Neural mechanisms in processing of emotion in real and virtual faces using functional-near infrared spectroscopy (fNIRS)

by

Dylan Rapanan

A thesis submitted to the

School of Graduate and Postdoctoral Studies

in partial fulfillment of the requirements for the degree of

**Masters of Science** in **Computer Science**

Faculty of Science Ontario Tech University Oshawa, Ontario, Canada

© Dylan Rapanan 2025

**Abstract**

As avatars permeate social media, gaming, and telecommunications, understanding how the brain reads emotions from virtual faces is increasingly important. We recorded func- tional near-infrared spectroscopy (fNIRS) data from adults viewing real photographs and matched computer-generated faces expressing Anger, Disgust, Fear, Joy, Sadness, Sur- prise, or Neutral (control). General-linear-model mapping revealed higher activation in virtual faces in the left occipital region, and higher activation in Neutral and Surprise compared to the other emotions in parietal and occipital regions. Functional-connectivity analysis revealed higher connectivity in real faces across the brain, and higher connectiv- ity across the brain in Anger and Fear compared to the other emotions. Collectively, the results demonstrate differences in activation in occipital areas, and differential process- ing of face and emotion types across the whole brain. These neural signatures provide quantitative targets for refining the realism and emotional efficacy of digital characters in virtual and augmented environments.

**Acknowledgements**

\*\* Put your Acknowledgements here. \*\*\*

**Contents**

|  |  |  |
| --- | --- | --- |
| [**Abstract**](#_bookmark0)  [**Acknowledgment**](#_bookmark1) |  | **ii**  **iii** |
| [**1 Introduction**](#_bookmark2) |  | **1** |
| [1.1 Functional Near-Infrared Spectroscopy (fNIRS)](#_bookmark3) | . . . . . . . . . . . . . . | 2 |

* 1. [Facial Emotion Perception](#_bookmark4) . . . . . . . . . . . . . . . . . . . . . . . . . . 4
  2. [Real vs. Virtual (Avatar) Face Perception](#_bookmark5) . . . . . . . . . . . . . . . . . 6
  3. [Objectives and Hypotheses](#_bookmark6) . . . . . . . . . . . . . . . . . . . . . . . . . . 8

1. [Methodology](#_bookmark7) 10
   1. [Participants](#_bookmark8) 10
   2. [Stimuli and apparatus](#_bookmark10) 12
      1. [Stimuli](#_bookmark11) 12
      2. [Apparatus](#_bookmark12) 12
   3. [Design and procedure](#_bookmark14) 15
      1. [Design](#_bookmark15) 15
      2. [Procedure](#_bookmark16) 15
   4. [Analyses](#_bookmark18) 16
      1. [fNIRS preprocessing](#_bookmark19) 17
      2. [Activation magnitude with General Linear Model (GLM)](#_bookmark21) 19

[Design Matrix](#_bookmark22) 19

[Contrast Computation](#_bookmark24) 20

* + 1. [Network mapping with Functional Connectivity Analysis](#_bookmark25) 21

[Paired Sample t-tests](#_bookmark26) 22

[ROI Chord Plots](#_bookmark27) 22

[Emotion Summary Ratio Plot](#_bookmark28) 23

[Region Summary Plot](#_bookmark29) 23

* + 1. [Memory Task Analysis](#_bookmark30) 24

1. [Results](#_bookmark31) 25
   1. [Neural activation magnitude](#_bookmark32) 25
   2. [Functional Connectivity Mapping](#_bookmark37) 29
   3. [Memory Task Results](#_bookmark43) 36
2. [Discussion](#_bookmark45) 38

[Bibliography](#_bookmark45) 38

[Appendix](#_bookmark46) 46

[A GLM Contrasts](#_bookmark99) 47

1. [Functional Connectivity Contrasts](#_bookmark101) 52
2. [Memory Task](#_bookmark105) 62
   1. [ANOVA Results](#_bookmark106) 62
   2. [Memory Task No Response Distribution](#_bookmark108) 63

**List of Tables**

* 1. [Table of significant contrast results from the GLM analysis.](#_bookmark100) 47
  2. [Ratio of positive to negative *t*-values for each contrast. The ratio is cal-](#_bookmark104) [culated as Ratio = (Number of Positive *t*-values - Number of Negative](#_bookmark104) [*t*-values) / (Number of Positive *t*-values + Number of Negative *t*-values).](#_bookmark104) [For condition1 *>* condition2, a positive ratio indicates that condition1 has](#_bookmark104) [more positive *t*-values than condition2, while a negative ratio indicates the](#_bookmark104)

[opposite.](#_bookmark104) 55

* 1. [Two-way ANOVA results for the effect of Face Type and Emotion and](#_bookmark107)

[their interaction on the correct responses.](#_bookmark107) 62

**List of Figures**

|  |  |  |
| --- | --- | --- |
| [2.1](#_bookmark9) | [Participant signal quality and inclusion flow](#_bookmark9) . . . . . . . . . . . . . . . . | 11 |
| [2.2](#_bookmark13) | [Montage and brain region map](#_bookmark13) . . . . . . . . . . . . . . . . . . . . . . . | 14 |
| [2.3](#_bookmark17) | [Experimental paradigm overview](#_bookmark17) . . . . . . . . . . . . . . . . . . . . . . | 16 |
| [2.4](#_bookmark20) | [Preprocessing steps for fNIRS data](#_bookmark20) . . . . . . . . . . . . . . . . . . . . . | 17 |
| [2.5](#_bookmark23) | [Sample design matrix for GLM](#_bookmark23) . . . . . . . . . . . . . . . . . . . . . . . | 20 |
| [3.1](#_bookmark33) | [GLM: Real vs. Virtual Faces](#_bookmark33) . . . . . . . . . . . . . . . . . . . . . . . . . | 25 |
| [3.2](#_bookmark34) | [GLM: Emotion vs. Neutral](#_bookmark34) . . . . . . . . . . . . . . . . . . . . . . . . . . | 26 |
| [3.3](#_bookmark35) | [GLM: Emotion vs. Surprise](#_bookmark35) . . . . . . . . . . . . . . . . . . . . . . . . . | 27 |
| [3.4](#_bookmark36) | [GLM: Face Type *×* Emotion Contrasts](#_bookmark36) . . . . . . . . . . . . . . . . . . . | 28 |
| [3.5](#_bookmark38) | [FC: Real vs. Virtual](#_bookmark38) . . . . . . . . . . . . . . . . . . . . . . . . . . . . . | 29 |
| [3.6](#_bookmark39) | [FC: Emotion Contrasts](#_bookmark39) . . . . . . . . . . . . . . . . . . . . . . . . . . . . | 30 |
| [3.7](#_bookmark40) | [FC: Summary of Contrasts by Emotion Pair](#_bookmark40) . . . . . . . . . . . . . . . . | 32 |
| [3.8](#_bookmark41) | [FC: Count of Significantly Different Channel Pairs by ROI](#_bookmark41) . . . . . . . . | 33 |
| [3.9](#_bookmark42) | [FC: Face Type *×* Emotion Contrasts](#_bookmark42) . . . . . . . . . . . . . . . . . . . . | 35 |
| [3.10](#_bookmark44) | [Correct Memory Task Responses by Face Type and Emotion](#_bookmark44) . . . . . . . | 37 |
| [B.1](#_bookmark102) | [FC: Additional emotion contrasts](#_bookmark102) . . . . . . . . . . . . . . . . . . . . . . | 52 |
| [C.1](#_bookmark109) | [Memory Task No Response Distribution](#_bookmark109) . . . . . . . . . . . . . . . . . . | 63 |

**Chapter 1 Introduction**

Our brains are evolutionarily primed to process faces, as they are crucial for social interac- tions ([Powell et al.](#_bookmark86), [2018](#_bookmark86)). A central aspect of this facial perception is the interpretation of emotional expressions, which underpins our interactions as social beings. Emotional expressions provide essential information about an individual’s internal state and inten- tions, allowing us to navigate complex social environments. Despite a rich literature examining human emotion perception, it remains unclear how the brain processes these emotional expressions definitively ([Barrett](#_bookmark47), [2006a](#_bookmark47)). This lack of clarity is particularly relevant as our interactions shift increasingly toward digital platforms featuring virtual representations of human faces.

There has been a dramatic increase in the use of avatars, computer-generated representations of humans, across a wide range of platforms, including social media, video games, virtual reality (VR), and augmented reality (AR). As interactions with avatars become more prevalent, particularly in applications involving communication and social interaction, their ability to convincingly express human-like emotions has become a subject of growing interest ([Kegel et al.](#_bookmark72), [2020](#_bookmark72)). The capacity of avatars to produce recognizable and appropriate facial emotional expressions is key to their social acceptance and utility.

Unlike real human faces, which convey emotion through a complex interplay of subtle muscle movements, virtual faces must rely on pre-programmed or algorithmically generated expressions. These expressions may differ in perceived realism, dynamicity, and authenticity, potentially altering how they are processed by the brain. This raises several important questions: Does processing virtual facial expressions engage the same neural mechanisms as processing real facial expressions? Do different emotional expressions elicit distinct neural responses in the brain? And finally, is processing emotional expressions in virtual faces fundamentally different from processing emotional expressions in real faces? To address these questions, this thesis investigates the neural mechanisms underlying facial emotion perception in both real and virtual faces using functional near-infrared spectroscopy (fNIRS).

* 1. **Facial Emotion Perception**

Emotion perception involves recognizing social cues such as facial emotional expressions, which is the focus of this thesis. Human facial emotion perception has been a central topic in affective neuroscience, with [Ekman](#_bookmark58) [and Friesen](#_bookmark58) ([1971](#_bookmark58)) identifying a set of basic emotions that are universally recognized across cultures. Ekman proposed six basic emotions: happiness, sadness, anger, fear

disgust, and surprise, with neutral faces serving as a baseline. These emotions are reliably associated with distinct facial configurations, known as the Facial Action Coding System (FACS), developed by [Ekman and Friesen](#_bookmark59) ([1978](#_bookmark59)), which provides a comprehensive taxonomy of facial muscle movements, known as action units, which underpin the visible expressions of emotion. FACS categorizes facial movements into action units (AUs), each corresponding to specific muscle contractions, such as the raising of the eyebrows or the curling of the lips. This system has informed both psychological research and the development of synthetic facial expression systems in virtual environments.

More recent approaches, however, advocate a constructionist view of emotion. Ac- cording to [Barrett](#_bookmark48) ([2006b](#_bookmark48)), emotions are not fixed categories but constructed experiences emerging from the brain’s interpretation of internal and external stimuli. [Lindquist et al.](#_bookmark77)

([2012](#_bookmark77)) argue that emotions arise from distributed and context-sensitive patterns of neural activity involving domain-general brain networks rather than discrete, emotion-specific regions. This debate has significant implications for avatar perception; if emotions are constructed rather than discrete, then the realism, dynamicity, and context of avatar expressions may critically shape how they are interpreted and processed neurologically.

There is increasing evidence supporting this involvement of domain-general networks in emotion perception. While multivariate pattern analysis (MVPA) studies have shown that both localized and distributed neural patterns can predict emotional states ([Kragel and LaBar](#_bookmark75), [2016](#_bookmark75)), findings remain inconsistent, especially in the prefrontal cortex (PFC) ([Westgarth et al.](#_bookmark96), [2021](#_bookmark96); [Bendall et al.](#_bookmark50), [2016](#_bookmark50)). Some fNIRS studies report increased PFC activation during facial emotion recognition (e.g., in the ventral and medial PFC), others find decreased or no significant changes in oxygenated hemoglobin (HbO) levels. Even studies using similar facial expression tasks report varying activation patterns depending on the specific emotion or cortical region involved. For instance, happy and fearful faces have been associated with increased right PFC activation, whereas sad faces tend to elicit decreased activation in the left PFC. These mixed findings echo fMRI research, which implicates a wide network, including the medial PFC, amygdala, fusiform gyrus, superior temporal sulcus, and insula in emotion perception, with specific emotions such as anger, disgust, and sadness engaging distinct cortical and subcortical areas. These findings underscore the complexity of emotion processing and highlights the need for more nuanced investigations of how different emotional expressions are represented in the brain.

* 1. **Real vs. Virtual (Avatar) Face Perception**

The increasing use of avatars has raised questions about how their facial expressions com- pare to real human faces. [Garc´ıa et al.](#_bookmark63) ([2020](#_bookmark63)) designed avatar facial expressions using the FACS that were validated by human observers, demonstrating the efficacy of using

FACS-based design principles to create reliable virtual human facial expressions. Studies have shown that expressions of happiness, anger, fear, and other basic emotions can be accurately interpreted from both static and dynamic virtual avatars ([Faita et al.](#_bookmark60), [2015](#_bookmark60); [Dyck et al.](#_bookmark57), [2008](#_bookmark57)). However, in some cases, avatars may convey emotional expressions more or less effectively than real human faces. [Dyck et al.](#_bookmark57) ([2008](#_bookmark57)) found that while disgust was challenging to convey using current avatar technology, virtual expressions of sadness and fear were recognized more accurately than their natural face counter- parts. [Hortensius et al.](#_bookmark69) ([2018](#_bookmark69)) provides guiding principles for designing avatars that can effectively convey emotional expressions, and found people are generally less accurate at recognizing emotions from robotic faces compared to human faces. However, virtual agents can be as effective as humans in conveying emotions, particularly when their facial muscle movements are clearly depicted. This work highlights the potential for avatars to convey facial/emotional expressions effectively, but it also raises questions about how these virtual faces are processed in the brain compared to real human faces.

A growing body of affective and cognitive neuroscience research suggests that virtual faces, while often processed like real faces, can still elicit distinct neural responses due to differences in perceived authenticity, dynamicity, and realism. [De Borst and](#_bookmark56) [De Gelder](#_bookmark56) ([2015](#_bookmark56)) note that since humans are highly attuned to perceiving real human faces, viewing avatars may engage different perceptual and neural processes, potentially leading to altered brain activity.

It is important to note, findings from studies using avatars and those using real human faces may not always be directly comparable and should be interpreted cautiously.

These perceptual discrepancies may partially stem from the so-called ”uncanny valley” phenomenon ([Mori et al.](#_bookmark80), [2012](#_bookmark80)), wherein highly realistic but imperfect virtual faces evoke a sense of unease or cognitive dissonance in observers. [K¨atsyri et al.](#_bookmark76) ([2017](#_bookmark76)) empirically tested this hypothesis using semirealistic computer-animated film characters and found that characters perceived as more realistic were rated as more ’eerie’, compared to the more cartoonish characters. The N170 is an event-related potential (ERP) component, commonly investigated in EEG studies, that is typically observed over occipitotemporal scalp regions and is associated with the early perceptual processing of faces. [Chen et al.](#_bookmark53) ([2024](#_bookmark53)) found a non-linear modulation of EEG responses to the realness of face images, suggesting that the brain’s processing of facial stimuli is sensitive to their perceived authenticity. Similarly, [Schindler et al.](#_bookmark89) ([2017](#_bookmark89)) tested six face-stylization levels varying from abstract to realistic and found that the N170 was generated more occipitally for abstract/virtual faces than for real faces. These findings suggest that small deviations from typical human facial expressions can lead to altered neural processing. This also highlights that the mechanisms underlying differences in neural processing of emotions of real and virtual faces remains unclear, and more neuroimaging studies are needed that systematically compares the physical characteristics of avatars and real faces.

* 1. **Functional Near-Infrared Spectroscopy (fNIRS)**

fNIRS is a non-invasive neuroimaging technique that measures brain activity by detecting changes in Blood Oxygenation Level Dependent (BOLD) signals, which are associated with neural activity, similar to functional magnetic resonance imaging (fMRI). fNIRS works by shining near-infrared light (760-850nm) through the scalp and measuring the amount of light that is absorbed by oxygenated (HbO) and deoxygenated hemoglobin (HbR) in the brain. This is possible through the Modified Beer-Lambert Law, which relates the concentration of hemoglobin to the absorption of light ([Kocsis et al.](#_bookmark74), [2006](#_bookmark74)). It is substantially more portable and cost-effective than MRI, tolerates moderate partic- ipant movement, and can be deployed in more ecologically valid or naturalistic settings ([Yu¨cel et al.](#_bookmark98), [2017](#_bookmark98)). Temporal resolution is moderate, on the order of seconds, which, although inferior to EEG’s millisecond fidelity, remains sufficient to capture the hemo- dynamic responses associated with emotional and cognitive processes. Despite these advantages, fNIRS remains limited to superficial cortical regions; it lacks sensitivity to

deeper subcortical structures such as the amygdala or insula, which play key roles in emotion processing ([Sato et al.](#_bookmark88), [2004](#_bookmark88)). Its spatial resolution is also lower than fMRI’s, and signal quality can be influenced by factors like hair density and skin pigmentation ([Holmes et al.](#_bookmark68), [2024](#_bookmark68)). Beyond systemic noise, fNIRS signals can also be affected by light in the recording environment and interference from participant hair; these issues can be minimized through careful preparation and room setup. These limitations are mit- igated through methodological refinements, such as high-density optode arrangements, short-separation channels ([Scholkmann et al.](#_bookmark90), [2014](#_bookmark90)), and motion correction techniques ([Fishburn et al.](#_bookmark61), [2019](#_bookmark61); [Bergmann et al.](#_bookmark51), [2023](#_bookmark51)).

Analysis methods such as the General Linear Model (GLM), and functional connec- tivity metrics allows for the identification of distributed activation/connectivity patterns within the cortical regions accessible to fNIRS. In a standard GLM approach to fNIRS analysis, brain activation by convolving the experimental design (onset and offsets of stimulus presentations modeled as a boxcar or impulse function) with a canonical hemodynamic response function (HRF) to estimate stimulus-evoked responses in cortical regions ([Tak and Ye](#_bookmark95), [2014](#_bookmark95)). Neural activity recorded using fNIRS tends to be noisy, correlated with physiological signals, is not independent across channels, and is non-uniformly distributed, however, the GLM approach is well suited for analyzing fNIRS data due to its ability to deal with this noise ([Huppert](#_bookmark71), [2016](#_bookmark71)). The GLM can then be used to estimate the activation of specific brain channels/regions in response to different stimuli, and to contrast these activations across conditions, such as real versus virtual faces or different emotional expressions. Functional connectivity analysis , on the other hand, examines the temporal correlations between different brain regions, providing insights into how these regions interact during emotional processing. Functional connectivity can be assessed using various methods, including coherence, phase-slope index, and Granger causality ([Bastos and Schoffelen](#_bookmark49), [2016](#_bookmark49)). The most common method for fNIRS functional connectivity analysis is the Wavelet transform coherence (WTC), having been employed in 90 fNIRS studies ([Hakim et al.](#_bookmark65), [2023](#_bookmark65)). WTC is calculated by convolving the signals with a wavelet function, such as the Morlet wavelet. WTC measures the strength of shared frequency components between signals in the time-frequency domain, allowing for the assessment of how connectivity patterns change over time, a key advantage when analyz- ing non-stationary physiological signals from fNIRS. Additionally, WTC can detect both in-phase and out-of-phase relationships between channels, which is particularly valuable for distinguishing neural signals from physiological noise.

* 1. **Objectives and Hypotheses**

This thesis aims to investigate the neural differences in how humans perceive emotional expressions in real versus virtual faces. We used fNIRS to measure brain activation and functional connectivity while participants viewed both real and virtual faces expressing

various emotions. We hypothesize that 1) there will be significant differences in activation patterns and functional connectivity profiles when comparing virtual faces to real faces,

2) different emotional expressions will elicit distinct activation patterns and functional connectivity profiles. We have no a priori predictions regarding the specific nature of these differences, as they may vary based on the emotional content and the realism of the faces. To date, no study has employed a fully crossed factorial design that systemat- ically examines all basic emotions across both real and virtual face types within a single experiment, particularly using fNIRS. Most prior research has concentrated on a limited subset of emotions (typically fear or anger) and often only with avatars. As a result, our understanding of the interaction between emotion type and face realism remains incom- plete across the full range of basic emotions. In doing so, this research contributes to the broader understanding of emotional cognition in the age of digital interaction and informs the design of emotionally expressive avatars for applications in education, mental health, and human-computer interaction.

Critically, fNIRS demonstrates strong sensitivity to the prefrontal cortex (PFC), a region heavily implicated in the perception, interpretation, and regulation of emotion ([Westgarth et al.](#_bookmark96), [2021](#_bookmark96); [Bendall et al.](#_bookmark50), [2016](#_bookmark50)). Although prior studies have examined facial emotion perception on real and avatar faces independently, few have explored differential neural processes within the same neuroimaging paradigm, with even fewer studies have employed fNIRS to do so. To our knowledge, no existing research has directly compared neural responses to emotional expressions in real versus virtual faces using a within-subject fNIRS design. This gap limits our understanding of how face realism and emotion interact to shape cortical activation patterns and functional connectivity during social perception. Addressing this gap will provide insights into how digital representations of human emotion are processed and perceived in the brain.