *16th IEEE International Conference on Robot & Human Interactive Communication*, 49–54.

Appendices

Appendix 1: Questionnaire used in the experiment (front page).

ID number: _____

Questionnaire

Thank you for participating in the experiment.
Please answer the questions below.

age: _____

gender: M _____ F _____

nationality: _____

education: first year student _____ Bachelor _____
Master _____ Ph.D. _____ other _____

handedness: L _____ R _____

vision: normal _____ corrected to normal _____

Appendix 2: Questionnaire used in the experiment (back page).

Do you have any idea what the experiment is about?

Did you notice anything particular with the faces ?

