| | |
|---|---|
| Title | Gaze and mirror : arousing emotional self-awareness through gaze-reflecting interactive device |
| Sub Title | |
| Author | ウ, 怡(Wu, Yi)<br>山岡, 潤一(Yamaoka, Junichi) |
| Publisher | 慶應義塾大学大学院メディアデザイン研究科 |
| Publication year | 2025 |
| Jtitle | |
| JaLC DOI | |
| Abstract | |
| Notes | 修士学位論文. 2025年度メディアデザイン学 第1180号 |
| Genre | Thesis or Dissertation |
| URL | https://koara.lib.keio.ac.jp/xoonips/modules/xoonips/detail.php?koara_id=KO40001001-00002025-1180 |

Master's Thesis
Academic Year 2025

# Gaze and Mirror: Arousing Emotional Self-Awareness through Gaze-Reflecting Interactive Device

Keio University
Graduate School of Media Design

Yi Wu

A Master's Thesis
submitted to Keio University Graduate School of Media Design
in partial fulfillment of the requirements for the degree of
Master of Media Design

Yi Wu

Master's Thesis Advisory Committee:

| | |
|---|---|
| Associate Professor Junichi Yamaoka | (Main Research Supervisor) |
| Project Senior Assistant Professor Giulia Barbareschi | (Sub Research Supervisor) |

Master's Thesis Review Committee:

| | |
|---|---|
| Associate Professor Junichi Yamaoka | (Chair) |
| Project Senior Assistant Professor Giulia Barbareschi | (Co-Reviewer) |
| Professor Kai Kunze | (Co-Reviewer) |

Abstract of Master's Thesis of Academic Year 2025

# Gaze and Mirror: Arousing Emotional Self-Awareness through Gaze-Reflecting Interactive Device

Category: Design


## Summary

Gaze is one of the most powerful nonverbal channels of human communication, conveying attention, intent, and emotion simultaneously. Humans can detect subtle changes in eyes within a few hundred milliseconds, but it is often difficult to realize the emotional subtext they convey. Friendly glances can promote trust and connection, while inappropriate or hostile glances can often trigger stress, discomfort and even harassment, especially for minority groups.

This study aims to investigate a smart wearable device that recognizes, detects and responds to the emotions conveyed by a gaze. When viewers cast different types of gazes, the device will provide instant feedback, exposing the emotions hidden in the gaze through different colors. This study hopes that by providing instant visual transformation functions (e.g., color shifting, specular reflection), it will open up a new avenue of research: they can serve as "social mirrors" that return to people's sight in a tangible form, thus prompting them to awaken their self-awareness or engage in self-reflection.



## Keywords:

gaze interaction, emotion recognition, emotion visualization, colour feedback



Keio University Graduate School of Media Design

Yi Wu

# Contents

# List of Figures

# List of Tables

# Acknowledgements

Foremost, I would like to express my deepest gratitude to my supervisor, Professor Junichi Yamaoka, who was a cornerstone of my research journey. His attentive guidance, professional insight, and exceptional patience supported me through every stage of the process. From programming and wiring to overcoming technical hurdles I was initially clueless about, his help was invaluable in navigating the many challenges of prototyping.

I am also sincerely grateful to Professor Giulia Barbareschi, whose support during the prototyping phase and fresh perspective on research inspired me to explore new directions.

To my friends and my cohort at KMD, thank you for helping me through the early, often confusing stages of my thesis. Your guidance, encouragement, and gentle corrections filled in the many knowledge gaps left by my ignorance and condescension.

Finally, I'd like to thank my families—especially my cat, Tony. Not only did he patiently endure the unsolicited endless petting I inflicted on him , but he also "assisted" in removing support structures from 3D-printed prototypes at critical times. He may know nothing about academic research, he has undeniably been my most plenteousl source of emotional support.

# Chapter 1
# Introduction

This research aims to develop a smart wearable device to recognize, detect, and respond to emotions conveyed under gaze. These materials will provide instant feedback when viewers cast a malicious gaze or vision and send their own glance back.

Gaze is one of the most powerful non-verbal channels in human communication, conveying attention, intention, and emotion simultaneously [1]. Although a friendly gaze can foster trust and bonding, an inappropriate or hostile gaze often triggers stress, discomfort, or even harassment, particularly for minority groups. Recent neuroethological studies demonstrate that humans can detect subtle changes in eye behaviour within a few hundred milliseconds, yet they are frequently unaware of the emotional subtext they project [2].

Conventional emotion-recognition systems focus on detecting emotions and logging data, but seldom provide real-time, in-situ feedback that helps users correct or regulate their own gaze. The smart materials that this research developes would offer a capability of instant visual transformation (e.g. colour shifts, mirror-like reflection) to open a new research avenue: they can act as a "social mirror," returning one's gaze in a tangible form and thereby prompting immediate self-reflection.

## 1.1. Background and Motivation

As I started to gain awareness of feminism, I came to understand how powerful different gazes can be while managing to understand contemporary feminist writings—particularly Laura Mulvey's "male gaze" essay and Bell Hooks' notion of the "oppositional gaze"—I noticed how often everyday interactions echoed those theories. On packed trainsin busy streets, or any confined setting, I observed that a prolonged gaze, though silent, could freeze the person being watched; women or people who did not conform to mainstream gender expression would shift positions or raise a phone as a barrier, yet the observer often seemed to be oblivious or precisely speaking, cavalier of the discomfort created. I have certainly been in this situation many times myself, and the transition from pretending I didn't know and suffering in silence to just staring back has undeniably required conscious effort and emotional labor. It also made me realize that even when staring back, some gazers will immediately divert their eyes, while others will not at all; and often, when that gazer doesn't stop staring, the moment our eyes meet, a comical and ridiculous competition about gaze begins.

These moments clarified a key feminist claim I once read: patriarchy is essentially the imbalance of power structures, a societal arrangement where one side dominates the other. It doesn't only refer to a society where males hold dominance, individuals of any gender can be victims of patriarchy. And male gaze is a mechanism of power, in which women are objectified in the process of being gazed at and become vessels of the gazer's desires. It is less an aesthetic than a practice of possession and domination [3]. Fear is the central mechanism by which the gaze operates, so the superior doesn't think that gaze will have much of an effect, subconsciously they default to the gaze as a birthright; thus an equivalent gazing by the inferior tends not to have an equivalent effect (discomfort, fear, shame, etc.) on the superior [4] [5]. In short, the problem was less about overt aggression and more about an *asymmetry of awareness*: the person looking seldom felt the emotional weight of their own gaze, whereas the recipient felt it instantly. That asymmetry resonated with Mulvey's claim that visual culture privileges the viewer's pleasure while objectifying the viewed [6]. It also highlighted a gap in current technology: abundant tools measure gaze for marketing or usability, yet few return any feedback to the user about how their gaze might affect others.

Public–space research over the past decade shows that *being watched* can itself constitute a form of harassment: large-scale surveys in five continents report that silent or lingering gaze is one of the most frequently cited sources of discomfort or fear, often ranking above verbal catcalling or physical proximity. A 2016 UK poll conducted by Stop Street Harassment and YouGov ( n2,000) found that 64

% of women—and 85 % of those aged 18–24—had experienced unsolicited stare-based harassment in everyday settings [7]. In parallel, ActionAid's multi-city "Safe Cities for Women" study (London, Delhi, Rio de Janeiro, Bangkok; 2016) reported prevalence rates between 75 % and 89 % for unwanted staring in public transport and streets [8]. Similar patterns emerge outside Europe: an Australian national survey (Australia Institute, 2015) shows that 65 % of women have been persistently stared at in public, and 40 % therefore avoid walking alone at night [9]. Even micro-contexts such as gyms are affected: a 2021 UK study of 2 000 female gym-goers indicates that 76 % modify clothing or behaviour to avoid intrusive gaze [10]. On a global scale, Gallup World Poll data (143 countries, 2011–14) reveal consistent gender gaps in perceived outdoor safety, with disproportionate anxiety linked to the possibility of hostile observation rather than physical assault [11].

From a feminist perspective, these numbers are unsurprising. Mulvey's classic *male gaze* thesis contends that mainstream visual culture privileges a distinct power asymmetry: men look, women are looked at [6]. Foucault's "panoptic" logic extends that asymmetry into a system of self-regulation, whereby the potential of constant observation disciplines behaviour [5]. Contemporary critics such as hooks highlight how marginalised bodies develop an "oppositional gaze" in resistance to the normative view [12]. Yet in daily interaction, individuals—especially those in majority positions—rarely recognise the emotional content embedded in their own eye behaviour. Advances in computer vision now allow millisecond-level decoding of fixation duration, saccade amplitude and pupil dynamics; however, most commercial eye-tracking applications focus on passive analytics or marketing, not on giving the user immediate, introspective feedback [2].

To conclude, gaze, or in other words, vision or ways of looking, as our primary sense, plays an important role in how we perceive and communicate with the world. Realizing and understanding emotions beneath it is essential for healthy communication. However, people often overlook the impact of their gaze on others. Visual interactions that convey unintended or harmful emotional signals often lead to discomfort, stress, or insecurity. And as said before, it is highly possible that some finds it difficult to notice their gaze, or realize the negative emotions or expressions that come along with them. Experiences and data above convinced me that a technical intervention was worthwhile only if it could (i) operate in real time, (ii) externalise the otherwise invisible emotional signal embedded in gaze, and (iii) do so in a way that fosters reflection rather than public shaming. The design challenge, therefore, is not merely to detect eye movements but to translate them into a subtle, self-facing cue—using smart materials—so that the wearer confronts their own visual behaviour in the moment it occurs. Therefore, the present study asks whether a discreet wearable device that *reflects* affects the affective tone of gaze back to the wearer - via instant color-coded visual cues - can close this awareness gap. By merging real-time emotion inference with smart materials, the project aims to help users perceive the emotional footprint of their gaze, and reduce inadvertent negative signalling in public encounters. Ultimately, the work seeks to contribute an evidence-based design strategy to foster more respectful and psychologically safe visual interaction.

## 1.2. Research Questions

To address the above motivations, the study is guided by two primary research questions:

- How can real-time, gaze-reflecting smart devices enhance emotional self-awareness through gaze-reflecting?
- Which color-emotion mapping strategy most effectively moderates negative or uncomfortable gaze and promotes respectful communication?

In this research, I'm using eye-tracking technology to detect and analyze emotions through a person's gaze. The smart material responds instantly by changing its color or texture based on the emotions identified. This creates a real-time feedback loop, where the material's visual changes help people become more aware of their emotions hidden in their gazes. Additionally, an important aspect of this research is exploring how these changes in the material might, in turn, influence the observer's emotional state—sometimes amplifying or moderating their reactions.

I hope this research can address the hidden emotions communicated through gazes to foster better emotional awareness and mutual respect.

## 1.3. Structure of the Paper

This thesis is organized into five chapters. Chapter 1 provides the background and motivation for this research, highlighting the importance of gaze and its emotional content in social interactions. It outlines the issues related to uncomfortable or harmful gazing, especially in public spaces, and discusses the potential of smart materials to enhance emotional self-awareness through gaze reflection.

Chapter 2 reviews related works in the areas of gaze theory, emotion recognition through gaze, color psychology and the use of smart materials for emotional feedback.

Chapter 3 presents the concept design, detailing the development of the gaze-reflecting system. This chapter covers the design process from its initial concept to the final prototype, including the development of gaze detection algorithms, the selection of appropriate smart materials, and the ethical considerations involved in providing real-time feedback to users.

Chapter 4 introduces the validation of the final prototype, including the objectives of the experiment, the design of the experimental setup, and the procedure for data collection. This chapter also discusses the results of the experiment and how they contributed to the final design iteration of the system.

Chapter 5 concludes the thesis by summarizing the findings of the investigation, discussing the implications for future work, and highlighting the limitations of the current study. Future research directions are also proposed to further explore the potential applications of gaze-reflecting systems and emotional feedback technologies.

# Chapter 2
# Related Works

## 2.1. Theories of Gaze

This section provides a summary of the current literature on the role and impact of gaze theory on people in social contexts. The definitions, emergence, and meanings of gaze are explored, as well as gaze behavior, gaze perception differences, and their relationship to factors such as social anxiety and persuasion.

### 2.1.1 Definitions of Gaze

The concept of gaze has been extensively explored and differently defined across various scholarly domains, primarily within sociology, anthropology, and perceptual psychology. A comprehensive understanding of gaze necessitates integrating these diverse disciplinary perspectives to establish a clear operational definition relevant to the present research.

From a social-anthropological point of view, gaze is primarily characterized by its social signaling capacity. Berger (1972) notably introduced gaze within the context of cultural critique, emphasizing how gaze functions as a socially constructed mechanism that embodies and mediates perceptions of dominance, power dynamics, and social interactions [4].

Argyle and Cook (1976) defined gaze in interpersonal interactions as "the direction of looking, usually defined by the line of sight; mutual gaze occurs when two people look directly at each other's eyes." They emphasized gaze's dual communicative function: acquiring visual information and signaling interpersonal intentions and emotions [13].

Emery (2000) expanded upon this by defining social gaze explicitly as "the direction, duration, and pattern of eye movements during social interaction, encompassing fixations, saccades, and periods of mutual or averted eye contact," highlighting its role in conveying attention, intention, and emotional states [1].

Gibson and Pick (1963) offered a perceptual definition of gaze as "any visible orientation of the eyes and head that specifies to an observer where attention is directed," emphasizing the observable, physical attributes of gaze [14].

Gobel, Kim, and Richardson (2015) further integrated the perceptual and cognitive views by describing gaze as simultaneously a perceptual mechanism for sampling visual information and a communicative social signal [15]. Cañigueral and Hamilton (2019) provided a precise operational definition by articulating gaze as comprising "fixation sequences and overt shifts (saccades) between them that people deploy to interpret and signal social information."

For practical measurement in experimental contexts, gaze is frequently defined through measurable eye-tracking metrics. Holmqvist et al. (2011) operationalized gaze as sequences of fixations and saccades, with fixations typically defined by eye stability within approximately 1-degree visual angle for at least 100–200 milliseconds [16].

Similarly, Tarnowski et al. (2020) characterized gaze data as timestamped coordinates of pupil centers on screens, producing identifiable fixations and rapid saccadic jumps [2].

Also, In Kim et al.'s study, gaze is defined operationally as the spatial distribution of a user's visual attention on a device screen, quantified via point-of-gaze (PoG) and dwell time in specific interface regions. No high-frequency fixation or saccade segmentation is used; rather, gaze features are limited to those easily accessible via commodity webcams or mobile cameras [17].

In summary, gaze definitions span a spectrum from macro-level social communication to micro-level quantitative eye movement metrics. This thesis adopts an integrative definition of gaze, grounded both in its social function and measurable physiological parameters, viewing it as sequences of fixations and saccades oriented toward socially meaningful visual targets (primarily another individual's eyes or face). This integrative definition aligns the social-anthropological understanding of gaze's communicative intent with the perceptual-cognitive psychology framework of measurable eye movement events, thereby providing both theoretical depth and empirical precision suitable for analyzing gaze within real-time emotional interaction scenarios.

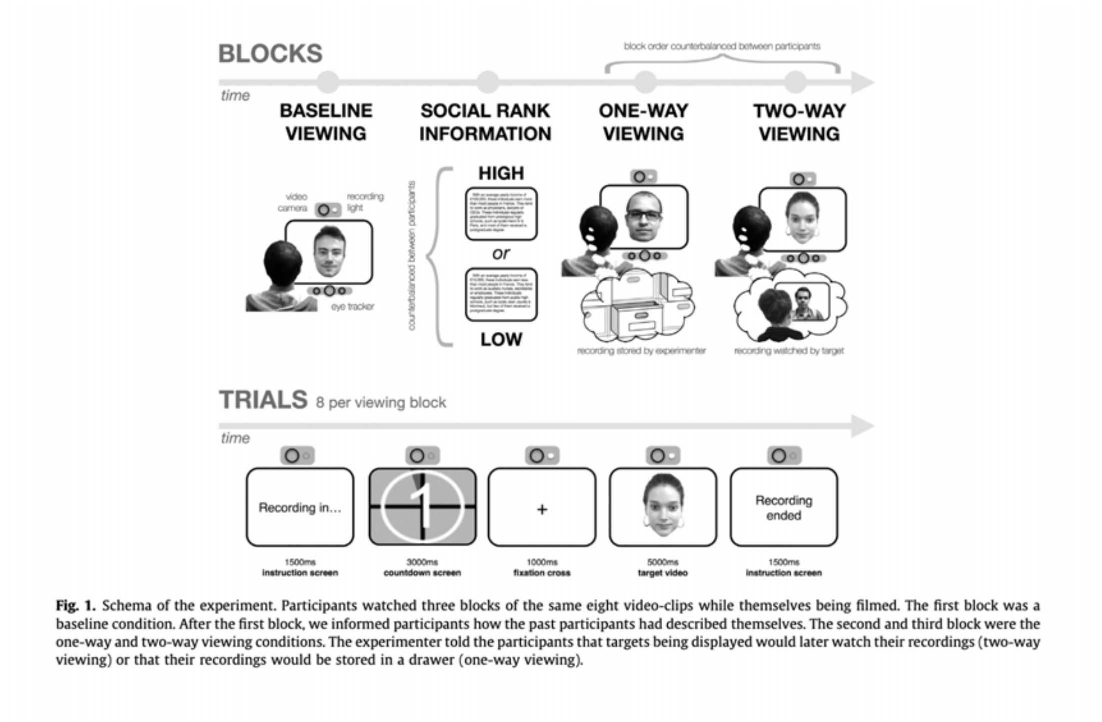
### 2.1.2 Role and Impact of Gaze in Social Contexts

The role of gaze in social interaction goes beyond just signaling interest or dominance; it is a complex and multifaceted tool, constantly shaped by the context in which it is used. Its ability to influence perceptions of power, intimacy, and even threat highlights its significance in everyday human encounters. While these dynamics make gaze a powerful tool in communication, they also introduce a layer of ambiguity in its interpretation, as the same gaze behavior may convey entirely different messages depending on the social context and the relationship between the individuals involved. Understanding this complexity is essential, as it can help design interventions, such as real-time gaze-feedback systems, that promote awareness and reduce unintended discomfort caused by misinterpreted gazes.

According to the seminal study of Gibson and Pick (1963), humans are highly sensitive to changes in the direction of the gaze of others, which makes it a vital component of social perception [14]. This sensitivity to gaze direction has been shown to be crucial in human social interactions, where gaze serves as a fundamental nonverbal channel of communication [15].

The white sclera in human eyes, a distinguishing feature of our species, has been suggested to facilitate the detection of eye movements, thus aiding in social learning and providing a reliable mechanism for referential communication, where individuals signal their attention to objects or events in their environment [1]. There are regions of the brain that specialize in the processing of gaze, such as the superior temporal sulcus (STS), the amygdala, and the orbitofrontal cortex, which play key roles in the processing of gaze and its socio-emotional significance assignment [18]. This ability to perceive gaze direction helps humans understand both the physical world and the social intentions behind others' behavior, creating a dual-purpose function of gaze: it provides information about the environment and signals the observer's intentions.

Direct gaze, in particular, is often linked to attention, attraction, and openness in social interactions. When maintained for longer periods, it can indicate a variety of social meanings, such as physical attraction, nonverbal dominance, or even an assertion of power. In some social contexts, prolonged eye contact is




(Source:The dual function of social gaze [15])

Figure 2.1 Schema of the experiment of Gobel, Kim and Richardson

perceived as a challenge or a threat, with the potential to evoke discomfort or conflict, especially when used in competitive or hostile situations [19]. In situations involving social hierarchies, when participants perceived that their gaze would be seen by higher-ranking figures in the video, they were less likely to look into the eyes of these higher-ranking figures, which could be a signal to avoid challenging their status [15]. Moreover, gaze plays an integral role in non-verbal intimacy behaviors. According to the equilibrium model of intimacy proposed by Argyle and Dean (1994), people maintain a certain balance of intimacy through behaviors such as eye contact, physical proximity, and smiling. Any change in one of these behaviors often triggers compensatory adjustments in others, ensuring that emotional distance remains within acceptable limits. When intimacy balance is disrupted, people compensate or reciprocally adjust through behaviors such as gaze [13].

Research on social anxiety related to gaze behavior and perceived differences also highlights the importance of the research topic. Individuals with higher levels of social anxiety tend to gaze less at faces during social interactions and significantly overestimate the extent to which they make eye contact with others. This kind of overestimation can exacerbate anxiety and distort social perceptions [18]. It has also been shown that individuals with high social anxiety have also been shown to pay more attention to the face during the first two seconds of a social scene, but there is no significant difference in gaze-following behavior. Although some studies have argued that socially anxious individuals show hypervigilance or avoidance of threatening social stimuli, the literature provided has limited support for these hypotheses in natural social situations, preferring to argue for the presence of an early bias for the opposite face and the possibility that the opposite face may be processed more quickly [20]. In addition, social anxiety was associated with reduced facial gaze during natural face-to-face interactions. This gaze avoidance was less pronounced when speaking than when listening, and more pronounced in low intimacy topics than in high intimacy topics. This suggests that socially anxious individuals consistently perceive the risk of negative evaluations in authentic social interactions that require self-disclosure, leading to facial gaze avoidance [21]. Also, constraining gaze in the eye region (i.e., forcing the eyes to look away) significantly increases activation in "social brain" regions, and this effect varies according to personality traits such as mood and anxiety, alexithymia, and autistic traits. Anxiety exacerbates the effects of this constrained gaze on emotional facial responses, particularly to anger, which represents a direct threat [22].

Gaze also plays an important role in therapies.A deeper understanding of differences in gaze perception could improve the efficacy of cognitive behavioral therapy (CBT) for social anxiety, as CBT can help individuals correct for cognitive biases, including their overestimation of eye contact. Gaze behavior is also a key biomarker for the diagnosis of disorders such as autism spectrum disorders (ASD) and schizophrenia, so research on gaze perception could also inform interventions for these disorders [18].

## 2.2. Gaze Behavior and Emotional Expression

As mentioned before, the meaning and function of gaze behavior are highly dependent on the specific social context. Research on gaze behavior during social interactions must also be an important foundation for this thesis.

### 2.2.1 The Belief in Being Watched

The study from Cañigueral R and Hamilton AFC demonstrates that when an individual believes their gaze can be perceived by others, often referred to as the "audience effect" or "being watched", their gaze behavior significantly changes to align with social norms or to convey specific signals. This "belief in being watched" is a critical modulator of gaze behavior and social cognitive processes. They use several cognitive theories attempt to explain how this belief influences human interaction, drawing upon the "Watching Eyes Model" and "Reputation Management Theory", as well as related findings [23]:

- The Watching Eyes Model proposes that direct gaze, acting as an ostensive communicative cue, automatically captures attention and enhances an individual's sense of self-involvement and self-awareness . This "self-referential processing" can then influence behavior, such as promoting prosocial actions. Critically, this model posits that it is not merely seeing a pair of eyes, but the belief that those eyes can see you, that triggers self-referential processing and self-awareness.
- Reputation Management Theory adopts an "other-focused perspective", suggesting that individuals modify their behavior when observed to maintain or enhance their social reputation. This includes exhibiting increased prosociality in economic games, particularly when the observer can reciprocate. The theory also encompasses adherence to social norms, such as "civil inattention," where individuals briefly acknowledge others in public but then avert their gaze to avoid implying excessive interest or initiating unwanted interactions. Gaze patterns are further modulated by social status, with individuals potentially gazing less at higher-ranked targets to avoid challenging their status, and more at lower-ranked targets to signal their own

superior position when observed.

- The Interpersonal Gaze Processing Model integrates the “sensing” and “signalling” functions of gaze within real-time social interactions. This model highlights how gaze patterns are dynamically adjusted to coordinate with other verbal and non-verbal signals like speech, turn-taking, and facial expressions. Gaze serves a ”monitoring function,” allowing individuals to track a partner’s attentional and emotional states to ensure mutual understanding and seek social approval. Mutual gaze is also linked to increased neural synchrony, facilitating information transfer and effective communication.

Implicit measures consistently suggest that direct gaze increases affective arousal and evokes a positively valenced affective reaction. However, explicit self-evaluations can be inconsistent, sometimes reporting less positive feelings, possibly due to heightened self-directed attention or discomfort from scrutiny.

### 2.2.2 Emotional Expression through Gaze

Studies that validate the correlation between gaze and emotion are crucial to this thesis.

There are researches consistently demonstrates that eye-tracking data, particularly features related to eye movements and pupil diameter, can serve as indicators of emotional states [24] [2] [25]. The pupil, for instance, reacts not only to light intensity but also to emotions [2] [1]. Studies have shown that pupil diameter increases when an individual feels excited [2] and generally with positive valence. Conversely, some research suggests that pupil diameter and fixation duration increase in proportion to negative valence [25]. Eye movement signals provide cues about subconscious behaviours and can reveal what is attracting a user’s attention, or what might be causing happiness or upset [24].

And as mentioned above, findings on explicit vs. implicit affective responses to gaze are also important [26] [23]. The results of studies on explicit affective evaluations of direct versus avoidance gaze are often inconsistent. Some studies report more positive subjective feelings towards direct gaze, while others observe the opposite, with contextual factors significantly influencing these perceptions. For instance, a direct gaze in a positive context (e.g., during a positive interview) might lead to more positive evaluations of the interaction and interviewer, while in a negative context, an averted gaze might be evaluated more positively. It has been suggested that the more realistic the stimuli resemble natural interaction, the less likely direct gaze is to evoke relatively more positive explicit feelings compared to averted gaze, possibly due to enhanced feelings of self-involvement or uneasiness from being watched. In contrast, implicit measures of affective responses to gaze tend to yield more consistent results. Direct gaze typically increases affective arousal and automatically evokes a positive affective reaction. Studies using the Implicit Association Test (IAT) have shown that individuals implicitly associate faces with direct gaze with positivity and averted gaze with negativity, even when faces display angry expressions or when the stimuli are non-facial. Furthermore, people are willing to exert more effort to view "direct gaze movies" compared to "averted gaze" or "object movies," suggesting a more positive affective reaction to direct gaze. The perception of direct gaze automatically activates more positive evaluations than seeing closed eyes [26].

In addition, the perception of emotions by ear and by eye from Gelder and Vroomen (2000) explored the perception of emotions through both visual and auditory stimuli, emphasizing the critical role of visual inputs, such as gaze, in emotional perception. Their study revealed that facial expressions, particularly those involving the eyes, are powerful conveyors of emotions and are often more reliable indicators of emotional state than auditory cues. It also underscores the significance of visual channels, especially eye movements, in emotional communication and supports the idea that smart materials can effectively use visual feedback to reflect the emotional state of the observer [27].

## 2.3. Eye-Tracking Technology for Emotion Recognition

Eye-tracking (ET) technology has been widely utilized across cognitive, psychological, and affective research disciplines due to its non-invasive nature and its ability to objectively quantify visual attention patterns and behaviors. Despite its popularity, its practical application in accurately recognizing emotional states remains a complex and often contested area within the research community.

### 2.3.1 Capabilities and Common ET Metrics

Eye-tracking methodologies typically use infrared-based corneal reflection techniques to precisely measure eye movements such as gaze, including fixations (periods of stable visual attention), saccades (rapid eye movements between fixations), and pupil dilation. Such eye-tracking systems have shown effectiveness in quantifying cognitive load, visual attention, and emotional arousal through fixation duration, pupilometry, gaze dispersion, and scanpath entropy.

Pupil dilation, in particular, is well documented to reflect emotional arousal and cognitive engagement. In addition, fixation and saccade patterns can help identify shifts in attentional focus, providing indirect indicators of affective states [16] [2].

The recent work by Skaramagkas et al. (2023) introduced the ESEE-D dataset for emotional state estimation based on eye-tracking data. The authors extracted a variety of gaze-related features—including fixation duration, saccade amplitude, blink rate, and pupil diameter—and used these as input to machine learning models for classifying emotions such as happiness, sadness, and anger. Their results demonstrated moderate classification performance, but also emphasized that high inter-subject variability and noise sensitivity limit the generalizability of such models [24].

### 2.3.2 Extensive Limitations of Single-Modality Gaze Detection

Despite some encouraging results, extensive literature consistently underscores critical limitations in using single-modality gaze detection methods for emotion recognition.

A core limitation lies in the inherent ambiguity of gaze behavior.Gaze behaviors can indicate general attentional shifts or arousal but cannot uniquely determine specific emotional experiences [2]. Gaze may function as a cue for mental state inference, but not as a definitive emotional signature. Holmqvist et al. (2011) explicitly stated that while fixation duration and saccadic behaviors might correlate loosely with broad emotional or attentional states, these metrics inherently lack the specificity required to identify discrete emotional categories [16].

Moreover, Cowen et al. (2020) argued through empirical studies that eye movements modulate and reflect affective states but cannot independently signal or classify emotions accurately. Their study emphasized that gaze patterns can vary significantly across individuals and situations, making generalized emotion detection via gaze alone highly unreliable [28].

Zhang et al. (2015), examining appearance-based gaze estimation, also highlighted substantial accuracy constraints of gaze tracking using webcams and other low-cost commercial devices. The authors confirmed that these methods are significantly influenced by environmental variables such as lighting, camera resolution, user posture, and calibration issues, limiting their practical application to coarse-grained attention detection rather than precise emotional classification [29].

Further limitations are noted by Tarnowski et al. (2020), who indicated that even sophisticated statistical models and machine learning algorithms achieve only moderate accuracy when relying solely on gaze features. They emphasized that high-quality emotion recognition requires combining gaze data with complementary expressive or physiological modalities, such as facial action coding systems (FACS), heart rate, electrodermal activity, or electroencephalography (EEG) [2].

Practical application of eye-tracking for emotion recognition also faces technical constraints. Scientific-grade ET systems (such as Tobii or EyeLink) require controlled laboratory environments, precise participant positioning, extensive calibration processes, and high sampling rates (100–1000 Hz). These demanding conditions greatly restrict their real-world applicability and ecological validity [16].

Webcam-based and low-cost consumer solutions, though more accessible, further amplify accuracy concerns, significantly limiting their efficacy in nuanced emotional recognition tasks [29] [17]. Thus, neither high-precision nor low-cost gaze-based methods alone satisfy the rigorous criteria needed for reliable emotional inference in diverse and naturalistic contexts.

Another challenge is the lack of consistent ground-truth labeling in gaze-emotion datasets. Emotion is inherently subjective, and labeling depends on either self-report (introducing bias) or observer inference (introducing ambiguity). This leads to difficulty in generalizing models across populations and contexts.

In conclusion, extensive academic literature underscores a fundamental consensus that single-modality gaze detection methods are inherently insufficient for reliable and specific emotion recognition. Although gaze metrics can offer valuable supplementary insights into emotional arousal and attentional dynamics, their standalone predictive capability remains limited.

## 2.4. Gaze and Emotion Detection on Smart Devices

Eye gaze analysis has increasingly been explored as a digital biomarker for emotional state detection, especially in the context of consumer-grade devices such as smartphones and webcams. However, the scientific literature consistently indicates that while gaze features—such as point-of-gaze location and dwell time—can offer valuable insights into user engagement and affective response, they are generally insufficient for robust and fine-grained emotion recognition when used in isolation.

A representative recent study by Kim et al. (2023) specifically investigated whether gaze-based features extracted from off-the-shelf smart devices (e.g., mobile phones and tablets) can effectively improve emotional state detection. In their controlled experiment, participants used Instagram while their eye gaze data was collected via the device's front-facing camera. The authors extracted simple gaze features—including point-of-gaze direction and dwell time in predefined screen regions—and trained a machine learning classifier to distinguish positive and negative emotional valence.

The key finding was that gaze-based features alone enabled binary emotion (valence) classification with up to 76.4% accuracy in their limited laboratory setting. However, the authors emphasized that this performance, while promising, does not generalize to nuanced, multi-class emotion detection or to more naturalistic, real-world settings. They explicitly acknowledged that the predictive power of gaze alone remains limited and that integrating additional physiological or facial features is essential for advancing emotion recognition on consumer devices.

> "Although our results suggest that gaze data can aid valence detection, it is not sufficient for accurate or reliable emotion recognition in more complex scenarios. Multi-modal approaches incorporating facial expressions and physiological signals remain necessary." [17, p. 380]

Methodologically, their study is highly relevant to the present work because it demonstrates that real-time gaze tracking with basic camera hardware is feasible and can yield moderate predictive performance in affective computing [17]. However, the study also reinforces a major limitation highlighted by broader reviews [16]: gaze features alone, without integration of facial muscle dynamics, pupilometry, or multi-modal signals, are insufficient for reliable emotion detection.

In summary, recent research on gaze-based emotion detection in consumer contexts demonstrates that while gaze can supplement affective computing systems, its utility as a sole indicator of emotional state remains limited. Only through the fusion of gaze features with facial expressions and, where feasible, physiological signals can robust emotion recognition be achieved on ubiquitous smart devices.

## 2.5. Color Psychology and Emotion Regulation

This section synthesises current research on the intricate relationship between color, particularly colored light, and human emotions, focusing on its potential for emotion regulation. Drawing from studies on human-robot interaction, general color psychology, and display space design, this review examines how different dimensions of color (hue, saturation, and brightness) influence emotional states such as pleasure, arousal, and dominance. It highlights specific applications where colored light acts as a feedback mechanism in robotics and a mood-setting element in visual display environments. Furthermore, the review explores individual differences, such as technical affinity, experience with robots, and gender, which can mediate color-emotion associations. While acknowledging the strong connection between color and emotion, the review also addresses methodological considerations and limitations in current research, identifying avenues for future investigation to enhance the understanding and application of color in regulating human affective states.

### 2.5.1 Foundations of Color and Emotion

Color has a profound impact on human psychological states, perceptions, and behaviors. The field of color psychology investigates the emotional and psychological effects of different hues, saturations, and brightness levels. Various studies have demonstrated how colors can evoke specific emotional responses and influence mood, which is crucial for emotion regulation.

One foundational study by Valdez and Mehrabian (1994) found that colors can be associated with a variety of emotional states, such as red being linked to aggression and blue to calmness. Green has been widely associated with relaxation, while yellow tends to evoke feelings of happiness or warmth. These associations form the basis for integrating color in environments to modulate emotional experiences and improve emotional regulation, particularly in contexts like human-robot interaction or emotional feedback devices [30].

Moreover, Xie et al. (2022) support the notion that color, especially when used dynamically in environments, can enhance emotional awareness and facilitate better emotional regulation. For example, blue is commonly used in therapeutic settings to reduce anxiety, while green has been found to decrease stress and promote relaxation [31].

Research by Lee (2019) explores the impact of various LED light colors, such as red, green, blue, yellow, and purple, on emotional states. The study demonstrates that LED light colors have a significant effect on human emotional behavior, supporting the use of colors in regulating emotional responses. For example, red light can evoke excitement or aggression, while blue can promote calmness or sadness. These color-emotion associations provide valuable insight for designing interactive feedback mechanisms that reflect and modulate emotions [32].

### 2.5.2 The Impact of Color Dimensions on Emotion

The dimensions of color, including brightness and saturation, have a significant and well-documented impact on emotional responses. These two dimensions affect how individuals perceive and react to colors in different emotional contexts.

- Pleasure: Brightness has a considerably stronger effect on pleasure than saturation. Brighter colors such as white and yellow are typically perceived as more pleasant. On the other hand, black and gray are often associated with less pleasant emotional responses. Also, highly saturated colors, such as red and orange, tend to evoke more intense emotional reactions, creating a sense of excitement or urgency [33].

- More saturated colors are strongly linked to higher levels of arousal. For example, colors like red and yellow can increase emotional intensity, while blue or green are associated with a calmer, more subdued emotional state.

  Darker colors, such as black, generally induce higher arousal compared to lighter colors. However, a U-shaped relationship has been observed, where arousal decreases with increasing brightness up to a certain threshold, beyond which it starts to increase again. This suggests that very bright colors, like white, may moderate emotional responses [33].

- Dominance: Less bright (darker) and more saturated colors induce greater feelings of dominance. The effect of brightness on dominance is considerably stronger than that of saturation. Dominance generally decreases as colors range from dark to light, levelling off for the lightest colors. Black elicits the highest level of dominance, while white elicits the lowest. Darker colors are associated with greater perceived power and competence [33].

These findings align with the commonly recognized distinctions between warm and cool colors. Warm colors like red, orange, and yellow are typically associated with higher energy and more active emotional states, while cool colors, such as blue, are linked to calmness and relaxation.

### 2.5.3 Contextual Applications in Emotion Regulation

The understanding of color-emotion relationships has important practical implications in various domains, particularly in human-robot interaction (HRI), interior design, and therapeutic spaces. Color serves as a simple yet effective means of feedback that can be easily utilized in various settings to regulate emotions and communicate information.

In the domain of human-robot interaction, colored light is frequently used as an effective tool to convey emotional states or guide emotional regulation. For instance, blue light is often used to signal a calm state, while red can indicate a more alert or high-energy emotional state. These color cues are intuitive, non-invasive, and easy to perceive, making them ideal for use in mobile robots or virtual assistants where emotional feedback is crucial for creating a connection with the user [32].

Research conducted by Russell (2003) has emphasized the importance of color in emotion regulation. His work highlights that different colors can either calm or activate individuals based on the emotional context. Blue, for example, is widely known for its calming effect, while red can heighten arousal. Such color-feedback mechanisms in emotional regulation are particularly valuable in interactive systems that aim to modulate user emotional states in real-time [34].

## 2.6. Gap Analysis and Conclusion

Collectively, these studies illustrate the fundamental position of gaze in social interaction, the significant role of eye-tracking and gaze analysis in emotion recognition and their potential application in gaze-reflecting interactive wearable devices. These technologies can provide immediate and accurate emotional feedback, facilitating the development of innovative materials that enhance privacy and promote respectful interactions.

And, there are three main gaps my research aims to fill:

1. Limitations of Existing Methods: Current emotional recognition systems mostly focus on analysis and data but lack a real-time, interactive component.

2. Reflective Interaction: My project stands out by providing real-time, visual feedback that not only detects but reflects emotions back to users, encouraging awareness and mutual respect. Unlike existing studies, which focus on observation, this project introduces an interactive element that actively engages the observer.

3. Emotional Regulation: Unlike traditional systems that just detect emotions, I hope my project can influence emotional self-regulation actively.

# Chapter 3
# Concept Design

## 3.1. Design Approach

The main purpose of the research is to provide a method for people to better understand and observe their own gaze while paying attention to the emotions underneath and to promote mutual respect by exploring the feasibility of wearable interactive devices to influence emotional self-awareness.

## 3.2. Conceptual Development

The conceptual development of this research aims to combine theoretical models with practical design approaches to create an innovative system that detects and responds to gaze-related emotional cues. This section outlines the key models and concepts that form the backbone of the design process, including the gaze-reflective feedback concept, the emotional categorization model, the gaze detection and emotion recognition algorithm, the logic behind color-emotion mapping, and ethical considerations in feedback design.

### 3.2.1 Definition of Gaze in This Research Context

The review in section 2.1.1 has pointed out that gaze possesses interdisciplinary attributes in modern human-computer interaction (HCI), social psychology (social cognition), and affective computing. Looking back at classic studies—from Gibson & Pick's experiments on directional sensitivity [14], to Argyle & Cook's systematic exposition of the dual functions of social signals [13], to Cañigueral & Hamilton's revelation of the regulatory role of self-referential processing through the "audience effect" model [23]—different disciplines focus on different aspects of the same behavioral phenomenon:

- Social psychology emphasizes the interactive significance of gaze: its signaling functions related to power, intimacy, and threat.
- Cognitive neuroscience analyzes the neural pathways underlying eye movements (such as the STS, amygdala, and orbitofrontal cortex) for rapid decoding of social meaning.
- HCI/wearable design focuses on the technical feasibility of detecting, calculating, and providing feedback on gaze.

This study is located at the intersection of these three fields: we use the social psychological framework to understand the bidirectional meaning of gaze, employ computer vision and deep learning technologies for real-time capture, and ultimately visualize it through material feedback in a wearable context, emphasizing gaze as a socially meaningful direction of visual attention detected through computational methods, providing a basis for immediate reflective feedback.

Building upon the theoretical integration discussed, we specifically define gaze in our study as follows:

Gaze refers explicitly to the sustained visual attention (duration exceeding 200 ms) directed by one participant toward the facial region (face and eye area) of another participant during naturalistic face-to-face interactions, as estimated by appearance-based computational methods [2, 29].

Unlike traditional eye-tracking paradigms which focus extensively on precise fixation durations (usually $\geq$ 100 ms) and microsaccades, our definition emphasizes the directional aspect of gaze as a socially communicative act rather than purely as a cognitive indicator.Our rationale is twofold:

**Technological Feasibility:** This research utilizes an appearance-based gaze estimation algorithm, leveraging a single-webcam solution running at approximately 30 frames per second (fps), providing an angular accuracy ranging between approximately 3 to 5 degrees. This approach is intentionally selected over traditional infrared-based eye-tracking systems (requiring calibration and higher sampling frequencies), aligning better with practical real-world wearability and ensuring applicability in everyday social interactions.

**Interactive Real-time Feedback:** In our detection framework, a gaze event is defined by continuous observation of another individual's facial region for a minimum threshold of approximately 200 milliseconds. This specific temporal criterion effectively filters out brief saccadic movements or incidental visual scanning and ensures that LED-based feedback is activated only in socially meaningful gaze interactions.

The operational criteria for gaze detection employed in our experiment include:

| **Criterion Type** | **Operational Measure** | **Description** |
|---|---|---|
| Spatial | Facial ROI inclusion | Gaze vector computed via appearance-based methods must fall within a predefined facial region (face/eye area). |
| Temporal | Gaze duration threshold | Continuous gaze fixation must last $\geq$ 200 ms to trigger system feedback. |
| Social-Cognitive | Social context annotation | Controlled experimental conditions establish the interactive scenario explicitly as "socially evaluative," reinforcing gaze as a meaningful communicative act rather than purely perceptual. |

Table 3.1 Operational criteria for gaze detection.

The proposed operational definition thus bridges established psychological insights and practical computational feasibility, ensuring robustness and ecological validity in experimental implementations.

### 3.2.2 Gaze-Reflective Feedback Concept

The Gaze-Reflective Feedback Concept is the innovative heart of this research, combining the principles of gaze detection with real-time emotional feedback through smart materials.

### Limitations of Traditional Wearable Devices

Despite significant advancements in wearable technology and affective computing, most existing systems remain limited by their unilateral approach to emotion detection or expression. Traditional devices frequently operate in isolation, either exclusively monitoring physiological signals such as galvanic skin response and heart rate variability or passively presenting emotional states without active user awareness or direct engagement. Such limitations hinder their effectiveness in real-time social interactions, as individuals remain largely unaware of their non-verbal communicative behaviors and lack the immediate self-awareness necessary for behavioral adjustment.

Furthermore, existing gaze-tracking devices generally focus on collecting and analyzing eye-tracking data for diagnostic or analytical purposes, such as attention monitoring in virtual environments or psychological research paradigms. However, few if any current solutions proactively reflect the user's gaze back to them in real-time, thereby omitting a crucial opportunity for self-awareness and behavioral modification during ongoing social interactions.

### Necessity of Arousing Emotional Self-Awareness through Gaze

From a socio-psychological standpoint, gaze serves as a critical non-verbal communication tool with profound implications in human social behavior. As extensively reviewed in Chapter 2, gaze is intricately linked to social dominance, intimacy equilibrium, and social anxiety. Argyle and Cook's equilibrium model of intimacy (1976) describes gaze as a key regulatory mechanism in social interactions, indicating that individuals actively manage eye-contact duration to achieve an optimal level of intimacy and comfort. Disruptions in gaze behaviors—either excessive staring or gaze avoidance—frequently lead to interpersonal discomfort or miscommunication, reinforcing the essential need for increased individual awareness and self-regulation of gaze behaviors.

Moreover, extensive evidence highlights the psychological phenomenon whereby individuals, particularly those with elevated levels of social anxiety, frequently misjudge their gaze behaviors—often believing they have maintained gaze longer or more intensely than in reality [?, 20]. These cognitive misinterpretations exacerbate anxiety and distort interpersonal interactions. Therefore, increasing real-time awareness of gaze direction and intensity is critically beneficial, potentially reducing social anxiety symptoms by correcting erroneous self-perceptions through immediate feedback.

Besides, the "dual function of social gaze," outlined by Gobel, Kim, and Richardson (2015), underscores gaze as simultaneously a perceptual mechanism (sensing social cues) and a communicative tool (signaling intentions).

Additionally, as discussed in Chapter 1, prior research and large-scale social surveys underscore a pervasive gap in users' awareness of the emotional impact of their own gaze. The extensive evidence and lived experience demonstrate that gaze, far from being a neutral visual act, frequently imposes emotional and psychological costs on recipients—especially in contexts marked by power asymmetry or gendered dynamics [5, 6, 8]. However, those responsible for sending the gaze seldom recognize its affective impact in the moment. This persistent asymmetry of emotional awareness underscores a critical design challenge: most technological interventions passively monitor or quantify gaze, but fail to deliver real-time, introspective feedback to the user.

This research therefore advances the field by introducing a reflective feedback paradigm: rather than punishing or broadcasting the detected gaze to others, the proposed system subtly returns affective cues to the wearer themselves—encouraging personal awareness and emotional reflection in situ. In this way, the device addresses the "asymmetry of awareness" highlighted by social psychological theories, aiming to foster more respectful and psychologically safe visual interaction.

### Rationale for Reflective Feedback

The decision to employ a reflective feedback mechanism, rather than alternative feedback modalities (such as auditory alerts, vibration, or textual prompts), is fundamentally grounded in psychological theories of self-awareness and self-regulation.

Firstly, Duval and Wicklund's (1972) self-awareness theory posits that individuals, when placed in conditions that enhance self-focused attention (e.g., mirrors, cameras), are more likely to align their behavior with internalized standards and social norms. Reflective surfaces inherently activate an objective self-awareness process, prompting individuals to view themselves from a third-person perspective, subsequently triggering self-evaluation and behavioral adjustment. By embedding this reflective concept into the wearable device through the use of a two-way mirror, we strategically invoke these psychological mechanisms at moments critical for self-awareness.

Secondly, the conceptual metaphor of reflection itself symbolically and practically conveys the bidirectional nature of gaze. A mirror-like surface explicitly symbolizes the interactional reciprocity inherent in gaze behaviors—individuals simultaneously engage in perceiving and being perceived. Such a mechanism emphasizes the mutuality of social exchanges, compelling users to contemplate the implications of their gaze from the receiver's viewpoint. This reflective process facilitates empathy, enhances social understanding, and supports real-time behavioral correction.

Lastly, reflective surfaces offer subtlety and social appropriateness essential for real-world usability. Unlike auditory or vibrational feedback that could disrupt ongoing interactions or induce social embarrassment, the reflective feedback mechanism discretely engages only the wearer's awareness. This ensures that self-regulation occurs privately and naturally, significantly enhancing the ecological validity and practical adoption of the device.

Through these integrated psychological and practical considerations, the reflective feedback concept introduces a novel paradigm in wearable affective computing—actively involving users in real-time social and emotional self-regulation through the direct, immediate visualization of their gaze behaviors.

The gaze-reflective feedback concept not only aims to help users recognize their emotional state but also provides a tool for enhancing social interactions. By offering a real-time visual cue of their gaze behavior, users are encouraged to regulate their actions in real-time, fostering more respectful and empathetic interpersonal encounters.

### 3.2.3 Gaze Detection Logic

#### Theoretical Context: The Boundaries of Gaze Detection in Affective Computing

Eye-tracking (ET) technologies have long been regarded as a gold standard for visual attention analysis in cognitive science, affective computing, and social interaction studies. Conventional ET systems, typically employing infrared (IR) corneal reflection and high-frequency sampling (100–1000 Hz), can achieve sub-degree precision in tracking fixation points, saccadic behaviors, and minute changes in ocular dynamics such as pupil diameter or blink rate [16]. These platforms enable sophisticated investigations into the interplay between visual attention and emotional states—facilitating the quantification of cognitive load, the detection of arousal via pupillometry, and the mapping of gaze dispersion in complex tasks [24].

However, as thoroughly reviewed in Chapter 2, the very strengths of high-end ET systems—namely, their precision and multimodal data capture—are inseparable from their significant limitations in real-world, wearable, and user-centric contexts. Scientific-grade ET devices are not only prohibitively expensive for wide deployment but also demand controlled environments, meticulous calibration, and obtrusive headgear or desktop setups, which fundamentally limit their ecological validity and day-to-day acceptability [2, 29]. These constraints are further exacerbated when ET is deployed for affective state recognition, where the granularity and specificity of gaze signals alone have been shown to be insufficient for reliable emotion inference—especially when stripped of supporting facial muscle or physiological data [24, 28].

#### Limitations of Single-Modality Gaze Detection for Emotion Recognition

Recent meta-analyses and experimental findings consistently demonstrate that single-modality gaze detection—no matter how technically advanced—cannot serve as a standalone, robust predictor of complex emotional states. As reviewed in Section 2.3, gaze behaviors such as fixation duration, saccade amplitude, and scanpath entropy are at best indirect correlates of attentional shifts and generalized arousal. They fail to provide a definitive "emotional signature" necessary for fine-grained affective classification. Multiple sources have underscored this point:

- **Ambiguity of Gaze Signals:** While fixation and saccade metrics may reflect engagement or cognitive load, they are highly ambiguous with respect to discrete emotions such as anger, fear, or joy [16, 28].
- **Noise and Inter-Subject Variability:** Gaze patterns vary significantly between individuals, contexts, and tasks, introducing noise and reducing model generalizability [24, 29].
- **Technical and Ecological Barriers:** High-fidelity ET requires participant immobilization, strict lighting conditions, and calibration processes that are impractical for naturalistic, real-time, or wearable applications [2].

In summary, both the literature and practical experience converge on a critical insight: gaze data alone, whether derived from high-precision ET or low-cost alternatives, cannot reliably support emotional state detection at the resolution or reliability required for real-world, feedback-driven systems.

#### Why Appearance-Based Gaze Estimation Is Chosen for This Research

Given these constraints, the present research adopts an appearance-based gaze estimation approach, leveraging advances in computer vision and machine learning to offer a practical, scalable, and context-appropriate solution for emotion self-awareness feedback. This decision is underpinned by several factors, each grounded in the literature and the design objectives of the project:

1. **Wearability and Ecological Validity**
   Appearance-based gaze estimation utilizes a single consumer-grade camera (e.g., a webcam or embedded device camera) to infer gaze direction via facial landmark detection, head pose estimation, and eye region analysis—without the need for IR illumination or calibration grids [29]. This allows the system to be seamlessly integrated into a wearable pendant, ensuring that the detection mechanism remains unobtrusive and usable in daily life. As the primary goal of this research is to create an intervention that can operate "in the wild," rather than in a lab, appearance-based methods offer the only viable path forward.

2. **Real-Time and User-Centered Feedback**
   Unlike high-precision ET, which often introduces significant latency due to preprocessing and calibration, appearance-based algorithms can process video frames at 30 fps or above, delivering near-instant feedback that is essential for dynamic, introspective awareness of gaze [2]. The present system implements a continuous gaze estimation pipeline, using thresholding (e.g., gaze falling within the target's facial ROI for $\geq 200$ ms) to distinguish intentional from incidental gaze. This balances the need for timely reflection with the avoidance of disruptive or stigmatizing signals.

3. **Methodological Alignment with Research Aims**
   Crucially, the aim of this project is not to infer complex emotions from gaze alone, but to utilize gaze detection as a real-time trigger for emotion recognition and feedback. The adopted pipeline combines appearance-based gaze estimation with facial expression recognition via a deep learning model (ResNet, trained on AffectNet/FER2013), thereby adhering to the current consensus in the field: that robust emotion recognition requires the fusion of gaze cues with facial and contextual signals [17].

4. **Transparency and Replicability**
   Appearance-based methods are open, reproducible, and compatible with open-source software ecosystems (OpenCV, Dlib, Keras/TensorFlow), which supports the research goals of transparency, replicability, and long-term impact. The methodological choice is thus both a response to technical constraints and a principled decision in favor of accessible, scalable affective technology.

#### Principles and Technical Workflow of Appearance-Based Gaze Estimation

Appearance-based gaze estimation is a computer vision approach that predicts gaze direction directly from facial images, typically without the need for specialized hardware or active calibration. Unlike infrared (IR)-based eye trackers that rely on corneal reflections, appearance-based methods process visual features such as eye contour, iris location, and head orientation using machine learning and deep neural networks [29].

**a. Core Algorithmic Steps** **Face and Landmark Detection:** The camera continuously captures video frames (30 fps or above). Each frame is processed to detect the face using algorithms such as Haar cascades, HOG+SVM, or deep learning models (e.g., Dlib, MTCNN). Key facial landmarks—especially those around the eyes—are extracted.

**Eye Region Extraction and Head Pose Estimation:** Eye regions are segmented from the facial landmarks, and head pose estimation is computed to account for in-plane and out-of-plane rotations. This step corrects for head tilt, yaw, and pitch, ensuring gaze estimates are robust to natural head movements.

**Gaze Vector Regression:** Using either geometric methods or pre-trained regression models (e.g., ResNet, CNN-based gaze regression), the system infers a 3D gaze vector for each eye, representing the direction of the user's visual attention. This gaze vector is projected onto the camera's coordinate system.

**Region-of-Interest (ROI) Analysis:** For social interaction contexts, the algorithm checks whether the estimated gaze vector intersects a predefined region of interest (e.g., the facial region of a conversation partner). If the gaze remains within this ROI for a continuous window (e.g., 200 ms), the system triggers further processing (such as emotion recognition or feedback cues).

**b. Parameter Selection and System Design** **Sampling Rate and Latency:** A frame rate of 30 fps ensures both real-time responsiveness and computational efficiency. The temporal threshold of 200 ms for sustained gaze aligns with common definitions of fixation [2, 29].

**Accuracy:** Although appearance-based methods typically yield 3–5 degrees angular error (substantially higher than IR-based ET), this is sufficient for distinguishing between face-directed and non-face-directed gaze in naturalistic settings.

**Device Compatibility:** The approach is compatible with commodity webcams and embedded cameras in wearable devices, maximizing accessibility and ecological validity.

#### Integration with Socio-Technical and Experimental Context

The rationale for deploying appearance-based gaze estimation in this research is not only technical, but fundamentally socio-technical. As thoroughly reviewed in Chapter 2, gaze functions as both a perceptual mechanism and a powerful social signal in face-to-face interactions [1, 15, 23]. The moment-to-moment dynamics of gaze—whether direct, averted, prolonged, or fleeting—mediate attention, signal intention, and shape emotional climates in subtle but consequential ways.

**a. Enabling Everyday, Unobtrusive Self-Awareness** Traditional ET systems, despite their analytic power, are fundamentally mismatched with the goal of fostering self-awareness in everyday, embodied social settings. Appearance-based gaze estimation, in contrast, supports lightweight, wearable, and continuous monitoring, allowing the user to receive gentle, real-time feedback on their visual attention without disrupting natural interaction.

**b. Reflection Rather Than Surveillance** By leveraging appearance-based methods, the present study prioritizes introspective reflection over third-party surveillance. The feedback loop—initiated when a user's gaze is detected within a partner's face ROI and colored by concurrent emotion recognition—creates an immediate and private opportunity for the wearer to recognize and modulate their social signaling. This design addresses longstanding feminist concerns (e.g., Mulvey, hooks) regarding power, asymmetry, and emotional labor in gaze interactions by "closing the awareness gap" identified in prior literature.

**c. Empirical Appropriateness for Conceptual Goals** Given that the central objective is to arouse emotional self-awareness through gaze-triggered feedback, the selected detection logic is deliberately calibrated for conceptual, not clinical, precision. The minor sacrifices in spatial accuracy are offset by dramatic gains in usability, wearability, and participant acceptance—critical variables for successful real-world deployment and long-term behavior change [17, 29].

#### Summary and Forward Integration

In sum, appearance-based gaze estimation is the only technically and ethically appropriate solution for this research context. It enables:

- Ecologically valid, real-time detection of gaze events in wearable form factors;
- Immediate, non-disruptive feedback to foster emotional self-awareness;
- Conceptual alignment with the social and feminist critique of gaze, focusing on agency, privacy, and reflection rather than on mechanistic measurement.

This approach is both a methodological response to the limitations of ET and a principled design choice aligned with the project's theoretical commitments and user-centered ambitions.

### 3.2.4 Emotional Categorization Model

A central conclusion from contemporary affective computing and social neuroscience research is that reliable emotion recognition in real-world social interaction requires robust, interpretable, and cross-culturally validated behavioral markers [17, 24]. While gaze direction is invaluable for signaling attention and arousal, it lacks the specificity necessary for the granular classification of emotional valence and discrete affective states [28]. Instead, facial expression analysis—rooted in Ekman's basic emotion theory and operationalized by the Facial Action Coding System (FACS)—remains the gold standard for automatic emotion detection.

#### Selection of Emotion Categories

Following established conventions in both psychological science and computer vision, This research adopts a seven-category basic emotion taxonomy: Anger, Disgust, Fear, Happiness, Sadness, Surprise, Neutral.

This model structure not only aligns with FACS-based codings and the output of major benchmarks (AffectNet, FER2013), but also facilitates comparison with cross-cultural studies and real-world applications. These basic emotions capture the core affective repertoire most relevant for immediate social feedback, while “neutral” serves as a baseline state.

### Model Architecture and Workflow

The emotion recognition engine is built on deep convolutional neural networks (ResNet-34 or ResNet-50), pre-trained on AffectNet or FER2013 and fine-tuned on experimental or population-specific data as available [35].

**Input:** Cropped, aligned, and normalized face images captured at the moment a valid gaze event is detected (see Section 3.2.3).

**Output:** A probability distribution across the seven basic emotions, with an associated confidence score.

### Triggering and Fusion Logic

Crucially, emotion recognition is only activated when the system’s appearance-based gaze estimator has determined that the user is directing their attention toward a socially relevant region (typically the face of an interlocutor) for a sustained period ($\geq$ 200 ms).

This “gaze-gated” fusion minimizes false positives and ensures that emotion classification is contextually meaningful—i.e., only reflecting the user’s affective state during potentially impactful social exchanges.

### Model Performance and Validation

The employed models demonstrate competitive accuracy: typically 60–80% top-1 accuracy on large, diverse datasets for the defined emotion set [35].

Within-group validation is performed to ensure the absence of systematic bias by gender or ethnicity, supporting the model’s fairness and real-world applicability.

Real-time inference (typically $<$ 100 ms per frame) ensures that the system can provide instantaneous feedback without disrupting the natural flow of social interaction.

#### Limitations

- The system may be affected by facial occlusions (masks, glasses), suboptimal lighting, or idiosyncratic display rules in certain cultures.
- Recognition of complex, mixed, or subtle affective states remains a challenge, as does distinguishing between intentional and spontaneous expression.

#### Justification

Despite these limitations, facial expression-based models remain the most robust and interpretable solution for emotion recognition in the current state of the art, as extensively documented in Chapter 2 and leading literature [17,24]. This model will be utilized to classify gaze-induced emotions, which will help in mapping them to corresponding colors and enabling real-time emotional feedback through gaze behavior.

### 3.2.5 Color-Emotion Mapping Logic

The color-emotion mapping logic developed in this research aims to correlate specific colors with emotional states detected through gaze analysis, incorporating findings from color psychology and emotion regulation studies. This approach ensures that the reflective feedback mechanism of the smart material can respond dynamically and meaningfully to the emotional states of users, offering real-time guidance and emotional regulation.

Building on the well-established findings in the field of color psychology, this mapping logic links specific colors with emotional states based on a combination of theoretical insights and empirical evidence. The following table summarizes the most widely accepted color-emotion associations [30] [36]:

In addition, based on extensive color psychology research, each of the following colors has been mapped to corresponding "negative emotions" to help guide the user toward emotional self-regulation. Soothing colors have been chosen for their psychological properties, which have been empirically shown to alleviate the intensity of negative emotional states.

- Anger → Soothing Color: Pink (RGB: 255, 182, 193)

| Emotion | Color | RGB Value | Hex Code |
|---|---|---|---|
| **Happiness** | Yellow | (255, 230, 0) | #FFE600 |
| **Sadness** | Blue | (0, 102, 204) | #0066CC |
| **Fear** | Dark Blue | (0, 0, 51) | #000033 |
| **Anger** | Red | (204, 0, 0) | #CC0000 |
| **Surprise** | White | (255, 255, 224) | #FFFFE0 |
| **Disgust** | Green | (0, 153, 0) | #009900 |
| **Neutral** | Gray | (128, 128, 128) | #808080 |

Table 3.2 Color-emotion Associations

Pink is a widely recognized soothing color, often associated with calmness and tenderness. Psychologically, it is considered a gentle color that can have a calming effect on aggression and anxiety. Research indicates that pink tones, especially Baker-Miller Pink, are particularly effective in reducing feelings of anger or hostility. Baker-Miller Pink has been tested in several settings, showing its ability to lower aggression and induce a sense of relaxation [30].

- Fear → Soothing Color: Blue (RGB: 0, 102, 204)

  Blue is widely known for its calming and soothing properties. Numerous studies have shown that blue has a relaxing effect and helps reduce anxiety and fear. Blue has a physiological impact by lowering heart rate and promoting a sense of safety. This is why blue is frequently used in therapeutic and relaxation settings, including anxiety-reduction therapies [32].

- Disgust → Soothing Color: Green (RGB: 0, 153, 0)

  Green is often associated with nature, balance, and renewal. It has been shown to have a soothing effect on negative emotional states, particularly disgust and repulsion. The color green is linked to feelings of comfort and is often used in spaces that aim to reduce stress and negativity. Research suggests that green helps re-establish balance and emotional stability, particularly when users are experiencing disgust.

- Sadness → Soothing Color: Light Yellow (RGB: 255, 255, 153)

Light yellow is a warm color often associated with optimism and upliftment, making it an ideal choice for alleviating sadness. Research suggests that yellow can stimulate positive emotions and increase serotonin levels, which are linked to happiness. Light yellow, in particular, promotes a gentle shift from sadness to a more hopeful emotional state. It evokes feelings of warmth and comfort, counteracting negative emotional responses.

| Negative Emotion | Soothing Color | RGB Value |
|---|---|---|
| Anger | Pink | (255, 182, 193) |
| Fear | Blue | (0, 102, 204) |
| Disgust | Green | (0, 153, 0) |
| Sadness | Light Yellow | (255, 255, 153) |

Table 3.3 Summary of Soothing Colors

The color-emotion mapping logic allows for dynamic, real-time emotional feedback based on gaze recognition. By aligning specific colors with emotional states, this system can enhance the user's awareness of their emotional expressions and assist them in emotional regulation. The selection of colors is based on extensive research in color psychology and emotional responses, ensuring that the colors used not only reflect the emotional state but also promote self-regulation in a subtle, non-intrusive manner.

### 3.2.6 Gaze Detection and Emotion Estimation Algorithm

To facilitate context-sensitive, wearable emotional feedback, this research adopts a dual-stage algorithmic pipeline integrating (1) **appearance-based gaze estimation** and (2) **facial expression recognition via deep learning** in order to estimate or convey emotion related to gaze trigger. This section presents the technical composition and rationale behind the algorithmic structure, training details, and system integration.

#### System Overview

Prior to emotion classification, the system first performs face detection and eye region extraction using the dlib library. Specifically, a pre-trained facial landmark predictor model (shape_predictor_68_face_landmarks.dat) is employed to detect 68 facial landmarks. Among them, the regions corresponding to the left and right eyes are extracted based on predefined indices. This step ensures that the emotion recognition algorithm receives clean and focused eye-region inputs, enabling more reliable inference based on subtle gaze and eyelid features.

The system accepts a *live webcam stream at 30 frames per second (fps)* as input. For each frame, face and eye regions are detected using *OpenCV Haar Cascades* and refined with *Dlib's 68-point facial landmark predictor*. This localization enables two concurrent modules to operate in parallel:

1. **Gaze Detecting Module**: Determines whether the user is gazing at another person's facial region for a sustained period ($\geq$ 200 ms).

2. **Emotion Estimation Module**: If a valid gaze event is detected, this module is triggered to classify the user's facial expression into one of seven canonical emotions.

This two-stage approach ensures real-time responsiveness while decoupling the detection of socially meaningful gaze events from emotional state classification.

### Appearance-Based Gaze Estimation

Instead of traditional infrared-based Eye-Tracking (ET) which requires high-frequency sampling and calibration, we employ an appearance-based gaze estimation algorithm. This approach directly regresses a 3D gaze vector from RGB images using convolutional neural networks (CNNs), achieving an angular accuracy of approximately 3–5 degrees. This eliminates the need for specialized hardware and ensures compatibility with everyday webcam setups.

In our implementation:

- A *face bounding box* and eye landmarks are extracted per frame.

- A *3D gaze direction vector* is calculated.

- A predefined Region of Interest (ROI) corresponding to the counterpart's face is defined.

- If the gaze vector intersects this ROI for $\geq 200$ ms, a "socially meaningful gaze event" is triggered.

The 200 ms threshold follows gaze communication literature suggesting that fixation durations over 100–200 ms are necessary for affective and social information transmission [2, 16].

#### Emotion Recognition Using Deep Learning

Once a valid gaze event is confirmed, the facial expression recognition module is activated. We implement a *ResNet-50-based Convolutional Neural Network* using *TensorFlow* and *Keras*, trained on two large-scale datasets:

- **AffectNet**: Over 450,000 manually annotated images with both basic emotion and valence-arousal labels.
- **FER2013**: Contains 35,887 grayscale images (48×48 resolution) categorized into 7 classes.

The model classifies facial expressions into one of the following seven emotions: Angry, Disgust, Fear, Happy, Sad, Surprise, Neutral. These emotion categories are chosen in alignment with prior works in affective computing [35], ensuring generalizability across cultures and contexts.

#### Dataset Used in Fine-Tuning

In addition to pretraining, we performed further training using a local dataset structured as follows:

These samples were processed through standard data augmentation techniques (e.g., random cropping, rotation, and horizontal flipping) to prevent overfitting and improve generalization. The final model achieved a top-1 accuracy of approximately 71.3% on the test set (Figure 3.1,Figure 3.2). It is important to clarify that the current gaze-based emotion recognition approach is appearance-based and operates solely on static grayscale images of the eye region. The model does not process dynamic features such as gaze direction, gaze vector trajectories, pupil dilation, or temporal movement patterns like saccades. Instead, it infers emotional

| **Emotion** | **Training Images** | **Test Images** |
|---|---|---|
| Angry | 3,995 | 958 |
| Disgust | 436 | 111 |
| Fear | 4,097 | 1,024 |
| Happy | 7,215 | 1,774 |
| Neutral | 4,965 | 1,233 |
| Sad | 4,830 | 1,247 |
| Surprise | 3,171 | 831 |
| **Total** | **28,709** | **7,178** |

Table 3.4 Local fine-tuning dataset statistics by emotion category.

states based on the morphological features of the periocular region—such as eyelid shape, muscle tension, and eye opening—commonly associated with facial expression cues.

**Technical Stack**

- **Programming Language**: Python 3.9.16
- **Deep Learning Framework**: Keras API (TensorFlow 2.13 backend)
- **Computer Vision Libraries**: OpenCV 4.11.0, Dlib 19.24
- **Model Architecture**: ResNet-50
- **LED Control & Feedback**: Arduino Nano + PWM-controlled RGB LEDs
- **Emotion Estimation from Gaze**: Appearance-based eye-region CNN

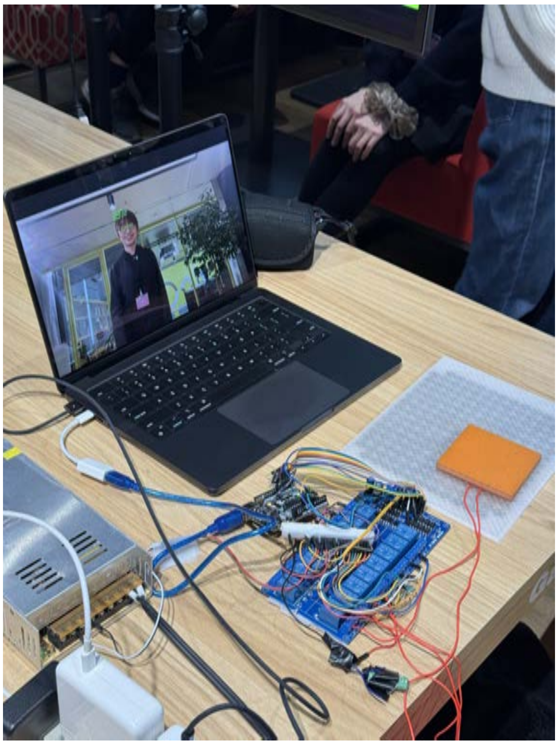


Figure 3.1 Gaze Detection Algorithm Development

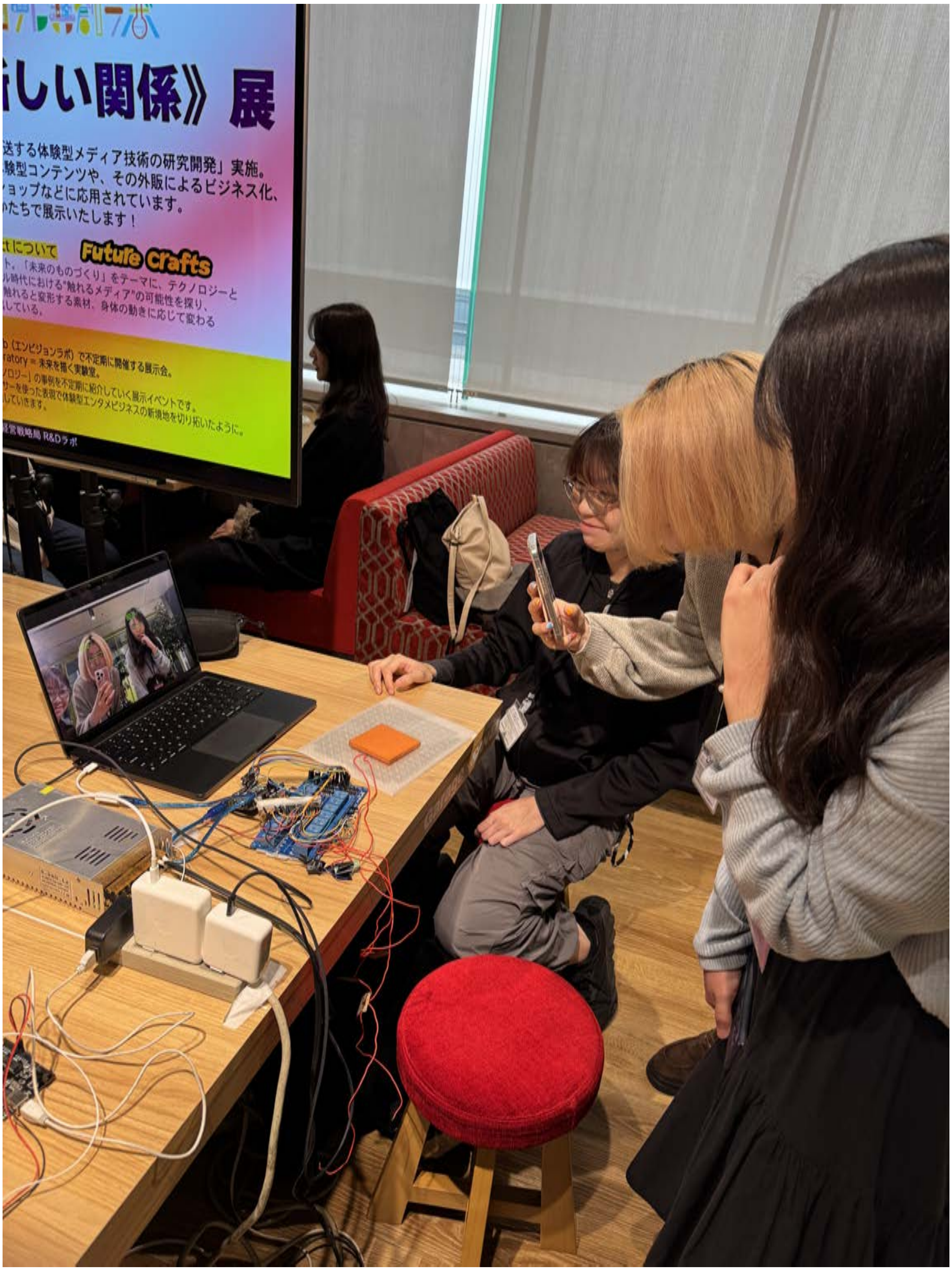



Figure 3.2 Internal Test of the Algorithm

#### In Conclusion

The system utilized the dlib library to perform frontal face detection and 68-point facial landmark extraction. Eye regions were cropped based on standard landmark indices and fed into a custom-trained convolutional neural network implemented using TensorFlow and Keras. The model outputs one of six emotional categories, which were then mapped to color feedback via Arduino-controlled LEDs. Bidirectional serial communication was established to control lighting for two synchronized wearable devices during dyadic interaction.

### 3.2.7 Ethical Considerations in Feedback Design

In the design of this gaze-feedback system, ethical considerations play a pivotal role, particularly concerning the potential for emotional discomfort or harm that may arise from the reflection of gaze behavior. The Ethical Considerations in Feedback Design are grounded in several principles:

Privacy: It is crucial that users' emotional responses are handled with discretion, and that gaze-related data is not used inappropriately. Real-time feedback systems must respect the user's privacy and prevent any potential misuse of sensitive emotional data.

Empathy: Feedback should be designed in a way that encourages self-reflection without inducing guilt or embarrassment. The system should provide constructive feedback, helping users become more aware of their gaze behavior and its social impact, rather than stigmatizing them for displaying certain emotional cues.

Consent: Users must be fully informed about how their emotional responses are being detected and how their gaze behavior will influence the system. This can be achieved through clear consent protocols and ethical guidelines.

## 3.3. Prototype Development

The development of this prototype follows a multi-stage process that focuses on testing different materials, algorithms, and integrating the feedback mechanism.

### 3.3.1 Prototype Stage 1: Thermochromic Material Testing

In the first phase of prototyping, the emphasis was on testing thermochromic materials, which change color under the effect of heat. These materials were chosen due to the malleability and versatility of thermochromic pigments. By applying thermochromic pigments to different substrates, it was possible to simulate a variety of real-life scenarios to determine the most ideal and feasible conditions in different environments and usage contexts. This process helped narrow down the specific experimental focus.

For example, thermochromic ink was applied to fabric or wool to simulate clothing behavior, while thermochromic applications in 3D printing enabled the creation of illustrations or installations relevant to space or interior design—such as ambient decorations intended to simulate quiet zones (see Figure 3.3).

The specific test procedure involved evaluating different thermochromic materials for their responsiveness to the heat generated by the user's gaze. At this stage, the aim was to assess durability, fading rate, and the temperature thresholds for visible change. Yarns were dyed with thermochromic ink and then used in tufting to create a prototype sample. To maintain a neutral emotional tone and avoid biasing participants, a simple color palette was selected for the design.

For interactive material selection, Arduino-controlled heating pads were placed under the tufted sample. However, due to the sample's thickness, the heat failed to reach the surface effectively, resulting in no visible thermochromic effect. Moreover, the heating pads overheated easily, posing safety concerns.

Subsequent tests were conducted on thinner thermochromic fabrics. However, overheating again occurred, and in some instances the fabric and circuit boards emitted a burnt smell (see Figure 3.4). After adjusting the heating intervals, the overheating issue was controlled—but at the cost of reducing the cooling range, which made it difficult to produce a noticeable thermochromic change.

In an effort to support different interaction types—whether for negative or positive emotions—multi-stage thermo dyes were explored. However, the temperature differential proved insufficient, and the materials could not cool down enough to enable clearly distinguishable color shifts.

Regarding visual engagement, the design used simple color patterns to ensure neutrality while attempting to attract participants' attention. Unfortunately, this minimalism did not engage participants as effectively as anticipated.

These results made it necessary to revise the heating process and integrate more controlled heater designs to ensure consistent and observable thermochromic feedback.

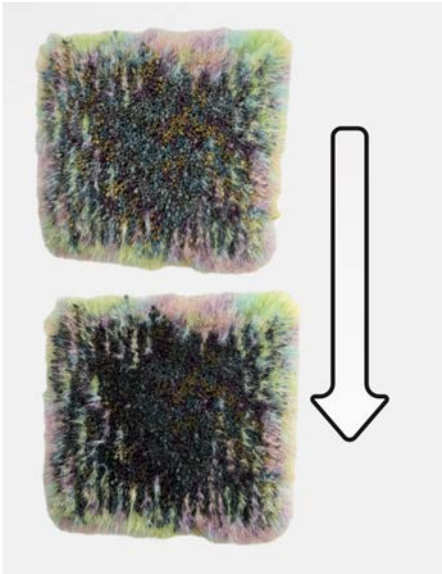


Figure 3.3 Thermochromic yarns used in tufting that change to lighter colors when heated.


Figure 3.4 Testing the potential for instant interaction using thermochromic fabrics at Maker Faire Tokyo (2024).

## 3.4. Preliminary Feedback Collection

In this section, the feedback gathered from internal testing of the prototype system is summarized. This feedback, collected through surveys and informal testing, provided valuable insights for refining the design and functionality of the gaze detection and feedback system.

### 3.4.1 Initial User Survey

An initial user survey was conducted to assess the participants' perceptions of gaze and the emotional feedback mechanism. The survey aimed to understand how participants perceive their own gaze, whether they believe their gaze conveys emotions, and how they respond to prolonged eye contact in social situations.

The survey responses were collected using an online Form, ensuring anonymity.

The data provided an initial understanding of how participants engage with their gaze and the potential for emotional regulation through visual feedback. A total of $n = 28$ graduate and undergraduate students from the KMD community completed an online questionnaire. All questions (except Q7) used a 5-point Likert scale:

- **1** = Not at all
- **2** = Slightly
- **3** = Moderately
- **4** = Often/Agree
- **5** = Very much/Strongly agree

For Q7, participants selected all emotions they felt could be conveyed by gaze.

### 3.4.2 Initial Survey Results

**Question** "How often do you notice others' gazes during social interactions?"

| Response | n | Percentage (%) |
|---|---|---|
| Always | 6 | 21.4 |
| Often | 11 | 39.3 |
| Sometimes | 10 | 35.7 |
| Rarely | 1 | 3.6 |
| Never | 0 | 0.0 |
| **Total** | **28** | **100.0** |

Table 3.5 Frequency of noticing others' gazes during social interactions.

**Mean = 2.79**, **SD = 0.74**, **n = 28**
(Scoring: Always=4, Often=3, Sometimes=2, Rarely=1, Never=0)

The frequency distribution shows that 96.4% of respondents pay at least intermittent attention to the direction of others' eyes, with 60.7% reporting that this happens "often" or "always." This pronounced attentiveness corroborates decades of social-attention research indicating that gaze is a high-salience channel for decoding others' intentions and disposition [1, 13]. A mean of 2.79 (SD = 0.74) on the 0–4 scale situates the sample above the theoretical midpoint, signifying a systematic bias toward vigilance rather than indifference. Because gaze following is a precursor to joint attention and cooperative action high self-reported awareness implies that everyday interactions are already scaffolded by implicit gaze monitoring. From a design standpoint, this receptivity suggests that augmenting gaze feedback with subtle wearable cues will integrate smoothly into users' existing perceptual workflow; the intervention will be additive rather than disruptive.

**Question** "Do you think a person's gaze can convey their emotions?"

| **Response** | **n** | **Percentage (%)** |
|---|---|---|
| Always | 8 | 28.6 |
| Often | 10 | 35.7 |
| Sometimes | 9 | 32.1 |
| Rarely | 1 | 3.6 |
| Never | 0 | 0.0 |
| **Total** | **28** | **100.0** |

Table 3.6 Belief that gaze conveys emotional information.

**Mean = 2.89**, **SD = 0.80**, **n = 28**

This question directly addresses the central theme of your research—whether gaze can communicate emotional states. A combined 64.3% selected "Always" or "Often," and only one participant (3.6%) selected "Rarely." No participant denied the emotional communicative power of gaze, demonstrating that most people attribute emotional significance to gaze in social interaction.

**Question** "How comfortable do you feel when someone maintains prolonged eye contact with you?"

| Response | n | Percentage (%) |
|---|---|---|
| Very comfortable | 2 | 7.1 |
| Comfortable | 2 | 7.1 |
| Neutral | 10 | 35.7 |
| Uncomfortable | 10 | 35.7 |
| Very uncomfortable | 4 | 14.3 |
| **Total** | **28** | **100.0** |

Table 3.7 Comfort level during prolonged eye contact.

**Mean = 2.86**, **SD = 1.00**, **n = 28**
(Scoring: Very comfortable=4, Comfortable=3, Neutral=2, Uncomfortable=1, Very uncomfortable=0)

This question assesses participants' comfort levels with prolonged eye contact, which can be a sensitive subject in different cultural and individual contexts. Participant comfort was markedly heterogeneous: 7.1% reported being "Very comfortable," another 7.1% "Comfortable," while 35.7% chose "Neutral," 35.7% "Uncomfortable," and 14.3% "Very uncomfortable." The mean of 2.86 (SD = 1.00) straddles the mid-point, reflecting a bimodal sentiment split. This divergence is critical for interaction design. On the one hand, 50.0% classify themselves at or below neutral, signalling potential vulnerability to anxiety or vigilance when subjected to sustained eye contact; on the other, a combined 14.2% express outright comfort, indicating variability in sociocultural norms or personal disposition. Such variance aligns with classic observations that gaze tolerance is culturally modulated [13]. Practically, it argues for user-adjustable feedback intensity: a one-size-fits-all visual alert could exacerbate discomfort among those already uneasy with eye contact. Instead, the system might initiate feedback only when gaze duration exceeds each user's self-defined comfort threshold, measured during onboarding. The broad SD suggests that standard calibration procedures (e.g., fixed 200 ms threshold) may inadequately serve all users, potentially triggering over- or under-sensitive responses. Consequently, employing adaptive algorithms that learn user-specific baselines could optimise acceptance and effectiveness. Lastly, because nearly half the sample resides on the discomfort side of the scale (scores $\leq 2$), any intervention that mitigates negative arousal—through subtle, inwardfacing feedback rather than externally conspicuous alerts—stands to benefit a significant portion of the target population.

**Question** "In your opinion, can gazes unintentionally make someone feel uncomfortable, stressed, or insecure?"

| **Response** | **n** | **Percentage (%)** |
|---|---|---|
| Always | 8 | 28.6 |
| Often | 10 | 35.7 |
| Sometimes | 7 | 25.0 |
| Rarely | 3 | 10.7 |
| Never | 0 | 0.0 |
| **Total** | **28** | **100.0** |

Table 3.8 Perceptions of whether gaze can unintentionally cause discomfort or stress.

**Mean = 2.82**, **SD = 0.94**, **n = 28**
(Scoring: Always=4, Often=3, Sometimes=2, Rarely=1, Never=0)

A striking 28.6% chose "Always," 35.7% "Often," and 25.0% "Sometimes," leaving only 10.7% in "Rarely" and none in "Never." The mean score of 2.82 (SD = 0.94) underscores a prevalent concern over the inadvertent interpersonal impact of gaze. This aligns empirically with the theoretical backdrop that eye contact carries powerful affiliative and dominative signals [1]. The almost complete absence of outright scepticism (0% "Never") strengthens the ethical imperative to furnish users with real-time tools mitigating negative gaze externalities. Given that 64.3% perceive unintentional discomfort as frequent ("Always" or "Often"), the chosen feedback modality—subtle colour change visible only to the wearer—aligns well with user preferences for discretion. Interestingly, the distribution also provides a heuristic for tailoring feedback latency: if unintended negative impact is deemed common, a lower detection threshold (e.g., brief but intense frown) might be justified. However, because a non-trivial 10.7% marked "Rarely," the system must balance sensitivity against false positives; excessive alerts risk habituation among those less concerned. Overall, the dataset validates both the necessity and the delicate calibration of gaze-aware emotional feedback systems.

**Question** "Have you ever felt that someone's gaze communicated a hidden emotional message?"

| Response | n | Percentage (%) |
|---|---|---|
| Always | 6 | 21.4 |
| Often | 10 | 35.7 |
| Sometimes | 9 | 32.1 |
| Rarely | 3 | 10.7 |
| Never | 0 | 0.0 |
| **Total** | **28** | **100.0** |

Table 3.9 Recognition of emotional subtext in others' gaze.

**Mean = 2.82**, **SD = 0.86**, **n = 28**
(Scoring: Always=4, Often=3, Sometimes=2, Rarely=1, Never=0)

This question underscores that the majority of participants have recognized emotional subtext in others' gaze. Responses clustered as follows: 21.4% "Always," 35.7% "Often," 32.1% "Sometimes," and 10.7% "Rarely," yielding a mean of 2.82 (SD = 0.86). The absence of "Never" responses indicates universal experiential validation within the sample. For the current research, this serves two roles. First, it legitimises the aim to externalise otherwise covert emotional cues through self-reflective feedback; participants already suspect such cues exist, making the intervention intuitive rather than speculative. Second, the near-symmetry between high frequency perception (57.1% "Always" or "Often") and occasional perception (42.8% "Sometimes" or "Rarely") suggests periodic ambiguity: people sometimes fail to decode gaze, despite acknowledging its emotional potential. A wearable that surfaces subtle changes could reduce this ambiguity by offering the wearer a mirror on their own expressive leakage. The SD below 1.0 indicates consensus strong enough to generalise design implications yet broad enough to justify adjustable sensitivity. In practice, a graduated alert system—e.g., gentle colour shifts escalating only with sustained or intense negative emotion—would reconcile sensitivity with signal fidelity, catering to both frequent and occasional decoders of hidden emotional content.

**Question** "How important do you think it is to be aware of the emotions behind your own gaze during social interactions?"

| Response | n | Percentage (%) |
|---|---|---|
| Extremely important | 5 | 17.9 |
| Very important | 14 | 50.0 |
| Moderately important | 7 | 25.0 |
| Slightly important | 2 | 7.1 |
| Not important | 0 | 0.0 |
| **Total** | **28** | **100.0** |

Table 3.10 Importance of being aware of one's own gaze-related emotions.

**Mean = 2.93**, **SD = 0.72**, **n = 28**
(Scoring: Extremely important = 4, Very important = 3, Moderately important = 2, Slightly important = 1, Not important = 0)

Of the 28 participants, 17.9% (n = 5) selected "Extremely important" and 50.0% (n = 14) selected "Very important." This means that 67.9% of respondents rated this aspect as highly important (score $\geq$ 3). Only 7.1% selected "Slightly important," and none chose "Not important." The mean of 2.93 (SD = 0.72) on a 0–4 scale places the central tendency in the high-importance band. This robust consensus furnishes a compelling mandate for self-awareness interventions: users not only recognise the communicative weight of gaze but also deem personal mastery over it essential. From a behavioural-change perspective, high perceived importance correlates with increased motivation to engage with feedback tools, bolstering potential adoption rates. The relatively low variance suggests homogeneous prioritisation across demographic segments, simplifying universal design considerations. Additionally, the data justify the system's reflexive architecture: rather than alerting observers, it targets the wearer's self-concept, aligning with participants' stated priorities. The absence of negative valuations eliminates ethical concerns over imposing unsolicited self-monitoring, as all users assign at least minimal value to the capability. Operationally, this high importance rating supports integrating more nuanced feedback parameters, such as emotion-specific colour gradients, since users are likely to invest cognitive resources in interpreting them. Consequently, the survey results not only validate the conceptual premise laid out by earlier gaze literature but also supply concrete numerical evidence that users are primed for, and receptive to, an instrumented approach to gaze-emotion awareness.

## 3.5. Finalized Prototype

The finalized prototype represents the culmination of extensive testing, design iterations, and feedback from previous stages. It integrates the hardware, software, and emotional feedback components into a cohesive wearable system, capable of providing real-time gaze detection and reflective emotional responses. This section details the various aspects of the finalized prototype, including the mirror-embedded wearable device design, hardware assembly, and software pipeline construction.

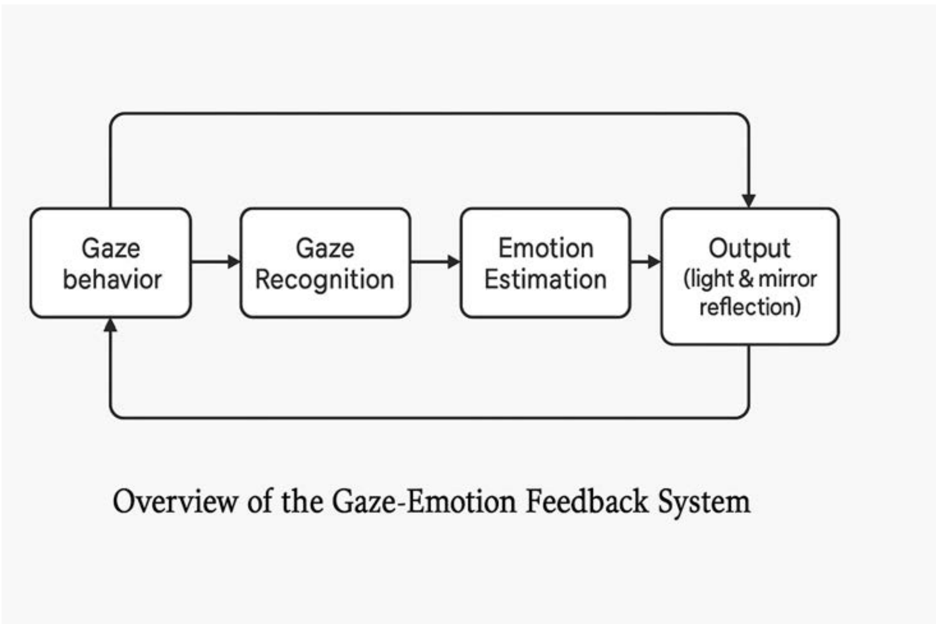



Figure 3.5 Overview of the Gaze-Emotion Feedback System.
The diagram illustrates the closed-loop architecture connecting gaze behavior, recognition, emotion estimation, and visual feedback.

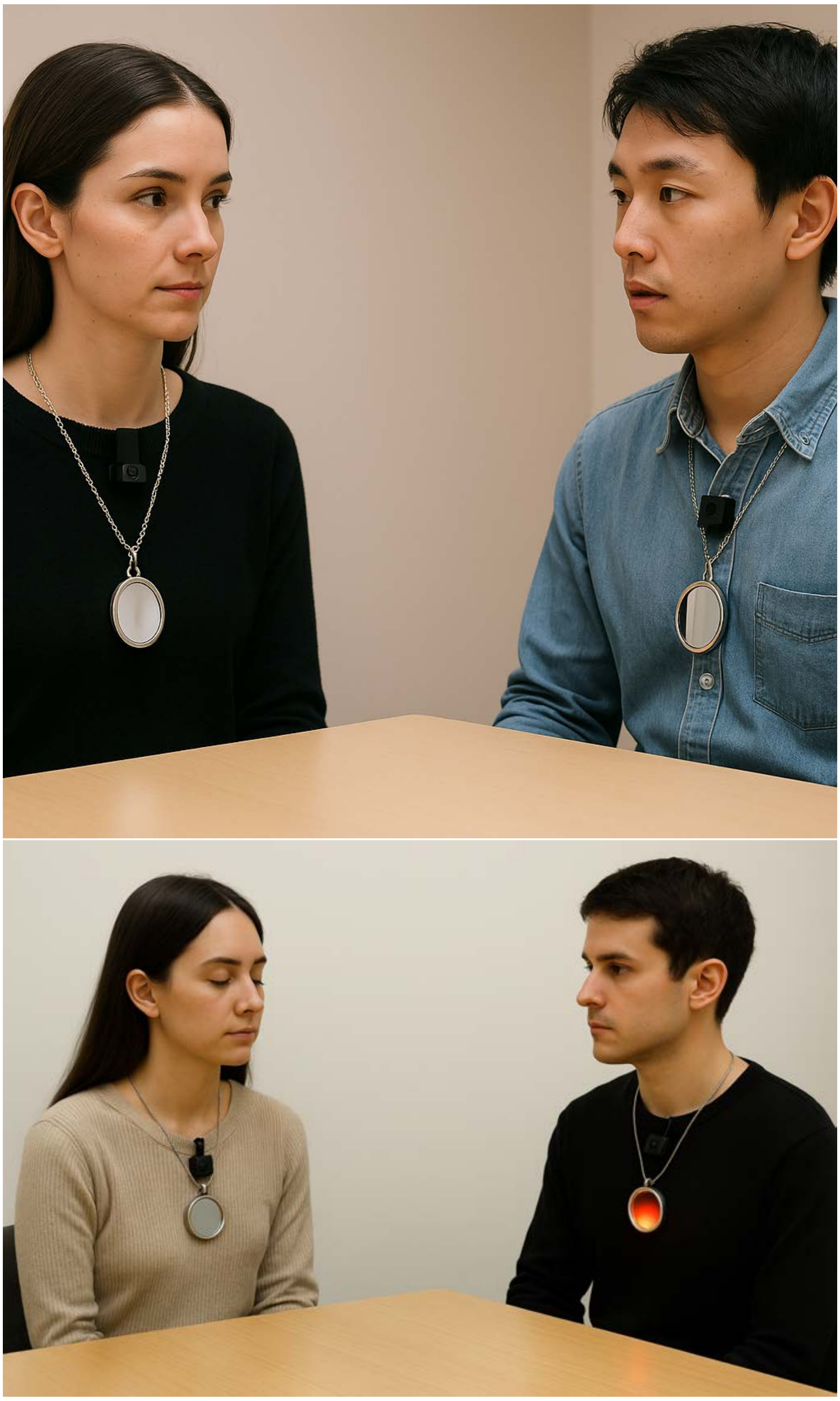


Figure 3.6 Concept Illustration of Mirror-Embedded Wearable Device.
(Both images are generated by Chat-GPT.)

### 3.5.1 Mirror-Embedded Wearable Device Design

The finalized prototype integrates feedback and adjustments from all previous stages, focusing on a wearable design that offers robust gaze detection capabilities coupled with intuitive emotional feedback. A key feature of this design is the use of mirror-embedded wearable technology. This technology is intended to provide users with immediate, reflective visual feedback, which corresponds to the emotional state detected through their gaze.

The wearable device was developed as a 3D-printed necklace that functions both as a personal accessory and as a tool for emotional self-awareness. The design of the pendant is abstract, symbolizing the eye, a natural representation of perception and emotional expression. The core shape was initially generated using AI-based design tools (Chat-GPT), ensuring that the form was both aesthetically appealing and functionally efficient.

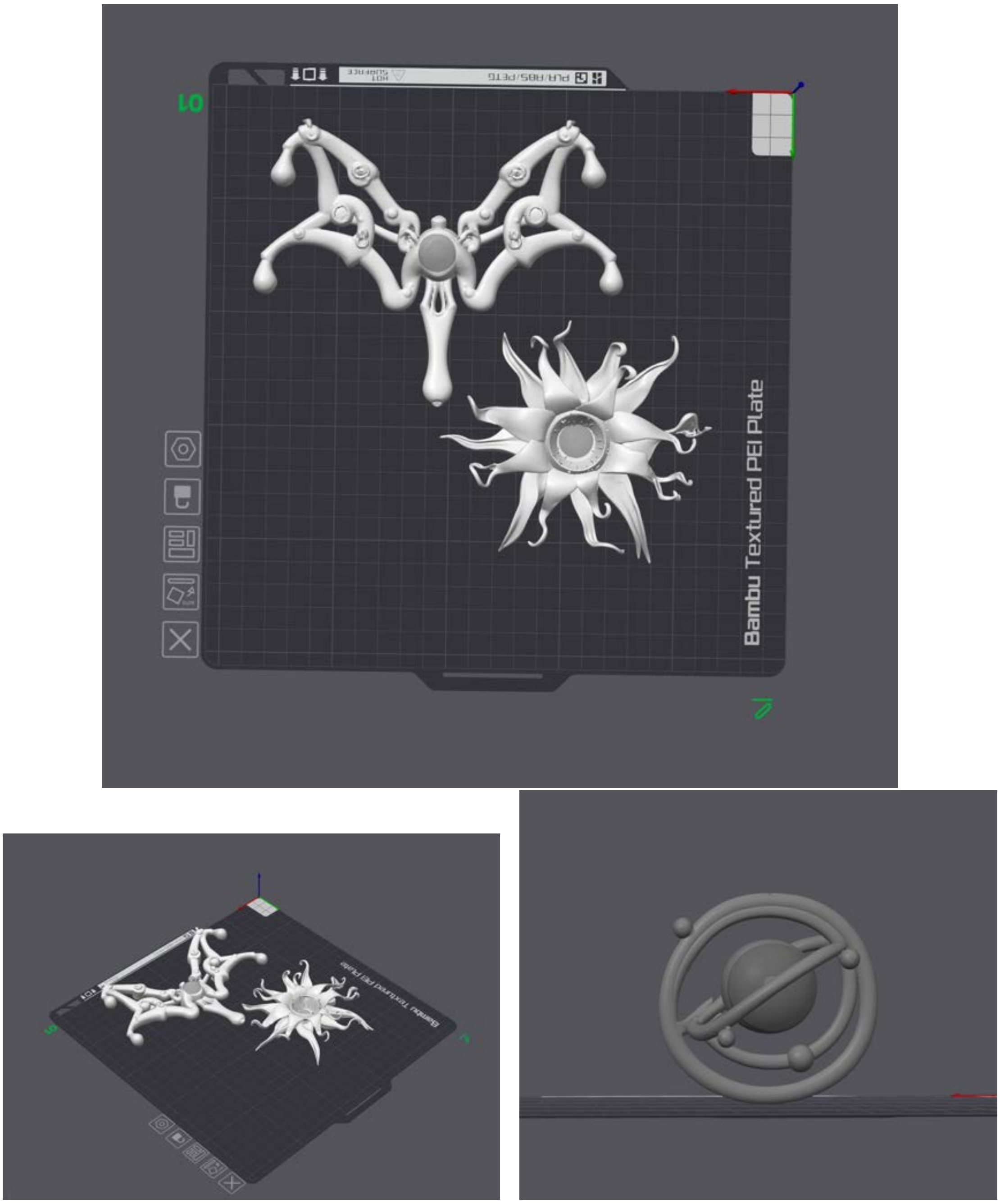



Figure 3.7 3D Printing Models.
*Shown are three structural prototypes developed during the design process.*

The pendant features an abstract eye shape that emphasizes the idea of self-reflection—literally and metaphorically. The choice of this form was driven by the desire to create a design that was non-invasive, neutral, and comfortable for a wide range of users. The pendant's central feature is a hollow hemisphere, which serves as a space to house the LED light strip.

The LED strip is housed within the hollow section of the pendant and is covered by a soft mirror-like material. This mirror surface provides a reflective quality, allowing the wearer to see their own gaze in the pendant when the LEDs are off. When the system detects negative emotions or inappropriate gaze behavior, the LEDs illuminate, changing color in response to the emotional state detected by the gaze tracking system.

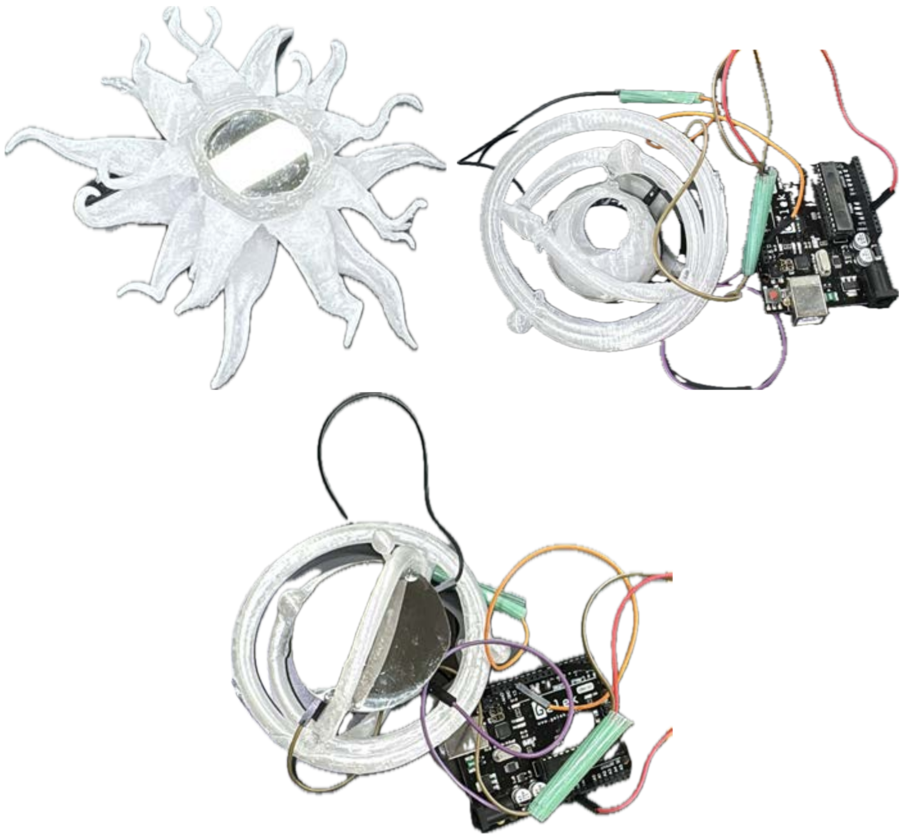


Figure 3.8 Final necklace prototypes with embedded LED cavities and soft mirror.

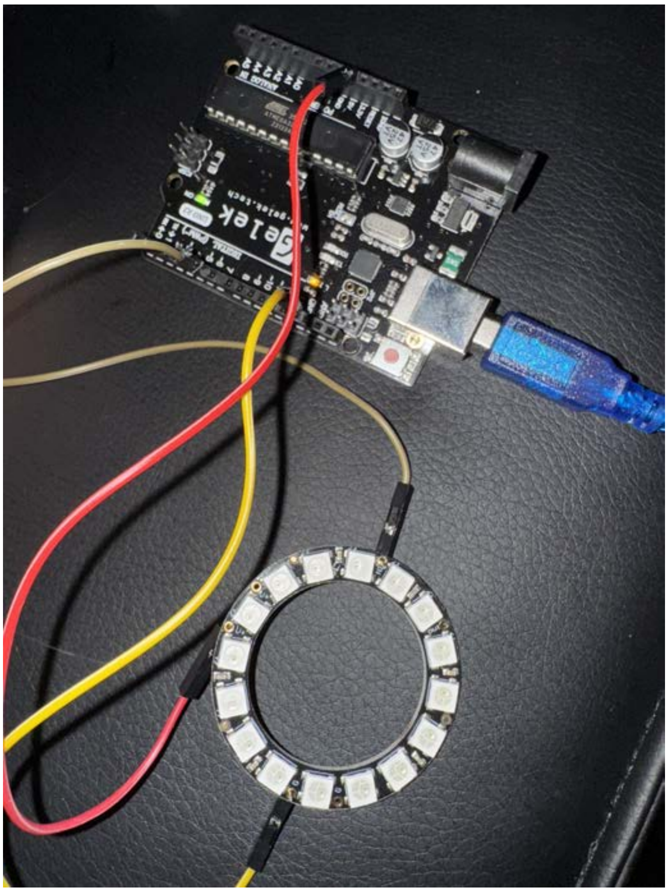


Figure 3.9 Connection of LED lights and the Arduino pad

### 3.5.2 Hardware System Assembly

The hardware system consists of multiple components that work seamlessly together to detect the user's gaze, process emotional data, and provide real-time visual feedback. The system components include Arduino microcontrollers, RGB LED strips, web cameras, and soft mirrors integrated into the wearable device.

The system includes two web cameras, which are critical for detecting and analyzing the user's gaze. Each of the camera are connected to an Arduino board, which processes the gaze data and sends it to a connected computer. The system operates in two configurations, with the Arduino board and cameras both connected to a central computer, where the data is processed and visual feedback is generated.

# Chapter 4
# Validation

## 4.1. Final Prototype Evaluation Experiment

### 4.1.1 Goals and Objectives

The primary objectives were to assess the system's accuracy in emotions estimation from facial expressions triggered by gaze, evaluate the effectiveness of color-emotion feedback, and determine how effectively the feedback influenced users' emotional self-awareness and behavioral adjustments.

### 4.1.2 Reliability Verification Test

**(1) Introduction and Purpose Clarification** This experiment is not intended to evaluate the emotional experience or subconscious gaze expression of the participants. Rather, it aims to verify whether the gaze-based emotion recognition model can correctly classify explicitly expressed emotional cues from the eye region, under controlled and instructed conditions.

This methodological distinction is essential in order to avoid potential confusion between deep-level affective communication and technical model validation. By establishing a strictly controlled testing environment, this study ensures that the evaluation of the model's classification accuracy is isolated from confounding variables related to social context, interpersonal dynamics, or emotional subtlety.

**(2) Experimental Rationale** The motivation for this experiment originated from a critical review pointing out the absence of a reliability assessment for the emotion recognition algorithm developed in Chapter 3. As the core algorithm uses appearance-based gaze information extracted from eye-region grayscale images, it is necessary to determine whether the system can accurately recognize basic emotional categories—particularly negative emotional states such as anger, fear, disgust, and sadness.

Because the model's input excludes broader facial configurations and dynamic gaze trajectories (such as saccades or pupil dilation), this experiment focuses solely on muscular cues surrounding the eye region, assuming participants are instructed to display a specific emotion. This allows for a clear comparison between the expected and predicted emotional labels.

Importantly, this test does not conflict with the overarching research interest in spontaneous gaze interactions. Rather, it serves as an isolated module to confirm that the employed machine learning system is not randomly assigning emotional labels, thereby enhancing the credibility of subsequent interactional findings.

**(3) Participants and Stimuli** A total of 8 participants (6 female, 2 male; aged 20–30) took part in this controlled validation experiment.

Each participant was asked to perform seven distinct facial expressions corresponding to the emotional categories mentioned in Chapter 3. Participants were verbally instructed which emotion to express in each trial. No sample images or facial prompts were provided to avoid bias from specific muscle mimicry. Participants were told to focus only on expressing the emotion using eye-region muscles without moving their heads or making overt lower-face gestures.

**(4) Procedure** Each participant performed all seven emotions, each repeated three times, for a total of 21 samples per participant (168 samples overall).

- Participants sat in front of a fixed webcam.
- Each trial began with a short verbal instruction from the researcher (e.g., "Please show anger").
- After 5 seconds, a snapshot of the eye region was taken using the webcam.
- The image was automatically fed into the trained emotion recognition system, which output one of the seven labels.

No further feedback or correction was provided to participants. Sessions were conducted in a quiet indoor environment with stable lighting.

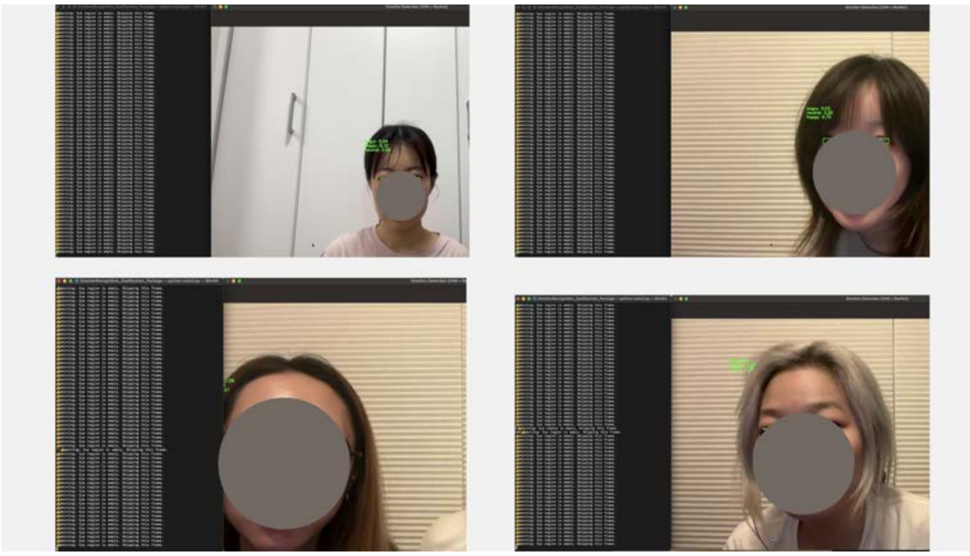


Figure 4.1 Reliability Verification Test

**(5) Evaluation Metric** To simplify analysis and avoid overloading the reader, we employed **Top-1 Accuracy** as the sole evaluation metric: the proportion of cases where the model's predicted label matched the instructed emotional category.

This metric is intuitive and commonly used in image classification benchmarking, especially when dealing with a small, balanced label set.

**(6) Results** The performance across the seven tested emotional states is summarized in Table 4.1. The highest accuracy was observed in the recognition of Happy (95.8%) and Surprise (87.5%), indicating that the model effectively captured distinctive eye-region features associated with these emotions. Conversely, Disgust (58.3%) and Neutral (62.5%) yielded lower accuracy, likely due to their relatively subtle ocular cues or overlap with other categories such as Sadness or Fear.

Overall, the model achieved a Top-1 Accuracy of 73.8%, suggesting a generally robust capability to classify explicit eye-region emotional expressions under controlled conditions. While this result does not represent performance in naturalistic or spontaneous settings, it provides initial quantitative support for the model's functional reliability in structured gaze-based feedback systems.

Table 4.1 Top-1 Accuracy per Emotion Category

| **Emotion** | **Correct Predictions** | **Top-1 Accuracy (%)** |
|---|---|---|
| Neutral | 15 / 24 | 62.5% |
| Happy | 23 / 24 | 95.8% |
| Angry | 18 / 24 | 75.0% |
| Fear | 16 / 24 | 66.7% |
| Disgust | 14 / 24 | 58.3% |
| Sadness | 17 / 24 | 70.8% |
| Surprise | 21 / 24 | 87.5% |
| **Overall Average** | – | **73.8%** |

This result, though not representing spontaneous emotion decoding, confirms the model's functional readiness to serve as a reactive component in emotion-aware interaction systems, justifying its application in subsequent dyadic and real-world scenarios (Parts 1–3). It is important to note that the aim of this experiment was not to infer subconscious states, but to validate the model's ability to classify explicitly posed emotional expressions from the eye region alone—thereby delimiting its current capability while providing a benchmark for future upgrades.

**(7) Summary and Integration** This verification procedure strengthens the experimental foundation of the thesis by providing empirical evidence for the accuracy of the core emotion recognition system. The Top-1 Accuracy of 73.8% indicates that the model reliably processes eye-region appearance features and maps them to standard emotional categories under instruction. This benchmark justifies its integration into the real-world interaction settings explored in Section 4.1.3 and supports subsequent affective feedback mechanisms based on gaze cues.

### 4.1.3 Feedback Consistency via AI-Face Induced Gaze

This experiment was designed to evaluate whether participants perceived the color-based emotional feedback generated by the wearable system as internally consistent, when responding to emotionally neutral AI-generated faces. The system's emotional prediction was based solely on the participant's gaze behavior and eye-region features. This setup provides an important verification of the symbolic feedback mechanism and addresses questions concerning whether the emotional feedback aligns with users' introspective impressions under controlled yet minimal social cues.

The objective of this experiment was to verify whether participants recognized the color feedback generated by the pendant device—corresponding to the model's interpretation of their gaze-induced emotion—as emotionally resonant and appropriate from their own subjective perspective. The experiment serves as a perceptual validation of the color–emotion mapping model used in the system.

Participants (n=8, same as in Section 4.1.2) were presented with AI-generated (Chat-GPT, Recraft, Zaohaowu) human face images. These faces were displayed on a computer screen, and were selected to be demographically balanced and culturally de-biased to reduce confounding effects related to identity, age, or race. Each participant viewed a unique AI face image.

Before the session began, participants were clearly informed of the emotion–color mapping logic used in the system (as defined in Section 3.2.5). This step ensured transparency regarding what emotional categories the colors represented.

During the task, each participant was instructed to fix their gaze at the eyes of the AI-generated face for a continuous duration of 30 seconds. No emotion was suggested or primed—participants were allowed to observe naturally. The pendant device affixed beneath the screen monitored the participant's gaze, captured static eye-region images in real time, and used the trained ResNet-based classifier to infer the dominant emotional state expressed through eye region appearance.

The resulting emotion label was then mapped to a corresponding color, which was presented through the pendant's RGB LEDs for approximately 5 seconds, starting at the end of the gaze period. Participants were then asked to rate the perceived appropriateness of the color feedback.

Immediately after each trial, participants rated the perceived consistency of the feedback color using a 5-point Likert scale, described as follows:

- 5 – Extremely consistent: The color felt exactly like what I was feeling.
- 4 – Somewhat consistent: The color roughly matched my internal state.

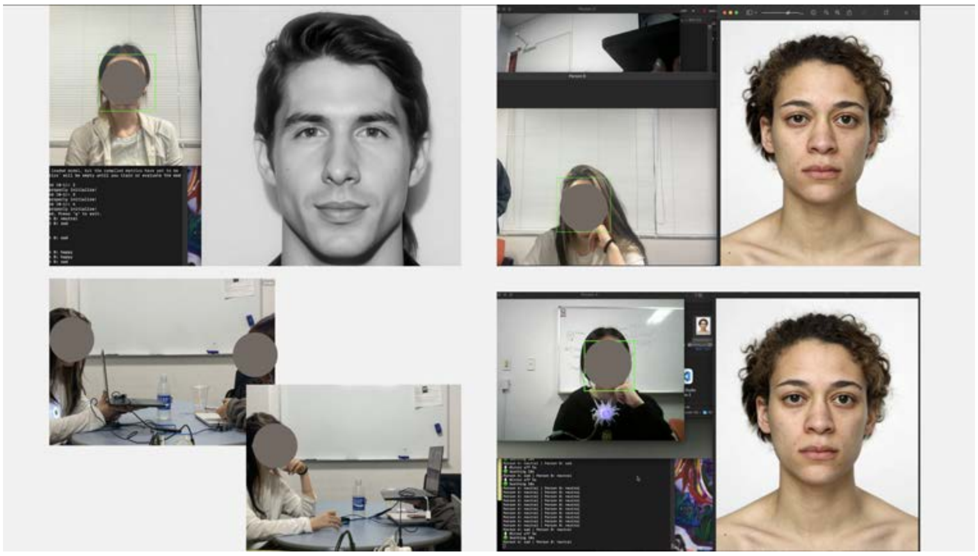


Figure 4.2 Reliability Verification Test 2

- 3 – Neutral: I could not judge or felt indifferent to the feedback.
- 2 – Somewhat inconsistent: The color felt slightly unrelated to my state.
- 1 – Very inconsistent: The color did not reflect my emotional state at all.

Each participant provided one rating based on their single session.

The collected Likert scores from all 8 participants were aggregated and summarized in Table 4.2. Although the AI face presented was emotionally neutral, the participants' gaze-related microexpressions still triggered emotional inference by the system. The average feedback consistency rating was 3.75, indicating a moderately high level of perceived alignment between internal emotional state and system output.

Table 4.2 Feedback Consistency Ratings (n = 8)

| **Participant ID** | **Likert Score (1–5)** |
|---|---|
| P01 | 4 |
| P02 | 3 |
| P03 | 4 |
| P04 | 5 |
| P05 | 3 |
| P06 | 4 |
| P07 | 3 |
| P08 | 4 |
| **Mean** | **3.75** |
| **SD** | **0.70** |

The findings suggest that even under emotionally neutral visual stimuli, participants were able to reflect upon and evaluate the symbolic emotional feedback generated from their gaze behavior. The generally positive ratings imply that the symbolic mapping of gaze-derived emotional inference to color cues is reasonably interpretable and accepted by users. It supports the utility of such feedback mechanisms for enhancing emotional self-awareness in real time, even when social cues are minimal.

This experiment also contributes a validation of perceptual resonance, complementing the classification reliability from Section 4.1.2. By assessing whether the system's emotional interpretation is accepted by users as personally relevant, this experiment strengthens the human-centered credibility of the feedback logic and positions it for broader deployment in future affective interfaces.

### 4.1.4 Hypothesis and Experimental Design

The overarching objective of this study is to investigate whether gaze-based real-time visual feedback can enhance emotional self-awareness in the emitter of gaze within naturalistic, dyadic interaction settings. I hypothesized that visual feedback through dynamic material changes (e.g., color shifts via hidden LEDs behind a two-way mirror) could help individuals become aware of their emotional state, especially when their gaze is perceived as negative or uncomfortable. By offering a reflective visual cue, the user would unconsciously adjust their emotional expression or gaze behavior. Rooted in social gaze literature and emotion-feedback interface studies, the experiment was designed to simulate diverse social scenarios in which gaze arises either spontaneously or under contextual modulation, while ensuring the experimental framework maintains ecological validity, analytical clarity, and methodological rigor.

To examine this research question in practice, the user study was divided into three structured parts: (1) Non-verbal gaze exchange between peers, (2) Verbal interaction with gaze feedback, and (3) Artificial face-based single-directional gaze simulation. Each part was carefully designed to probe a different modality or condition of gaze behavior and introspection, reflecting the multidimensional nature of gaze as both a social cue and an emotional signal.

### The Logic of Part 1–3

Each of the three parts was designed with a distinct purpose, and collectively they map onto a conceptual continuum from mutual, naturalistic gaze to isolated, reflexive gaze experiences.

**Part 1: Silent Mutual Gaze between Peers** In the first segment, two participants were seated face-to-face for two minutes without verbal communication. No specific instruction was given beyond maintaining a relaxed posture and facing each other naturally. This condition aimed to replicate implicit, unstructured gaze interactions as they occur in real-life casual encounters. The silent environment allowed spontaneous gaze to emerge based solely on social dynamics, without linguistic prompts or performative expectations. Real-time LED feedback was active, yet unobtrusive, designed to provide subtle visual cues without interrupting the natural rhythm of the interaction.

**Part 2: Verbal Interaction and Gaze Feedback in Dialogue** The second segment introduced a dynamic conversational element. Participants engaged in a free or thematically guided conversation for five minutes while seated in the same configuration. This part of the experiment aimed to simulate a more naturalistic interaction where participants might have experienced a broader range of emotional responses, such as curiosity, discomfort, or positive engagement. In other words, this phase was intended to examine how gaze-related feedback functions in cognitively engaging, emotionally expressive, and socially performative contexts.

The feedback mechanism remained active, enabling the study to observe how participants' emotional modulation or awareness shifted in more complex interpersonal exchanges.

**Part 3: Single-Directional Gaze toward AI-Generated Face** In the third segment, participants alternated roles interacting with a projected or printed AI-generated face. The face was produced using AI (Chat-GPT, Recraft, Zaohaowu) that ensured absence of cultural, gender, or emotional bias in its expression. Participants wore the device while facing the synthetic face, enabling the system to detect their emotional state through gaze-based cues and deliver visual feedback. This condition was designed to strip away interpersonal familiarity, social reciprocity, or identity-related influence, thereby simulating a gaze environment directed at a socially unfamiliar "other."

This phase also served to test the system's ability to induce reflective reactions even in low-context settings, and to contrast the emotional impact of interacting with a "stranger-like" artificial face as opposed to a human peer. Importantly, this design reflects literature that highlights increased emotional awareness in interactions with emotionally ambiguous stimuli.

To conclude, Part 1 was designed to simulate unstructured, non-verbal gaze interactions between two individuals without constraining their social relations, while Part 3 introduced a one-directional gaze scenario with unfamiliar, AI-generated facial features to isolate reflexive responses. Part 2, in contrast, explored natural gaze emergence during verbal interaction, capturing emotional leakage under conversational conditions. Together, these three parts provide a spectrum of interaction conditions—silent, verbal, and synthetic—offering a triangulated view of how gaze feedback affects emotional introspection under varying social cognitive loads and contextual expectations.

### Clarification of Experimental Role Assignment

While the study's core research question is focused exclusively on the gaze sender, both participants were equipped with the pendant feedback device in all dyadic phases (Parts 1 and 2). This design decision was made deliberately, and does not imply a shift in analytical focus toward mutual gaze interaction or receiver perception.

**Rationale for Bilateral Device Deployment** Equipping both individuals with devices served three key purposes:

- **Ecological Validity**: In real-life interactions, gaze roles are fluid and often simultaneous. Wearing devices on both participants allowed them to alternate spontaneously between being gaze senders and recipients. However, only the sender role was analytically focused on.
- **Avoidance of Experimental Bias**: Asymmetric device deployment—i.e., giving one participant the device while the other remains untreated—would introduce visible role asymmetry and potentially distort gaze behavior (Argyle & Cook, 1976). Participants might act differently knowing they are being observed or not, triggering social desirability effects or observer pressure. Symmetric deployment minimized such bias.
- **Role-Specific Data Coding**: Even though both participants wore the device, all recorded data, observational logs, and questionnaire responses were coded with participant identifiers (A and B) and separated for role-specific analysis. This enabled a sender-centric analysis while still preserving ecological realism.

This approach mirrors best practices in social gaze research where symmetric exposure is used to reduce reactivity while still permitting asymmetric hypothesis testing. The system, in essence, was deployed symmetrically, but the analysis and interpretation remained asymmetrical.

#### Survey Design and Role-Specific Classification

To accompany the behavioral observation and LED feedback tracking, participants completed post-interaction questionnaires designed to capture subjective emotional reflections. Crucially, the questionnaire was not shared or co-written by dyads; each participant filled out a separate form to minimize mutual influence and role confusion.

The key question relevant to emotional awareness was:

> *"Did you sense any hidden emotions in the eyes of the person you were experimenting with?"*

The structure of the question intentionally positioned the responder as an active interpreter of gaze while maintaining introspective framing. Participants were also asked about their own emotional states, and how the feedback (if noticed) influenced their mood, comfort level, and self-reflection.

**Justification for Role-Based Separation in Survey** Responses were labeled and analyzed based on sender-receiver roles at the time of the gaze emission event, enabling comparisons across conditions. This allowed the study to answer nuanced questions such as:

- Do participants more easily detect hidden emotions when they are the ones gazing versus being gazed at?
- Does real-time feedback improve introspection only when the individual is in the sender role?

This structure also serves to mitigate interpretive asymmetry bias—i.e., the fact that one individual may overinterpret while the other underperceives the same event. By separating roles and evaluating individual reflections, the study preserved analytic clarity despite the complexity of dyadic design.

### 4.1.5 Setup and Testing Procedure

This section outlines the detailed procedure for the experiment, designed to assess how the gaze-reflecting wearable device influences emotional awareness through real-time feedback. The study involves three main parts, focusing on capturing gaze data, emotional responses, and real-time feedback. The experimental design ensures both controlled and naturalistic conditions to validate the system's ability to detect and respond to emotional states based on gaze behavior.

1. **Part 1: Baseline Data Collection (2 minutes)**
   The first phase of the experiment aimed to establish a baseline for the participants' gaze and emotional responses in a neutral setting. During this phase, two participants were seated face-to-face without engaging in verbal communication. The participants wore the necklace devices equipped with real-time eye-tracking capabilities, which monitored their gaze and detected any emotional shifts based on their eye movement patterns.

   This phase served as a control period where no active interaction occurred between the participants, ensuring that the recorded gaze data reflected only their natural and undisturbed emotional states. The goal was to collect baseline gaze data and assess initial emotional responses to the mere act of looking at another person without any additional stimuli.

   In the first segment, two participants were seated face-to-face without engaging in verbal communication. Participants were instructed to remain relaxed and natural, facing one another without explicitly being told to maintain eye contact.

   The pendant devices were activated, and the LED feedback system responded to real-time emotion inference based on eye-region features. Each participant's video feed was independently analyzed using a pre-trained python program, and corresponding emotional categories were mapped to color outputs (see Section 3.2.5).

   No tasks were assigned other than sitting still for 2 minutes. This allowed organic gaze patterns to emerge, reflecting the subtle emotional dynamics that naturally occur in peer-to-peer settings without overt communication.

2. **Part 2: Interactive Conversation (5 minutes)**
   Following the baseline period, the same two participants, still wearing the necklace devices, engaged in a 5-minute conversation. The conversation

could have been structured, with a predefined topic such as a current event or general discussion, or it could have been spontaneous, allowing the participants to interact freely. The format of the conversation was counterbalanced to ensure variability in the interaction dynamics. This part was designed to simulate socially embedded gaze contexts, where gaze is accompanied by verbal cues, turn-taking, and emotional tone modulation. The pendant devices remained active throughout the session. Emotional states were continuously inferred, and the LED feedback was displayed dynamically based on gaze direction and intensity.

During this phase, the gaze and estimated emotional data were continuously recorded. The real-time feedback mechanism embedded in the necklace device responded accordingly to any significant changes in gaze behavior, providing immediate visual feedback through color shifts in the device.

3. **Part 3: AI-Generated Face Mask** The third and final part of the experiment introduced a non-human interaction object to simulate the gaze directed at strangers in everyday situations: an AI-generated face. Participants took turns interacting with the AI face while wearing active devices. Each participant gazed at the image for 30 seconds.

   During each trial, the system processed eye region features, inferred emotional signals, and responded via LED feedback based on its trained classification model.

   As emphasized earlier, this setup created a one-way gaze scenario, simulating an encounter with a stranger lacking verbal or emotional interaction. The objective was to control for social familiarity bias present in the first and second parts and to assess the consistency and clarity of feedback perception in a standardized gaze scenario.

The system will analyze the participants' gaze and emotional expressions in real-time using eye-tracking technology connected via Python code. The feedback mechanism will adjust the LED lighting on the wearable device based on the detected emotional state:

- **Neutral or Positive Gaze:** If the gaze is neutral or positive, no visible change will occur in the device's lighting, or a gentle color illumination will be triggered to reinforce positive emotional interaction.

- **Negative or Hostile Gaze:** If a negative or hostile gaze is detected, the LEDs concealed behind the two-way mirror will illuminate, shifting to warmer colors, thus providing visual feedback intended to moderate emotional expression.

Emotional and behavioral responses will be recorded through video, photo, and brief post-session interviews. Each session will last approximately 5-15 minutes. Participants are informed that they can withdraw from the experiment at any time without providing reasons, ensuring their autonomy and comfort throughout the study.

## 4.2. Results of the Experiment

This section presents a thorough analysis of the data collected from the experiment, which aimed to evaluate the effectiveness of the real-time gaze feedback system in promoting emotional self-awareness and emotional regulation.

The experiment involved five groups of participants, each consisting of two individuals. In addition to the primary experimental procedures, a pre-experiment questionnaire was administered to all participants (n=10) to assess their baseline attitudes toward gaze and emotional interpretation. All participants identified as female, aged between 20 and 25 years, and reported moderate to high sensitivity to others' gazes in daily interactions. Specifically, 90% affirmed that they often or always noticed others' gaze, and 90% also believed that gaze can convey emotional states. Furthermore, 70% of participants reported feeling either "uncomfortable" or "very uncomfortable" when subjected to prolonged eye contact.

Interestingly, 80% of participants recalled having experienced instances where someone's gaze communicated a hidden emotional message, despite the absence of verbal cues. When asked about their current emotional state prior to the experiment, all participants reported feeling neutral. Regarding color associations, blue was most commonly associated with calmness (60%), while red was overwhelmingly identified as a source of tension or discomfort (70%).

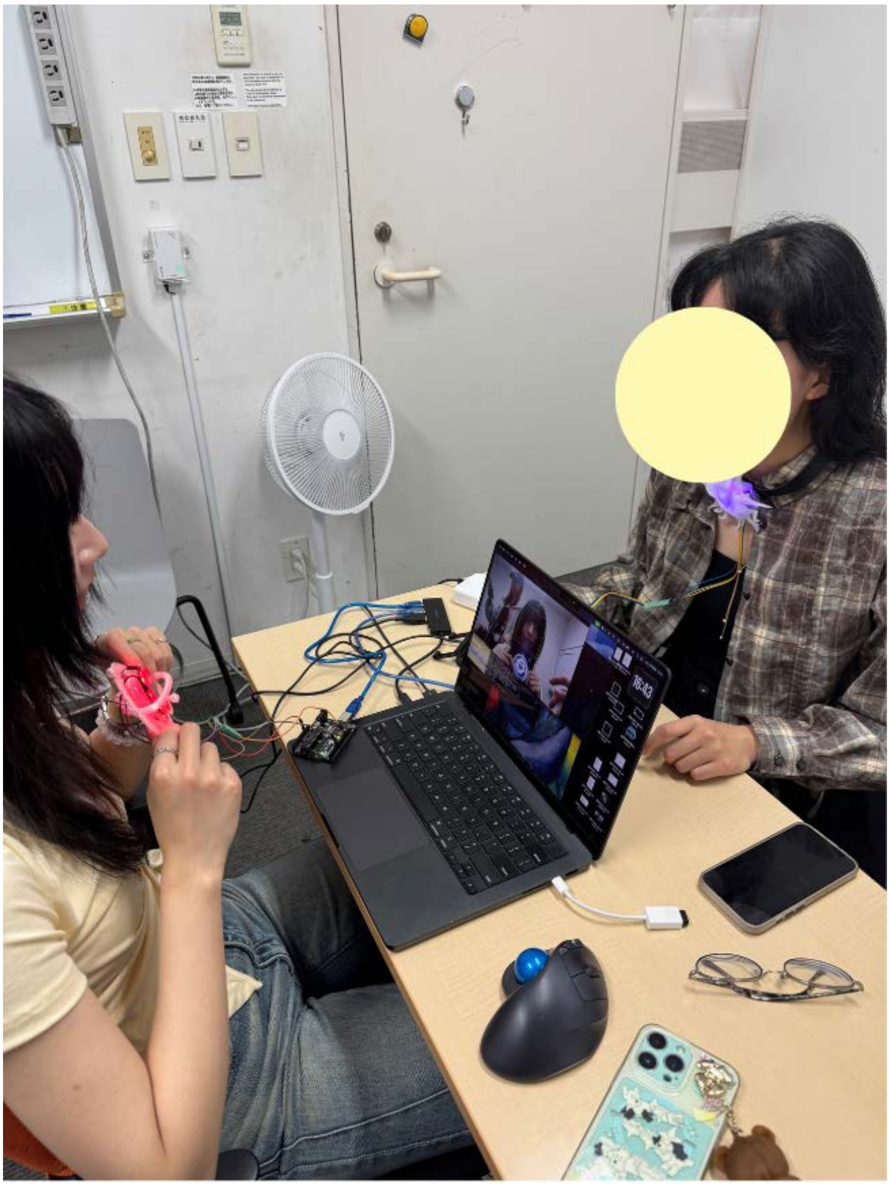


Figure 4.3 Evaluation Experiment

These insights serve to contextualize participants' predispositions toward gaze sensitivity and emotional interpretation. Although not formally analyzed as part of the core evaluation, this preliminary data supports the experimental design by confirming the psychological salience of gaze-related interactions in this co-

Figure 4.4 Evaluation Experiment

hort. It also justifies the use of color-emotion feedback in the device, grounded in participants' subjective associations between emotional states and visual stimuli.

Besides, based on the main research question—whether gaze-based feedback triggers self-awareness in the gaze sender—the analysis in this section does not distinguish between the roles of sender and receiver. Although all participants alternated between being gaze senders and gaze receivers in the experimental protocol, the analysis method combines the responses of all participants. This strategy aims to strictly maintain consistency with the theoretical focus of the study (i.e., arousing emotional self-awareness on the sender's side) and avoid interpretive ambiguity arising from role distinctions, as discussed in Section 4.1.4.

### 4.2.1 Emotion Estimation Performance

This subsection aims to evaluate the consistency between the system's emotion recognition output and participants' subjective perception of their own emotional states, as reported in the post-task questionnaires. The results presented here address RQ1, which asks whether the gaze-reflecting system can effectively detect and reflect the emotional valence conveyed through gaze in real-time social interactions. To this end, we analyzed data from three experimental phases, focusing on the Top-1 Emotion–Color Match Rate—defined as the percentage of participants whose self-reported main emotion (per After Phase questionnaire) matched the system-detected top emotional category associated with the corresponding color illumination.

#### Phase-Wise Accuracy Comparison

The experiment was conducted in three distinct stages: Part 1 (Silent Eye Contact), Part 2 (Conversation Task), and Part 3 (One-on-One AI Face Observation). Each stage presented distinct interaction dynamics and social gaze conditions. The corresponding Top-1 match rates were as follows:

- **Part 1**: 6 out of 10 participants (60.0%) showed consistent Top-1 matches.
- **Part 2**: 7 out of 10 participants (70.0%) showed consistent Top-1 matches.
- **Part 3**: 8 out of 10 participants (80.0%) showed consistent Top-1 matches.

These figures indicate a gradual improvement in recognition accuracy across phases. A possible explanation is that increasing engagement and clearer emotional stimuli in later parts (e.g., conversational content or the emotionally ambiguous AI face) may have enhanced the salience of gaze-based expressions, facilitating more accurate emotion detection by the algorithm.

A common trend across mismatches was a systemic confusion between *disgust* and *sadness*, particularly in low-light or low-gaze-variation conditions. Visual inspection of the facial landmark detection logs indicated that, in these instances, subtle movements around the eyes and eyebrows may have been underrepresented due to partial occlusion or angle distortion.

**Phase-Specific Insights**

**Part 1** presented the most variable results, likely due to the absence of verbal or affective cues and the relatively short duration (2 minutes). Many participants reported neutral or confused emotional states, increasing classification ambiguity.

In contrast, **Part 2** introduced verbal interaction, producing clearer emotional experiences (e.g., "happy," "curious," or "nervous"), which were more accurately picked up by the system. This supports the hypothesis that emotion-enriched contexts enhance classification reliability.

In **Part 3**, although only one participant wore the device, the surreal AI-generated face stimulus elicited intense affective reactions. Participants frequently reported feelings such as discomfort, "eeriness," or empathy, which were reflected in the system's detection of "disgust" or "sadness" states. These findings suggest that emotionally provocative stimuli enhance detection accuracy even in one-way interactions.

**Triangulation with Observational Notes**

Experimenter observation logs supported the accuracy patterns. High Top-1 match sessions were often accompanied by stable gaze direction, visible microexpressions, and prolonged eye contact. In contrast, mismatched sessions involved frequent gaze aversion, blinking, or flattened affect, reducing detection reliability.

**Technical Limitations**

The classification algorithm combined gaze direction and facial expression estimation using pretrained models. However, ambient lighting, mirror surface reflection, head pose angle, and camera resolution introduced noise. These factors may explain decreased performance in ambiguous or low-motion conditions.

In addition, the prototype device's limited field of view and processing latency (especially during head movements) likely contributed to missed cues.

### 4.2.2 Emotional Self-Awareness Analysis

This section investigates how the gaze-triggered visual feedback influenced participants' emotional self-awareness under different experimental conditions. To address RQ1—*How can real-time, gaze-based feedback arouse the emotional self-awareness of the sender?*—we analyzed four Likert-type questions covering key perceptual and reflective aspects of feedback effectiveness.

**Measured Items** The four core statements were evaluated using 5-point Likert scales (1 = Strongly disagree, 5 = Strongly agree):

- Did you notice any change in your gaze?
- Did the LED color feedback affect your emotions?
- Did the visual feedback make the interaction feel more natural or awkward?
- Do you think the visual feedback helped you regulate your emotions or awareness of your gaze?

The first three items were administered after each experimental part, while the fourth was part of the final post-experiment questionnaire.

Table 4.3 Mean and Standard Deviation of Feedback Perception Scores

| **Measure** | **Part 1** | **Part 2** | **Part 3** |
|---|---|---|---|
| Gaze Awareness | $4.00 \pm 0.67$ | $4.00 \pm 0.65$ | $3.60 \pm 0.97$ |
| Emotion Impact | $3.80 \pm 0.79$ | $3.80 \pm 0.86$ | $3.40 \pm 1.07$ |
| Interaction Naturalness | $3.70 \pm 0.82$ | $3.40 \pm 1.06$ | $3.30 \pm 1.25$ |

**Descriptive Statistics** The final reflection item—*"Do you think the visual feedback helped you regulate your emotions or awareness of your gaze?"*—received highly positive evaluations:

- Mean = 4.6, Standard Deviation = 0.52, $n = 10$
- 6 participants selected "Definitely", and 4 selected "Somewhat"
- No participant selected neutral or negative responses

**Phase-Based Analysis** Gaze awareness remained consistently high in Part 1 and Part 2 ($M = 4.0$), suggesting that the visual feedback was effective even during silent, non-verbal interaction. A slight decrease in Part 3 ($M = 3.6$) may be attributed to reduced emotional investment in the AI-generated face condition.

Emotion impact followed a similar trend, with Parts 1 and 2 both scoring $M = 3.8$, and Part 3 showing a reduction to $M = 3.4$. Interview data suggested that dynamic social interaction (Part 2) elicited more emotionally engaging reactions.

Interaction naturalness showed a declining trajectory:

- Part 1: $M = 3.7$
- Part 2: $M = 3.4$
- Part 3: $M = 3.3$

This may reflect a tension between increased awareness and reduced intuitive flow, possibly due to the novelty or artificiality of the device.

**Qualitative Reinforcement**

**Summary of Findings** Real-time visual feedback successfully promoted emotional self-awareness, particularly in interactive (Part 2) and mutual gaze (Part 1) contexts. Although perceived naturalness declined, participants consistently reported enhanced self-monitoring and affect regulation. The data support the hypothesis that gaze-driven color feedback functions as a third-person reflection tool, helping individuals mindfully adapt their expressive behavior.

### 4.2.3 Color Mapping Feedback Effectiveness

The effectiveness of the color feedback mechanism was evaluated based on how well participants responded to different colors and how these colors influenced their emotional states. The system was designed to map specific colors to emotional responses, with the goal of reinforcing emotional regulation through visual feedback. In the post-experiment questionnaire, participants were asked two direct questions:

- "Which LED feedback color made you feel calm?"
- "Which color feedback made you feel uncomfortable?"

Table 4.4 Subjective Evaluation of Color Feedback

| **Color** | **Soothing / Relaxing (n)** | **Uncomfortable / Confusing (n)** |
|---|---|---|
| Blue | 4 | 1 |
| Light Yellow | 5 | 0 |
| Green | 1 | 0 |
| Red | 0 | 8 |
| Gray | 0 | 1 |

These results reveal several notable trends:

- **Light Yellow** and **Blue** emerged as the most calming colors, each selected by 5 and 4 participants as contributing to emotional relaxation. Their high acceptability validates the system's post-feedback soothing logic, particularly for *Fear* and *Anger*, which were often followed by Blue or White in the feedback loop.
- In stark contrast, **Red** was selected by 8 out of 10 participants as the most uncomfortable or confusing feedback color, making it the most emotionally provocative color in the system. While effective in attracting attention, Red may risk overexposure and psychological stress when used in high frequency or without contextual explanation.
- **Gray**, used sparingly in the system, was selected once as discomforting, indicating some ambiguity in its affective interpretation.

The color feedback system was most effective when using calming colors like Blue and Green and Light Yellow. Participants in both groups overwhelmingly selected these colors as the most comforting and relaxing. Red was the color most frequently associated with discomfort and tension. This indicates that the color feedback system is effective in using color to influence emotional states. The overwhelming preference for Blue and Green demonstrates the potential of these colors to induce calmness and relaxation, while the discomfort induced by Red highlights the system's capacity to evoke emotional awareness and regulation.

## 4.3. Conclusion

The findings from this experiment provide meaningful insights into the application of real-time gaze detection and emotional feedback systems. Through an analysis of participants' emotional responses, gaze behaviors, and feedback on the effectiveness of the system, this study contributes to understanding how technology can enhance emotional self-awareness and regulate emotions in real-time interactions.

The real-time gaze feedback system proved to be effective in influencing participants' emotional awareness and self-regulation. The results from 4.2.1 Emotion Estimation Performance indicate that the system accurately detected shifts in gaze behavior and correlated these shifts with emotional states. However, the experiment also highlighted that emotional awareness, while triggered by external feedback, also relies on the participants' initial level of self-awareness and familiarity with their own emotions.

Participants demonstrated a strong connection to the feedback system. This suggests that the system was not only accurate but also perceptible to participants, offering valuable insights into how emotional cues can be externalized and recognized through gaze. And it proves the potential of gaze detection systems as tools for developing self-awareness, a key aspect of emotional regulation, in real-time interactions.

The analysis on color feedback underscores the significant role that color plays in shaping emotional states. Blue and Green were consistently identified as calming colors, reinforcing the well-established psychological association between these colors and relaxation. On the other hand, the negative emotional reactions triggered by Red align with previous research on color psychology, where Red is often linked to arousal and negative emotional states such as anger or stress.

What is particularly noteworthy is the participant feedback on color preferences, which suggests that individual emotional responses to colors may vary based on context and prior experiences. Although the system's color feedback proved effective in influencing emotional regulation, these findings suggest that there may be further opportunities to personalize color feedback systems to suit individual emotional needs or contexts. Future iterations of the system could incorporate adaptive color mappings that respond to individual emotional baselines or user preferences.

In addition, the ability of the system to prompt behavioral adjustments and enhance emotional self-awareness is a critical takeaway from the experiment. The results of emotional self-awareness analysis indicate that the visual feedback system helped participants become more aware of their emotional responses and modify their gaze behavior. One insight that emerged is the dual nature of emotional self-awareness: participants were not only aware of the emotional impact of their gaze on others but also the potential for their own emotional regulation. This is significant for future applications in therapeutic and social contexts, where understanding the emotional impact of one's non-verbal cues (e.g., gaze) can help improve interpersonal communication and emotional intelligence. Another key reflection from the data is the complex interplay between technology (gaze detection, color feedback) and human emotional responses. The experiment demonstrated that while real-time feedback can influence emotional states and behavior, the effectiveness of this feedback is shaped by individual differences. This suggests that the same technology may have varying impacts on different users, depending on their emotional sensitivity, previous experiences, and willingness to engage with the feedback system. The experiment also highlighted the importance of personalization in emotional feedback systems. Although the color-emotion mapping logic used in this study was based on common psychological associations, the emotional responses to color may differ across individuals and contexts. Therefore, adaptive feedback mechanisms, which take into account the emotional baseline of each user, could enhance the system's effectiveness in real-world applications, such as therapy or emotional regulation training.

In conclusion, the experimental results demonstrate the effectiveness of the wearable gaze detection and emotional feedback system in enhancing emotional self-awareness and regulating emotional responses. By providing real-time, color-based feedback based on gaze behavior, the system facilitated emotional regulation and promoted behavioral adjustments, highlighting the potential of this technology for applications in therapeutic and social contexts.

# Chapter 5
# Discussion and Future Works

## 5.1. Limitations

### 5.1.1 Technical Limitations of Webcam-Based Systems

While webcam systems are quick and easy to implement, there are some challenges. Many studies and the results of the final experiment have shown that they are generally much less accurate than infrared eye-tracking devices [37]. Furthermore, webcam-based eye tracking performs poorly, if at all, in low light. Many gaze estimation methods have always required cumbersome explicit personal calibration procedures to estimate subject-dependent eye-movement parameters. This can degrade the user experience.

One critical limitation of the current system lies in the underlying assumptions and constraints of the emotion recognition model. Although the project frames itself as targeting gaze-induced emotional awareness, the computational model used for inference does not directly analyze dynamic gaze features such as gaze direction, saccadic movement, or gaze duration—key indicators often emphasized in affective computing and social gaze research.

Instead, the model receives as input static grayscale images of the eye region and infers emotional states based primarily on the morphological features of periocular facial muscles. These features include subtle cues in the upper cheeks, brow tension, eyelid curvature, and eye openness—all of which are tightly associated with facial expressions rather than gaze behavior per se. As such, the model operates closer to a constrained facial expression recognition system than a true gaze intention decoder.

This distinction is crucial because it affects both the accuracy and interpretability of the feedback provided to users. While the system performs reasonably well in controlled settings—achieving a Top-1 accuracy of 70% in validation tasks—its reliance on static features makes it inherently limited in detecting emotions conveyed through more nuanced or culturally modulated gaze patterns (e.g., averted gaze signaling discomfort or dominance). Consequently, participants may receive feedback that reflects surface-level expressions rather than the deeper intentions or emotions communicated through their eye movements.

Furthermore, this architecture may introduce a mismatch between users' internal affective states and the system's output, especially in emotionally ambiguous scenarios. This mismatch becomes more pronounced in contexts where gaze is not accompanied by overt expression, such as social withdrawal, passive observation, or subtle intimidation. Without integrating dynamic visual input or gaze trajectory modeling (e.g., LSTM-based spatiotemporal encoding), the current system risks oversimplifying the emotional semantics of gaze.

This technical constraint should be viewed not only as a computational bottleneck but also as a conceptual limitation that constrains the generalizability of the findings. Future iterations of this work could benefit from the integration of multimodal gaze features, including real-time gaze vectors, fixation maps, and blink patterns, to more faithfully decode emotional intention embedded in gaze and better align system output with the subjective experience of users.

### 5.1.2 Limitations in Participant Demographics and Sample Size

The sample size used in this study was relatively small, consisting of only five groups, each containing two participants (Group A and Group B). The limited number of participants restricts the generalizability of the results to broader populations. Larger sample sizes would help to account for variability in emotional responses, gaze behaviors, and feedback preferences, providing a more comprehensive understanding of the system's applicability to different user groups.

Moreover, the study did not account for a diverse demographic of participants, such as those from different cultural backgrounds, age groups, or individuals with varying emotional regulation abilities. Previous research suggests that emotional expression and regulation can differ significantly across cultures, meaning that the findings from this study may not be fully applicable to populations with different cultural norms or emotional experiences.

Another challenge is the need for personal calibration in many eye-tracking methods. These systems often require cumbersome, subject-dependent calibration procedures to account for individual differences in gaze behavior. This can degrade the user experience by introducing additional setup time and reducing the system's reliability. Furthermore, appearance-based webcam methods suffer from poor generalization performance. Models trained on specific datasets often fail to generalize to new subjects, head poses, or environments due to the limited variation in training data. This means that the system may struggle to maintain consistent performance when applied to a broader range of real-world scenarios.

### 5.1.3 Potential Biases in Self-Reported Data

This study relied on self-reported data from participants through surveys and questionnaires, which introduces the potential for response bias. Participants may have been influenced by social desirability or may have interpreted emotional cues differently based on their own prior experiences or expectations. For instance, they may have been more likely to report that they noticed emotional changes in their gaze or were affected by the color feedback simply because they were aware of the experiment's goals.

Additionally, the self-reported data may not fully capture the subtlety of emotional shifts that occurred during the experiment. Participants may have failed to notice or accurately describe how their emotions were affected by the feedback, particularly in cases where the emotional shift was not overt or consciously acknowledged.

### 5.1.4 Limitations in Data Analysis

While the experiment provided valuable insights into emotional self-regulation, the data analysis methods used in this study were relatively simple, focusing primarily on quantitative measures such as emotional state ratings and gaze behavior. More sophisticated statistical models could be employed in future studies to provide deeper insights into the relationships between gaze behavior, emotional feedback, and long-term emotional regulation. Advanced data analysis techniques, such as machine learning models, could be used to identify patterns and correlations in the data that might not be immediately apparent through basic analysis.

Moreover, while the system was successful in tracking gaze behavior and color feedback, cross-validation with other emotional state measurement tools (e.g., physiological data such as heart rate or skin conductance) could provide a more holistic understanding of emotional responses and offer more robust validation of the feedback system.

## 5.2. Future Work

### 5.2.1 Enhancing Gaze Tracking Accuracy

Future research should aim to improve the accuracy of gaze tracking, particularly for webcam-based systems. As noted earlier, infrared systems provide more precise tracking, and combining webcam systems with infrared technologies could offer an affordable yet accurate alternative. Additionally, reducing calibration time and improving its efficiency will be essential for enhancing user experience, particularly for non-expert users.

### 5.2.2 Personalization of Feedback

The system could be made more personalized by incorporating adaptive algorithms that take individual differences into account. For example, users could undergo an initial emotional baseline assessment where the system learns the user's emotional triggers and preferences. Over time, the system could adjust its feedback based on the user's emotional responses, providing more accurate and effective interventions. This would be particularly useful for individuals with specific emotional regulation challenges, such as anxiety or mood disorders.

### 5.2.3 Longitudinal and Cross-Cultural Studies

To better understand the long-term benefits of the system, longitudinal studies should be conducted, tracking users over extended periods of time to observe how the feedback system influences their emotional regulation and behavior in real-world settings. Additionally, cross-cultural studies should explore how people from different cultural backgrounds respond to color feedback and gaze cues, as emotional expressions and interpretations of gaze can vary significantly across cultures. Given that this study primarily assessed emotional responses to gaze without addressing gender-specific nuances, it would be valuable to explore how gendered expectations of gaze influence the emotional feedback system. Another avenue for future research is the exploration of how power dynamics, including surveillance and control, are enacted through gaze in mediated environments like wearable devices.

### 5.2.4 Integration with Other Feedback Mechanisms

Future systems could incorporate multimodal feedback (e.g., auditory or haptic feedback) in addition to visual cues. Combining visual color feedback with auditory or tactile sensations could provide a more comprehensive emotional regulation tool. For instance, combining calming colors with soothing sounds or vibrations may enhance the overall emotional impact and promote greater emotional self-regulation.

## 5.3. Conclusion

This study has provided valuable insights into the use of wearable gaze detection and emotional feedback systems for enhancing emotional self-awareness and regulation. The experimental results demonstrated that the system successfully influenced participants' emotional responses, helping them become more aware of their gaze behaviors and emotional states. While the system showed promise, limitations in accuracy, calibration, and environmental adaptability were identified.

Future work should focus on improving the accuracy of gaze tracking, personalizing feedback mechanisms, and exploring long-term effects through longitudinal studies. By refining these aspects, the system has the potential to become a powerful tool for emotional regulation and could have broad applications in therapy, social skill training, and everyday emotional well-being.

# References

[1] Nathan J. Emery. The eyes have it: The neuroethology, function and evolution of social gaze. *Neuroscience and Biobehavioral Reviews*, 24:581–604, 2000. doi:10.1016/S0149-7634(00)00025-7.

[2] Paweł Tarnowski, Marcin Kołodziej, Andrzej Majkowski, and Remigiusz Jan Rak. Eye-tracking analysis for emotion recognition. *Computational Intelligence and Neuroscience*, 2020(2909267), 2020.

[3] Chizuko Ueno. *Onnagirai: Nippon no misojinī = Misogyny*. Kinokuniya Shoten, 2010. URL: https://nla.gov.au/nla.cat-vn5019874.

[4] John Berger. *Ways of Seeing*. Penguin Books, London, 1972.

[5] Michel Foucault. Discipline and punish: The birth of the prison. 1977.

[6] Laura Mulvey. Visual pleasure and narrative cinema. *Screen*, 16(3):6–18, 1975. doi:10.1093/screen/16.3.6.

[7] Stop Street Harassment YouGov. Unsafe and harassed in public spaces, 2016. URL: https://www.stopstreetharassment.org/unsafe-harassed-in-public-spaces-report-2016/.

[8] ActionAid. Safe cities for women: Global report, 2016. URL: https://actionaid.org.uk/safe-cities-for-women.

[9] Australia Institute. Polling: Safety in public places, 2015. URL: https://www.tai.org.au/research-paper/safety-public-places.

[10] MyProtein. Women in the gym survey, 2021. URL: https://www.myprotein.com.

[11] Gallup. Law and order index, 2014. URL: https://news.gallup.com.

[12] Bell Hooks. The oppositional gaze: Black female spectators. In Peg Zeglin Brand Weiser and Carolyn Korsmeyer, editors, *Feminism and Tradition in Aesthetics*, pages 142–159. 1995.

[13] Michael Argyle, Mark Cook, and Duncan Cramer. Gaze and mutual gaze. *British Journal of Psychiatry*, 165(6):848–850, 1994. `doi:10.1017/S0007125000073980`.

[14] James J. Gibson and Herbert L. Pick. Perception of another person's looking behavior. *Psychological Review*, 70:184–196, 1963. `doi:10.1037/h0043806`.

[15] Daniel C. Richardson Matthias S. Gobel, Heejung S. Kim. The dual function of social gaze. *Cognition*, 136:359–364, 2015.

[16] Kenneth Holmqvist, Marcus Nyström, and Richard et al. Andersson. *Eye Tracking: A Comprehensive Guide to Methods and Measures*. Oxford University Press, 2011. URL: `https://global.oup.com/academic/product/eye-tracking-9780198738596`.

[17] Jiwan Kim, Doyoung Lee, Jaeho Kim, and Ian Oakley. Can eye gaze improve emotional state detection on off the shelf smart devices. *2023 IEEE International Conference on Big Data and Smart Computing (BigComp)*, pages 378–380, 2023.

[18] Paniz Azizi. *Assessing Associations Between Gaze Perception Discrepancies and Social Anxiety Symptoms*. PhD thesis, 2024. - Database copyright ProQuest LLC; ProQuest does not claim copyright in the individual underlying works; - 2024-08-26. URL: `https://kras.lib.keio.ac.jp/dissertations-theses/assessing-associations-between-gaze-perception/docview/3086483114/se-2`.

[19] M. et al. Chen. In the eye of the beholder: Eye contact increases resistance to persuasion. *Psychological Science*, 24(8):1298–1303, 2013. `doi:10.1177/0956797613480804`.

[20] Nicola J. Gregory, Helen Bolderston, and Jastine V. Antolin. Attention to faces and gaze-following in social anxiety: Preliminary evidence from a nat-

uralistic eye-tracking investigation. *Journal of Anxiety Disorders*, 58:67–74, 2018. doi:10.1016/j.janxdis.2018.06.019.

[21] Jiemiao Chen, Esther van den Bos, Julian D. Karch, and P. Michiel Westenberg. Social anxiety is related to reduced face gaze during a naturalistic social interaction. *Journal of Anxiety Disorders*, 82:16–23, 2022. doi:10.1016/j.janxdis.2022.102464.

[22] N. Hadjikhani, N. R. Zurcher, A. Lassalle, L. Hippolyte, N. Ward, and J. Åsberg Johnels. The effect of constraining eye-contact during dynamic emotional face perception—an fmri study. *Social Cognitive and Affective Neuroscience*, 12(7):1197–1207, 2017. doi:10.1093/scan/nsx046.

[23] R. Cañigueral and A. F. Hamilton. The role of eye gaze during natural social interactions in typical and autistic people. *Frontiers in Psychology*, 10:560, 2019. doi:10.3389/fpsyg.2019.00560.

[24] Vasileios Skaramagkas, Emmanouil Ktistakis, Dimitrios Manousos, Eleni Kazantzaki, Nikolaos S. Tachos, Evanthia Tripoliti, Dimitrios I. Fotiadis, and Manolis Tsiknakis. esee-d: Emotional state estimation based on eye-tracking dataset. *Brain Sciences*, 13(4):589, March 2023. URL: https://doi.org/10.3390/brainsci13040589, doi:10.3390/brainsci13040589.

[25] Jia Zhen Lim, James Mountstephens, and Joon Teo. Emotion recognition using eye-tracking: Taxonomy, review and current challenges. *Sensors*, 20(8):2384, 2020. URL: https://doi.org/10.3390/s20082384, doi:10.3390/s20082384.

[26] Jari K. Hietanen. Affective eye contact: An integrative review. *Frontiers in Psychology*, 9:1587, 2018. URL: https://doi.org/10.3389/fpsyg.2018.01587, doi:10.3389/fpsyg.2018.01587.

[27] Beatrice de Gelder and Jean Vroomen. The perception of emotions by ear and by eye. *Cognition and Emotion*, 14(3):289–311, 2000.

[28] Alan S. Cowen, Dacher Keltner, Florian Schroff, Brendan Jou, Hartwig Adam, and Gautam Prasad. Sixteen distinct facial expressions occur in natural social interactions. *Science*, 369(6507):1519–1524, 2020.

[29] Xucong Zhang, Yusuke Sugano, Mario Fritz, and Andreas Bulling. Appearance-based gaze estimation in the wild. In *2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 4511–4520, 2015. doi:10.1109/CVPR.2015.7299081.

[30] Patricia Valdez and Albert Mehrabian. Effects of color on emotions. *Journal of Experimental Psychology: General*, 123(4):394–409, 1994. doi:10.1037//0096-3445.123.4.394.

[31] Xin Xie, Jie Cai, Hao Fang, Xiaodong Tang, and Tomoharu Yamanaka. Effects of colored lights on an individual's affective impressions in the observation process. *Frontiers in Psychology*, 13:938636, December 2022. URL: https://doi.org/10.3389/fpsyg.2022.938636, doi:10.3389/fpsyg.2022.938636.

[32] Heejin Lee. *Effects of Light-Emitting Diode (LED) Lighting Color on Human Emotion, Behavior, and Impression*. PhD thesis, Michigan State University, 2019.

[33] Andrew J. Elliot and Markus A. Maier. Color and psychological functioning: a review of theoretical and empirical work. *Frontiers in Psychology*, 5, 2014. URL: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3980918/, doi:10.3389/fpsyg.2014.00368.

[34] James A. Russell. Core affect and the psychological construction of emotion. *Psychological Review*, 110(1):145–172, 2003. doi:10.1037/0033-295x.110.1.145.

[35] Ali Mollahosseini, Behzad Hasani, and Mohammad H Mahoor. Affectnet: A database for facial expression, valence, and arousal computing in the wild. *IEEE Transactions on Affective Computing*, 10(1):18–31, 2017. doi:10.1109/TAFFC.2017.2740923.

[36] T.M. Sutton and J. Altarriba. Color associations to emotion and emotion-laden words: A collection of norms for stimulus construction and selection. *Behavior Research*, 48:686–728, 2016. URL: https://doi.org/10.3758/s13428-015-0598-8, doi:10.3758/s13428-015-0598-8.

[37] Wang Kang. *Robust and Accurate Eye-gaze Tracking and Its Applications.* PhD thesis, 2019. Database copyright ProQuest LLC; ProQuest does not claim copyright in the individual underlying works; 2023-06-21. URL: https://kras.lib.keio.ac.jp/dissertations-theses/robust-accurate-eye-gaze-tracking-applications/docview/2272213045/se-2.

# Appendices

## A. Arduino Code for LED Control in the Final Prototype

a

```
    int redPin = 9;
int greenPin = 10;
int bluePin = 11;

void setup() {
  Serial.begin(9600);
  pinMode(redPin, OUTPUT);
  pinMode(greenPin, OUTPUT);
  pinMode(bluePin, OUTPUT);
}

void loop() {
  if (Serial.available() > 0) {
    String data = Serial.readStringUntil('\n');
    int r, g, b;
    sscanf(data.c_str(), "%d,%d,%d", &r, &g, &b);
    analogWrite(redPin, r);
    analogWrite(greenPin, g);
    analogWrite(bluePin, b);
  }
}
```

b

```
    #include <Adafruit_NeoPixel.h>

#define PIN 6
#define NUMPIXELS 16

Adafruit_NeoPixel pixels(NUMPIXELS, PIN, NEO_GRB + NEO_KHZ800);

void setup() {
  Serial.begin(9600);
  pixels.begin();
  pixels.show();
}

void loop() {
  if (Serial.available() > 0) {
    String data = Serial.readStringUntil('\n');
    int r, g, b;
    sscanf(data.c_str(), "%d,%d,%d", &r, &g, &b);
    for(int i=0; i<NUMPIXELS; i++){
      pixels.setPixelColor(i, pixels.Color(r, g, b));
    }
    pixels.show();
  }
}
```

# B. Python Code for Gaze Detection in the Final Prototype

```
import cv2
import serial
import time
import numpy as np
```

```
from tensorflow.keras.models import load_model
from tensorflow.keras.preprocessing.image import img_to_array

# Load model (make sure it's trained on RGB images)
model = load_model("resnet_emotion_model.h5")

# Emotion labels
emotion_labels = ['angry', 'disgust', 'fear', 'happy', 'sad', 'surprise', 'neutral']

# Calming color mapping per emotion
calming_colors = {
    'angry': (255, 182, 193),    # pink
    'disgust': (0, 153, 0),      # green
    'fear': (0, 102, 204),       # blue
    'sad': (255, 255, 153),      # light yellow
}

# Normal emotion display colors
emotion_colors = {
    'happy': (255, 230, 0),
    'surprise': (255, 255, 224),
    'neutral': (128, 128, 128),
    'angry': (204, 0, 0),
    'disgust': (0, 153, 0),
    'fear': (0, 102, 204),
    'sad': (0, 102, 204),
}

# Connect serial ports (replace with your ports)
serA = serial.Serial('/dev/tty.usbmodem21301', 9600)  # Arduino UNO PWM
serB = serial.Serial('/dev/tty.usbmodem21401', 9600)  # GELEK UNO NeoPixel

# Helper function to send color to Arduino
def send_color(ser, color):
    r, g, b = color
    ser.write(f"{r},{g},{b}\n".encode())
```

```
    time.sleep(0.05)

# Haar cascade face/eye detection
face_cascade = cv2.CascadeClassifier(cv2.data.haarcascades + 'haarcascade_frontalface_d
eye_cascade = cv2.CascadeClassifier(cv2.data.haarcascades + 'haarcascade_eye.xml')

# Auto-detect cameras
max_tested = 5
available_cams = []
for i in range(max_tested):
    cap = cv2.VideoCapture(i)
    if cap.isOpened():
        available_cams.append(i)
        cap.release()

if len(available_cams) < 2:
    print(" Cannot find two cameras!")
    exit()

camA = cv2.VideoCapture(available_cams[0])
camB = cv2.VideoCapture(available_cams[1])
print(" System fully initialized. Press 'q' to exit.")

# Main loop
while True:
    retA, frameA = camA.read()
    retB, frameB = camB.read()
    if not retA or not retB:
        print(" Failed to grab frames")
        break

    def detect_emotion(frame):
        face_gray = cv2.cvtColor(frame, cv2.COLOR_BGR2GRAY)
        faces = face_cascade.detectMultiScale(face_gray, 1.3, 5)
        for (x, y, w, h) in faces:
            roi_color = frame[y:y+h, x:x+w]
```

```
            eyes = eye_cascade.detectMultiScale(roi_color)
            if len(eyes) > 0:
                ex, ey, ew, eh = eyes[0]
                eye = roi_color[ey:ey+eh, ex:ex+ew]
                eye_resized = cv2.resize(eye, (48, 48))
                eye_array = img_to_array(eye_resized) / 255.0
                eye_array = np.expand_dims(eye_array, axis=0)
                preds = model.predict(eye_array, verbose=0)[0]
                label = emotion_labels[np.argmax(preds)]
                return label, (x, y, w, h)
        return "neutral", (0, 0, 0, 0)

    emotionA, boxA = detect_emotion(frameA)
    emotionB, boxB = detect_emotion(frameB)

    # Draw rectangles
    if boxA[2] > 0:
        cv2.rectangle(frameA, (boxA[0], boxA[1]), (boxA[0]+boxA[2], boxA[1]+boxA[3]), (
    if boxB[2] > 0:
        cv2.rectangle(frameB, (boxB[0], boxB[1]), (boxB[0]+boxB[2], boxB[1]+boxB[3]), (

    # Display real-time frames
    cv2.imshow("Person A", frameA)
    cv2.imshow("Person B", frameB)

    print(f"Person A: {emotionA} | Person B: {emotionB}")

    # Main logic for light control
    negative = ['angry', 'fear', 'disgust', 'sad']
    if emotionA in negative or emotionB in negative:
        # 2s red flash
        for _ in range(4):
            send_color(serA, (204, 0, 0))
            send_color(serB, (204, 0, 0))
            time.sleep(0.25)
            send_color(serA, (0, 0, 0))
```

```
            send_color(serB, (0, 0, 0))
            time.sleep(0.25)
        # 5s black mirror
        send_color(serA, (0, 0, 0))
        send_color(serB, (0, 0, 0))
        print("Mirror off 5s")
        time.sleep(5)
        # 10s calming color
        calmingA = calming_colors.get(emotionA, (128, 128, 128))
        calmingB = calming_colors.get(emotionB, (128, 128, 128))
        send_color(serA, calmingA)
        send_color(serB, calmingB)
        print(" Soothing 10s")
        time.sleep(10)
    else:
        # Positive emotion: dim color, no delay
        colorA = emotion_colors.get(emotionA, (128, 128, 128))
        colorB = emotion_colors.get(emotionB, (128, 128, 128))
        dimA = tuple(int(c*0.5) for c in colorA)
        dimB = tuple(int(c*0.5) for c in colorB)
        send_color(serA, dimA)
        send_color(serB, dimB)

    # Exit
    if cv2.waitKey(1) & 0xFF == ord('q'):
        break

camA.release()
camB.release()
cv2.destroyAllWindows()
serA.close()
serB.close()
```

# C. Photos of AI-Generated Face Mask

Figure C.1 AI-Generated Face (Set 1).

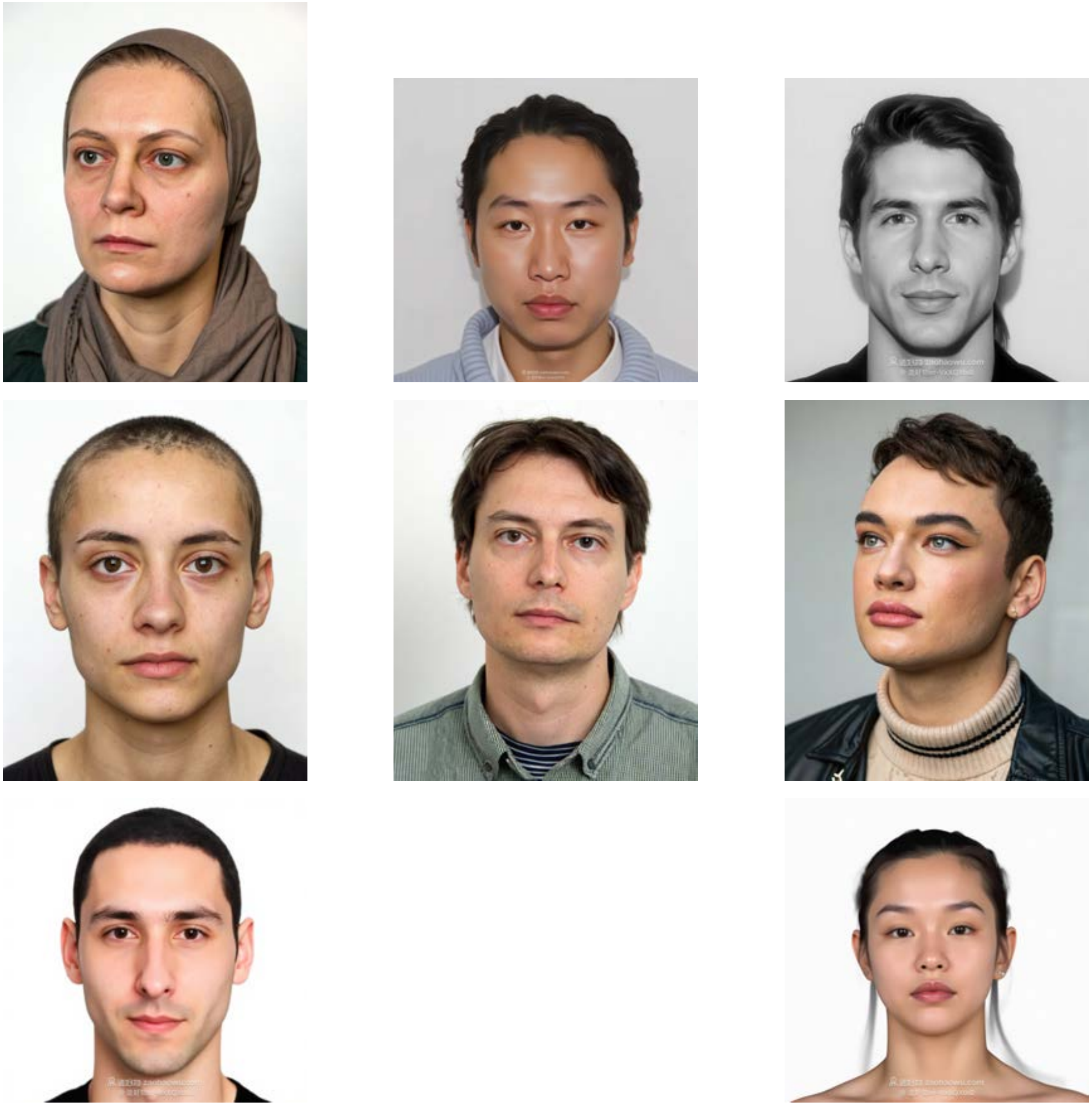


Figure C.2 AI-Generated Face (Set 2).

# D. Survey of the final Prototype

# Pre-Experiment Questionnaire 実験前アンケート

**The purpose of this survey is to gather insights into how people perceive gaze and the emotions behind it. Your responses will help us develop Gaze-Reflecting Interactive Materials based on Emotion Recognition through Eye-Tracking, which reflect emotions and enhance emotional awareness.The study also examines how these material changes influence the observer's emotional state, providing new ways to engage with and understand emotions.Thank you very much for your participation.**

**このアンケートにご参加いただきありがとうございます。この調査の目的は、人々が視線をどのように認識し、その背後にある感情をどのように理解するかについての洞察を得ることです。あなたの回答は、視線追跡による感情認識に基づき、感情を反映して感情認識を高める「視線を反映するインタラクティブ素材」の開発に役立てられます。また、この研究では、これらの素材の変化が観察者の感情状態にどのような影響を与えるかを検証し、感情に関与し理解を深める新しい方法を模索します。**

* 必須の質問です

*This survey is part of a research study. The data collected will be used solely for academic purposes. Unauthorized use, reproduction, or sharing of this form or its contents is prohibited.

*本調査は研究の一環として実施されています。収集されたデータは、学術目的にのみ使用されます。このフォームおよびその内容の無断使用、複製、共有は固く禁じられています。

1。 **What is your gender? 性別を教えてください。** *

*1 つだけマークしてください。*

- Female 女性
- Male 男性
- Other その他
- Prefer not to say 回答したくない

2。 **How often do you notice others' gazes during social interactions? 他人の視線に気付く頻度はどのくらいですか?** *

*1 つだけマークしてください。*

- Always 常に
- Often よくある
- Sometimes 時々
- Rarely ほとんどない
- Never 全くない

3。 **Do you think a person's gaze can convey their emotions? 人の視線が感情を伝えると思いますか?**

*1 つだけマークしてください。*

- Yes, definitely はい、確実に
- Somewhat ある程度
- Not really あまりそう思わない
- No いいえ

4。 **How comfortable do you feel when someone maintains prolonged eye contact with you? 他人と長時間視線を合わせることにどのくらい快適さを感じますか?** *

*1 つだけマークしてください。*

- Very comfortable とても快適
- Comfortable 快適
- Neutral 普通
- Uncomfortable 不快
- Very uncomfortable とても不快

5。 **In your opinion, can gazes unintentionally make someone feel uncomfortable, stressed, or insecure? 視線が意図せず他人を不快にさせたり、ストレスを与えたり、不安を感じさせたりすることがあると思いますか?** *

*1 つだけマークしてください。*

- Yes, very often はい、とてもよくある
- Occasionally 時々ある
- Rarely まれにある
- No いいえ

6。 **Have you ever felt that someone's gaze communicated a hidden emotional message? 誰かの視線が隠れた感情を伝えていると感じたことがありますか?** *

*1 つだけマークしてください。*

- Yes, frequently はい、頻繁に
- Sometimes 時々
- Rarely まれに
- Never 全くない

7。 **What is your current emotional state? 現在のあなたの気分はどのようなものですか?** *

*1 つだけマークしてください。*

- ◯ Happy 嬉しい
- ◯ Surprise 驚き
- ◯ Neutral 特に何もない
- ◯ Disgust 嫌悪
- ◯ Sad 悲しい
- ◯ Angry 怒り
- ◯ Fear 恐れ
- ◯ その他: ____________

8。 **Which color makes you feel calm?　どの色を見ると気持ちが落ち着きますか?** *

*1 つだけマークしてください。*

- ◯ Red 赤
- ◯ Yellow 黄色
- ◯ Blue 青
- ◯ Green 緑
- ◯ Purple 紫
- ◯ Pink ピンク
- ◯ Gray グレー
- ◯ White 白
- ◯ その他: ____________

9。 **Which color makes you feel tense or uncomfortable? どの色を見ると緊張や不快感を感じますか?** *

*1 つだけマークしてください。*

- Red 赤
- Yellow 黄色
- Blue 青
- Green 緑
- Purple 紫
- Pink ピンク
- Gray グレー
- White 白
- その他: ____________

このコンテンツは Google が作成または承認したものではありません。

Google フォーム

# Phase 1/2/3

**The purpose of this survey is to gather insights into how people perceive gaze and the emotions behind it. Your responses will help us develop Gaze-Reflecting Interactive Materials based on Emotion Recognition through Eye-Tracking, which reflect emotions and enhance emotional awareness.The study also examines how these material changes influence the observer's emotional state, providing new ways to engage with and understand emotions.Thank you very much for your participation.**

**このアンケートにご参加いただきありがとうございます。この調査の目的は、人々が視線をどのように認識し、その背後にある感情をどのように理解するかについての洞察を得ることです。あなたの回答は、視線追跡による感情認識に基づき、感情を反映して感情認識を高める「視線を反映するインタラクティブ素材」の開発に役立てられます。また、この研究では、これらの素材の変化が観察者の感情状態にどのような影響を与えるかを検証し、感情に関与し理解を深める新しい方法を模索します。**

*This survey is part of a research study. The data collected will be used solely for academic purposes. Unauthorized use, reproduction, or sharing of this form or its contents is prohibited.

*本調査は研究の一環として実施されています。収集されたデータは、学術目的にのみ使用されます。このフォームおよびその内容の無断使用、複製、共有は固く禁じられています。

1。 **How comfortable do you feel when someone maintains prolonged eye contact with you? 他人と長時間視線を合わせせることにどのくらい快適さを感じますか?**

*1 つだけマークしてください。*

- Very comfortable とても快適
- Comfortable 快適
- Neutral 普通
- Uncomfortable 不快
- Very uncomfortable とても不快

2。 **Did you sense any hidden emotions in the eyes of the people you were experimenting with in the same group? 同じグループで実験していた人の視線に、何か隠された感情を感じましたか?**

*1 つだけマークしてください。*

- Yes, frequently はい、頻繁に
- Sometimes 時々
- Rarely まれに
- Never 全くない

3。 **What do you think was your main emotional state during this experiment stage?**

**今回の実験段階で主に感じた感情は何と思いますか。**

*1 つだけマークしてください。*

- Happy 嬉しい
- Surprise 驚き
- Neutral 特に何もない
- Disgust 嫌悪
- Sad 悲しい
- Angry 怒り
- Fear 恐れ
- その他: ______

4。 **Did you notice any change in your gaze?**

**自分の視線に変化があったことに気付きましたか?**

*1 つだけマークしてください。*

| 0 | 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|---|
| ◯ | ◯ | ◯ | ◯ | ◯ | ◯ |

5。 **Did the LED color feedback affect your emotions? LEDの色のフィードバックはあなたの感情に影響しましたか?**

*1 つだけマークしてください。*

| 0 | 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|---|
| ◯ | ◯ | ◯ | ◯ | ◯ | ◯ |

6。 **Which color felt most comforting or helpful?**

**どの色が一番心地よく感じましたか?**

*1 つだけマークしてください。*

- Red 赤
- Yellow 黄色
- Blue 青
- Green 緑
- Purple 紫
- Pink ピンク
- Gray グレー
- White 白

7。 **Did the visual feedback make the interaction feel more natural or awkward?**

**視覚的なフィードバックによって、相互作用がより自然に感じましたか、それともぎこちなく感じましたか?**

*1 つだけマークしてください。*

| | 0 | 1 | 2 | 3 | 4 | 5 | |
|---|---|---|---|---|---|---|---|
| Ver | ◯ | ◯ | ◯ | ◯ | ◯ | ◯ | Very natural |

8。 **Other feelings**

# Post-Experiment Questionnaire 実験後アンケート

**The purpose of this survey is to gather insights into how people perceive gaze and the emotions behind it. Your responses will help us develop Gaze-Reflecting Interactive Materials based on Emotion Recognition through Eye-Tracking, which reflect emotions and enhance emotional awareness.The study also examines how these material changes influence the observer's emotional state, providing new ways to engage with and understand emotions.Thank you very much for your participation.**

**このアンケートにご参加いただきありがとうございます。この調査の目的は、人々が視線をどのように認識し、その背後にある感情をどのように理解するかについての洞察を得ることです。あなたの回答は、視線追跡による感情認識に基づき、感情を反映して感情認識を高める「視線を反映するインタラクティブ素材」の開発に役立てられます。また、この研究では、これらの素材の変化が観察者の感情状態にどのような影響を与えるかを検証し、感情に関与し理解を深める新しい方法を模索します。**

* 必須の質問です

*This survey is part of a research study. The data collected will be used solely for academic purposes. Unauthorized use, reproduction, or sharing of this form or its contents is prohibited.

*本調査は研究の一環として実施されています。収集されたデータは、学術目的にのみ使用されます。このフォームおよびその内容の無断使用、複製、共有は固く禁じられています。

1。 **Overall, how do you feel after completing the experiment?** *

**実験を終えて全体的にどのような気持ちですか?**

*1 つだけマークしてください。*

- Happy 嬉しい
- Surprise 驚き
- Neutral 特に何もない
- Disgust 嫌悪
- Sad 悲しい
- Angry 怒り
- Fear 恐れ
- その他: ______

2。 **Which led feedback color makes you feel calm? どのLEDの色が一番リラックスできましたか?** *

*1 つだけマークしてください。*

- Red 赤
- Yellow 黄色
- Blue 青
- Green 緑
- Purple 紫
- Pink ピンク
- Gray グレー
- White 白

3。 **Which color feedback made you feel uncomfortable?** *

**どの色のフィードバックがあなたを不快にさせましたか?**

*1 つだけマークしてください。*

- Red 赤
- Yellow 黄色
- Blue 青
- Green 緑
- Purple 紫
- Pink ピンク
- Gray グレー
- その他: ______________________

4。 **Do you think the visual feedback helped you regulate your emotions or awareness of your gaze?** *

**視覚的フィードバックが自分の感情や視線への意識を調整するのに役立ったと思いますか?**

*1 つだけマークしてください。*

- Definitely はい、明らかに役立った
- Somewhat ある程度役立った
- Not really あまり役立たなかった
- Not at all 全く役立たなかった

5。 **Are there any additional thoughts or concerns you have? 他に何か意見や懸念はありますか?**

______________________________________________

______________________________________________

______________________________________________

______________________________________________

______________________________________________