NEURAL AND COGNITIVE UNDERPINNINGS OF TRUSTWORTHINESS IN THE PREFRONTAL CORTEX VIA OPTICAL BRAIN IMAGING

ALİ TAYLAN BİLİK

BOĞAZİÇİ UNIVERSITY

2019

NEURAL AND COGNITIVE UNDERPINNINGS OF TRUSTWORTHINESS IN THE PREFRONTAL CORTEX VIA OPTICAL BRAIN IMAGING

Thesis submitted to the

Institute for Graduate Studies in Social Sciences

in partial fulfillment of the requirements for the degree of

Master of Arts

in

Cognitive Science

by

Ali Taylan Bilik

Boğaziçi University

2019

Neural and Cognitive Underpinnings of Trustworthiness

in the Prefrontal Cortex via Optical Brain Imaging

This thesis of Ali Taylan Bilik

has been approved by:

Assist. Prof. İnci Ayhan
(Thesis Advisor)

Assist. Prof. Tuna Çakar
(Thesis Co-Advisor)

Prof. Albert Ali Salah

Assist. Prof. Güneş Ünal

Assoc. Prof. Şefik Şuayb Arslan
(External Member)

July 2019

DECLARATION OF ORIGINALITY

I, Ali Taylan Bilik, certify that

- I am the sole author of this thesis and that I have fully acknowledged and documented in my thesis all sources of ideas and words, including digital resources, which have been produced or published by another person or institution;
- this thesis contains no material that has been submitted or accepted for a degree or diploma in any other educational institution;
- this is a true copy of the thesis approved by my advisor and thesis committee at Boğaziçi University, including final revisions required by them.

Signature 

Date 20.08.2019

ABSTRACT

Neural and Cognitive Underpinnings of Trustworthiness

in the Prefrontal Cortex via Optical Brain Imaging

Human face has been considered to be an indicator of various attributes including attractiveness, calmness, and trustworthiness. Even a single look for a short duration might have an influence on one's attitude. This empirical study aimed to understand the neural correlates of trustworthiness judgments on face photos via an optic brain imaging (fNIRS) method. The experimental design has been developed to categorize the presented photos after watching them for 4000 milliseconds. The responses were collected within a 3-point Likert scale. The obtained findings indicate that there is a significant level of neural activation difference between positive and negative responses among a specific optode (Optode-3). Moreover, there are also significant levels of neural activation differences for the neutral responses versus others among 3 optodes (Optodes 12, 14, and 16). The results were also controlled regarding the saliency levels of the photos as well as the emotional states of the individuals in the photos via facial activation coding analysis. Overall results imply that trustworthiness level might be understood at a certain level.

ÖZET

Bireylerin Güven Temelli Karar Verme Süreçlerinin Beyinsel İzdüşümlerinin Optik Beyin Görüntüleme Sistemi Kullanılarak İncelenmesi

İnsan yüzü, çekicilik, sakinlik ve güvenilirlik gibi çeşitli özelliklerin bir göstergesi olarak kabul edilmiştir. Kısa süreliğine tek bir bakış bile birinin tutumunu etkileyebilir. Bu deneysel çalışma, optik beyin görüntüleme (fNIRS) yöntemi ile yüz fotoğraflarındaki güvenilir yargıların sinirsel ilişkilerini anlamayı amaçlamaktadır. Deneysel tasarım, sunulan fotoğrafları 4000 milisaniye boyunca izledikten sonra kategorilere ayırmak için geliştirilmiştir. Katılımcıların yanıtları 3 puanlık Likert ölçeğinde toplandı. Elde edilen bulgular, belirli bir optod (O-3) arasında pozitif ve negatif cevaplar arasında anlamlı düzeyde nöral aktivasyon farkı olduğunu göstermektedir. Ayrıca, birkaç farklı optod arasında diğerlerine karşı nötr tepkiler için üç optodda (12, 14 ve 16) istatistiksel olarak anlamlı seviyede nöral aktivasyon farklılıkları saptanmıştır. Sonuçlar aynı zamanda yüz aktivasyon kodlama analizi ile fotoğrafların belirginlik seviyelerinin yanı sıra fotoğraftaki bireylerin duygusal durumlarına göre de kontrol edilmiştir. Sonuç olarak elde edilen veriler, güvenilirlik seviyesinin belirli bir düzeyde anlaşılabileceğini göstermektedir.

ACKNOWLEDGEMENTS

My initial thesis advisor Prof. Albert Ali Salah was incredible support throughout the years, when I was trying to find enough time focus on my research. I gratefully acknowledge that his encouragement made the initiation of this research possible. Their friendship, as well as guidance, was priceless.

I would like to express my appreciation to my Thesis Advisor Ass. Prof. İnci Ayhan and Co-advisor Ass. Prof. Tuna Çakar for accepting and guiding me. Their advises and support are invaluable.

I would also like to thank ThinkNeuro Research Company for allowing me to use their fNIRS device and their warm welcome.

Finally, I wish to thank for the patience and warm support of my family, friends and colleagues who supported me during the times when I lost my hope.

TABLE OF CONTENTS

CHAPTER 1: INTRODUCTION AND THEORETICAL BACKGROUND……..... 1

1.1 Evaluating trustworthiness through faces ……..……. 3

1.2 Neurobiology of trust ……..…...…… 4

1.3 Neural correlates …….……...……..…….. 6

1.4 Face database ……...…….……. 7

1.5 Research scope, questions and hypotheses …….……..…….. 8

CHAPTER 2: METHODOLOGY ……..……..……. 11

2.1 Optic brain imaging …….……..…… 12

2.2 fNIRS device ……...……...……..……...…….…….. 13

CHAPTER 3: EXPERIMENTATION ……..……..……. 15

3.1 Experimental design …….……...…..…… 15

3.2 Participants ……..……...…….. 15

3.3 Procedure and stimuli ….……..…..…….. 16

3.4 Data analysis …….……...……. 16

CHAPTER 4: RESULTS ……..…….……..…….…… 18

4.1 Behavioural findings …….……..……...…….. 18

4.2 Reaction time …….……..……..…..……. 19

4.3 Neural activations …….……..……..…… 21

4.4 Saliency analysis …….……..……..……. 24

4.5 Emotional analysis via deep learning toolbox …….……...…. 26

CHAPTER 5: DISCUSSION …….……...…. 27

5.1 Correlation between masculinity and trustworthiness …….…… 27

5.2 Reaction time difference between positive and negative responses 28

5.3 Neural activation pattern differences in prefrontal cortex ……..… 28

5.4 No significant difference observed related to brain reward system 29

5.5 Limitations due to experimental design and methodology ……... 30

CHAPTER 6: CONCLUSION …….……..…… 31

APPENDIX A: FACE IMAGES…….…….. 32

APPENDIX B: CONSENT FORM…….…….…….. 39

REFERENCES ….…….……..…….. 40

LIST OF FIGURES

Figure 1. Chromophore within 700-900 wavelength range and having the highest light absorption characteristic used in the optical brain imaging is hemoglobin molecule…….…….…….....13

Figure 2. fNIR sensor placed on the forehead (upper left-hand side figure), approximate measuring points in the brain surface image (upper right-hand side figure), fNIR sensor is able to measure 16 different brain areas by the help of 4 light sources and 10 photodetectors (lower figure)……..…….… 14

Figure 3. Flowchart of the experimental procedure……..…….…… 16

Figure 4. The distribution of different responses of trustworthy judgments for each of the face photos (1: negative, 2: neutral, 3: positive)…….. 19

Figure 5. The mean reaction times related to responses (1: positive, 2: neutral, 3: negative) with the error bars of 95% confidence intervals……..…...….... 20

Figure 6. The mean reaction times for each of the participant…….…….…… 21

Figure 7. High versus low masculinity levels related to HBR levels for each optode (for positive and negative responses)…….…… 22

Figure 8. The response-wise HBR levels for each optode (for positive and negative responses)……..……..…..… 23

Figure 9. Sample saliency maps for the masculine (right hand side) and non-masculine (left hand side) photos.……...…….. 25

Figure 10. Sample attended location maps for the masculine (left hand side) and non-masculine (right hand side) face photos…….…... 26

CHAPTER 1

INTRODUCTION AND THEORETICAL BACKGROUND

Trust is a term, which people use frequently in their daily lives, whereas it is also a topic, which has been widely evaluated and examined by social sciences. People often refer to another person or an entity when they use the word “trust” and they tend to mean a combined feature of “trustee” which is related with honesty, integrity, accountability, fairness and predictability. For interpersonal relationships, trust might be accepted to have a foundational role, since it seems to convey a huge societal relevance. Most of our decisions, including buying decisions, are affected by trust that we feel towards the person with whom we have been communicating (Zak & Knack, 2001; Dimoka et al., 2010). Moreover, people with chronic problems related to trust might have some psychological or neurological disorders that can affect them in their social relations. Given its survival value, even brief durations of exposure to faces can trigger trustworthiness judgments. Here, we have been mainly interested in this decision-making process of judging a person as trustworthy or not.

Different disciplines might provide different conceptualizations for trust, for instance, trust is perceived and accepted as a measure of belief in honesty and fairness in sociology and psychology, whereas it appears as a term related to reliability in transactions in economics (Zak & Knack, 2001). Trust could be conceptualized as a complicated emotional state to provide a way of coping with the complexities and uncertainties of the external world. In this sense, it has also been a significant concept in relation to our attribution to other people. In an experimental design of buying items from sellers, two dimensions were emphasized as the social constructs of trust: credibility and benevolence (Dimoka et al., 2010). In this study,

Dimoka et al. aimed at understanding the neural correlates of evaluating the process of credibility while looking at a person's face photo. On the other hand, there have been other behavioral and neuroscientific studies that involved the facial features such as beauty and masculinity with respect to trustworthy judgments.

For practical purposes, understanding the potential differences in neural activations is important for further studies that might be adapted to different contexts like trust/distrust in political campaigns or any other domain. It should also be mentioned that our attributions related to trust are influenced by several different factors including facial features, social status, and behavioral tendencies. This thesis has the main objective to understand the implicit judgments related to trust (as trustworthiness judgments) regarding the facial appearance of people. There is a considerable literature on the effect of facial features on the perception of trustworthiness (Dimoka, 2010; Engell et al., 2007; Todorov et al., 2008; Winston et al., 2002). The second interest of this thesis to understand the neural correlates of trustworthy/untrustworthy judgments, particularly in the prefrontal cortex (PFC), via an optical brain imaging method within the context of existing literature (Winston et al., 2002; Todorov & Engell, 2008; Engell et al., 2007).

This section includes the general framework in relation to the literature survey with the following subsections: (1) evaluation of trustworthiness through faces, (2) neurobiology of trust and distrust, (3) neural correlates of trustworthy judgments, (4) the Face Database, (5) research scope, questions and hypotheses. This theoretical background is provided to explain the logic behind the proposed hypotheses which are provided at the end of this section.

1.1 Evaluating trustworthiness through faces

Faces are processed over a wide range of dimensions that could be used to predict various outcomes including the voting behavior (Ulemanet al., 2005; Todorov et al., 2005; Ballew & Todorov, 2007) or sentencing decisions (Blair et al., 2004; Eberhardt et al., 2006). Even brief exposures as short as 100 milliseconds can affect one's judgments about the characteristics of a person such as competence and aggressiveness, as well as trustworthiness (Willis & Todorov, 2006). In fact, there are empirical studies which suggest that exposure durations as short as 35-100 milliseconds can be sufficient for the facial processing of some features, particularly linked to traits that have high survival value such as threat (Bar et al., 2006; Willis & Todorov, 2006). While interpreting these brief exposure effects, however, it is important to note that different characteristics, upon which judgments can be made are highly correlated with each other. These correlated judgments generally tend to direct a positive or negative valence (Rosenberg et al., 1968). This type of overall evaluation has been used to produce the outcome like approach-avoidance behavior.

Trustworthiness is a commonly studied trait in humans. Impressions on trustworthiness are often formed rapidly and spontaneously from minimal information (Todorov & Uleman, 2003; Uleman, Blader, & Todorov, 2005). One main source of such information is the facial appearance. In fact, there is abundant research about the effects of facial appearance on social outcomes (Blair, Judd, & Chapleau, 2004; Eberhardt, Davies, Purdie-Vaughns, & Johnson, 2006; Hamermesh & Biddle, 1994; Hassin & Trope, 2000; Langlois et al., 2000; Montepare & Zebrowitz, 1998; Zebrowitz, 1997). Whereas inferences on competence, based solely on facial appearance, predict the outcomes of the U.S. congressional (Todorov, Mandisodza, Goren & Hall, 2005) and gubernatorial elections (Ballew & Todorov,

2007; Hall, Goren, Chaiken, & Todorov, 2009), for example, inferences on dominance predict military rank attainment (Mazur, Mazur, & Keating, 1984; Mueller & Mazur, 1996). Previous studies have also shown that trustworthiness judgments from facial appearance approximate the general valence evaluation of faces (Oosterhof & Todorov, 2008) and are made even after as little as 100 milliseconds of exposure to novel faces (Willis & Todorov, 2006). Thus, it might be concluded that even brief durations of exposure to faces initiate certain neural activations and behavioral tendencies.

1.2 Neurobiology of trust

Integrating the neural correlates of trust and distrust shows a stark contrast in the activated brain areas. Using functional neuroimaging data, a study by Dimoka (2010) showed that trust and distrust may be distinct constructs that span over different brain areas. Whereas trust is associated with the brain areas that are linked to reward, prediction, and uncertainty, distrust is rather associated with the areas responsible for intense emotions and fear for loss (Dimoka et al., 2010; Riedl & Javor, 2012). Distrust is associated with the amygdala, a critical area triggered by intense and sudden emotional states, particularly negative ones (LeDoux, 2003), and the insular cortex (Dimoka et al., 2010). In Dimoka's et al. study, in a context of IT-related impersonal item buying setting, the theoretical distinction between trust and distrust was made based upon the two main dimensions that are the credibility and malevolence. The results indicated that the neural correlates of trust include the brain's reward system, and areas associated with prediction, and uncertainty including orbitofrontal cortex, caudate nucleus, anterior PCC, and putamen (Dimoka et al., 2010).

The asymmetrical bias of distrust, demonstrating a bigger effect size than trust can be partly explained by the nature of distrust, which seems to have more emotional valence than trust (McKnight, Kacmar, & Choudhury, 2004); thus, they are more salient and accessible (Luhmann, 1979). From an evolutionary perspective, this could be due to the potential activation of the alarm systems that would protect the organism from the potential threats. While a variety of stimuli are used in the trustworthiness literature, including occasions/situations, as well as inspecting faces (facial features), we have been especially interested in the implicit processing of trust, thus, used faces as experimental stimuli set in this thesis.

Another study, using the *f*MRI methodology to question the trustworthiness evaluations of online-offers considered user personality traits such as impulsiveness (Hubert et al., 2018). In this study, participants were asked to evaluate the trustworthiness of the online-offers while their brain activations were recorded via *f*MRI. The results showed that the brain activation patterns for different personality traits such as high (hedonic) versus low (prudent) degrees of impulsiveness were mostly similar. However, there were some dissimilarities in terms of magnitude. Participants with high level of impulsiveness, for example, tended to have higher level of neural activations within the brain regions known to be significant for trust and impulsiveness such as dorsal striatum, anterior cingulate, the dorsolateral prefrontal cortex, and the insular cortex. Thus, in Hubert et al. study, whereas the activated brain regions were mostly the same for both groups, the magnitudes of brain activations were shown to differ to a certain extent.

1.3 Neural correlates

Findings in the relevant literature indicate that several different factors are influential on the decisions related to trust. For instance, the cause of uncertainty is one of the key factors for the decision-making process. It has been shown that different brain/neural regions are activated depending on whether the uncertainty is caused by another agent or by nature (Aimone et al., 2014). In an fMRI study, Aimone et al. showed the role of betrayal avoidance in trusting decisions, where they defined the betrayal avoidance as "a desire to avoid negative emotions that arise from learning one's trust was betrayed". They demonstrated that trusting decisions in the presence of betrayal potential were mostly modulated by the activity in the anterior insula, distinctively from those made under uncertain environments with a random nature of chance (Aimone et al., 2014). The anterior insula region has also been known to take a role during the emotional processing of socially invoked external information including disgust (Aimone et al., 2014).

Behavioral findings that show trustworthiness judgments reflect the emotional valence of faces are consistent with studies suggesting that amygdala, a subcortical brain structure critical for fear conditioning (LeDoux, 2000; Phelps & LeDoux, 2005) and detection of threats in the environment (Amaral, 2002; Davis & Whalen, 2001; Whalen, 1998), plays a key role in such judgments (Todorov & Engell, 2008). For example, Adolphs, Tranel, and Damasio (1998) showed that patients with bilateral amygdala damage are impaired in their discrimination of trustworthy- and untrustworthy-looking faces. Consistent with the Adolphs et al. (1998) findings, functional neuroimaging studies showed that untrustworthy-looking faces evoke stronger activity in the amygdala than trustworthy-looking faces (Engell, Haxby, & Todorov, 2007; Todorov, Baron, & Oosterhof, 2008; Winston, Strange,

O'Doherty, & Dolan, 2002). Thus, convergent evidence from behavioral and functional neuroimaging studies suggest that judgments of trustworthiness from faces (a) approximate valence evaluation of faces; (b) are rapidly and spontaneously computed; and (c) have an identifiable neural signature.

As opposed to fMRI, some neuroimaging systems such as optical brain imaging (fNIRS) does not allow to monitor the neural activity across the midbrain. Thus, fNIRS would only allow us to study the neural underpinnings of trust/distrust across the human prefrontal cortex. As well as the overall activity, cortical hemispheric asymmetry has also been reported to be one of the indicators of emotional reactivity and temperament (Davidson, 1995).

Our judgments on trustworthiness are affected by several features including the physical appearance and specifically, facial attractiveness (that is referred to as beauty). In other words, people tend to perceive physically attractive people as more competent, intelligent, and trustworthy. (Todorov et al., 2015). These findings have been replicated by studies using computer-generated faces, as well as the natural photographs (Todorov et al., 2015). The level of facial masculinity, on the other hand, has been shown to have a negative impact on the perceived trustworthiness (Todorov et al., 2015). Thus, here, one of our aims is to investigate the neural correlates and cognitive underpinnings of the relationship between trustworthiness and masculinity.

1.4 Face database

There are more than ten thousand face photographs in the 10k US Adults Faces Database and 2,222 of these faces have different measures including psychological attributes like happiness, anger, sadness and memorability scores obtained from

several participants. The photograph files have 72 pixels/in resolution and 256-pixel height. There are plenty of academic studies on this database including intrinsic memorability of face images (Bainbridge et al., 2013) and how this memorability can be exposed to change (Khosla et al., 2013). This database was based on a scientific research that includes high resolution and standardized photographs of people from different ages, gender, and ethnicities. For each individual model, extensive level of data was collected to provide a reliable normalization (Bainbridge et al., 2013). This normalization covers subjective ratings for physical attributes as well as subjective ratings for judgments like attractiveness. In this database, photos have ratings on several different dimensions such as typical, boring, calm, catch, cold, common, confident, egoistic, intelligent, introverted, kind, responsible, weird, aggressive, attractive, emotional, happy, humble, interesting, mean, sociable, familiar, and the negated forms of these adjectives (Bainbridge et al., 2013). This face database consists of more than thirty thousand of face photos that are collected and distributed for scientific research. The main database includes face photos of 2.222 people from various origins.

1.5 Research scope, questions and hypotheses

To understand the potential effect of any factor, it is important to match different groups with respect to their other physical attributes. For instance, two groups with diverging masculinity features/scores should be highly similar in terms of their physical attractiveness/beauty etc. to have a reliable comparison. Faces with higher level of perceived masculinity would be rated as lower in terms of trustworthiness if the other conditions were kept the same. Thus, the face photos from the database have been grouped as “trustworthy” and “untrustworthy” faces.

The research scope of this thesis includes the inspection of trustworthy/untrustworthy judgments mainly related to the masculinity features of the presented faces. It is also of our interest to investigate the neural activations related to these trustworthiness judgments. Thus, this thesis work aims to investigate the neural correlates of the perceived trustworthiness with fNIRS system targeting the prefrontal cortex (PFC). The findings from the related academic literature show that medial (central) prefrontal cortex is activated with the onset of human brain reward system. The related research questions are as follows: (1) Is there any difference related to response times of the positive and negative judgments for the subjects following the presentation of the face photos? (2) Do these judgments cause different neural activation patterns across the prefrontal cortex? The studies in the literature have already shown that trust is correlated with the brain regions in the brain reward system. Thus, it might be argued that the face photos rated as high trustworthy could cause a higher activation across the medial prefrontal cortex (mPFC) that has been suggested to be the end region of the brain reward system (Tzschentke, 2000). Thus, different neural activation patterns might be used for predictive purposes in the later phases.

H_1: The positive responses (trustworthy judgments) will be responded quicker than the negative responses (untrustworthy judgments) since negative processing possibly requires higher level of mental effort. It may be that the negative judgments for trustworthiness are more difficult to process than the positive ones. This is related to the delayed processes of negated items.

As an auxiliary assumption: The level of perceived masculinity is negatively correlated with the level of perceived trustworthiness.

H_2: There will be higher level of medial prefrontal activation during the positive responses (trustworthy judgments) as opposed to the negative responses (untrustworthy judgments).

H_3: It is also expected that the hemispheric lateralization (asymmetry) among the prefrontal cortex (PFC). This asymmetry will be monitored as the neural activity difference in left versus right prefrontal cortex (PFC). Also, there will be higher level of lateral prefrontal activation for the negative responses when compared to the positive responses. Because it is expected that the negative responses will need higher level of mental effort (more challenging) during the processing interval.

CHAPTER 2

METHODOLOGY

The aim of this thesis work is to understand the neural underpinnings of trustworthiness judgments from facial appearance via optic brain imaging (fNIRS). Due to the technical limitations of the fNIRS technology, which does not allow us to study subcortical structures, the focus will be mainly on the neural activations in the prefrontal cortex. The empirical evidence in the relevant literature indicates that the evaluation of the briefly presented faces is in line with the trustworthiness judgments regarding the facial appearance (Oosterhof & Todorov, 2008). The neuroscientific studies, on the other hand, indicate that there is a significant distinction in the activated brain areas for trust versus distrust judgments (Dimoka, 2010). Interestingly, the neural correlates of trust have been associated with several brain functions including reward processing, prediction, and uncertainty. Whereas, distrust has been indicated with the neurophysiological findings related to intense emotions and fear for loss. The neural patterns of trust and distrust being different suggest that these may be distinct constructs (LeDoux, 2003; McKnight, Kacmar, & Choudhury, 2004; Luhmann, 1979). Moreover, neuroimaging findings indicate that the level of neural activation in the amygdala is higher for the untrustworthy faces than for the trustworthy faces that is probably related to the valence of the experienced emotion (Engell, Haxby, & Todorov, 2007; Todorov, Baron, & Oosterhof, 2008; Winston, Strange, O'Doherty, & Dolan, 2002).

2.1 Optic brain imaging

In this study, the neuroscientific methodology was the optic brain imaging that is based on near-infrared technology. Jobsis (1977) showed for the first time that brain functions could be monitored by near-infrared spectroscopy method. Later in the 1990s, by indicating that this technology would be a safe, noninvasive and more affordable alternative compared to the other alternatives to be used for monitoring human brain functions, its use became widespread (Villringer et al., 1993; Strangman et al., 2002; Chance et al., 1993; Villringer & Chance, 1997). In the 2000s, this method gained even a wider use in studies testing children, as well as adults thanks to the advancements in both hardware and software technologies (Obrig et al., 2000; Ayaz et al., 2012).

The working principle of the functional near-infrared spectroscopy (fNIRS) is based on the absorption characteristics of infrared light inside the tissue. As illustrated in Figure 1, infrared rays between 700 and 900 nm wavelengths are mostly absorbed by hemoglobin molecules in the red blood cells among other molecules in a tissue. It is possible to monitor changes in the hemoglobin quantity by fNIRS method. By the inspection of red blood cells (hemoglobin), photons are partly absorbed and partly reflected in the reaction between the light and the tissue. The absorption level of the light directly varies with the oxy-hemoglobin quantity. The oxygen, required for metabolizing the glucose during the generation of energy necessary for the functioning of nerve cells, is carried to the cells through the hemoglobin molecules in the blood. By the higher rate of functioning of human brain, oxygen demand and therefore the supply of clean blood increases. As the number of oxy-hemoglobin molecules is increased depending on the increase in blood flow rate and volume, more infrared rays are absorbed. Thus, fNIR enables the

quantitative monitoring of changes in the brain functions by sending a signal proportional to the functioning of nerve cells.

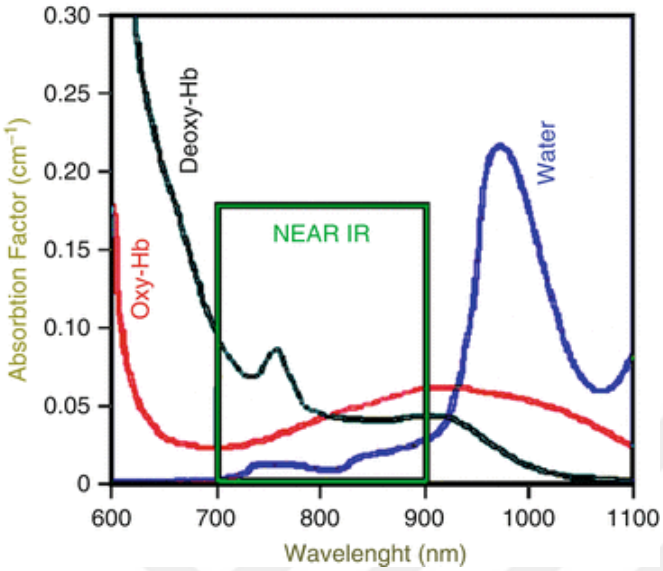


Fig. 1 Chromophore within 700-900 wavelength range and having the highest light absorption characteristic used in the optical brain imaging is hemoglobin molecule.

2.2 fNIRS device

fNIRS system used in this study was first described by Chance (1993) and then it was developed in the Drexel University (Philadelphia, PA) and manufactured by the fNIR Devices, LLC (Potomac, MD; www.fnirdevices.com). As seen in Figure 2, by the aid of fNIRS sensors placed on the forehead, it is possible to obtain data having 2 Hz frequency from 16 different areas in the prefrontal cortex region. The spatial resolution of these measuring points named as voxels or optodes is at the centimeter scale, with a depth sensitivity of approximately 1.5 cm. When Brodmann map for the left-brain cortex layer is taken into consideration, these areas correspond to the brain regions numbered as 9, 10, 46 and 45.

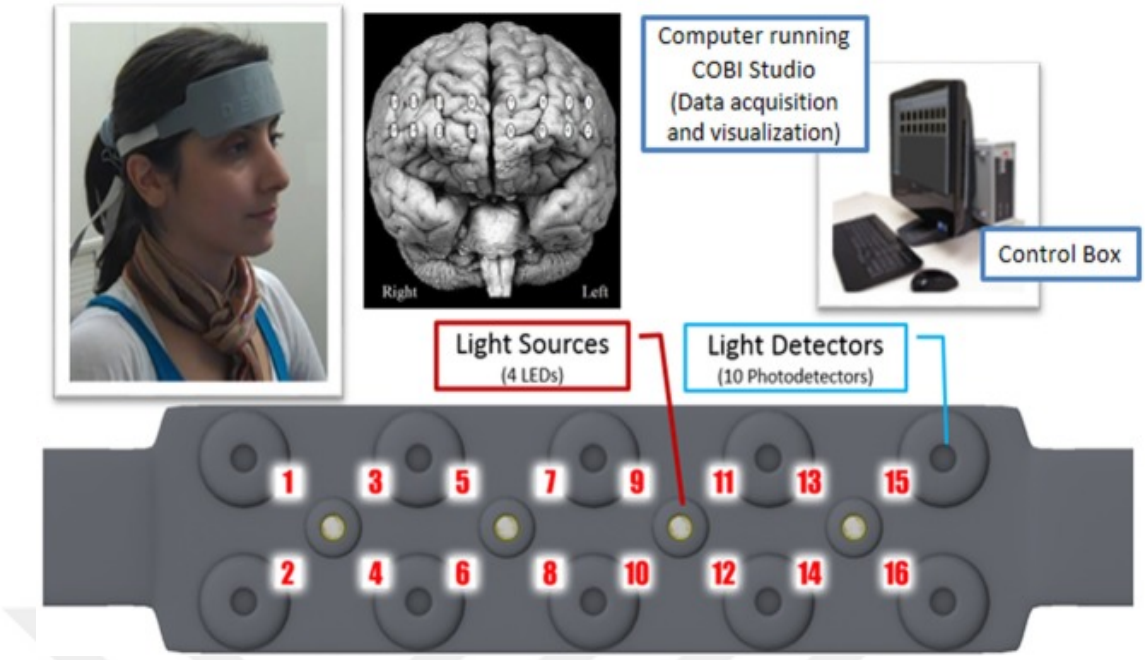


Fig. 2 fNIR sensor placed on the forehead (upper left-hand side figure), approximate measuring points in the brain surface image (upper right-hand side figure), fNIR sensor is able to measure 16 different brain areas by the help of four light sources and 10 photodetectors (lower figure)

Compared to the functional magnetic resonance imaging (fMRI) method that enables the monitoring of hemodynamic actions of the brain in all depths and at a millimeter scale, fNIRS offers a relatively lower spatial resolution. However, as it better tolerates the head movements occurring during the measurement taking, it offers a higher temporal resolution. fNIRS is also portable, and therefore can be used on a desk, even in the markets, thus, fNIRS method provides significant advantages in comparison to the fMRI methodology.

CHAPTER 3

EXPERIMENTATION

3.1 Experimental design

In the experiment, participants were shown a series of photographs with male faces for an interval of four seconds. During each experimental block, participants were then asked to rate whether the presented person is trustworthy, untrustworthy or neutral on a 3-point Likert scale (forced-choice task). Following the decision screen, presented for three seconds (fixed duration), a fixation spot appeared on the screen for eight seconds as the resting phase. There were a total of 62 pictures (Appendix A). Hereby, each trial consisted of three phases (visual image presentation, decision screen, and fixation screen). Thus, the total duration for the experiment was about 922 seconds (15 minutes 22 seconds). With the preparation and the instruction periods, the total duration of the experiment for each participant was estimated to be around 25 minutes.

3.2 Participants

30 participants in the age range of 27-42 years old have participated in the experiments. The mean age was 33.31 years with a standard deviation of 5.08. All participants were male in order to avoid potential gender biases. They were selected from the participant database of a market research company. Participants were paid 35 TL for their participation. They were selected based on the criteria that they should not have taken any psychiatric medication specifically in the last six months. They also confirmed that they had not participated in a similar experiment before. Before the experiment, all participants signed the consent form (Appendix B).

3.3 Procedure and stimuli

First, participants were informed about the content and details of the experiment. They were then asked to sign consent forms if they were to declare that their participation is voluntary. In the experimental phase, they were presented with 62 face photos. They would see each face photo for 4,000 milliseconds and were expected to respond whether they found the person as trustworthy or not. The choices were tertiary (trustworthy, untrustworthy and neutral) and forced-choice. They were then given 3000 milliseconds to provide their response. The missed responses were not included in the data analyses. The response screen would be followed by fixation of eight seconds. The flowchart for each block/trial is shown below.

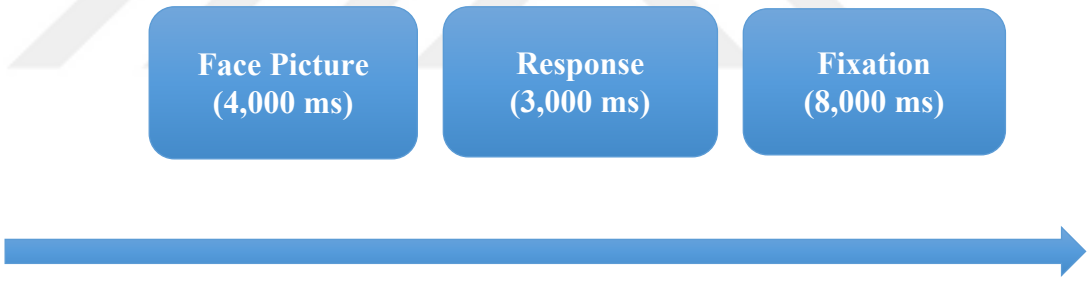


Fig. 3 Flowchart of the experimental procedure

3.4 Data analysis

3.4.1 Preprocessing step

The tertiary decisions, reaction times, and the raw fNIRS measures were obtained for each participant. None of the participants was outlier thus no exclusion was provided before the data analyses regarding the standardization (*z* scores). A low-pass filter (a finite impulse response, linear phase filter with order 20 and cut-off frequency of 0.1

Hz) were applied on the raw fNIRS data (16 optodes × 2 wavelengths) to attenuate the high-frequency noise because of the respiration and cardiac cycle effects (Ayaz et al., 2012). There was an analog-to-digital converter limit for excluding the saturated channels with respect to the light intensity at the detector. Finally, a motion artifact filter was applied to the preprocessed dataset to remove the parts with excessive motion artifact (Ayaz, 2011). For the extraction of the final dataset for processing, time synchronization markers were used to obtain continuous data during the task period as well as the resting stages (with fixation points). Finally, for each period, fnirSoft Software was used to calculate the modified Beer-Lambert Law (Ayaz, 2010).

3.4.2 Calculations for each block (face photo)

Neural activations during the image presentation and decision blocks were calculated with respect to the level of oxygenated and deoxygenated hemoglobin levels. In line with the literature, the last four seconds of the previous fixation durations were taken as the baseline. Thus, the accumulative effect of the biosignals was minimized. The oxygenation and deoxygenation levels were then averaged over the blocks. Thus, in the end, a mean value was obtained for each block. For each optode, these mean values were compared via an independent samples t-test regarding the responses of the participants. The fNIRS data of tow participants had problems. Thus, two participants' data were excluded from the analyses and the final results included the neuroimaging data of 28 (of 30 in total) participants.

CHAPTER 4

RESULTS

The experiment consisted of 3 phases: (1) demonstration of a face photo for four seconds, (2) decision screen for three seconds, and (3) fixation screen for eight seconds. Each trial had a total duration of 15 seconds. The results are presented in three subsections: behavioral findings, reaction time, neural activations.

4.1 Behavioral findings

Behavioral findings were composed of the responses of participants. Since this was a forced-choice task, there were three options - either trustworthy or not, or neutral. As seen in Figure 4, overall, 54.1% of the responses were negative, 43.5% of the responses were positive, and the rest were neutral. A chi-square test showed that untrustworthy judgments are significantly higher than trustworthy judgments ($\chi 2$ (2, $N = 28$) = 7.639, $p < .001$). The rest of the responses were omitted either due to the late-response or no response.

The statistical analyses for the relationship between the response type and the item masculinity category are as follows: There were categories for 54 face photos (27 each), on which the statistical analyses were performed. The results of a chi-square test indicated that people tend to respond positively for the face photos with low masculinity rather than those with high masculinity ($\chi 2$ (1, $N = 28$) = 9.864, $p < .001$). Table 1 shows the distributions of the decisions for yes/no for masculine with label-1 and non-masculine with label-0.

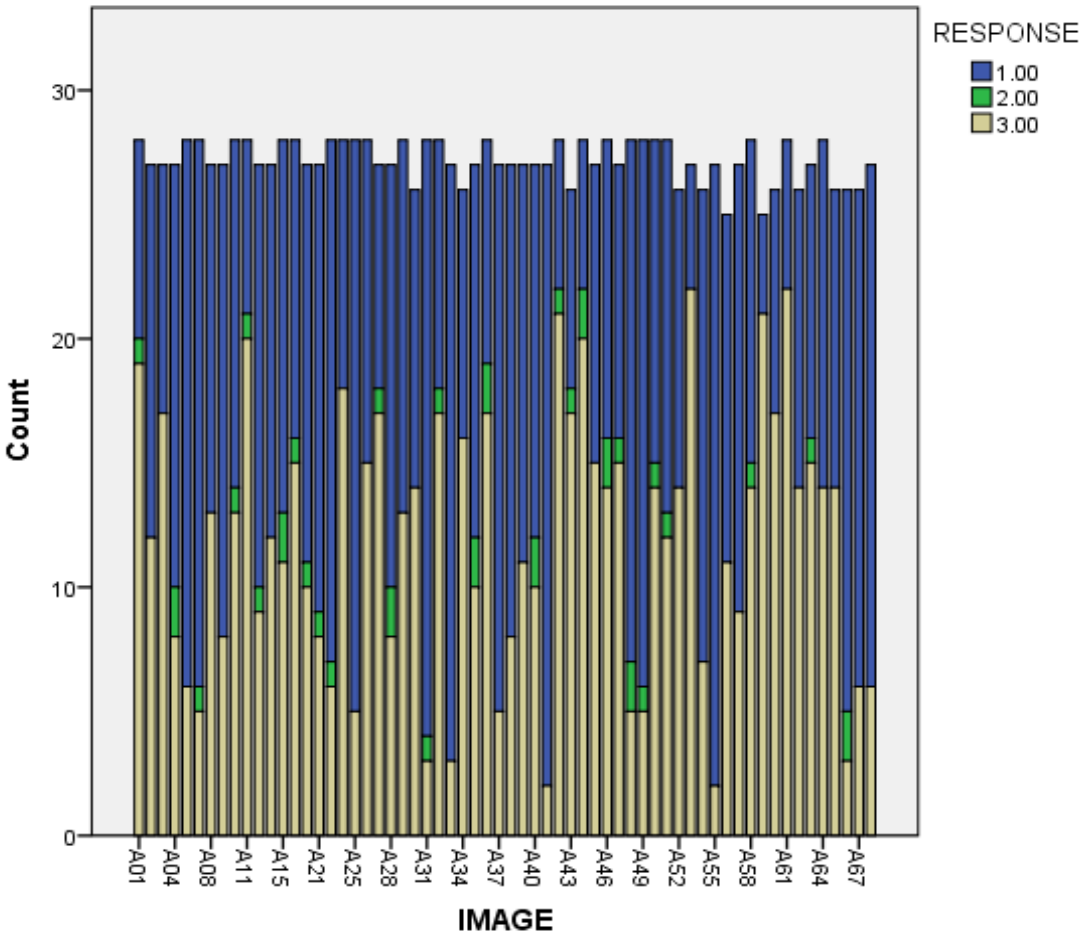


Fig. 4 The distribution of different responses of trustworthy judgments for each of the face photos (1: negative, 2: neutral, 3: positive)

Table 1. The Distribution of the Responses for Trustworthy and Non-Trustworthy Decisions for High-Masculine (with Label-1) and Low-Masculine (with Label-0) Face Photos.

	No	**Yes**
0	376	345
1	434	285

4.2 Reaction time

As well as the decisions with respect to trustworthiness judgments, response times for each stimulus were also collected. Here the hypothesis was that participants

would make trustworthy decisions more quickly than untrustworthy decisions. However, as seen in the Figure 5, the mean value for positive responses (yes meaning that "trustworthy") was 724.77 (*SE* = 17.27) milliseconds, similar to that for negative responses (no meaning that "not trustworthy"), which was 732.79 (*SE* = 15.70), and for the in-between responses the mean response time was 585.65 (*SE* = 61.28) milliseconds. The statistical analyses performed by one-way ANOVA demonstrated that there was no significant effect related to the type of responses on response times ($F(2, 28) = 2.057$, $p > .05$). Figure 6 illustrates the participant-wise mean reaction times, which indicate high variance.

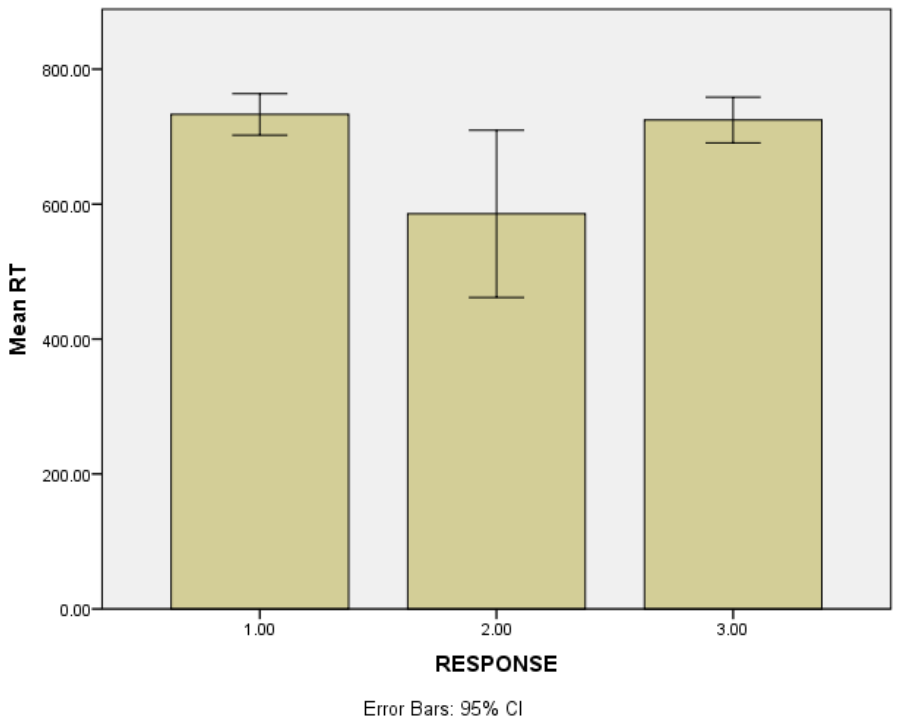


Fig. 5 The mean reaction times related to responses (1: positive, 2: neutral, 3: negative) with the error bars of 95% confidence intervals

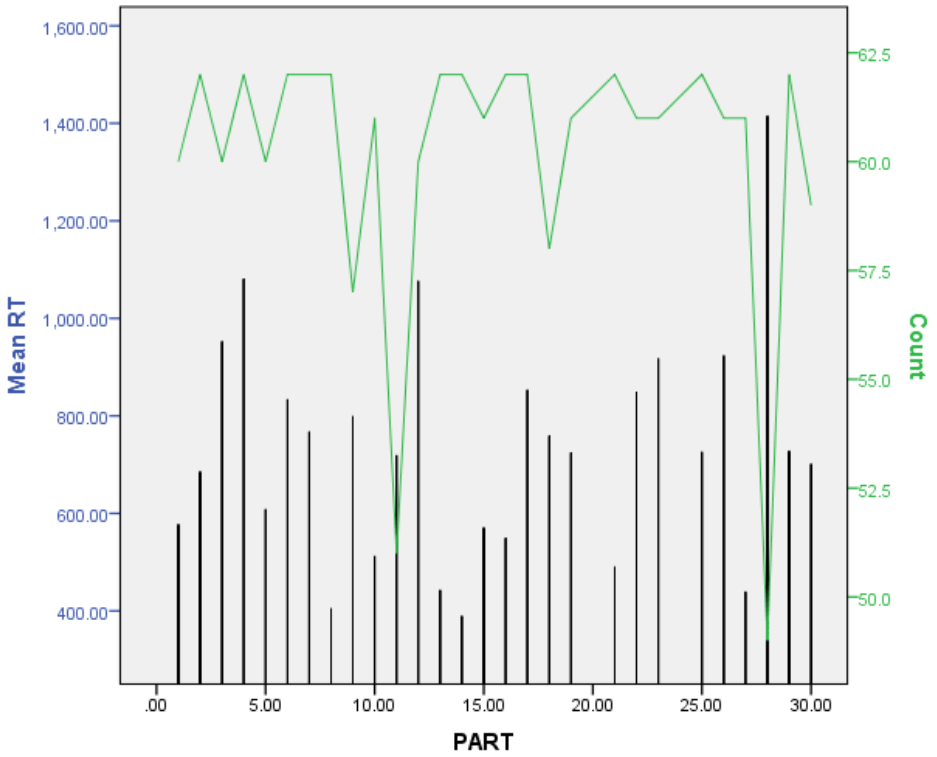


Fig. 6 The mean reaction times for each participant

4.3 Neural activations

For the statistical analyses, the mean values for 4 neural metrics (Oxygenated Hemoglobin (HBO), De-oxygenated Hemoglobin (HBR), Total Hemoglobin (HBT) and Oxygen change (OXY)) were calculated for 16 optodes. There were 64 different dependent variables for 4 (metrics) x 16 (optodes). The statistical analyses were done both for item-based and response-based blocks. Pearson-correlation was used to check the correlation between the mean values for each optode and the average ratings across these dimensions.

4.3.1 Category-wise analyses

The items in this experiment were the face photos, which were already rated across different dimensions including attractiveness and masculinity. The first analysis was

done for the potential effect of masculinity feature. There was a reference dataset for masculinity ratings of the same face photos (54 of 62 photos). As seen in Figure 7, the mean value for each categorized face photo was calculated for all of the optodes.

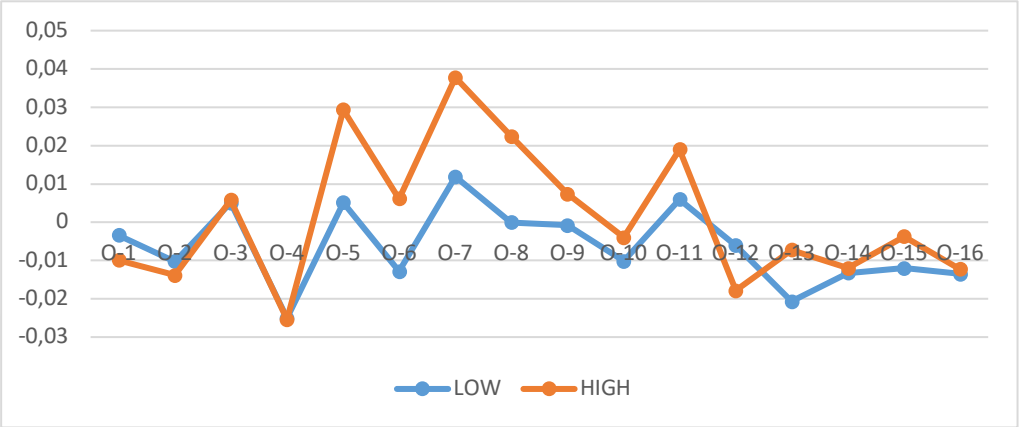


Fig. 7 High versus low masculinity mean levels related to HBR levels for each optode (for positive and negative responses)

The results of the statistical analysis showed that there is a significant difference in the signal read by Optode-6 and Optode-13 for masculine face photos for HBR (t (52) = -2.136, $p < .037$, η_p^2=.003; t (52) = -2.159, $p < .035$, $\eta_p^2 = .003$). The results also showed that there is a marginal difference for Optode-4, Optode-7, and Optode-8 for HBR for high versus low masculinity (t (52) = -1.716, $p < .092$; t (52) = -1.765, $p < .083$; t (52) = -1.712, $p < .093$, $\eta_p^2 = .001$). Moreover, the results also showed that there is a marginal increase for masculine face photos for Optode-1, Optode-8, and Optode-9 for HBO (t (52) = -1.816, $p < .075$; t (52) = -1.692, $p < .097$, $\eta_p^2 = .001$; t (52) = -1.776, $p < .082$, $\eta_p^2 = .001$). These results indicate that there is a higher activation tendency for masculine photos than for non-masculine photos for several optodes.

4.3.2 Response-wise analyses

As seen in Figure 8, the response-based analyses included a comparison of positive and negative responses related to brain activations. The mean values for each optode for negative and positive values were calculated and compared via independent samples *t*-test for HBO, HBR, HBT, and OXY measures. There were significant neural activations for HBO, HBT, and OXY values in neither of the optodes but for the HBR results, there were differences for several optodes. The response-wise analyses for the HBR showed that there is a significant difference in the neural activations for positive responses than for negative responses for one of the optodes (Optode-3, F (2, 28) = 3.119, $p < .05$, $\eta_p^2 = .002$), which is on the left prefrontal cortex. A pairwise t-test with adjusted-Bonferroni showed that for the positive responses as opposed to the negative ones, the neural activation in the left prefrontal cortex is relatively higher than the right ($M = .054$, $SE = .022$, $p < .05$, $\eta_p^2 = .001$).



Fig. 8 The response-wise HBR levels for each optode (for positive, negative and in-between responses)

One-way ANOVA tests showed that there is a significant difference for neutral judgments for the following channels of HBR: Optode-12 (F (2, 28) = 4.466,

$p < .012$), $\eta_p^2 = .006$, Optode-14 (F (2, 28) = 6.354, $p < .002$), and Optode-16 (F (2, 28) = 4.039, $p < .018$, $\eta_p^2 = .005$). The pairwise comparisons indicated that the mean hemodynamic activation for neutral judgments is lower than the positive and negative judgments. For HBO levels, a similar neural pattern and mean difference have been observed for the two of the three channels: Optode-12 (F (2, 29) = 3.197, $p < .041$, $\eta_p^2 = .003$), Optode-14 (F (2, 28) = 5.592, $p < .004$, $\eta_p^2 = .008$). The pairwise comparisons imply that the neutral judgments cause significantly lower level of HBO level when compared to positive and negative judgments.

4.4 Saliency analysis

Studies in the relevant literature on human visual system indicate that people tend to fixate more on the parts of visual scenes with high resolution rather than low resolution (Montepare & Zebrowitz, 1998). In this realm, visual attention has been considered to have a key role in the processing of visual information. Saliency detection has been an influential way for understanding the human visual attention system. The saliency maps have widely been used for detecting and/or recognizing objects in a scene that could be provided as an image or a video. There are two common methods for saliency detection that are bottom-up and top-down methods. Bottom-up methods focus on pre-attentive features to extract saliency in a quicker manner whereas top-down methods are known to be highly task-dependent and goal-driven for saliency extraction with a slower manner. Itti's bottom-up model is used for detecting image saliency focuses on the center-surround features of various features including color, intensity, and orientation (Montepare & Zebrowitz, 1998**).**

Saliency analysis in this study was conducted upon a pre-preparation procedure, where the frames and hairs of the images obtained from the database were

cropped so as to leave the facial region only. SaliencyToolbox was used for the analysis. (Walther & Koch, 2006). As seen in Figure 9 and Figure 10, the outputs of saliency analysis indicated that the eyes are generally captured regardless of the masculinity of the individuals. This can be accepted as a universal feature for the saliency analysis of the faces. Moreover, moustache was one of the other regions that were captured in the photos that might exist both in the masculine as well as non-masculine photos. Eyebrows were another region of interest that were captured for both of the groups. The other captured face regions included mouth, nose, beard, forehead, teeth, and cheek. From an overall perspective, none of these facial features seemed to correspond to a difference regarding the masculinity level of the faces. Thus, it might be reasonable to claim that the salient features of the face photos studied in this thesis work did not seem to cause any bias in the trustworthiness judgments.

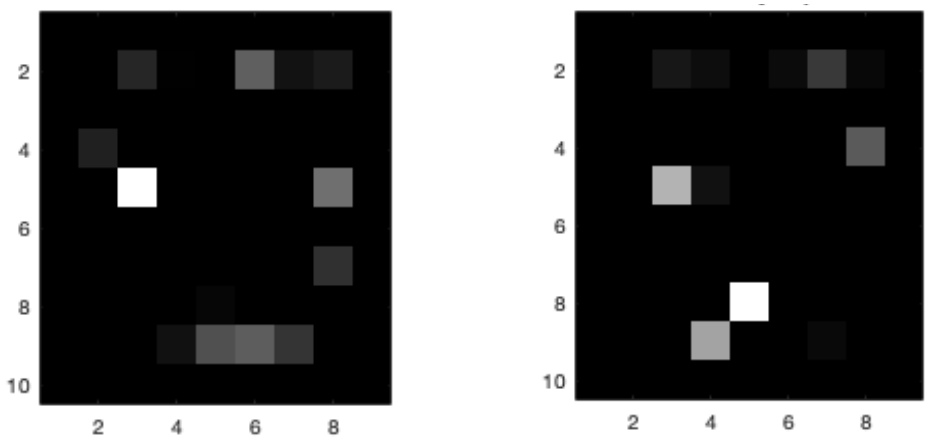


Fig. 9 Sample saliency maps for the masculine (right hand side) and non-masculine (left hand side) photos

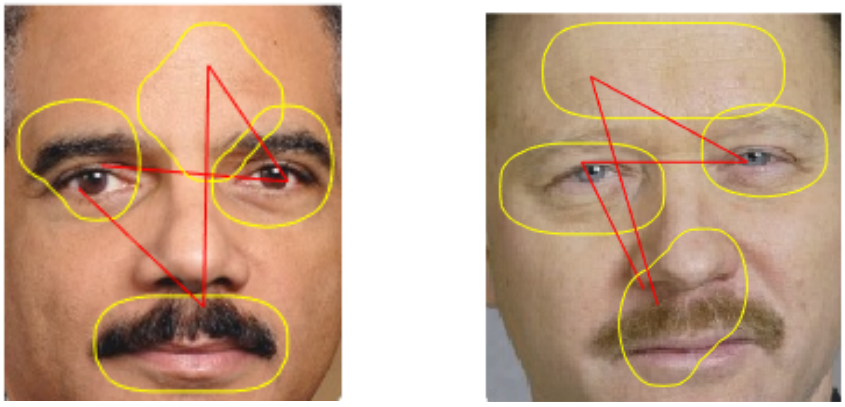


Fig. 10 Sample attended location maps for the masculine (left hand side) and non-masculine (right hand side) face photos

4.5 Emotional analysis via deep learning toolbox

The emotional analyses of the face photos have been performed via Microsoft FaceAPI. Facial expressions especially at an emotional level might also be crucial for one's perceived trustworthiness. In other words, two different face photos of the same person might induce different emotions due to her perceived emotional state. For instance, one with a happy face might trigger higher level of trustworthiness rather than a photo with an angry face. 56 of 62 face photos have been analyzed to be neutral-dominance. The results of the Pearson-correlation test showed that any of the emotional dimensions do not have any effect on the subjective responses or the level of neural activations in the prefrontal cortex measured via fNIRS ($r = 0.0734$, $p >$ .05). Thus, it is possible to come up with the possibility that the emotional states in face photos might not be influential for the people's judgments related to trustworthiness. To drive this conclusion another study should be carried out to have a comprehensive comparison including the same person with different emotional states.

CHAPTER 5

DISCUSSION

Facial appearance provides a means for social cognition for several social implications including attractiveness, intelligence, aggressiveness and other factors. This empirical study aimed at focusing on the cognitive and neural underpinnings of trustworthiness of facial appearance using photos of faces. The neural activation patterns in the prefrontal cortex have been investigated during both the presentation of the faces and the trustworthiness judgments. The obtained empirical results demonstrated four main empirical findings to be discussed further: (1) correlation between masculinity and trustworthiness, (2) reaction time difference between the positive and negative rating trials, (3) higher neural activation in left prefrontal cortex during the positive ratings, and (4) tendency for higher prefrontal activation (HBR) for the face photos with higher masculinity level.

5.1 Correlation between masculinity and trustworthiness

The literature has already provided evidence for the negative correlation between masculinity and trustworthiness. To be more specific, the more a face has masculine features, the less trustworthy the person on the photo is perceived (Todorov et al., 2007). Here, our results are also in parallel with the previous findings that participants tend to rate people with higher masculinity scores as more untrustworthy. This line of research has been shown to have a prediction power for the election results or the potential aggressive behavior (Carre et al., 2009). Thus, perceived masculinity might be assessed as a reliable correlate for trustworthiness judgments. This situation might have evolutionary or cultural origin that the facial

masculinity might be perceived as sign of danger and thus, the perceiver might prefer to be precautious while interacting with that person.

5.2 Reaction time difference between positive and negative responses

The second issue was about reaction times for different ratings of trustworthiness (neutral, positive (trustworthy) and negative (untrustworthy)). The results showed no significant difference between the reaction times in the positive and negative rating trials. This might be due to our experimental paradigm design, where faces were already shown for a relatively longer period of time (4000 milliseconds) before the decision screen is presented for 3000 milliseconds. Thus, participants could already decide upon their ratings during the first presentation phase. In fact, the mean of the response times was found to be around 720 milliseconds. The primary reason for this experimental design (duration times) was for observing neural activations. And it was important to provide a saturation duration for the neural activations related to the nature of neurophysiological signals observed by fNIRS as in fMRI method. In other words, unlike EEG/ERP method, longer durations are required to observe the peak of these neural signals.

5.3 Neural activation pattern differences in prefrontal cortex

The neural activations (HBR level) in the positive rating trials were observed to be higher in the left prefrontal cortex than in the right prefrontal cortex. This can be explained with the effect of positive emotions (pleasantness) for the left PFC rather than the right PFC (Davidson, 2001). This difference in neural activation patterns might also be assessed with frontal asymmetry approach that is frequently used in the neuroscientific literature (Davidson, 1983; Davidson, 1995). Frontal asymmetry

approach is about the level of neural activation difference among the different hemispheres of frontal cortex for two discrete emotions. In other words, the higher activation in left frontal cortex than right frontal cortex has been accepted as a neural indicator of a positive emotional state and vice versa.

Similar empirical findings have been presented in the literature, which shows an increase in the activation of rostromedial prefrontal cortex during the experience of positive emotions (Kreplin & Fairclough, 2013). There is also an increase in the activity of the left dorsolateral prefrontal cortex (Optode-2). Left regions of the prefrontal cortex (especially the lateral regions, also referred to as left dorsolateral PFC that corresponds to Optode-1, Optode-2, and Optode-3) is associated with the working memory and strategical decision-making processes (Ogawa et al., 2014).

5.4 No significant difference observed related to brain reward system

Our second experimental hypothesis was that the neural activations within the medial prefrontal cortex will be higher when the participants evaluate a photo of a person who was labeled as more trustworthy than untrustworthy. The medial prefrontal cortex is taken as part of the brain reward system. This reward system has been observed to be active when the person experiences a sense of pleasure or a positive emotion such as eating a delicious food, having sexual intercourse, or even making a purchasing (as opposed to no-purchase) decision (Berridge & Kringelbach, 2008). In fact, pleasantness is particularly associated with the activity in the medial region of a specific Brodmann area (BA10), an area also known to be a part of the human brain reward system. However, the obtained results did not show any indication of the impact of human brain reward system, since there was no significant difference in the neural activation patterns when the choices of trustworthy/untrustworthy are

compared to each other. This might be due to the fact that this type of judgment did not require any activation among the related brain areas since this reward system has not been activated. On the other hand, if this were a judgment related to beauty then it might have been the case that a significantly bigger neural activation would be observed for the positive than for the negative judgments.

5.5 Limitations due to experimental design and methodology

One of the main limitations of this study was the use of a forced-choice task. Participants were able to respond either positively, neutrally or negatively but they had no other option. Forced-choice paradigm was introduced to avoid indecisions, which had happened in the earlier pilot experiments. However, in forced-choice tasks the participants are forced to select among given options but on some occasions the participant might prefer to an intermediate state, or a continuous scale might be more appropriate. Moreover, -participants were presented with the face photos for 4000 milliseconds, during which they were not allowed to give any response. This, however, was necessary to provide sufficient time for a measure of an accumulated hemodynamic activity in the prefrontal cortex. These factors should be evaluated as a limitation of this neurophysiological methodology.

CHAPTER 6

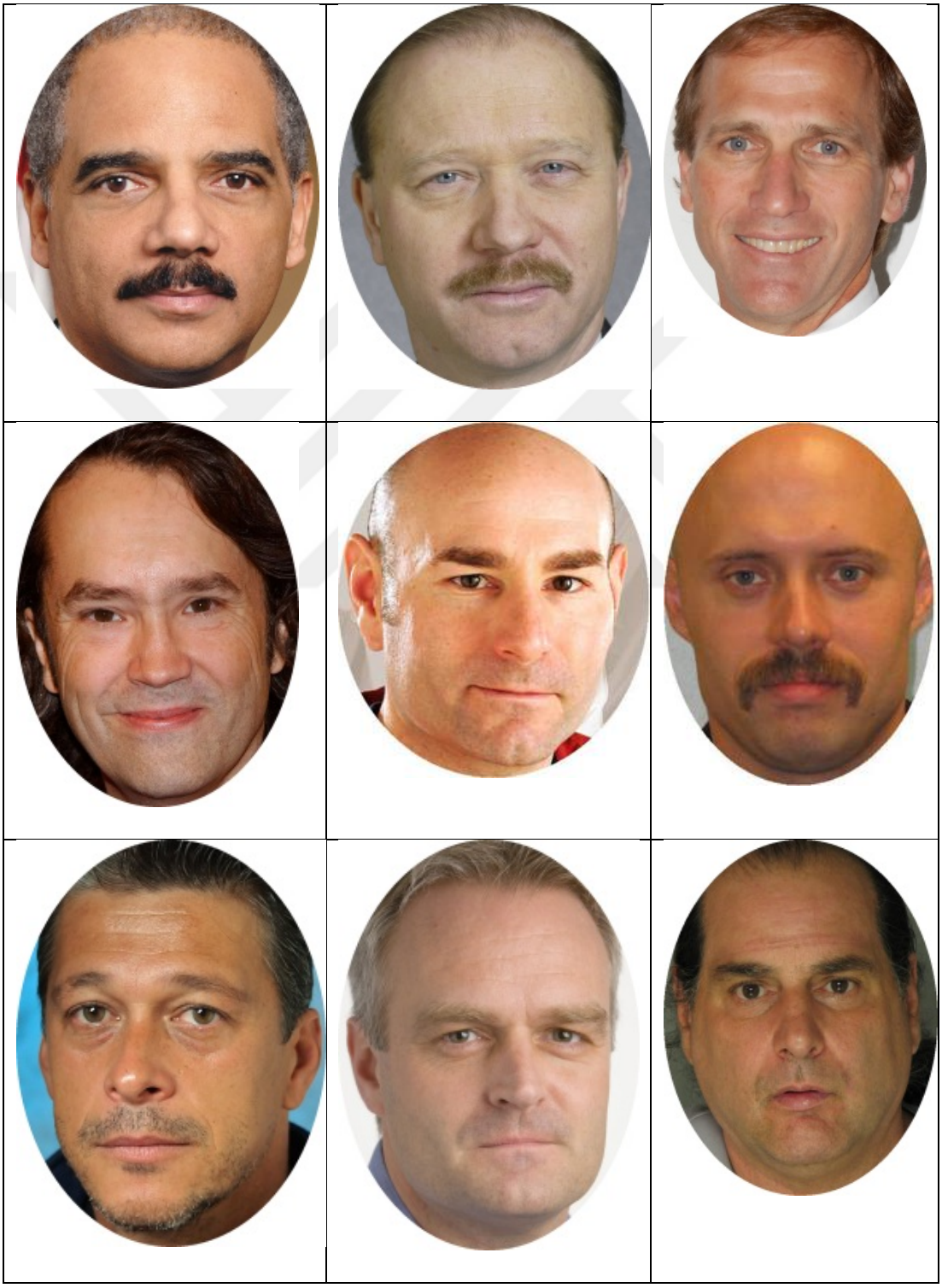
CONCLUSION

This thesis work has initially been intended to investigate the potential use of measuring the trustworthiness perception/judgments of the face photos via an optic brain imaging method (fNIRS). For this aim, a collection of face photos were derived from a scientifically valid database. Face photos were selected with respect to their ratings related to the masculinity dimension.

The main contribution would be about understanding the neural mechanisms within the prefrontal cortex during the trustworthiness judgments for the presented face photos. The obtained results indicated that the positive (trustworthy) judgments were associated with a higher level of neural activation in the left and left-medial prefrontal cortex when compared to the negative (untrustworthy) judgments. This is most probably due to the frontal asymmetry because of the contribution of approach/withdrawal impact (also known as motivation/positiveness metric) during the positive judgments than negative ones (Davidson, 1995). In other words, left frontal cortex is activated more than the right frontal cortex during the experience of positive emotions. Despite the limitations due to experimental setting including forced-choice and relatively long display durations, the obtained empirical findings indicate that useful practical applications might be developed with further research in the near future. For future research, the 3-point scale might be converted into 5-point or 7-point Likert scale with the same photos presented. Thus, it might be more convenient to compare the distribution of the more/less untrustworthy face photos. Also, it might be interesting to observe the patterns of women participants as well as the effect of faces from different origins to see potential cross-cultural impacts.

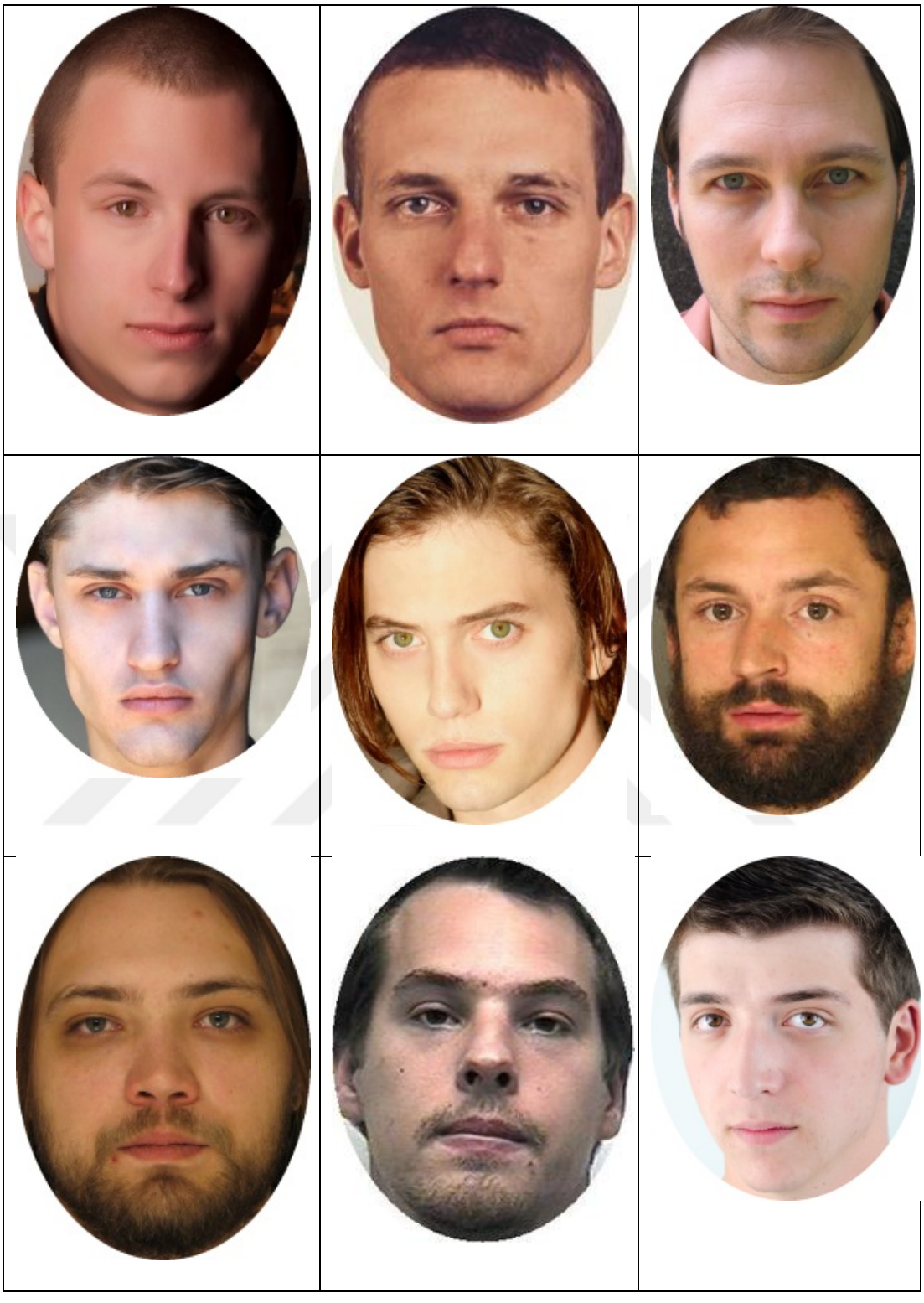
APPENDIX A

FACE IMAGES















APPENDIX B

CONSENT FORM

Bilgilendirilmiş Gönüllü Katılım Onay Formu

Projenin Adı: **Bireylerin Güven Temelli Karar Verme Süreçlerinin Beyinsel İzdüşümlerinin Optik Beyin Görüntüleme Sistemi Kullanılarak İncelenmesi**

Proje yürütücüsünün adı ve iletişim bilgileri: **Taylan Bilik (taylan@artipr.com.tr)**

Proje bütçesine katkı ile ve/veya diğer şekillerde projeyi destekleyen kuruluşlar: **Projeye bütçesel anlamda katkı sağlayan bir kurum yoktur.**

Süreç: Gönüllü katılım sağladığınız takdirde, testimizde 62 insan yüz resmi göreceksiniz ve bu resimdeki kişiyi güvenilir bulup bulmadığınız konusunda yanıt vermeniz gerekmektedir. Test sonrasında sizden bazı tutumlarınız ile ilgili soruları cevaplandırmanız beklenmektedir.

Gizlilik: Katılımcıların isimleri paylaşılmayacaktır. Araştırma sonrasında katılımcıların bilgilerinin gizli tutulacağı yazılı olarak alınan onay formunda da belirtilecektir. Kişisel bilgiler analizler öncesinde imha edilerek vakalar anonimleştirilerek analize sokulacaktır.

Gönüllü Katılım: Bu projeye katılım tamamen gönüllülük temelindedir. Projeden herhangi bir zamanda çekilme hakkınız vardır ve katılımı reddetmek herhangi bir zarara uğramanıza sebep olmayacaktır.

Bu formda anlatılan araştırmanın etik yönleriyle ve/veya araştırma detaylarıyla ilgili sorularınız, sorunlarınız veya önerileriniz varsa lütfen MEF Üniversitesi İnsan Araştırmaları Etik Kurulu ile (212) 395 36 15 nolu telefondan temasa geçiniz.

Yukarıda sözü geçen **Bireylerin Güven Temelli Karar Verme Süreçlerinin Beyinsel İzdüşümlerinin Optik Beyin Görüntüleme Sistemi Kullanılarak İncelenmesi** adlı araştırma projesinin detaylarını okudum ve bu proje ile ilgili sorularım cevaplandı. Bu çalışmaya gönüllü olarak katılıyorum.

_____ _____

Ad Soyadı Tarih, İmza

REFERENCES

Adolphs, R., Tranel, D., & Damasio, A. R. (1998, June 4). The human amygdala in social judgment. *Nature,* 393, 470-474.

Aimone, J. A., Houser, D., & Weber, B. (2014). Neural signatures of betrayal aversion: an fMRI study of trust. *Proceedings of the Royal Society B: Biological Sciences, 281(1782)*. doi: 10.1098/rspb.2013.2127.x

Amaral, D. G. (2002). The primate amygdala and the neurobiology of social behavior: Implications for understanding social anxiety. *Biological Psychiatry, 51(1)*, 11-17.

Ayaz, H., Shewokis, P. A., Bunce, S., Izzetoglu, K., Willems, B., & Onaral, B, (2012). Optical brain monitoring for operator training and mental workload assessment. *Neuroimage, 59(1)*, 36-47.

Bainbridge, W.A., Isola, P., & Oliva, A. (2013). The intrinsic memorability of face images. *Journal of Experimental Psychology: General, 142*(4), 1323-1334.

Ballew, C. C., Todorov, A. (2007). Predicting political elections from rapid and unreflective face judgments. *Proceedings of the National Academy of Sciences of the USA, 104(46)*, 17948-53. doi:10.1073/pnas.0705435104.x

Bar, M., Neta, M., & Linz, H. (2006). Very first impressions. *Emotion, 6(2),* 269-278.

Berridge, K. C., Kringelbach, M. L. (2008). Affective neuroscience of pleasure: reward in humans and animals. *Pyschopharmacology, 199(3)*, 457-480.

Blair, I. V., Judd, C. M., & Chapleau, K. M. (2004). The influence of Afrocentric facial features in criminal sentencing. *Psychological Science, 15(10)*, 674–679.

Cerine, L. & Leahy, M.J. (2016). An updated review of methods and advancements in microvascular blood flow imaging. *Microcirculation, 23(5)*, 345-363.

Chance, B., Alfano, R. (1993). photon migration and imaging in random media and tissues. Proceedings of Society of Photo-optical Instrumentation Engineers, Biomedical Optics Society, 1888. In *OE/LASE'93: Optics, Electro-optics, and Laser Applications in Science and Engineering*. Los Angeles, CA: The International Society for Optics and Photonics.

Davidson, R.J. (1983). Hemispheric asymmetry and emotion. In R.J. Davidson & P. Ekman (Eds.), *Questions about emotions (pp. 39- 57)*. Massachusetts, MA: MIT Press.

Davidson, R.J. (1995). Cerebral asymmetry, emotion, and affective style. In R.J. Davidson & K. Hughdahl (Eds.), *Brain asymmetry* (pp. 361-387). Massachusetts, MA: MIT Press.

Davis, M., & Whalen, P. J. (2001). The amygdala: Vigilance and emotion. *Molecular Psychiatry, 6(1)*, 13-34.

Díaz-Gutiérrez, P., Alguacil, S., & Ruz, M. (2017). Bias and control in social decision-making. In Ibanez, Agustin, Sedeño, Lucas & García, Adolfo, *Neuroscience and social science: The missing link (pp. 47-68)*. Springer.

Dimoka, A. (2010). What Does the Brain Tell Us About Trust and Distrust? Evidence from a Functional Neuroimaging Study, *MIS Quarterly*, *34(2)*, 373-396.

Eberhardt, J. L., Davies, P. G., Purdie-Vaughns, V. J., & Johnson, S. L. (2006). Looking deathworthy: Perceived stereotypicality of black defendants predicts capital-sentencing outcomes. *Psychological Science, 17(5)*, 383–386.

Engell, A. D., Haxby, J. V., & Todorov, A. (2007). Implicit trustworthiness decisions: Automatic coding of face properties in human amygdala. *Journal of Cognitive Neuroscience, 19(9)*, 1508-1519.

Hall, C. C., Goren, A., Chaiken, S., & Todorov, A. (2009). Shallow cues with deep effects: Trait judgments from faces and voting decisions. In E. Borgida, J. L. Sullivan, & C. M. Federico (Eds.), *The political psychology of democratic citizenship (pp. 73-99)*. New York, NY: Oxford University Press.

Hamermesh, D. S., & Biddle, J. E. (1994). Beauty and the labor market. *American Economic Review, 84(5)*, 1174-1194.

Hassin, R,. & Trope, Y. (2000). Facing faces: studies on the cognitive aspects of physiognomy. *Journal of Personality and Social Psychology 78(5)*, 837–852.

Hubert, M., Linzmajer, M., Riedl, R., & Kenning, P. (2018). Trust me if you can–neurophysiological insights on the influence of consumer impulsiveness on trustworthiness evaluations in online settings. *European Journal of Marketing, 52*(1/2), 118-146.

Jobsis, F. F. (1977). Noninvasive, infrared monitoring of cerebral and myocardial oxygen sufficiency and circulatory parameters. *Science, 198*(4323), 1264-1267.

Khosla, A., Bainbridge, W.A., Torralba, A., & Oliva, A. (2013). Modifying the memorability of face photographs. In *ICCV '13: Proceedings of the 2013 IEEE International Conference on Computer Vision*, (pp. 3200-3207). Sydney, Australia: Institute of Electrical and Electronics Engineers Computer Society. Retrieved from http://people.csail.mit.edu/

Langlois, J. H., Kalakanis, L., Rubenstein, A. J., Larson, A., Hallam, M., Smoot, M., (2000). Maxims or myths of beauty? A meta-analytic and theoretical review. *Psychology Bulletin, 126(3)*, 390–423.

LeDoux, J. E. (2000). Emotion circuits in the brain. *Annual Review of Neuroscience, 23(1)*, 155-184.

LeDoux, J. (2003) The Emotional Brain, Fear, and Amygdala. *Cellular and Molecular NeuroBiology, 23(4-5)*, 727-738.

Luhmann, N. (1979). Trust and power. Chichester, UK: Wiley.

Mazur, A., Mazur, J., & Keating, C. (1984). Military rank attainment of a West Point class: Effects of cadets' physical features. *American Journal of Sociology, 90(1)*, 125-150.

McKnight, D. H., Kacmar, C., & Choudhury, V. (2004). Dispositional trust and distrust distinctions in predicting high and low risk internet expert advice site perceptions. *E-Service Journal, 3(2)*, 35-58.

Montepare, J. M., Zebrowitz, L. A., (1998). Person perception comes of age: The salience and significance of age in social judgments. In Zanna M., *Advances in Experimental Social Psychology* (pp. 93-163). San Diego, CA: Academic Press.

Mueller, U., & Mazur, A. (1996). Facial dominance of West Point cadets as a predictor of later military rank. *Social Forces, 74(3)*, 823-850.

Obrig, H., Wenzel, R., Kohl, M., Horst, S., Wobst, P., Steinbrink, J. & Villringer, A. (2000). Near-infrared spectroscopy: Does it function in functional activation studies of the adult brain*? International Journal of Psychophysiology, 35(2),* 125-142.

Ogawa, Y., Kotani, K., & Jimbo, Y. (2014). Relationship between working memory performance and neural activation measured using near-infrared spectroscopy. *Brain Behavior, 4(4)*, 544-551.

Oosterhof, N. N., & Todorov, A. (2008). The functional basis of face evaluation. *Proceedings of the National Academy of Sciences of the United States of America, 105(32)*, 11087-11092.

Phelps, E. A., & LeDoux, J. E. (2005). Contributions of the amygdala to emotion processing: From animal models to human behavior. *Neuron, 48(2),* 175-187.

Riedl, R. & Javor, A., (2012). The biology of trust: Integrating evidence from genetics, endocrinology, and functional brain imaging. *Journal of Neuroscience, Psychology, and Economics, 5(2),* 63-91.

Rosenberg, S., Nelson, C., & Vivekananthan, P. S. (1968). A multidimensional approach to the structure of personality impressions. *Journal of Personality and Social Psychology, 9(4)*, 283–94.

Strangman, G., Boas, D. A., & Sutton, J. P., (2002). Non-invasive neuroimaging using near-infrared light. *Biological Psychiatry, 52*(7), 679-693.

Todorov, A., & Uleman J. S. (2003). The efficiency of binding spontaneous trait inferences to actors' faces. *Journal of Experimental Social Psychology, 39(6),* 549–62.

Todorov, A., Baron, S., & Oosterhof, N. N. (2008). Evaluating face trustworthiness: A model based approach. *Social, Cognitive and Affective Neuroscience, 3(2),* 119-127.

Todorov, A., Mandisodza A. N., Goren, A., & Hall, C. C. (2005). Inferences of competence from faces predict election outcomes. *Science, 308(5728)*, 1623–1626

Todorov, A., & Engell, A. (2008). The role of the amygdala in implicit evaluation of emotionally neutral faces. *Social, Cognitive and Affective Neuroscience, 3(4),* 303-312.

Tzschentke, T. M. (2000). The medial cortex as a part of the brain reward system. *Amino Acids, 19(1)*, 211-219.

Uleman, J. S., Blader, S. L., & Todorov, A. (2005). Implicit impressions. In R. R. Hassin, J. S. Uleman, & J. A. Bargh (Eds.), *Oxford series in social cognition and social neuroscience. The new unconscious* (pp. 362-392). New York, NY: Oxford University Press.

Villringer, A., & Chance, B., (1997). Non-invasive optical spectroscopy and imaging of human brain function. *Trends in Neurosciences, 20(10)*, 435-442.

Villringer, A., Planck, J., Hock, C., Schleinkofer, L., & Dirnagl, U. (1993). Near infrared spectroscopy (NIRS): a new tool to study hemodynamic changes during activation of brain function in human adults. *Neuroscience Letters, 154(1-2)*, 101-104.

Walther, D. B., & Koch, C. (2006). Modeling attention to salient proto-objects. *Neural Networks, 19(9)*, 1395-1407.

Whalen, P. J. (1998). Fear, vigilance, and ambiguity: Initial neuroimaging studies of the human amygdala. *Current Directions in Psychological Science, 7(6)*, 177-188.

Willis, J., & Todorov, A. (2006). First impressions: making up your mind after a 100-ms exposure to a face. *Psychological Science, 17(7)*, 592–98.

Winston, J., Strange, B., O'Doherty, J., & Dolan, R. (2002). Automatic and intentional brain responses during evaluation of trustworthiness of faces. *Nature Neuroscience, 5(3)*, 277-283.

Zebrowitz, L. A. (1997*). Reading faces: Window to the soul?* Boulder, CO: Westview Press.