CAN THE RECOGNITION OF EMOTIONAL EXPRESSIONS BE ENHANCED BY MANIPULATING FACIAL SKIN COLOR?

Monika Żak, Bruno Laeng
Dept. of Psychology, University of Oslo
Oslo, Norway

Joschua T. Simon-Liedtke
Gjøvik University College
Gjøvik, Norway

Trichromacy might have evolved in primates for the purpose of emotion recognition [1]. Emotional states change the quantity of blood and its oxygenation under the skin, which cause subtle changes in skin color to become visible, especially in the face. We manipulated photos of basic emotional expressions so that skin color was either congruent or incongruent (according to [1]) with the expression. Thirty-five participants selected emotional labels for each expression. We predicted that enhanced but congruent colors would improve performance, whereas enhanced but incongruent colors should worsen discriminating emotions correctly. The findings weakly supported the hypotheses. Finally, the Farnsworth-Munsell 100 Hues Test revealed that errors correlated negatively with accuracy in the task, supporting the idea that color ability benefits the recognition of emotions.

Keywords: ***color vision, skin color, facial expression***

I. INTRODUCTION

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 [2-6], 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 [7-8]. Therefore they passed the color vision to their descendants and eventually to us, a specific group of hunter-gatherer hominids who would have supposedly benefited from feeding on fruits and vegetables.

The above, dominant, evolutionary account has been strongly criticized by Mark Changizi and colleagues [1] 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 [1], 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. Other studies indicate that facial expressions may cause drainage of blood in compressed regions of the face, as well as hemoglobin perfusion into other facial areas [9]. Moreover, using thermography, large temperature differences have been found between the nose and forehead during anger or laughter [10].

Changizi, Zhang and Shimojo [1] also revealed connection between bare skin and color vision using photographs of different species of primates. Specifically, they found a positive correlation between the amount of facial skin exposed (i.e., 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 richer colored world tend to have larger bare skin regions in the face. This relationship supports the theory that the purpose of color vision may be found in the ability to perceive color changes and therefore to recognize emotional states. Hence, Changizi et al.'s account provides an appealing and logically convincing, alternative view on the evolution of primates' color abilities. However, it lacks direct empirical evidence showing that seeing colors of the face would actually assist the proper categorization of the emotion expressed in the face.

II. THE PRESENT STUDY

The present study's goal was to investigate whether facial color information affects emotion recognition and, more specifically, if the relationship between emotions and skin colors proposed in [1] can be empirically supported by studying human perception. An experiment was designed where participants viewed faces on a screen and chose – as fast and as accurately as possible – the correct emotional expression by pressing one of 5 keys labeled on the keyboard. The stimuli could be either original or color-enhanced images of facial expressions. We expected that participants' performance would be improved or disrupted by skin colors that were either congruent (e.g., anger = enhanced red) or incongruent (e.g., sad = enhanced red) with the displayed emotional state [1]. We used a standard face database (the Karolinska Directed Emotional Faces), where the recognition of expressions is typically high. We collected accuracy rates and response times

2015 Colour and Visual Computing Symposium (CVCS)
978-1-4799-1765-5/15/$31.00 ©2015 IEEE

(RTs) as well as pupillary data. Specifically, we predicted that accuracy will be higher for congruently enhanced images than for original ones, whereas RTs will be longer for original pictures than for color-enhanced images. In contrast, incongruent skin colors should yield lower accuracy and longer RTs than original images.

III. METHODS

A. Participants

Forty 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). 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. We counterbalanced the labels on the keyboard across every other participant in each group by inverting the sequence HaNeSaAnAf to AfAnSaNeHa. All the participants were tested at the University of Oslo and were compensated with 100 NOK at the end of the session.

B. 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, according to a validation study by Goeleven, De Raedt, Leyman and Verschuere [11]. All stimuli, except for neutral expressions, were edited in *Adobe Photoshop CS5 Extended* to enhance the color of the skin. 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. We first piloted the level of enhancement sot that skin color was manipulated in a subtle way and participants could not notice the change or find the faces odd-looking. Indeed, during debriefing none of the participants claimed to notice anything unusual in the facial skin color. Specifically, we adjusted color levels for the congruent versions (according to [1]) of images to be as 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 1).

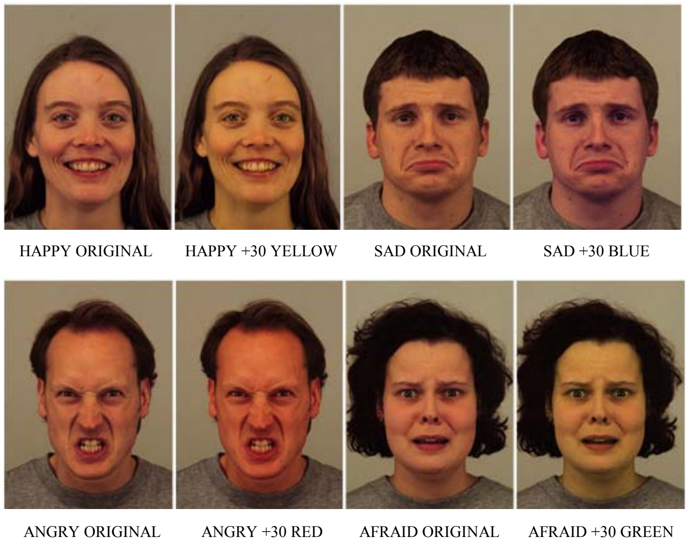

Fig. 1. Examples of original and color-enhanced stimuli (shown side-by-side).

An incongruent version of the same images was also created by modulating color information in the pictures in such a way that color enhancement was always incongruent with the emotion expressed in a facial image and to the same degree: +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.

Finally, two sets of stimuli were used in the experiment: one version 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 the other 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.

C. Apparatus

Participants viewed all stimuli on a Dell P2213 VGA LCD monitor (size = 18.5” or 47 cm) at a distance of 60 cm. The display resolution was set to 1680 x 1050 pixels, and held constant throughout the procedure. The sequence of stimuli and collection of responses were controlled with software Experiment Center 3.2 by SMI* 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™. Stickers representing sketchy faces for each expression were applied over the keyboard buttons in the

*We used an eye-tracker to record pupillary responses and use eye fixations to trigger the onset of the face images; however, we do not report data acquired with the eye-tracker in this report since the analyses failed to reveal differences between the conditions.

following order: happy – G button, neutral – H button, sad – J button, angry – B button, afraid – N button or reversed order as mentioned earlier. All the participants performed the task under typical room illumination kept constant.

D. Procedure

The experiment was conducted in the Cognitive Laboratory at the Department of Psychology, University of Oslo, in a quiet, windowless room. Subjects were instructed to sit comfortably but keeping as stable as possible, by using a chinrest, during the entire duration of the experiment. Detailed task instructions were displayed on the screen at the start of the test. Both congruent and incongruent versions of the experiment consisted of 60 trials in total. The order of the trials was fixed according to a pseudo-random sequence, so that it was unpredictable what will be seen next, and to prevent that the same expression (despite shown by different individuals) happened to be displayed back-to-back. 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 (which triggered the next image via Experiment Center 3.2). Then, an image of a face was displayed for 2000 ms and the participant decided, as quickly as possible, what was the correct label for the emotional expression by pressing the corresponding key. The task was self-paced and the total duration of the experiment was of approximately 15 min.

In the last phase of the experiment, participants were tested with the Farnsworth-Munsell 100 Hue Color Vision Test (FM100), which is a standardized and normed test. It consists of four trays containing a total 85 removable color reference caps spanning the visible spectrum. Participants were asked to order shown color caps in their correct order according to color similarity. The results of the FM100 Hue Test were calculated using the scoring program provided by Béla Török on his web-based platform (www.torok.info/colorvision/fm100.htm).

IV. Results

A. Accuracy

Participants' key presses were extracted using BeGaze™ analysis software provided by SMI. Then, mean accuracy percentage was calculated for each participant for 5 different emotional expressions (neutral, happy, sad, angry, afraid) and the 2 different Manipulation conditions (original, enhaced).

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 also analyzed separately. Therefore, the factor Expression in the above 2-way ANOVA had a total of four levels (afraid, angry, happy, sad).

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 was 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 2), 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 to be non-significant, $F(1, 34) = 1.210$, $p = .278$. 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.

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

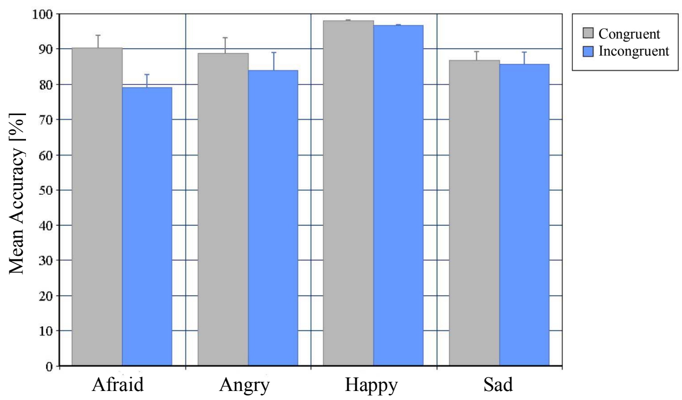

Fig. 2. Mean accuracies and Standard Errors for emotional expressions in congruent and incongruent version of the experiment.

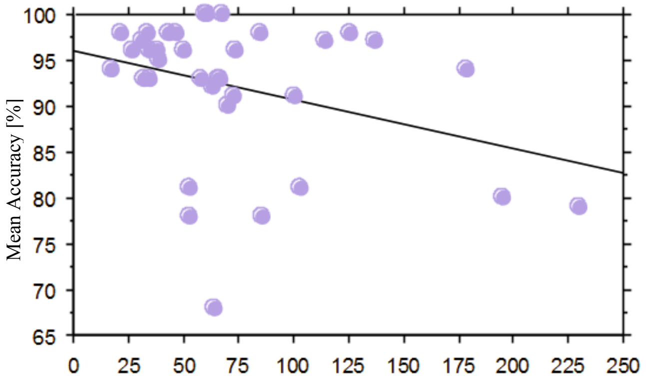

Fig. 3. Regression plot presenting the relationship between Total Error Score (FM100) on the X axis and Mean % Accuracy on the Y axis.

in accuracy, $t(1,34) = -2.052$, $p = .048$ (see Figure 3). 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.

B. *Response Times*

RTs of incorrect answers were excluded from the following analyses as well as outliers (by removing RTs that exceeded 3 standard deviations from the mean). The average values of the remaining response latencies were calculated for each participant and for each of the 5 different emotional expressions (afraid, angry, happy, neutral, sad) for both versions of the experiment (congruent, incongruent).

Two separate two-way ANOVA tests examined the influence of color manipulation and emotional expression on RTs in the congruent and incongruent versions. The analysis on the congruent skin colors revealed no main effect of Manipulation, $F(1, 16) = .682$, $p = .421$. There was 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

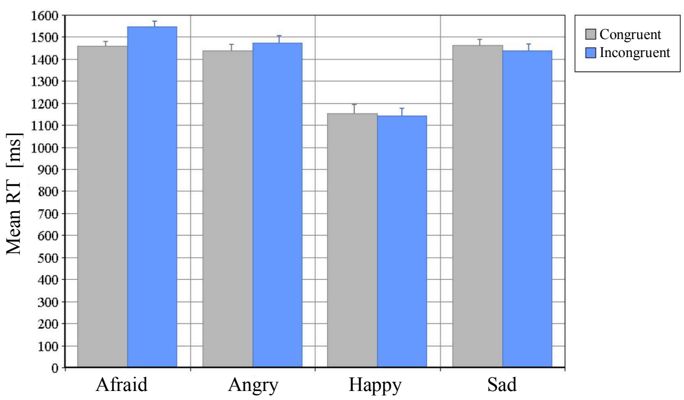

Fig. 4. Mean response times and Standard Errors (bars) for emotional expressions in the congruent and incongruent conditions.

t-tests revealed that participants responded significantly faster to happy facial images than to other emotional expressions (see Figure 4).

The ANOVA for incongruent version showed a main effect of Expression, $F(3, 15) = 21.18$, $p < .0001$, but no main effect of Manipulation, $F(1, 17) = 0.275$, $p = .549$. Interestingly, the interaction of Manipulation*Expression approached significance, $F(3, 15) = 2.437$, $p = .077$. As shown in Figure 5, RTs for the afraid and angry expressions tended to be slower in the incongruent version than in the congruent version. 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 non-significant, $F(1, 34) = 2.011$, $p = .166$. Thus, we assume that the 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.

V. Discussion

The purpose of this study was to investigate whether making facial skin color more salient could modulate emotion recognition, as one would predict on the basis of Changizi and colleagues' account [1]. In the experiment, we manipulated photos by enhancing the skin color on the face of individuals presenting 4 different but basic emotional expressions. The images were used in a task where participants had to identify the correct label for each emotional expression. In the incongruent version of the experiment, colors were enhanced incongruently (according to [1]).

The results revealed no significant difference in accuracy between original and enhanced trials in the congruent version of the experiment. Therefore, the results failed to support the notion that a more salient but correct facial skin color provides information about emotional states of other persons. 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. However, other experiments using KDEF images also found no difference in accuracy for different emotional expressions [12].

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 other persons' emotional states. Therefore, we hypothesized that fewer errors in the color test should correlate with better ability to recognize the 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, in the incongruent version of the task, the more negative and difficult to decode expressions, i.e. afraid and angry, tended to be slower than in the congruent version, which seems consistent with the idea that subjects' performance should worsen when exposed to incongruent images.

VI. Limitations and Recommendations for Future Studies

There are several 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 may be to use more subtle facial expressions in order to increase task difficulty.

Another important limitation of our experiment is that the stimuli presented fake emotional expressions played by actors. As professionals, these actors could convincingly pretend to be happy and fake a smile, but we note that pupil size cannot be consciously controlled by most people [13] 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 [14]. 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 some relevant information (i.e. pupil size).

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) might be more powerful than the effect of color manipulation per se. 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 not provide a sufficient manipulation for decreasing accuracy or RTs to a visible 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 warrants some caution.

Lastly, it is possible that data were collected with an insufficient sample size to achieve a respectable level of statistical power.

VII. 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 other individuals can either improve or disrupt emotion recognition. Although our hypotheses were not explicitly confirmed, the results were somewhat encouraging since we observed suggestive differences in participants' performance between the 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. 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. RTs also indicated a possible disruptive effect of incongruent coloring of the skin of the face. By improving the present experimental design, we might gain in a future study a better understanding of the role of trichromatic color vision in the recognition of emotional expressions.

References

[1] M.A. Changizi, Q. Zhang, and S. Shimojo, “Bare skin, blood and the evolution of primate colour vision,” Biology Letters, vol. 2, pp. 217-221, 2006.

[2] P.M. Kainz, J. Neitz, and M. Neitz, “Recent evolution of uniform trichromacy in a New World monkey,” Vision Research, vol. 38, pp. 3315-3320, 1998.

[3] F. Pichaud, A. Briscoe, and C. Desplan, “Evolution of color vision,” Current Opinion in Neurobiology, vol. 9, pp. 622-7, 1999.

[4] P. Lucas et al., “Evolution and function of routine trichromatic vision in primates.” Evolution, 57(11), pp. 2636-2643, 2003.

[5] A.K., Surridge, D. Osorio, and N.I. Mundy, “Evolution and selection of trichromatic vision in primates,” Trends in Ecology and Evolution, 18(4), pp. 198-205, 2003.

[6] D. Osorio and M. Vorobyev, “A review of the evolution of animal colour vision and visual communication signals,” Vision Research, vol. 48, pp. 2042-2051, 2008.

[7] D. Osorio and M. Vorobyev, “Colour vision as an adaptation to frugivory in primates,” Proceedings of the Royal Society of London B, 263, pp. 593-99, 1996.

[8] P. Sumner and J.D. Mollon, “Did primate trichromacy evolve for frugivory or folivory?,” in Normal and Defective Colour Vision, J.D. Mollon, J. Pokorny, and K. Knoblauch, Eds. Oxford University Press, 2003, pp. 21-30.

[9] Jimnez et al., “A practical appearance model for dynamic facial color,” ACM Transactions on Graphics, 29(6), pp. 1-10, 2010.

[10] T. Yamada and T. Watanabe, “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, 2007, pp. 49–54.

[11] E. Goeleven, R. De Raedt, L. Leyman, and B. Verschuere, “The Karolinska Directed Emotional Faces: A validation study,” Cognition & Emotion, 22(6), pp. 1094-1118, 2008.

[12] E. Sucksmith, C. Allison, S. Baron-Cohen, B. Chakrabarti, and R.A. Hoekstra, “Empathy and emotion recognition in people with autism, first-degree relatives, and controls,” Neuropsychologia , 51(1), pp. 98–105, 2013.

[13] B. Laeng, S. Sirois, and G. Gredebäck, “Pupillometry: A window to the preconscious?,” Perspectives on Psychological Science, 7(1), pp. 18-27, 2012.

[14] K.E. Demos, W.M. Kelley, S.L. Ryan, F.C. Davis, and P.J. Whalen, “Human Amygdala Sensitivity to the Pupil Size of Others,” Cerebral Cortex, vol. 18, pp. 2729-2734, 2008.