Doctoral Dissertation

The Effects of Appearance-Mediated Social Presence Regulation on Emotion and Performance

GONG Ziting

Supervisor: KANAI Hideaki

Division of Advanced Science and Technology
Japan Advanced Institute of Science and Technology
[Knowledge Science]

March 2026

Abstract

As virtual and augmented environments become increasingly central to everyday social interaction, supporting emotionally sustainable and effective communication in these spaces has emerged as a critical design challenge. Prior research has predominantly focused on how individuals' own avatars influence behavior, while the role of interaction partners' appearance in shaping emotional experience and performance remains underexplored. Addressing this gap, this dissertation adopts an individual-level experimental approach to examine how adjusting the appearance of others can regulate interaction experiences in mediated contexts.

Two empirical studies form the foundation of this work. Study 1 investigates high-pressure oral communication in video-mediated interviews, examining how different forms of interaction partner appearance—familiar faces, stylized representations, and neutral strangers—affect perceived social presence, state anxiety, attentional engagement, verbal behavior, gaze patterns, and physiological responses. Results show that familiar appearances reduce anxiety and promote deeper engagement, whereas stylized faces disrupt emotional interpretation and offer limited relief from evaluative stress.

Building on these findings, Study 2 extends the investigation to embodied interaction through augmented reality (AR) exercise companions. Across three experiments, the results demonstrate that emotionally resonant AR companions enhance enjoyment, reduce perceived pressure, and improve physical performance. Familiar avatars foster psychological safety through emotional bonding, while stylized companions support motivation and comfort when aligned with users' aesthetic preferences. Qualitative interviews further reveal that users assign distinct social and functional roles to avatars based on appearance, favoring companions that convey supportive presence without surveillance or judgment.

Together, these studies identify appearance (familiarity and liking as affective channels) mediated modulation of social presence as a unifying mechanism linking emotional experience and performance across cognitive and physical tasks. By extending social presence research beyond self-avatar effects to interaction partner appearance, this dissertation contributes theoretical insight into interpersonal regulation in mediated interaction. The findings also offer practical design implications for communication platforms, AR fitness systems, and future AI-driven companions, pointing toward more adaptive, empathetic, and psychologically supportive virtual communities.

Keywords: Avatar appearance; Social presence; Augmented reality; Human–computer interaction; Emotional experience; Performance.

Acknowledgements

I would like to express my sincere and deepest gratitude to my supervisor, Professor KANAI Hideaki, for his unwavering guidance and support throughout both my master's and doctoral studies. From the very beginning of my academic journey under his supervision, he has patiently guided me in choosing research topics, conducting research, and completing this thesis. Throughout these years, Professor KANAI has been exceptionally generous with his time, knowledge, and encouragement. When I encountered difficulties or felt uncertain about my research direction, he helped me carefully reflect on my goals and guided me to approach problems thoughtfully and systematically. He taught me to move forward step by step and to remain resilient in the face of challenges. The numerous books he recommended greatly broadened my perspective and opened new ways of thinking that enriched my understanding of the field. His mentorship extended far beyond academic research and has had a lasting influence on both my personal and professional development. I am deeply grateful for his trust, unwavering support, and constant encouragement throughout this journey. Without his guidance and belief in my potential, this thesis would not have been possible.

I am deeply grateful to all members of KANAI Lab who have helped me tremendously in many ways. Whether it was helping me set up experiments, troubleshooting equipment issues, or simply lending a hand when I needed it, their kindness and helpfulness made my research journey much smoother. I must apologize for any inconvenience caused by noise during my experiments and sincerely appreciate their patience and understanding throughout the process.

Thank you to the defense committee members (Koike sensei, Nishimoto sensei, Nishimura sensei, Yuizono sensei) for their valuable comments and constructive suggestions, which greatly helped improve this thesis. I am also grateful to my second supervisor (Nishimoto sensei) and advisor for minor research project (Yuizono sensei) for their patient guidance and support throughout my studies.

I appreciate all the participants who generously volunteered their time to take part in my studies. Their willingness to contribute made this research possible.

Finally, I want to express my heartfelt gratitude to my family for their unconditional love, understanding, and unwavering support. Their belief in me sustained me through challenging times and made this achievement possible.

Contents

List of Figures

List of Tables

List of Publications

Papers published in journals:

1. Gong, Ziting, and Hideaki Kanai. "Effects of appearance modifications on oral presentation anxiety in video conferencing." *Scientific Reports* 15.1 (2025): 30090. ***Corresponding to Chapter 3 of this dissertation.***

Papers submitted to journals:

1. Gong, Ziting, and Hideaki Kanai. "Role of virtual audiences in self-regulated exercise: effects on pressure and performance." *Scientific Reports*, Under Review. ***Corresponds to Chapter 4, Section 4.2: Preliminary Experiments of Study 2.***

Papers published in international conference proceedings:

1. Gong, Ziting, and Hideaki Kanai. "Comparison of Two Methods for Altering the Appearance of Interviewers: Analysis of Multiple Biosignals." *International Conference on Human-Computer Interaction.* Cham: Springer Nature Switzerland, 2023. ***Pilot evidence corresponding to Chapter 3.***

2. Gong, Ziting, and Hideaki Kanai. "Impact of Audience Presence on Pressure and Running Performance: The Potential of AR Presence." *Human Factors in Design, Engineering, and Computing* 159.159 (2024). ***Pilot evidence from Experiment 1 in Section 4.2 of Chapter 4.***

3. Ziting GONG and Hideaki KANAI. 2026. Running With a Virtual Character: How AR Companions Support Comfort and Motivation. In The Augmented Humans International Conference 2026 (AHs 2026), March 16–19, 2026, Okinawa, Japan. ACM, New York, NY, USA, 6 pages. https://doi.org/10.1145/3795011.3795027 ***Corresponds to Chapter 4, Section 4.3: Validation Experiment of Study 2.***

Chapter 1

Introduction

1.1 Overview

The Introduction chapter presents the conceptual framework of the dissertation (Fig 1.1), illustrating the motivation, theoretical grounding, and research structure of the present work. Starting from a community-level objective of supporting healthy and psychologically sustainable virtual communities, the framework identifies two prevalent individual-level interaction challenges: anxiety in video-mediated oral communication and motivational difficulties in self-regulated physical activity. Drawing on social facilitation theory, social presence is introduced as a theoretically grounded construct that may shape emotional experience, motivation, and task performance in mediated interaction. Prior work on emotional attachment to virtual others (LOVE) provides empirical inspiration by suggesting that increased perceived realness can enhance presence-related effects. Building on these foundations, the dissertation hypothesizes that adjusting the appearance of interaction partners through emotional channels (e.g., familiarity and liking) can modulate perceived social presence, thereby influencing emotional and performance outcomes. This hypothesized mechanism is empirically examined through two individual-level experimental studies situated in complementary interaction contexts.

Community-Level Objective

Healthy, harmonious, and psychologically supportive virtual communities

↓ **research motivation**

Individual-Level Interaction Challenges

- High anxiety in video-mediated oral presentation
- Motivational difficulties in self-regulated physical activity

↓ **theoretical grounding**

Social Presence as a Theoretical Construct

- Social facilitation: Social presence influences performance
- Mediated interaction: Presence effects also occur in mediated contexts

↓ **prior empirical inspiration**

Prior Evidence

LOVE → Increased perceived realness → Social facilitation effects

↓ **research hypothesis**

Hypothesized Design Mechanism

Interaction partner appearance

└── Emotional channels: familiarity, liking Virtual Communities

1. Modulation of perceived social presence
2. Emotional and performance outcomes

↓ **empirical investigation**

Experimental Studies

- Study 1: Video-mediated interviews
- Study 2: AR-based self-regulated physical activity

Figure 1.1: **Conceptual framework**

1.2 Human Interaction in an Increasingly Virtual World

Human interaction is increasingly mediated by digital technologies. Video conferencing systems, augmented reality (AR) applications, and AI-driven virtual agents have become integral to how people communicate, collaborate, and participate in social life. These mediated environments are no longer temporary substitutes for face-to-face interaction; instead, they are evolving into persistent social spaces that support ongoing relationships, evaluation, and collective activity. As such environments scale and stabilize, they increasingly resemble virtual communities whose long-term health depends not only on technical reliability, but also on whether everyday interactions within them are emotionally supportive, motivating, and psychologically sustainable.

A central challenge for the design of future virtual communities is therefore not merely enabling interaction, but ensuring that mediated interactions promote well-being, engagement, and effective participation over time.

1.3 Individual-Level Interaction Challenges in Mediated Contexts

As virtual interaction becomes routine rather than exceptional, recurring difficulties in everyday mediated encounters have begun to shape users' long-term participation in virtual communities. Experiences such as heightened anxiety during video-mediated interviews or declining motivation in self-regulated digital exercise are not isolated usability issues; they influence whether users feel willing, confident, and sustained in engaging with mediated social environments over time. Addressing such recurring interaction-level difficulties is therefore central to supporting the broader social viability of virtual communities.

Against this backdrop, the present work focuses on two representative interaction contexts:

1. Video-mediated oral communication and interview anxiety: High-stakes oral communication tasks conducted via video conferencing, such as presentations and interviews, often induce elevated anxiety. Users may experience social evaluation pressure, self-consciousness, or nervousness, negatively affecting both emotional well-being and performance.

2. Self-regulated physical activity and motivation challenges: Individuals engaging in self-regulated exercise, including home-based workouts with virtual companions, often struggle to maintain motivation and adherence. Without social facilitation cues, users may

experience lower engagement and reduced enjoyment, limiting the effectiveness of such interventions.

These contexts serve as empirically tractable micro-level scenarios, allowing the investigation of how modulating social presence can influence emotion and performance.

1.4 The Role of Social Presence in Mediated Interaction

Social presence—the sense of being with another socially salient entity—has been widely studied as a determinant of both emotional experience and behavioral performance. According to social facilitation theory, the mere presence of others can enhance performance on simple or well-learned tasks while sometimes impairing performance on complex or novel tasks [166, 14]. In mediated interactions, perceived social presence has been shown to influence task engagement, motivation, and anxiety levels [7]. These findings provide theoretical and empirical justification for targeting social presence as a mechanism: modulating perceived social presence may be able to influence emotional and performance outcomes in specific interaction contexts.

1.5 Design Insight: Appearance Adjustment via Emotional Channels

Prior research on emotional attachment to virtual others (e.g., LOVE) suggests that people can develop meaningful emotional connections with virtual agents, which increases the perceived realness of these virtual others and facilitates social engagement [49]. This body of work inspires the current research: if social presence can affect performance and attachment increases perceived realness, then manipulating the appearance of interaction partners via emotional channels (e.g., familiarity, liking) may provide a practical mechanism for modulating social presence and its downstream effects on emotion and behavior.

Building on this insight, this dissertation investigates how emotional channels—specifically familiarity and liking—can be leveraged to adjust the visual appearance of interaction partners. The hypothesis is that by enhancing familiarity or positive affect toward a partner's appearance, users may experience greater perceived social presence. In turn, changes in perceived social presence are expected to influence users' emotional experiences (e.g., reducing anxiety, increasing comfort or enjoyment) and potentially support improved behavioral engagement (e.g., motivation or communicative performance). The proposed studies empirically examine these

mechanisms in controlled yet ecologically valid settings, providing foundational evidence for how appearance-based design can support adaptive, socially facilitative virtual interactions.

1.6 Motivation

Because existing research has largely focused on self-representation, a critical gap remains in our understanding of how the appearance of others—such as interaction partners, audiences, or virtual agents—shapes emotional experience, social perception, and performance outcomes. Yet in everyday interactions, individuals are often more influenced by how others appear than by how they represent themselves.

As virtual environments increasingly incorporate diverse social actors, including AI-driven agents and hybrid human–agent interactions, this gap limits both our theoretical understanding of mediated social presence and the development of effective design practices. Addressing this academic gap therefore constitutes a central motivation for the present research.

Moreover, given the growing prevalence of virtual presence and virtual environments in future societies, it is crucial to understand how the roles and visual representations of avatars or agents—whether representing oneself or others—shape users' perceptions and influence interaction processes. Promoting more effective communication and interaction on virtual platforms, thus represents another key motivation for pursuing this line of research.

1.7 Augmented Reality as a Methodological Lens

Augmented reality provides a uniquely suitable platform for this investigation. Unlike fully virtual environments, AR preserves real-world social cues—such as facial expressions, gaze, and physical context—while enabling targeted visual augmentation. This hybrid quality allows appearance to be manipulated without stripping away the richness of embodied interaction.

By leveraging AR, this dissertation examines how subtle adjustments to others' appearance can recalibrate social presence in everyday interaction contexts, offering both experimental control and ecological validity.

1.8 Overview of the Two Studies

To explore appearance-mediated social presence across different forms of interaction, this dissertation presents two empirical studies situated in complementary contexts.

Study 1 examines high-pressure oral communication in video-mediated interviews. By augmenting the interviewer's appearance using AR face filters, the study investigates whether reducing or reshaping perceived social presence can alleviate anxiety and improve communicative performance.

Study 2 extends this inquiry to embodied physical activity, examining AR-based exercise companions. Here, the appearance of a virtual agent is designed to enhance social presence, with the goal of increasing motivation, engagement, and physical performance.

Although these studies differ in task demands and interaction goals, they are unified by a common mechanism: manipulating the appearance of others as a means of regulating perceived social presence and, in turn, shaping user experience and performance. By connecting these individual-level mechanisms to broader community-level objectives, the research aims to provide insights for designing virtual environments that are both socially and emotionally supportive.

1.9 Contributions and Scope

Together, these studies advance understanding of how interaction partner appearance operates as a key yet understudied mechanism in mediated interaction. By extending social presence research beyond self-avatar effects, this dissertation addresses a significant theoretical gap and offers practical design insights for communication platforms, AR applications, and future AI-driven companions.

Ultimately, this work lays foundational groundwork for designing virtual and augmented environments that are not only functional but also emotionally intelligent, adaptive, and supportive of human well-being.

1.10 Thesis Structure

This dissertation is organized into six chapters (see Fig 1.2), each addressing a specific component of the research program and collectively building toward a comprehensive understanding

of how the appearance of others influences user experience in augmented and virtual communication environments.

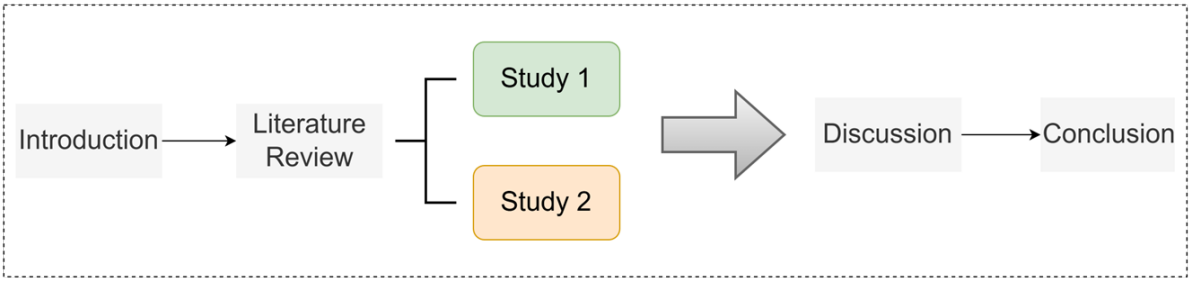


Figure 1.2: **Thesis structure**

Chapter 1: Introduction

This chapter introduces the dissertation's conceptual framework, outlining the research motivation, theoretical foundations, and overall structure of the study. It describes the key interaction challenges addressed in the dissertation and positions social presence as the central theoretical construct guiding the subsequent empirical investigations.

Chapter 2: Literature Review

This chapter reviews prior research that provides the theoretical foundation and design rationale for the two empirical studies in this dissertation. The literature is organized around the role of social presence in shaping emotional, motivational, and performance-related outcomes in mediated interaction. The chapter is structured in two parts: the first examines social presence and anxiety in evaluative communication contexts, while the second reviews social facilitation and mediated companionship in physical activity. On this basis, the chapter also articulates a conceptual model proposing that strategic engagement of emotional appearance channels can mitigate anxiety in evaluative communication and support motivation in self-regulated physical activity.

Chapter 3: Study 1 — Relieving User Anxiety in Video Interviews

This chapter details the first empirical study, which investigates the framework's predictions in a high-stakes evaluative communication context by examining how augmenting the interviewer's

appearance through AR face filters affects user anxiety, comfort, and communication performance in video interviews. It presents the study design, methodology, experimental procedure, and quantitative analyses. The chapter concludes with a discussion of key findings and their implications for anxiety reduction in high-stakes digital interactions.

Chapter 4: Study 2 — Enhancing User Engagement in Physical Exercise

Study 2 extends these predictions to embodied physical activity, examining how the appearance of an AR-based virtual agent influences user engagement, motivation, and behavioral adherence during physical exercise. The chapter describes the agent design, experimental setup, data collection procedures, and analytic methods. Results are presented alongside an interpretation of how augmented agent appearance shapes user perception and activity participation.

Chapter 5: General Discussion

This chapter integrates findings from two empirical studies to examine how the appearance of others regulates emotional experience and behavioral performance in mediated interaction. By comparing familiarity, realism, and stylization across contexts, the chapter synthesizes how appearance cues calibrate psychological distance, shape appraisals of social evaluation or support, and influence affective and performance outcomes. Then develops a cross-study synthesis highlighting common mechanisms, contextual differences, and implications for theory and design.

Chapter 6: Conclusion

This chapter concludes the dissertation by summarizing its motivation, research approach, and contributions, and by discussing its significance, limitations, and future directions.

1.11 Summary

As virtual and augmented environments become increasingly central to everyday social interaction, supporting healthy, effective, and emotionally sustainable communication within these spaces has become a critical challenge. Rather than approaching this problem at the community level directly, this work takes an individual-level experimental approach as a first step toward understanding how interaction experiences can be shaped and regulated. Across two studies

situated in distinct yet representative contexts—high-pressure oral communication and self-regulated physical exercise—we examine how adjusting the appearance of interaction partners can modulate perceived social presence. By demonstrating that appearance-based social presence regulation influences both users' emotional experiences and performance outcomes, this dissertation extends prior social presence research beyond self-avatar appearance effects and provides empirical grounding for new design strategies. Together, these findings point toward more adaptive and psychologically supportive interaction environments, offering early insights for the design of future virtual communities.

Chapter 2

Literature Review

2.1 Overview

This chapter reviews prior research that informs the design and theoretical grounding of the two empirical studies in this dissertation. Rather than treating appearance, social presence, and performance as independent factors, the reviewed literature is organized around how the visual representation of others shapes users' emotional and behavioral responses in mediated interaction. The chapter is structured in two parts. The first part focuses on appearance-based modulation of anxiety in evaluative communication contexts, examining how familiarity, stylization, and visual abstraction influence perceived safety and evaluation pressure. The second part reviews research on social facilitation and virtual or augmented companions in physical activity, highlighting how different forms of mediated presence affect motivation, engagement, and performance. Together, this review establishes appearance as a central yet underexplored mechanism for regulating social presence across cognitive and embodied interaction domains.

2.2 Appearance and Social Presence in Mediated Interaction

Social presence—the sense of being with another social entity—has long been recognized as a central factor shaping interaction quality in mediated environments. Prior research demonstrates that social presence is not determined solely by communication bandwidth or system fidelity, but is strongly influenced by how interaction partners are visually represented. Appearance cues shape expectations of agency, evaluative intent, and emotional availability, thereby affecting how users interpret and respond to mediated interaction.

In virtual and augmented environments, appearance becomes a designable variable. Avatars,

face filters, and embodied agents allow designers to systematically manipulate realism, stylization, familiarity, and expressiveness, offering new opportunities to regulate perceived social presence. Understanding how these appearance manipulations influence users' emotional experience and performance is therefore essential for both theory and design.

2.3 Asymmetry in Prior Work: From Self-Avatars to Others' Appearance

A substantial body of research has examined how users' own avatars influence their attitudes and behavior, most notably through the Proteus effect. Studies consistently show that avatar attributes such as attractiveness, height, and role-related cues can alter self-perception, confidence, and task performance. While this work has established the importance of appearance in mediated interaction, it has predominantly focused on self-representation.

In contrast, far less research has systematically examined how the appearance of others—interaction partners, evaluators, audiences, or virtual agents—shapes emotional responses, social judgments, and behavioral outcomes. Yet in everyday interaction, individuals are often more sensitive to how others appear than to how they represent themselves. This imbalance represents a critical gap in the literature on social presence and mediated interaction.

2.4 Appearance-Based Regulation in Evaluative Communication Contexts

One line of relevant research examines how the appearance of audiences or evaluators influences anxiety and performance in public speaking and presentation tasks. Prior work demonstrates that audience realism, facial expression, and visual abstraction can significantly affect speakers' emotional responses.

Study A: Girondini et al. [51] – Human vs. Cartoon avatar (Fig 2.1)

This study investigated how variations in graphical realism (human-like vs. cartoon-like avatars) and acoustic realism (human vs. synthetic voices) within a virtual audience influence users' perceptions during public-speaking tasks. Participants engaged in multiple virtual public-speaking scenarios, each featuring different combinations of avatar and voice realism.

The findings indicated that higher graphical realism—specifically the use of human-like avatars—substantially increased participants' sense of presence and perceived realism. Conversely, presenting audiences with less realistic, stylized avatars (e.g., cartoon figures) helped mitigate anxiety during the speaking task.

Relevance. These results suggest that stylized representations create emotional distance and reduce evaluative pressure, aligning with Study 1's argument that anime-style visuals function as a psychological buffer.

Figure 2.1: **Cartoon avatar from related work [51]**

Study B: Monteiro et al. [102], Barrett et al. [10]

However, the role of familiarity is less consistent. This study examined how audience familiarity influences anxiety during foreign-language speech tasks across both virtual and real-world settings. Participants delivered speeches to two types of audiences: (1) familiar individuals, such as teachers or researchers, and (2) unfamiliar individuals. Measures of anxiety and performance were collected across conditions.

The results showed that the presence of familiar faces did not reduce participants' speaking anxiety, nor did it enhance their performance.

Relevance. These findings suggest that the expected benefits of audience familiarity may be diminished when the familiar individuals occupy evaluative or authority roles.

Study C: Leong et al. [88] – AR Filters in Video Presentations, see Fig 2.2

Augmented reality studies further suggest that selectively modifying others' appearance may offer a practical means of anxiety regulation. In this study, participants delivered video-conference presentations while applying augmented-reality (AR) filters to the audience's faces. These filters included cosmetic enhancements, accessories, and fully non-human (e.g., animal) facial transformations.

Findings from semi-structured interviews indicated that these face filters reduced perceived intimidation and increased presenters' comfort during the task, suggesting that such visual modifications can serve as an effective anxiety-management tool.

Relevance. This study illustrates the potential of privately altering others' facial appearances to mitigate evaluation anxiety, providing conceptual support for the AR-filter design used in Study 1.

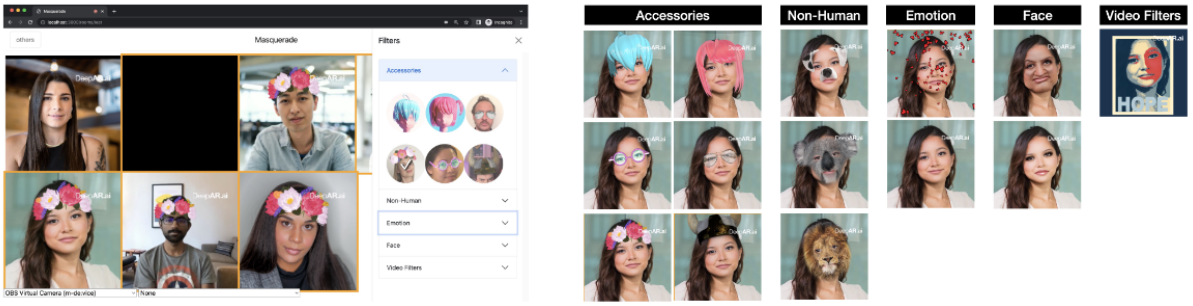


Figure 5: Left: The user interface for a user in the *on-others* filter mode. The filter-menu on the right-hand side of the screen is opened with the flower-crown filter option selected. The video-grid is populated with the user's video positioned at the top-left of the screen, followed by the study coordinator's video (hidden), and virtual audience members with the selected filter applied. Right: The filter options available in the app, as applied on a person's face.

Figure 2.2: **AR face filters from related work [88]**

Study D: Yoneyama et al. [163] – Anime vs. Real Faces in AR Conversation, see Fig 2.3

This study explored the effects of augmented-reality face overlays during casual, non-evaluative conversations. Participants interacted with partners whose faces were modified using three overlay types: human-like faces, anime-style faces, and smiling-face augmentations.

Participants reported warm and positive feelings toward anime-style faces. However, evidence for anxiety reduction was inconclusive, potentially due to practice effects and uncontrolled facial-expression variables.

Relevance. The findings suggest that anime-style avatars may foster comfort, but controlled

evaluative contexts are needed to assess their anxiolytic potential—precisely the gap addressed by Study 1.

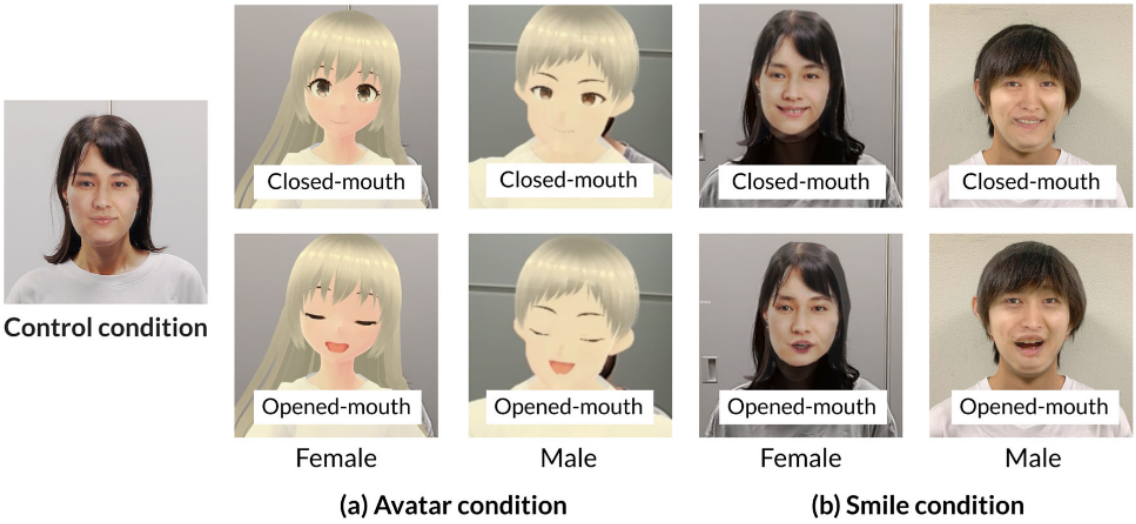


Figure 2.3: **Visual effects from related work [163]**

Synthesis and Gaps

Collectively, these studies reveal that others' appearance influences evaluative anxiety, but through inconsistent and insufficiently understood mechanisms:

Visual abstraction shows promise but lacks clarity:

- Stylized avatars (Study A) and AR filters (Study C) reduced perceived intimidation and anxiety
- However, the specific features responsible—reduced realism, altered expressions, or novelty effects—remain confounded
- Study D's findings on anime faces were inconclusive due to uncontrolled variables and non-evaluative contexts

Familiarity effects are contradictory:

- Familiar faces did not reduce anxiety when occupying evaluative roles (Study B)

- Yet other studies suggest familiarity can provide comfort
- This contradiction suggests that familiarity's effect depends on perceived relational role—whether the familiar person is perceived as supportive or evaluative.

Critical limitations of prior work:

- Confounded variables: Most studies manipulate multiple appearance features simultaneously (realism + expression + familiarity), making it impossible to isolate specific mechanisms
- Lack of controlled evaluative contexts: Studies either use non-evaluative conversations or fail to systematically vary evaluation intensity
- Insufficient theoretical grounding: No studies explicitly test whether effects operate through social presence modulation, cognitive distancing, or affective safety pathways
- Neglect of individual differences: Trait anxiety and personal preferences are rarely measured, despite likely moderating appearance effects

Research gap: A systematic investigation is needed to disentangle familiarity from stylization, test their effects under controlled evaluative pressure, and identify the psychological mechanisms (social presence, evaluation threat appraisal) through which appearance regulates anxiety and performance.

2.5 Virtual Social Presence in Physical Activity

A second body of research relevant to this dissertation concerns social facilitation and the presence of others during physical activity. Classic social facilitation theory predicts that the presence of others enhances performance on simple or well-learned tasks, while potentially impairing performance on complex tasks. However, evidence for such effects in virtual environments remains mixed.

Study A: Sterna et al. [139] – Social Facilitation in Virtual Environment

This review examined prior research on how virtual social presence influences task performance across domains such as public speaking and search tasks. The studies typically explored classic social facilitation patterns—enhanced performance on simple tasks and diminished perfor-

mance on complex tasks. However, the overall evidence was inconsistent, with only one study fully demonstrating the expected facilitation–inhibition effects.

Relevance. These findings indicate that virtual presence alone often fails to reproduce the robust social facilitation effects observed with real audiences, highlighting the need for more effective designs for mediated spectators.

Study B: Haller et al. [59] – VR Cycling with Applauding Crowd

This study examined how an active virtual audience influences physical performance during high-intensity interval training (HIIT) on a cycling ergometer. Participants completed the exercise within a virtual environment under two conditions: (1) a virtual crowd delivering synchronized applause and rhythmic cues, and (2) no audience present.

Results showed that the active virtual crowd elicited higher cycling speeds and increased heart rates, indicating elevated physical effort. However, it remained unclear whether these performance gains were attributable specifically to the presence of the crowd, the rhythmic cues, or the combination of both.

Relevance. This work suggests that virtual audiences can enhance performance, but it also highlights the need to disentangle social effects from rhythmic or auditory entrainment mechanisms.

Study C: Xu et al. [160] – Encouraging vs. Neutral Virtual Audiences in Exergames, see Fig 2.4

Participants engaged in gesture-based virtual reality (VR) exergames while being exposed to virtual audiences exhibiting distinct behaviors. The type of audience feedback was systematically manipulated, consisting of either encouraging feedback, such as cheering and clapping, or neutral feedback, in which no response was provided. The results indicated that exposure to encouraging audiences significantly enhanced participants' game performance, overall experience, and levels of physical exertion.

Relevance: The findings suggest that behavioral feedback from virtual audiences exerts a significant influence on participants; however, it remains unclear whether this effect is primarily driven by the audience's visual behaviors or by auditory cues, such as cheering.

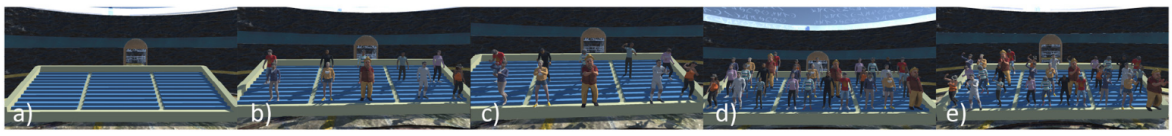


FIGURE 2
Conditions tested in our experiment: **(a)** *ES*: empty stadium; **(b)** *SSNF*: small NPC spectator group size without feedback to the player; **(c)** *SSWF*: small NPC spectator group size with feedback to the player; **(d)** *LSNF*: large NPC spectator group size without feedback to the player; **(e)** *LSWF*: large NPC spectator group size with feedback to the player. This figure only displays one spectator stand used in our game scene. There are three more spectator stands placed around the playfield or stadium.

Figure 2.4: **Audience feedback and size conditions from related work [160]**

Study D: Hamada et al. [60] – Smartglasses Running Partner, see Fig 2.5

More recent work has explored partner-like virtual and augmented companions. In this study, participants engaged in running while wearing smartglasses that displayed a virtual running partner. The experimental conditions consisted of jogging alone, jogging with a human partner, and jogging with a virtual partner, with variations in the representation of the virtual partner. The results demonstrated that running with virtual partners significantly influenced participants' motivation and behavior compared to running alone. Moreover, the specific representation of the virtual partner shaped how participants perceived and interacted with the agent.

Relevance: These findings indicate that partner-like virtual agents can effectively support exercise engagement and may serve as a valuable tool for promoting sustained physical activity.

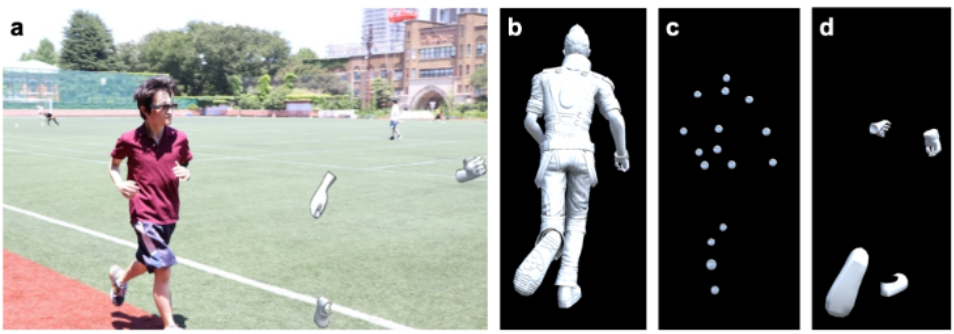


Figure 1: (a) Joggers can see a virtual runner which has each of three body visibilities: (b) full-body, (c) point-light, and (d) limb-only.

Figure 2.5: **AR runner from related work [60]**

Study E: Nam et al. [107] – AR Fitness Dog, see Fig 2.6

In this study, participants exercised while accompanied by an augmented reality (AR) virtual dog that mirrored their actions. The experimental conditions included exercising with a mimicking virtual dog, exercising with a randomly behaving virtual dog, and exercising without a virtual dog. These conditions were tested in both solo and paired group exercises. The results indicated that the mimicking virtual dog significantly enhanced the exercise experience for solo participants by fostering a stronger sense of companionship.

Relevance: These findings suggest that AR companions can provide emotional support and promote a feeling of "not exercising alone," thereby potentially increasing exercise engagement and adherence.

(a) Solo (b) Group (Pair)

Fig. 1: The proposed AR Fitness Dog. User-mimicking interactive virtual dog performing movements in (a) Solo and (b) Group (Pair) exercise mode.

Figure 2.6: **AR fitness dog from related work [107]**

Synthesis and Gaps

While these studies demonstrate that virtual and AR companions can influence exercise motivation and performance, several critical questions remain unanswered:

1. Appearance is underexplored:
    - Studies focus primarily on behavior (encouraging vs. neutral, Study C) or synchrony (mimicking vs. random, Study E)

- The visual appearance of companions has not been systematically manipulated

2. Mechanisms are unclear:
    - Prior work conflates social presence with behavioral feedback (cheering, clapping)
    - It remains unknown whether visual presence alone, without behavioral cues, can modulate motivation
3. Pressure vs. support tradeoff is unexamined:
    - Real audiences boost performance but also increase pressure (classic social facilitation)
    - Virtual companions may reduce pressure, but can they maintain motivational benefits?
    - No study has directly compared real spectators, realistic virtual agents, and stylized companions to identify optimal presence calibration

Research gap: Prior work confounds companion appearance with behavioral feedback (cheering, synchrony), preventing a clear assessment of whether visual design alone can modulate motivation and pressure. This gap is critical because if appearance can calibrate perceived social presence, it offers a scalable design strategy for AR fitness systems.

2.6 Research Gaps and Dissertation Approach

The literature reviewed above reveals critical gaps in understanding how appearance regulates social presence across interaction contexts. While existing work demonstrates that social presence is not uniformly beneficial—it can enhance performance through motivation (physical activity) or impair it through evaluation anxiety (public speaking)—research remains fragmented. Evaluative communication studies examine anxiety reduction but rarely manipulate appearance systematically. Physical activity research focuses on behavioral feedback but neglects visual design. No work systematically compares how the same appearance manipulations operate across cognitive-evaluative versus embodied-motivational tasks, nor accounts for individual differences such as trait anxiety.

This dissertation addresses these gaps through two complementary studies:

- Study 1 systematically manipulates interviewer appearance (familiar, unfamiliar, stylized) in high-pressure video interviews to isolate appearance effects on anxiety and communi-

cation performance.

- Study 2 extends this investigation to exercise, comparing real audiences, video spectators, and AR companions with varying appearances to test whether appearance can provide motivational support while reducing evaluation pressure.

By examining others' appearance as a regulator of social presence across both contexts, this work establishes appearance as a generalizable design mechanism, providing both theoretical clarification and practical guidance for adaptive virtual interaction systems.

2.7 General Framework

Conceptual Model: Appearance-Mediated Social Presence Regulation

Building on prior work in social presence, mediated interaction, and digital embodiment, this dissertation adopts a conceptual model in which the appearance of interaction partners functions as a primary design lever for regulating perceived social presence, thereby shaping users' emotional experience and behavioral performance. Within this model, appearance is not treated as a static visual attribute, but as an interpretable social signal that conveys cues about evaluative intent, emotional availability, and relational distance.

Specifically, variations in appearance—such as familiarity, stylization, and aesthetic alignment—are proposed to modulate how intensely others are perceived as socially present. These modulations, in turn, influence how users appraise an interaction situation, including perceived safety, pressure, and support. Rather than assuming social presence to be uniformly beneficial, this model emphasizes its context-dependent and bidirectional role: elevated social presence may increase motivation and engagement in supportive contexts, while reduced or softened social presence may alleviate anxiety and evaluative stress in high-pressure situations.

Downstream, these appearance-mediated changes in perceived social presence are expected to affect both subjective outcomes (e.g., anxiety, comfort, enjoyment, perceived support) and objective or behavioral outcomes (e.g., attentional engagement, verbal fluency, persistence, and physical performance). Importantly, the model also accommodates individual differences, such as trait anxiety, which may moderate how appearance cues are interpreted and how social presence is experienced.

Together, this conceptual model provides a unifying framework for examining how targeted appearance design can systematically regulate social presence and interaction outcomes across

distinct domains. It motivates the two empirical studies in this dissertation, which investigate this mechanism in evaluative communication and embodied physical activity, respectively, offering a foundation for more adaptive and psychologically supportive mediated interaction design, shown in Fig 2.7.

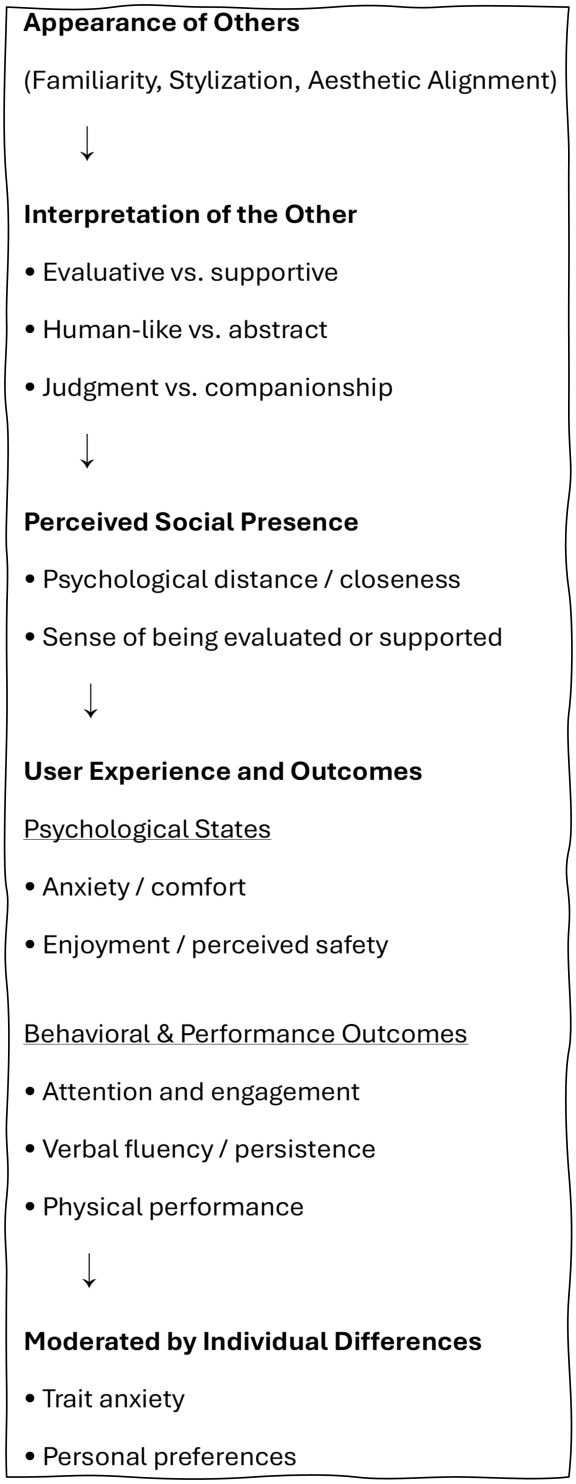


Figure 2.7: **Appearance-Mediated Social Presence Framework**

2.7.1 Research Objectives

Validate the proposed appearance-based social presence regulation mechanism for addressing real-world challenges in mediated interaction.

Objective 1 — Appearance → Perception

To examine how variations in others' appearance (e.g., familiarity, stylization, and likability) influence users' perception of social presence and evaluation in mediated interactions.

Objective 2 — Emotion Regulation

To investigate how these appearance cues regulate users' emotional states, particularly anxiety reduction in evaluative cognitive tasks and affective up-regulation in physical activities.

Objective 3 — Performance Consequences

To assess how appearance-driven emotional and perceptual changes translate into behavioral and performance outcomes in both communication and embodied action tasks.

Objective 4 — Integration & Design

To integrate findings across contexts in order to develop a generalizable framework of appearance-mediated social presence and derive design implications for AR-based communication systems and virtual companions.

2.7.2 Overview of Empirical Studies

- Study 1 focuses on reducing evaluative anxiety by softening social presence in video-mediated interviews.
- Study 2 examines how enhancing supportive social presence through AR companions facilitates motivation and performance in physical exercise.

2.8 Summary

This chapter reviewed prior research on appearance-based modulation of social interaction across evaluative communication and embodied physical activity, highlighting how visual representations of others shape emotional experience, engagement, and performance in mediated contexts. The literature on public speaking and video-mediated interviews suggests that modifying evaluators' appearance—through familiarity or stylization—can influence anxiety by altering perceived safety and evaluative pressure, yet existing studies often confound familiarity, expressiveness, and task context. Research on social facilitation and virtual companions demonstrates that the presence of others can enhance physical performance and motivation, but frequently at the cost of increased pressure, and with limited attention to how appearance and emotional closeness shape these effects.

Building on these findings, this chapter introduced a General Framework that is explicitly predictive. The framework proposes that the appearance of others functions as a controllable social cue that systematically modulates perceived social presence, which in turn predicts emotional and behavioral outcomes. Specifically, it generates testable predictions that appearance cues associated with intimacy increase psychological safety and reduce evaluative anxiety, whereas cues associated with distance or stylization reduce perceived scrutiny by weakening evaluation-related signals. In action-oriented contexts, the same mechanism predicts that socially supportive appearances enhance motivation and performance without eliciting excessive pressure.

Importantly, the framework further predicts that these effects are context- and person-dependent, with task demands (evaluative vs. action-oriented) and individual differences such as trait anxiety moderating how appearance-based social presence is interpreted. Rather than assuming social presence to be uniformly beneficial or harmful, the framework treats it as a regulative variable whose optimal level depends on both situational goals and user characteristics.

The framework carries broader implications for the design and sustainability of future virtual and augmented communities. As social interaction increasingly takes place in persistent digital spaces, individuals will be exposed to a growing volume of mediated social presence. By conceptualizing others' appearance as a mechanism for regulating social presence rather than merely increasing it, this work highlights a pathway toward more harmonious and psychologically supportive virtual communities. The framework supports a vision of future virtual communities in which interaction is not only engaging and effective, but also emotionally sustainable—promoting inclusion, well-being, and long-term social harmony.

Together, this literature review and predictive framework motivate the two empirical studies

presented in the following chapters. Study 1 tests the framework's predictions in a high-stakes evaluative communication context by examining how interviewer appearance modulates anxiety and performance in video interviews. Study 2 extends these predictions to embodied physical activity, examining how the appearance of augmented companions regulates motivation and engagement. Collectively, these studies provide a systematic empirical test of appearance-mediated social presence as a mechanism for regulating emotion and performance in mediated interaction.

Chapter 3

Study 1: Relieving User Anxiety in Video Interviews

3.1 Overview

Chapter 3 investigates how different forms of mediated social presence influence emotion regulation during an evaluative cognitive task. It opens with a theoretical rationale connecting social facilitation and cognitive reappraisal, framing interviewers using face-filter as affective stimuli that can either heighten or relieve presentation anxiety. The method section details a three-condition design—Friend (familiar), Anime (stylized figure), and Stranger (neutral)—with measures of state anxiety (STAI-S), perceived performance, word count, gaze behavior, and physiological indices (nasal temperature, prefrontal blood flow). The results section reports significant reductions in anxiety and improved perceived performance in the Friend condition, but limited differences for Anime. Behavioral and eye-tracking data reveal that emotional familiarity promotes engagement, whereas detachment weakens social focus. These findings highlight how mediated appearance modulates evaluative stress: familiar representations promote mediated affective co-regulation, thereby alleviating anxiety, whereas anime-style stylization facilitates cognitive distancing by reducing perceived realism—but the reduced feedback cues can also introduce a sense of unease.

3.2 Introduction

Anxiety is generally considered a feeling of worry, unease, or apprehension about an uncertain outcome [36]. From a psychological point of view [77], it is considered an emotional state

manifested through sensations of tension, worrying thoughts, and physical changes [147]. According to cognitive-behavioral theory [11], anxiety often stems from negative thoughts and irrational beliefs. The anxiety sensitivity theory [127] suggests that anticipatory anxiety occurs before an event, as individuals fear the potential occurrence of negative outcomes. Meanwhile, the evolutionary paradigm [109] views anxiety as an adaptive response that helps enhance alertness when facing challenging environments. Anxiety permeates various aspects of our daily lives, including communication anxiety, test anxiety [147], programming anxiety [165], math anxiety [70, 151], computer anxiety [28], among others.

Oral presentations have long been regarded as one of the most anxiety-inducing events [38, 96, 138, 102]. Given that prolonged anxiety can lead to various adverse effects, such as limitations in personal social interactions, erosion of self-esteem and confidence, and negative impacts on work performance [165, 70, 54, 67], various measures have been proposed to alleviate such anxiety. For instance, previous studies have leveraged virtual reality (VR) to help them overcome presentation anxiety by being repeatedly exposed to public speaking scenarios [55, 133, 63, 123, 145, 134]. Additionally, one study explored the usage of speech-based feedback to address anxiety. It first recorded the speaker's voice under stable emotional conditions to capture its characteristic features. During the presentation, the speaker receives the modified voice which reflects the characteristics of their stable-state voice, effectively creating the illusion that they were composed and confident aiming to reduce the speaker's anxiety [72].

Oral presentations are typically considered to take place in face-to-face settings. However, with the rapid development of computer-mediated communication technology in recent years, giving presentations through video conferences has become very common. However, there is limited research on the anxiety experienced by speakers in video conferences which is based on computer-mediated communication platforms.

In our investigation of video conferences, we found that although it has revolutionized communication by providing a platform for efficient and instantaneous connections across various settings, this technology also presents challenges, including video conferencing fatigue, which is perceived as more mentally taxing than face-to-face interactions [75]. Previous studies have suggested that heightened anxiety levels during video conferences may stem from platform characteristics, particularly the sense of close proximity [19, 69]. The intimacy created by the proximity between users and screens can make people feel uncomfortable when interacting with unfamiliar individuals, thereby increasing anxiety. According to interpersonal distance norms, the personal distance for interactions among close friends or family falls within a close range of 0.46 to 0.76 meters. Maintaining a social distance of 1.2 to 2.1 meters is generally considered more comfortable when interacting with strangers [58]. However, during video conferences, the physical distance between participants and the unfamiliar faces on the screen typically shrinks

to about 0.5 meters, leading to discomfort and anxiety due to the significant deviation from expected social norms.

Additionally, the limited environmental cues in video conferencing settings amplify the prominence of participants' gazes [4, 37, 62, 116]. In traditional face-to-face conferencing scenarios, various environmental factors and body language cues help distribute attention and reduce the intensity of direct eye contact. However, in video conferencing scenarios, the absence of these contextual elements makes participants' gazes more pronounced. This increased eye contact not only creates discomfort from being constantly observed but also leads to the anxiety of maintaining it, thus causing anxiety [66]. In summary, performing a presentation in video conferences triggers inherent anxiety not only from the event itself but also due to the platform of video conferences.

Video conferences offer a distinct advantage of modifying visual contents on the screen that may influence our perception. Virtual background selection plays a vital role in shaping others' perceptions. Previous research revealed that concealing the original setting using virtual backgrounds effectively mitigated extreme ratings of speakers' personality traits [69]. A commonly recommended strategy for reducing anxiety during oral presentations is to mentally reframe the audience in a less intimidating and more familiar context. For instance, speakers are often advised to imagine their audience in humorous or mundane situations—such as picturing them wearing mismatched socks or having holes in their socks—to reduce the perceived threat of negative evaluation [12, 88, 87]. Such mental reframing techniques aim to foster a more relaxed internal environment by lowering the audience's perceived authority or formality. Extending this idea into the digital domain, computer-mediated communication platforms like video conferencing now offer technological means—such as Augmented Reality (AR) filters—to visually transform the audience's appearance. These filters provide an opportunity to create a presentation context that is less intimidating and more psychologically comfortable. Specifically, comfort may arise through two complementary pathways: familiarity and visual abstraction. Familiar faces, such as those of close friends, can evoke a sense of trust and safety, reducing anticipatory anxiety by minimizing fear of negative judgment [91]. On the other hand, cartoon-like or anthropomorphic avatars introduce a layer of abstraction that distances the speaker from the realism of social evaluation. These non-threatening, stylized visuals reduce perceived interpersonal intensity, creating an emotionally buffered environment that supports more confident speaking behavior [88, 163]. A previous study surveyed 100 participants to investigate their perceptions of this private AR face filter technology [88]. The results revealed a high level of acceptance, even when the filter was privately applied to themselves. One primary reason for using the filter was to reduce negative emotions such as anxiety and fear, thereby creating a more relaxed atmosphere. Additionally, applying filters to others' faces was often motivated by

the desire to diminish the sense of intimidation posed by others.

Building upon these insights and the psychological mechanisms underlying audience perception, the present study investigates whether visually altering the appearance of the audience can help reduce oral presentation anxiety in one-on-one video conferencing settings. Specifically, we examine two types of visual modifications: replacing the audience's face with that of a close friend (to evoke familiarity), and with an attractive anime character (to introduce visual abstraction). To assess their effects, we conducted a between-subjects experiment with 45 participants, analyzing physiological signals, user behaviors, and self-reported perceptions. This approach aims to provide empirical evidence on how different visual strategies may alleviate presentation-related anxiety and inform the design of supportive communication technologies in virtual settings.

3.3 Face-filter proposal

In addition to audience behavior [13, 27, 74, 73], attitude [105, 27], presence [46], and size [104, 133, 149], the appearance of the audience [10, 102] has also been found to influence the oral presentation anxiety levels in virtual environments. In the context of video conferencing, privately applying AR filters to modify the appearance of the other party on one's own screen holds great potential for alleviating users' social anxiety, especially in anxiety-inducing situations such as interviews or public speaking [88, 163].

3.3.1 Audience familiarity

Audience familiarity is considered an important factor in eliciting anxiety [10, 5, 91, 21, 99]. Familiar friends as audience are often perceived as more tolerant and understanding, with a lower likelihood of giving negative evaluations [91]. Consequently, speakers are likely to anticipate a pleasant speech environment, allowing them to deliver speeches in a more relaxed manner. When the actual situation aligns with this expectation, familiar friends as audience evoke the lowest levels of tension. However, if the expectation is not met and the audience behaves unpleasantly, familiar audiences may provoke anxiety levels similar to or even greater than those caused by unfamiliar audiences [18, 48]. This observation was further supported by MacIntyre and Thivierge's study [91], in which participants were asked to imagine giving speeches in front of audiences that varied in familiarity and emotional expressions. Their findings indicated that anxiety was lower when speaking in front of friends, as a high familiarity. However, the effect of familiarity was weaker compared to that of pleasant facial expressions. Future studies should

focus on designing experiments that measure actual emotional responses to minimize potential biases arising from imagined scenarios. In particular, it is crucial to control for the influence of emotional feedback when assessing the effect of familiarity on anxiety. Additionally, the study required participants to imagine an audience size of 20, raising some questions, such as 20 distinct or 20 identical faces. This ambiguity could impact the interpretation of familiarity effects. Despite some limitations, this study provides preliminary evidence that highly familiar audiences can be effective in alleviating anxiety during oral presentations.

Previous study [102, 10] examined how familiar and unfamiliar audiences influence speakers' anxiety levels during foreign language speeches in both VR and real-world settings. In the VR environment, familiar audiences were found to induce higher anxiety compared to unfamiliar ones. The authors attributed this to the lower realism of computer-generated unfamiliar audiences, which might have reduced the sense of presence and, consequently, the anxiety they provoked. In contrast, familiar audiences were represented using scanned 3D images. However, the authors later challenged their explanation by referencing prior literature [115, 46] suggesting that photorealism is not a precursor to the sense of co-presence in virtual environments. In the real-life setting, no significant effect of audience familiarity on anxiety was observed. Additionally, the results indicated that individuals with a moderate fear of public speaking experienced reduced anxiety when presenting in front of virtual audiences composed of familiar faces. The small sample size of the experiment (10 participants) and the choice of familiar faces, such as teachers and researchers, who are often perceived as authoritative figures, who could potentially heighten anxiety [71], may have influenced the effect of audience familiarity.

To bridge the gap in empirical research on the effects of highly familiar audiences, we propose using the appearance of good friends as a facial filter to alleviate oral presentation anxiety during online conferences. Moreover, the choice is likely to minimize the presence of authoritative audiences, thereby reducing their potential impact on anxiety levels.

Hypothesis 1: Modifying the interviewer's appearance to resemble a good friend will reduce the speaker's anxiety in one-on-one video conferences.

3.3.2 Non-human entity

Leong et al. [88] found that compared to makeup or accessory-based modifications, most participants preferred applying non-human AR filters, such as lions, to the audience's faces during presentations. These full-face filters completely masked the real human appearance, reducing the sense of realism in the online public speaking environment and thereby alleviating the anxiety of being judged. Additionally, these non-human characters were perceived as cute and

friendly, which helped speakers feel more confident and contributed to reducing presentation anxiety during video conferences [88]. While animal appearances can help mitigate the anxiety of negative evaluations, they also introduce additional challenges. One major issue is the lack of facial expression cues, which may lead to uncertainty and discomfort for the speaker. A promising solution to this problem is the use of anthropomorphic characters. Their human-like appearance provides a sense of familiarity and facial expression cues [163], while their non-human nature can still reduce realism and alleviate the anxiety associated with being judged. This idea is supported by the study conducted by Girondini et al. [51], which found that delivering a speech in front of real humans induced higher anxiety levels than speaking in front of cartoon characters. Moreover, using anime characters to cover the face of one's conversation partner is considered similar to positive facial feedback, making interactions feel easier for individuals with social anxiety [162].

Yoneyama et al. [163] explored the impact of modifying a conversation partner's appearance into either an anime character or a real person using AR head-mounted displays (HMD) during real-life interactions. Their experiment found that anime characters did not significantly reduce anxiety during casual interview sessions. This result might be influenced by a practice effect, as participants repeated the experiment across control, avatar, and smile conditions. Furthermore, in the smile condition, the visual appearance remained the same as in the control condition (a real human face), which may have introduced familiarity effects. Additionally, both the control and avatar conditions featured positively biased facial expressions, masking the actual expressions of the communication partner. This may also explain why anxiety levels in the smile condition were also not significantly lower than in the control condition. Although there is no clear evidence that anime characters alleviate anxiety due to their results being mixed with positive facial feedback and facial familiarity, researchers suggest that anime characters create a sense of warmth, enhancing the comfort of interpersonal interactions. Considering their research findings and the advantages of anthropomorphic characters, we are highly confident in the potential of using attractive anime characters as facial filters to alleviate speakers' anxiety, particularly after minimizing practice effects and the influence of positive facial expressions.

Hypothesis 2: Modifying the interviewer's appearance to an attractive anime character will reduce the speaker's anxiety in one-on-one video conferences.

3.4 Validation experiment

We compared the effects of the two aforementioned face filters on speakers' anxiety during presentations in a real video conference scenario. We employed a between-group design to minimize practice effects. The participants were randomly divided into one of three groups (conditions):

- The Good Friend Interviewer (GFI) group, where the interviewer's appearance is modified to resemble that of the speaker's good friend.
- The Anime Character Interviewer (ACI) group, where the interviewer's appearance is modified to resemble that of an attractive anime character.
- The Stranger Interviewer (SI) group, where the interviewer appeared as a default stranger, as the control condition.

3.4.1 Experiment design

We designed an experiment in which participants delivered a 3-minute presentation in front of an interviewer who used existing software to apply face filters, masking his real appearance. We utilized Avatarify Desk (https://github.com/alievk/avatarify-desktop) to manipulate the appearance of the interviewer as seen by the participants. This software maps real-time facial and head movements onto a subject in a photograph, enabling characters in still images to exhibit natural expressions and bodily movements, closely resembling those in actual video interactions.

We invited a 30-year-old male PhD candidate from another unit to act as our interviewer (practical assistant). Before each experiment, the interviewer conducted thorough tests to ensure that his facial expressions (neutral but slightly negative) and movements were accurately mapped onto the corresponding positions in the photograph. During the presentation task, the assistant was instructed to maintain eye contact with the participant throughout the session, except for 1–2 instances of looking down to simulate note-taking. If a participant finished their presentation early, the assistant would remain in the note-taking posture until the video conference automatically ended.

To better simulate an authentic interview environment, neutral reading tasks and quiet waiting periods were arranged before and after meeting the interviewer. These tasks served two purposes: first, to ensure participants were alone in the room during the video conference, giving the experimenter time to leave and re-enter; and second, to reduce pre-interview anxiety

by redirecting participants' attention to the reading material. The detailed task schedule was as follows:

- 1-minute neutral reading task: Participants were instructed to read the content displayed on the screen, which included experimental instructions and an introduction to the automatic transitions.
- 3-minute presentation task with the interviewer: The screen displayed a live video of the interviewer via Cisco Webex App version 42.6 (https://www.webex.com). To enhance immersion, the small window showing the participant's video and the menu bar were disabled, allowing the interviewer's video to occupy the entire screen.
- 30 seconds post-presentation reflection: A prompt displayed the message, "Presentation completed, please wait quietly for the experimenter to return," simulating a brief period of reflection after the presentation.

The transitions between tasks were implemented by pre-loading the relevant materials in the corresponding sequence and setting them to automatically close at specific times. The detailed settings were as follows:

- Full-screen reading material (top layer): Automatically closed after 60 seconds.
- Adjusted video conferencing software window (middle layer): Automatically closed after 240 seconds.
- Image with concluding text (bottom layer): Automatically closed after 270 seconds.

Based on the prior experiment [53], the preparation time for the presentation was set at 7 minutes, which was sufficient for participants to outline their main points. The speech topic was assigned as: "Please introduce your strengths and/or weaknesses," as this is one of the most common interview questions. Participants could choose to focus on either one aspect or both, aiming to reduce potential anxiety from prematurely ending their presentation.

For the ACI condition, we selected five male anime character images from the internet. These images were evaluated by 20 students based on familiarity, anxiety, and likability. The most highly rated character, which scored highest in likability and familiarity, and lowest in anxiety, was chosen as the ACI interviewer (See Fig 3.1). For the GFI condition, participants provided a frontal photo of one of their good friends. For the SI condition, the practical assistant's own frontal photograph was used as the stranger's appearance. Thus, in the ACI and SI groups, the audience appearance was predetermined and consisted of a single character or

face, whereas in the GFI group, the appearances varied. To standardize the visual presentation across conditions, we pre-adjusted the images to ensure similar head proportions and positioning on the screen. Using Adobe Photoshop, we also ensured uniformity in clothing, head size, and background, maintaining a light grey backdrop for all three conditions. These appearances are illustrated in Fig 3.8, showcasing the visual representations of the interviewers across the conditions in eye-tracking results.

Figure 3.1: **Experimental setup under the Anime Character Interviewer (ACI) condition.** An anime character with the highest likability rating was displayed as the interviewer during the presentation.

The experimental room layout is shown in Fig 3.2, highlighting the equipment worn by participants, including the eye-tracking device (Tobii Pro Glasses 3) and the functional near-infrared spectroscopy (fNIRS) device (Hot-2000). Additionally, the setup features an infrared thermal imaging camera (InfReC R450). Furthermore, we administered questionnaires, including the State-Trait Anxiety Inventory (STAI), the Perception of Speech Performance Scale (PSP), and a self-designed questionnaire, to collect subjective data.

3.4.2 Measures

Eye-tracking We utilized the Tobii Pro Glasses 3 (Tobii AB, Danderyd, Sweden), a wearable eye-tracking device resembling regular eyewear, designed for precise recording of participants' eye movements. The system comprises a head unit, a recording unit, and a controller application, which can be installed on a computer for data management. The glasses feature removable lenses tailored to correct near-sightedness and far-sightedness, with lens options ranging from +3 to -8.0 diopters for each eye, adjustable in 0.5 diopter intervals. With a sampling rate of 100

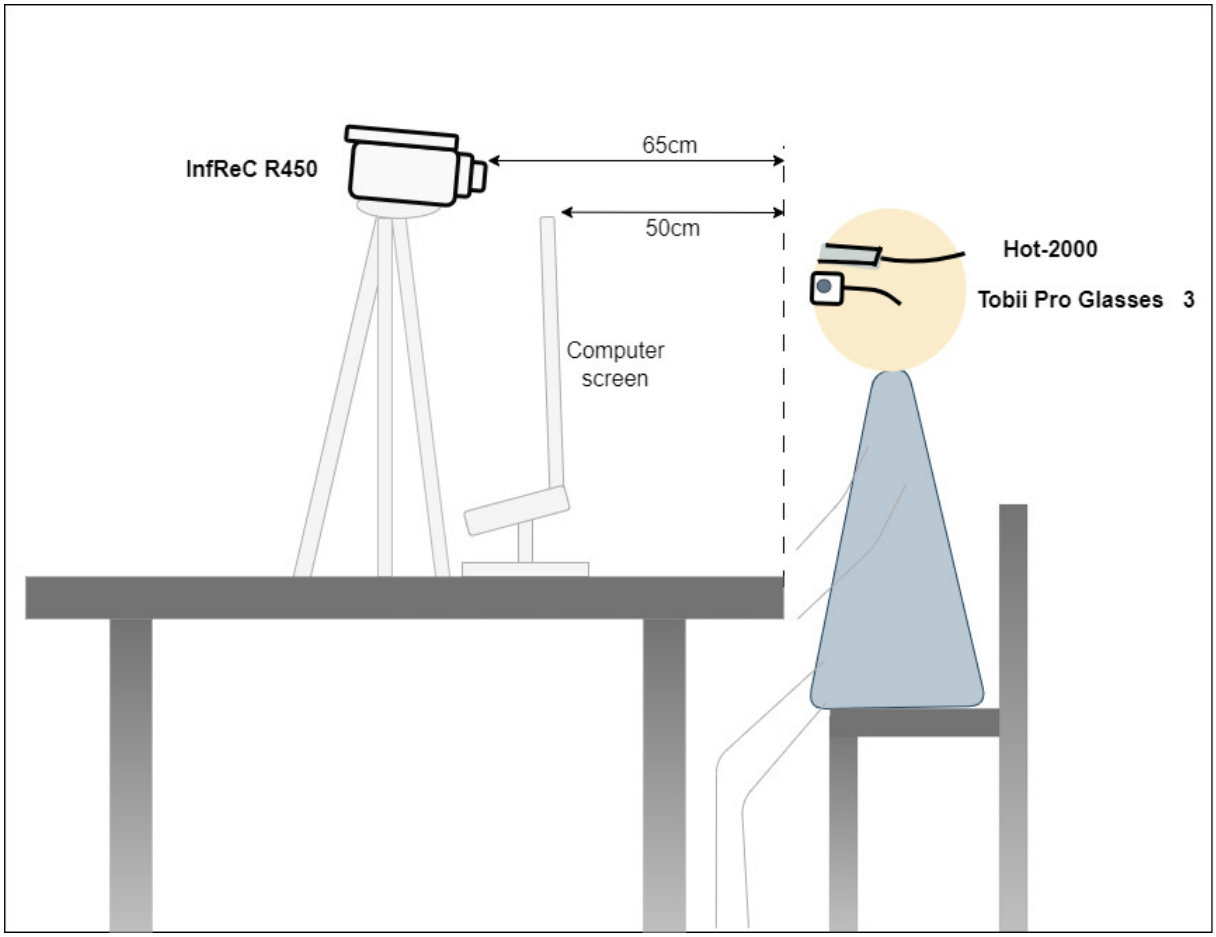


Figure 3.2: **Layout of the experimental room and devices.** Shows the participant's position and measurement devices, including eye-tracking glasses, fNIRS, and thermal imaging.

Hz, the device was configured and calibrated for optimal accuracy in tracking eye movements.

The device recorded participants' eye movements during the experimental tasks. This setup enabled us to analyze participants' visual behaviors during the presentation task, offering valuable insights into their focus and attentional patterns.

Cerebral blood flow changes We utilized the Hot-2000 device (NeU Corp, Tokyo, Japan) to measure changes in cerebral blood flow. The device featured two source-detector (SD) pairs positioned on the left and right sides of the forehead, each consisting of a 1-cm SD pair and a 3-cm SD pair. The device emitted infrared light at a wavelength of approximately 800 nm and recorded data at a sampling rate of 10 Hz. The 3-cm SD pair functioned as a long-separation detector, capturing changes in blood flow linked to neural activity, while the 1-cm SD pair acted as a short-separation detector to filter out interference from the scalp.

Increased activity in the prefrontal cortex (PFC), particularly in the left PFC, has been shown to occur during the regulation of negative stimuli [114, 156]. This device enabled us to analyze changes in participants' cerebral blood flow before and after seeing the interviewer, potentially offering insights into the neural responses elicited by the interviewer's appearance.

Thermal infrared imaging We used thermal imaging with the InfReC R450 device (Nippon Avionics Co., Ltd, Yokohama, Japan) to measure changes in participants' nasal temperature. This non-contact method allows the detection of emotional states by capturing changes in facial temperature [50, 111, 113]. The InfReC R450 recorded facial temperatures at a rate of one frame per second, using a 480×360 detector pixel array with spectral sensitivity ranging from 8 to 14 micrometers.

Alterations in nasal tip temperature are recognized as reliable indicators of anxiety, with lower temperatures correlated with heightened anxiety levels [113, 152, 43]. Fig 3.3 displays the extracted nose region from the thermal images, which were intentionally blurred to preserve privacy. We employed the open-source project TIPA [26] to automatically track the nasal region of interest. The device was positioned 70 centimeters from the participants, with blackbody calibration conducted prior to recording. This setup allowed us to monitor participants' nasal temperatures during the presentation task, potentially providing insights into their physiological responses.

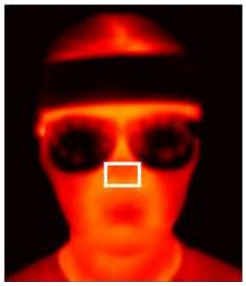


Figure 3.3: **Extracted nasal region with privacy blur.** Depicts the tracked nasal region used for thermal analysis while maintaining participant anonymity.

STAI-S The STAI [137] consists of two scales: the State Anxiety Scale (STAI-S) and the Trait Anxiety Scale (STAI-T). The STAI-S assesses an individual's feelings and emotions at the present moment, whereas the STAI-T measures more general and enduring anxiety tendencies. For our experiment, we utilized the Japanese STAI-S (Form-JYZ), which comprises 20 items (e.g., "I feel tense") that are presented in a 4-point Likert scale format. The total score for state anxiety ranges from 20 to 80. Lower scores (20-44) indicate lower levels of state anxiety, reflecting calmness or a lack of temporary anxiety. Moderate scores (45-54) suggest moderate

levels of state anxiety. Higher scores (55-80) indicate higher levels of state anxiety, reflecting greater temporary anxiety. After completing the presentation task, participants self-reported their current emotional state using the STAI-S. Previous studies have demonstrated strong internal consistency for the STAI-S [25, 31]. In our study, the STAI-S exhibited exceptional internal consistency, with α=.95.

PSP The PSP [125] evaluates speech performance using 17 items, which cover both general aspects (5 items, e.g., "Appeared Nervous") and specific aspects (12 items, e.g., "Voice Quivered") of the speech. The participants were asked to rate their agreement with various statements regarding their speech performance on a scale ranging from 0 ("not at all") to 4 ("very much"). Previous studies have consistently shown strong internal consistency when using the PSP scale for self-evaluation [31, 129, 117, 155]. In our current study, the internal consistency of the total scale score for participant ratings is excellent, with α=.87.

Self-designed questionnaire We designed an 8-item questionnaire using a 5-point Likert scale (items are listed in Table 3.1), with response options ranging from 1 (strongly disagree) to 5 (strongly agree). The purpose of this questionnaire was to evaluate the extent to which various factors potentially influencing anxiety levels were effectively controlled during the experiment. Rather than measuring a single latent construct, the questionnaire captured a broad range of participants' subjective experiences across multiple dimensions, including perceived familiarity, perceived kindness, and distraction. Additionally, an open-ended section was included to allow participants to freely share their thoughts and feelings, providing further qualitative insights.

Table 3.1: **Self-designed questionnaire items.**

1. The topic of the speech is difficult.
2. I am accustomed to video interviews.
3. I am familiar with the face of the interviewer.
4. The interviewer was kind.
5. I have a strong desire to speak to the interviewer.
6. The expression of the interviewer is stern.
7. The face of the interviewer is distracting.
8. The fear of interviews is alleviated.

3.4.3 Experiment procedure

Participants We recruited participants by sending an invitation email to all master's and doctoral students at our university. The email outlined the experiment content, which is a real video interview experiment where participants deliver a presentation to an interviewer who uses a facial filter. A total of 45 participants, consisting of 33 males and 12 females, aged 22 to 29 years, were evenly distributed across three groups. Participants assigned to the GFI group were informed of submitting a frontal photo of a good friend. While the sample size of 45 participants may appear modest for a between-subjects design involving three conditions, it aligns with standards in the human-computer interaction (HCI) field. Based on a survey by Caine [24], the average number of participants per condition in CHI publications is approximately 13. Our study exceeds this benchmark with 15 participants per condition, providing reasonable confidence in the robustness of the findings within this research context.

Considering that the interview language in the region was primarily Japanese, Japanese was designated as the language for oral presentations in the experiment. Among the participants, nine were native Japanese speakers. The remaining non-native Japanese participants possessed a proficiency level of N2 or above on the Japanese Language Proficiency Test, a standardized assessment certifying non-native speakers' Japanese language skills. A passing grade of N2 or above demonstrates strong abilities in comprehending complex texts, engaging in advanced conversations, and writing at a high level, making it widely recognized for non-native speakers working, studying, or living in Japan. We evenly distributed the participants' Japanese proficiency levels across the three groups. All participants had experience in online Japanese-language interviews, whether for employment, academic advancement, or simulation training.

On the day of the experiment, participants were provided with a verbal introduction by the experimenter, following the details outlined in a written document. This introduction covered the purpose and content of the experiment, instructions on the types of equipment to be used, details about the data to be collected, its storage duration, and the compensation rate set by the university. They were also informed that the interviewer's appearance in the video call had been modified using a facial filter and that behind the filter was a real person from outside the university who would evaluate their presentation. The experimenter also emphasized that participants could stop the experiment at any time if they felt uncomfortable. Once participants fully understood the provided information and had no further questions, they voluntarily signed a consent form and completed a personal information sheet to receive compensation.

Subsequently, we introduced a separate document containing screen transition demonstrations. This document included the content for the neutral reading task, simple diagrams representing the interviewer shown during the presentation task, and the quiet waiting instructions

displayed at the end. We emphasized that any errors or omissions during the reading task would not affect the experiment, as the data collection focused on the presentation task. This served to highlight the importance of the presentation task while reducing potential pressure during the reading task. Participants were given five minutes to familiarize themselves with the reading task, which also included instructions to minimize head movement and avoid ending their presentation prematurely. Then, participants received a sheet of paper with the assigned speech topic and blank space to draft notes or prepare their presentations during the 7-minute preparation time.

Afterward, participants were directed to sit in front of the computer screen. The experimenter replaced the necessary lenses and connected the head and recording units as outlined in the Tobii Pro Glasses 3 manual. Participants wore eye-tracking glasses continuously, even during rest periods, to adapt to their weight (76.5 grams, including the connecting cable). Participants then positioned the fNIRS device on their dry foreheads, ensuring no interference from hair, as per the Hot-2000 manual. Once the placement was confirmed, device measurements began. The experimenter also activated the infrared camera at this stage. After a 3-minute rest, the experimenter calibrated the eye-tracking glasses according to the Tobii Pro Glasses 3 User Manual. During calibration, a calibration card was placed on the computer screen, and participants were instructed to focus on its center. Calibration was initiated via the application interface, and upon successful completion, eye-tracking data recording commenced. Fig 3.4 illustrates the operational status of these devices during the experimental process, while Fig 3.5 presents a photo of a participant wearing both the eye-tracking and fNIRS devices. Privacy concerns were addressed by applying image blurring to the photo.

After calibration, the experimenter left the room, and the computer screen displayed the reading task. Participants independently completed the following tasks shown on the screen. The room temperature was controlled between 18-20°C, which was considered a comfortable temperature by the participants. The interviewer was located in a room on the other side of the campus, and his side of the video call was set to mute to act as the audience. As a result, participants were in a quiet room while delivering the presentation to the interviewer.

After 250 seconds, the experimenter returned and shut down the types of equipment. Participants then completed the questionnaires on the computer screen using a mouse. In the final section of the self-made questionnaire, they were encouraged to verbally describe their experiences and express their feelings. The experimenter recorded their responses through audio. Finally, the experimenter engaged in a brief conversation with them to address any questions they had. The entire experiment lasted approximately 40 minutes, and participants received a compensation of 1,000 JPY.

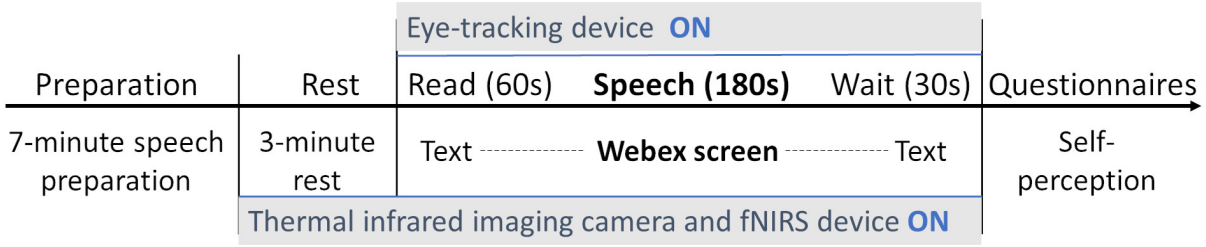


Figure 3.4: **Overview of the experimental procedure.** Illustrates task order and device usage during the presentation experiment.

Figure 3.5: **Participant wearing measurement devices.** Displays the setup of eye-tracking glasses and the fNIRS device during data collection.

Ethics and Consent Statement The study was approved (Approved No. 04-013) by the Japan Advanced Institute of Science and Technology Life Science Committee on July 27, 2022. All methods were performed in accordance with the relevant guidelines and regulations. The recruitment period for the validation experiment was from August 1, 2022, to November 1, 2022. Participation was voluntary and signed written informed consent forms. Additionally, participants were informed that they could withdraw or terminate the experiment at any time without any negative consequences. Pre-counseling and post-counseling sessions were conducted before and after the experiment, respectively, to alleviate potential discomfort caused by the measurement devices and to mitigate the potential impact of the experiment.

All figures containing identifying information in this manuscript were obtained with the informed consent of the respective participants for the publication of their identifying information/images in the online open-access publication.

3.4.4 Experiment results

Due to technical issues, three participants required additional data and were subsequently excluded from the analysis. Thus, the statistical analysis included data from 45 participants, consisting of 33 males and 12 females between the ages of 22 and 29.

Because the data were not normally distributed (Shapiro-Wilk, $p<.05$), we employed non-

parametric tests for the statistical analysis. For each dependent variable, we performed a Kruskal-Wallis omnibus test followed by Mann-Whitney U tests for pairwise group comparisons when appropriate, and Bonferroni correction for multiple comparisons was applied as implemented in IBM SPSS Statistics version 27 (https://www.ibm.com/products/spss-statistics). However, correction was not applied across the full set of all p values reported in the study.

Perception

STAI-S Fig 3.6 shows the distribution of scores on the self-reported current anxiety state scale among the three groups, with higher scores indicating higher anxiety levels. A Kruskal-Wallis test revealed a significant difference in state anxiety scores across the three groups, H(2)=7.81, p=.020. The median scores were 31 for GFI, 50 for ACI, and 48 for SI. Follow-up Mann-Whitney U tests with Bonferroni correction showed that the GFI group reported significantly lower anxiety than both the SI group (Z=-2.41, adjusted p=.048, r=.36) and the ACI group (Z=-2.43, adjusted p=.045, r=.36). No significant difference was observed between the ACI and SI groups (Z=.03, adjusted p=1, r=0).

PSP Fig 3.7 shows the PSP scores of the participants' self-evaluations in the three groups, with higher scores indicating a poorer perceived performance. A Kruskal-Wallis test revealed a significant difference in self-perceived performance across the three groups, H(2)=10.72, p=.005. The median scores were 21 for GFI, 34 for ACI, and 36 for SI. Post hoc Mann-Whitney U tests with Bonferroni correction indicated that participants in the GFI group rated their performance significantly better than those in the SI group (Z=-3.07, adjusted p=.006, r=.46) and the ACI group (Z=-2.53, adjusted p=.035, r=.38). No significant difference was found between the ACI and SI groups (Z=-.54, adjusted p=1, r=.08).

Self-designed questionnaire Item 3 – Familiarity with the Interviewer's Face. A Kruskal-Wallis test revealed a significant group difference in perceived familiarity with the interviewer's face, H(2)=29.45, p<.001. Follow-up Mann-Whitney U tests with Bonferroni correction showed that participants in the GFI group rated the interviewer as significantly more familiar than both the SI group (Z=4.23, adjusted p<.001, r=.63) and the ACI group (Z=5.06, adjusted p<.001, r=.75). No significant difference was found between the ACI and SI groups (Z=-.83, adjusted p=1, r=.12).

Item 4 – Perceived Kindness of the Interviewer. A Kruskal-Wallis test indicated a significant difference in perceived kindness across groups, H(2)=8.06, p=.018. Post hoc comparisons revealed that participants in the GFI group perceived the interviewer as significantly kinder than

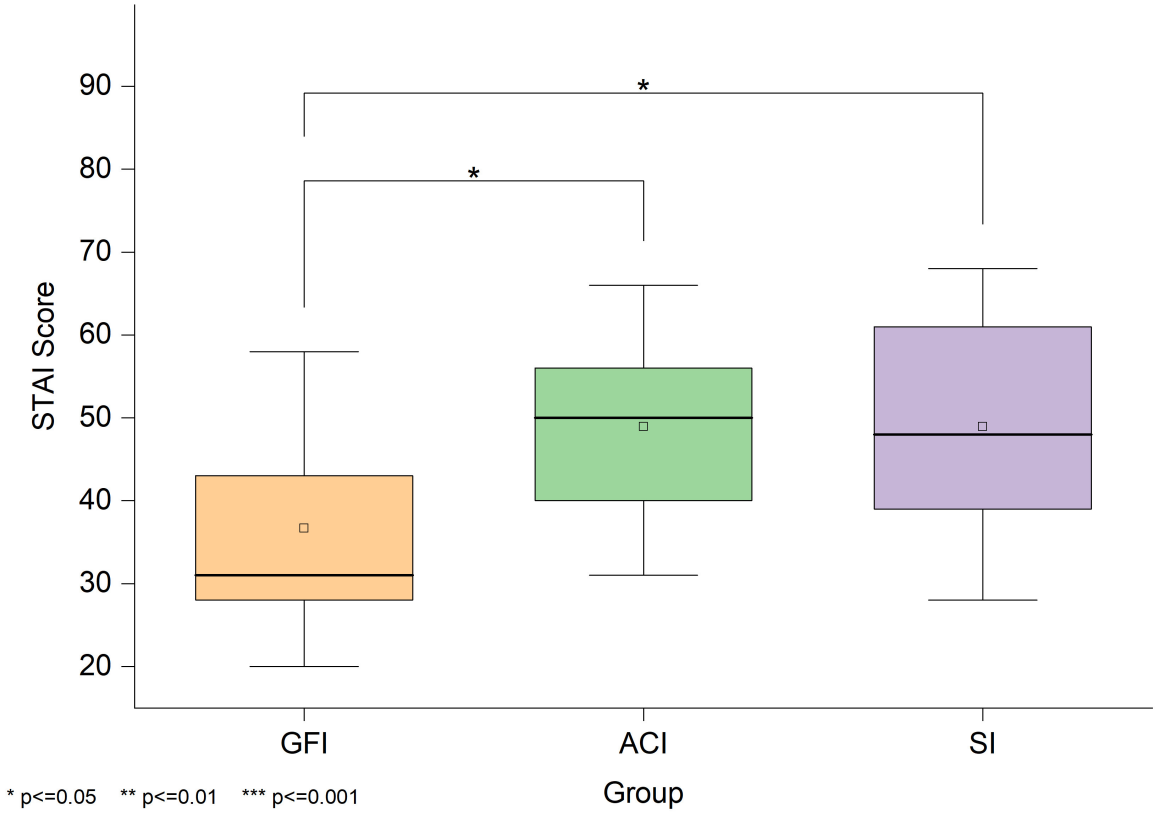


Figure 3.6: **State-Trait Anxiety Inventory-State (STAI-S) scores across the three groups.** Shows anxiety level differences among the Good Friends Interviewer (GFI), Anime Character Interviewer (ACI), and Stranger Interviewer (SI) groups, with lower scores indicating lower anxiety.

those in the ACI group (Z=2.67, adjusted p=.023, r=.40). No statistically significant differences were observed between GFI and SI (Z=2.18, adjusted p=.088, r=.32) or between ACI and SI (Z=-.49, adjusted p=1.00, r=.07).

Item 8 – Fear Reduction During the Interview. A Kruskal-Wallis test found a significant difference in interview-related fear reduction, H(2)=9.45, p=.009. Pairwise tests indicated that the GFI group reported significantly more reduction in fear compared to the SI group (Z=3.06, adjusted p=.007, r=.46), while no significant differences were found between GFI and ACI (Z=1.76, adjusted p=.236, r=.26) or between ACI and SI (Z=1.31, adjusted p=.575, r=.20). No significant differences were observed between groups for the other items.

Several noteworthy points emerged from the spontaneous expressions of the participants. In the GFI group, participants mentioned experiencing dissonance when seeing foreign expressions or angles shown on familiar faces. In the ACI group, many participants expressed difficulty receiving feedback from the anime character, which caused unease. In the SI group, some

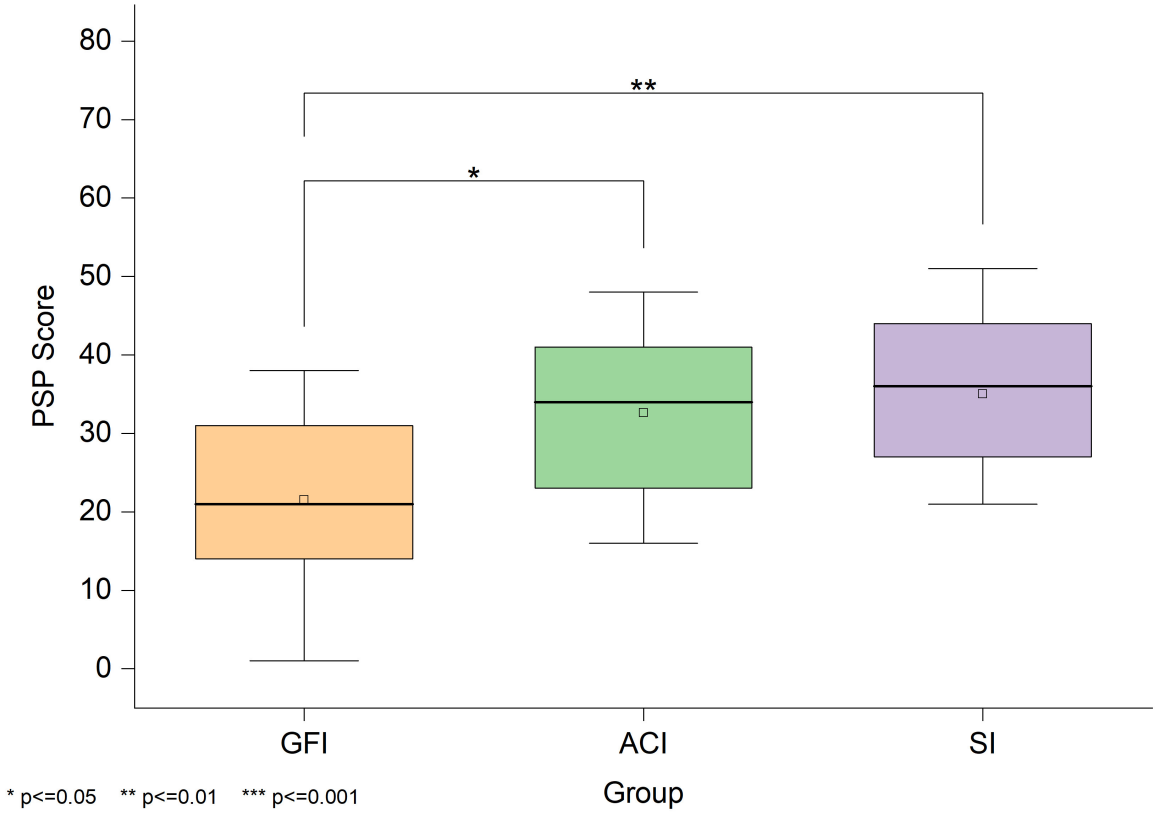


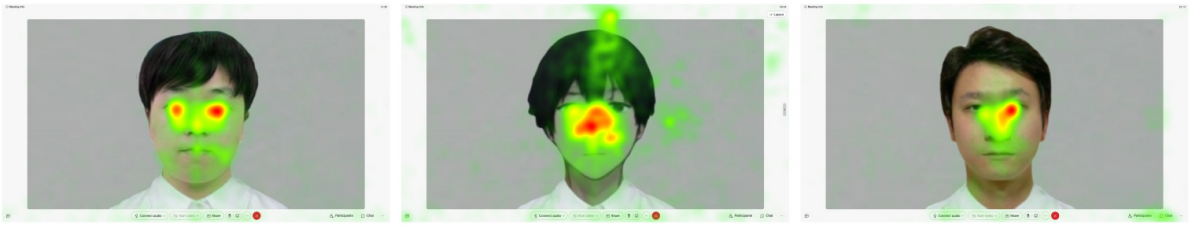
Figure 3.7: **Perception of Speech Performance (PSP) scores across the three groups.** Shows perceived speech performance in the Good Friends Interviewer (GFI), Anime Character Interviewer (ACI), and Stranger Interviewer (SI) groups, where lower scores correspond to better self-assessed performance.

participants reported that the interviewer was familiar because it aligned with their expectations of an interviewer, despite never having met him in real life.

Behavior

Eye-contact The recorded data were analyzed using Tobii Pro Lab version 1.207 (https://www.tobii.com/products/software/ behavior-research-software/tobii-pro-lab) to identify and quantify gaze interactions. We focused on the area of interest (AOI) around the interviewer's eyes and analyzed the participants' gaze duration and frequency within this area. To accurately capture participants' gazes in dynamic situations, we utilized a dynamic AOI that can be repositioned to match targets in the video. We employed an automatic mapping system within the software to map the positions of an AOI from dynamic videos to static photos. Subsequently, frame-by-frame visual inspection and adjustments were made to ensure accurate correspondence. To ensure consistent exposure to the interviewer's gaze, we analyzed the first 120 seconds of the

presentation task; two participants finished their speech slightly over two minutes, and the interviewer maintained a downward gaze during the remaining time. We applied the Tobii I-VT attention filter with a threshold of 100 degrees per second, calculating both the duration and frequency of the participants' gazes in the dynamic AOI. The heat maps in Fig 3.8 depict the overall distribution of the eye movement data for each group. In the GFI, we focused the data on the specific photo, which is shown in a). Images a), b), and c) are actual screenshots of the screen participants faced during the presentation task. The gradient from green to yellow to red represents the continuum of fixation duration, which ranges from short to long.



(a) GFI (b) ACI (c) SI

Figure 3.8: **Overlaid heat maps of gaze fixation during the presentation task.** Shows the superimposed gaze patterns from the participants under the Good Friends Interviewer (GFI), Anime Character Interviewer (ACI), and Stranger Interviewer (SI) groups during presentations. Colors from green to yellow to red indicate fixation durations from short to long.

The total duration of the whole fixation within the AOI under the three groups is shown in Fig 3.9. A Kruskal-Wallis test revealed a significant difference in the total duration of fixations on the interviewer's eyes across the three groups, H(2)=9.15, p=.010. The median scores were 54.53 for GFI, 12.10 for ACI, and 31.33 for SI. Follow-up Mann-Whitney U tests with Bonferroni correction showed that participants in the GFI group maintained significantly longer gaze duration than those in the ACI group (Z=3.01, adjusted p=.015, r=.45). No significant differences were observed between GFI and SI (Z=1.22, adjusted p=1.00, r=.18) or between ACI and SI (Z=-1.72, adjusted p=.516, r=.26).

The number of whole fixations within the AOI under the three groups is shown in Fig 3.10. A Kruskal-Wallis test indicated a significant difference in the number of fixations across conditions, H(2)=8.68, p=.013. The median scores were 103 for GFI, 50 for ACI, and 136 for SI. Follow-up comparisons indicated that participants in the GFI group made significantly more fixations on the interviewer than those in the ACI group (Z=2.69, adjusted p=.043, r=.40). No significant differences were found between the GFI and SI groups (Z=.22, adjusted p=1.00, r=.03) or between the ACI and SI groups (Z=-2.38, adjusted p=.103, r=.36).

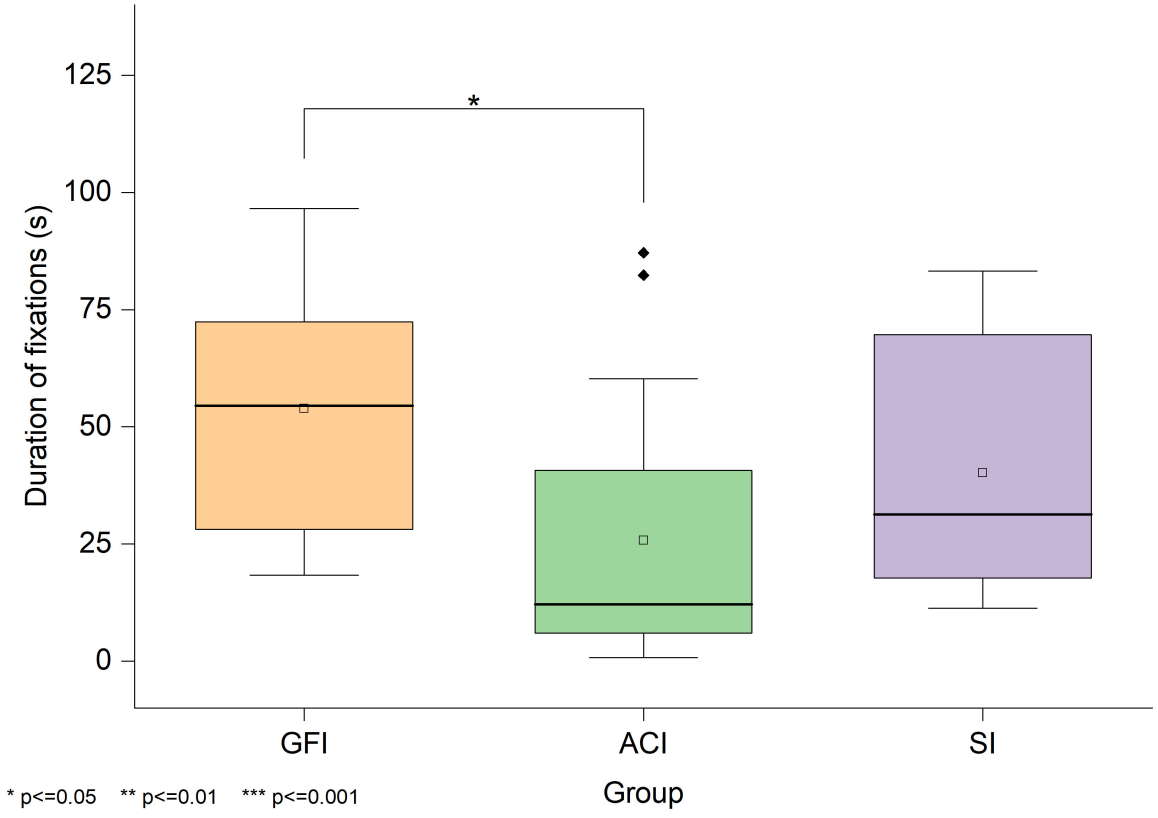


Figure 3.9: **Fixation duration within the eye region of interest.** Shows the total time participants spent fixating on the interviewer's eye region under the Good Friends Interviewer (GFI), Anime Character Interviewer (ACI), and Stranger Interviewer (SI) groups.

Word count We analyzed the number of complete words spoken by the participants during the 3-minute interview sessions. We employed Whisper software (https://github.com/openai/whisper) to transcribe the participants' spoken content into text and subsequently quantified it. A Kruskal-Wallis test revealed a significant difference in word count across the three groups, H(2)=6.06, p=.048. The median scores were 517 for GFI, 512 for ACI, and 441 for SI. Post hoc comparisons indicated only a trend toward a higher word count in the GFI group compared to the SI group (Z=2.43, adjusted p=.090, r=.36).

Correlation

We utilized Spearman correlation analysis to individually compare the following variables in pairs: STAI_S, PSP, fixation duration, fixation number, and word count (see Table 3.2). The results revealed a strong positive correlation between fixation duration and fixation number. Additionally, a strong positive correlation was observed between the STAI_S and PSP scales,

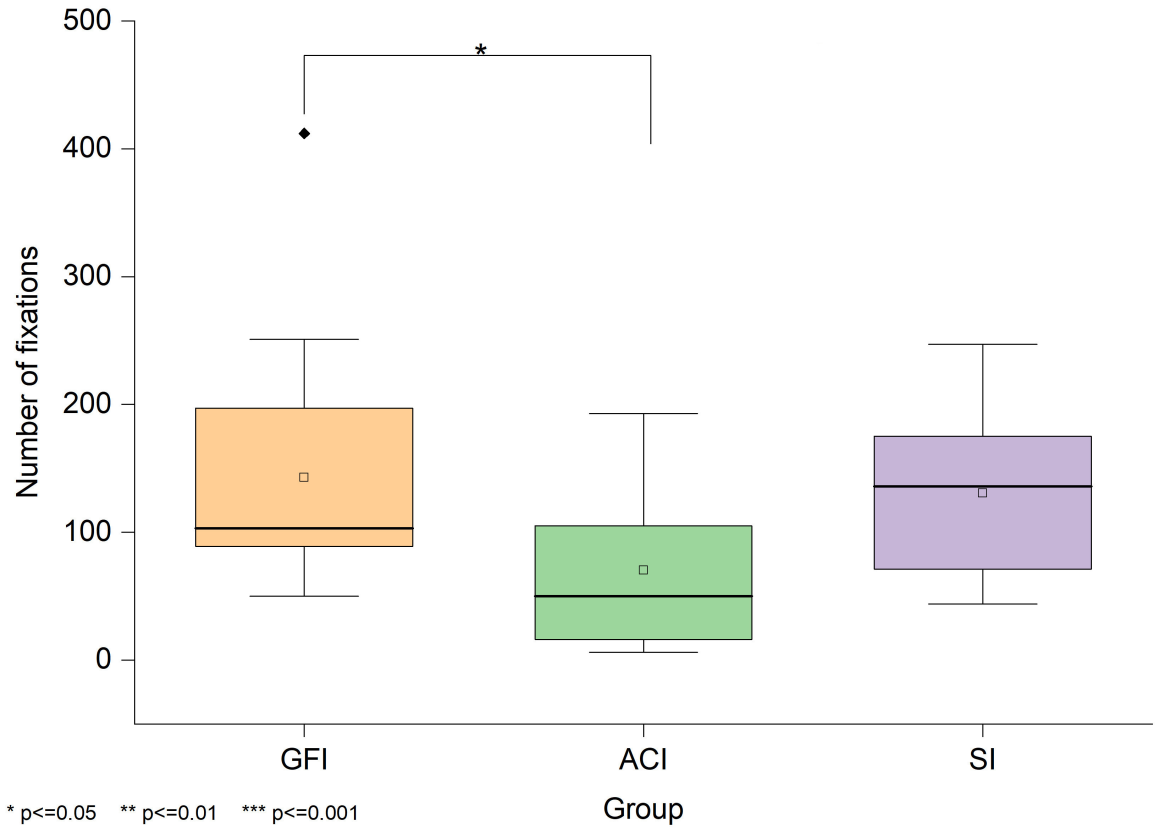


Figure 3.10: **Fixation count within the eye region during the presentation task** Shows the total number of fixations on the interviewer's eye region across the Good Friends Interviewer (GFI), Anime Character Interviewer (ACI), and Stranger Interviewer (SI) groups.

indicating that higher self-reported anxiety levels were associated with poorer perceived performance. Both scales displayed a weak negative correlation with word count, suggesting that speaking less during the presentation was somewhat associated with higher self-perceived anxiety and lower self-perceived performance.

Physiological data

Nasal temperature To minimize individual differences, we first calculated the nasal temperature changes by subtracting the temperature recorded during a continuous 120-second resting period from that recorded during the first 120-second presentation task period, using the same rationale as in the eye-contact analysis. To capture the psychological changes triggered by the gaze of the interviewer, we further subtracted the average nasal tip temperature from the last 30 seconds of the reading task. The last 30 seconds of the reading task were selected to minimize the influence of the experimenter's presence, as participants were alone in the room during this

Table 3.2: **Correlations among variables in the validation experiment.**

	STAI_S	PSP	FD	FN	Word_count
STAI_S					
Spearman correlation	1	0.83^{**}	-0.111	-0.040	-0.301^{*}
Sig. (2-tailed)		0.001	0.467	0.793	0.044
PSP					
Spearman correlation	0.83^{**}	1	-0.143	-0.058	-0.352^{*}
Sig. (2-tailed)	0.001		0.347	0.705	0.018
Fixation_duration (FD)					
Spearman correlation	-0.111	-0.143	1	0.792^{**}	0.046
Sig. (2-tailed)	0.467	0.347		0.001	0.763
Fixation_number (FN)					
Spearman correlation	-0.04	-0.058	0.792^{**}	1	-0.013
Sig. (2-tailed)	0.793	0.705	0.001		0.932
Word_count					
Spearman correlation	-0.301^{*}	-0.352^{*}	0.046	-0.013	1
Sig. (2-tailed)	0.044	0.018	0.763	0.932	

* Correlation is significant at the 0.05 level (2-tailed).
** Correlation is significant at the 0.01 level (2-tailed).

period after the experimenter had left, ensuring a consistent environment for baseline recording.

As shown in Fig 3.11, which displays the average nasal temperature variations across participants for each group during the presentation task, the SI group appears to exhibit a more substantial decrease, while the other two groups show little difference. However, statistical analysis did not reveal any significant differences in nasal temperature changes among the three groups during this period.

Brain activity We calculated the changes in cerebral blood flow between -2 and 4 seconds from the start of the interviewer-interviewee meeting, aligning with the observation time in the previous study [82]. However, no significant differences were observed across conditions. Fig 3.12 illustrates the mean change in cerebral blood flow over time in the left PFC. The right side of the dashed line corresponds to the change in cerebral blood flow after the interviewer's appearance, corrected by the mean value 2 seconds before the interviewer's appearance (left side of the dashed line). In our experiment, the observed increase in cerebral blood flow within the left PFC in the SI group could be related to the modulation of anxiety states experienced when facing the audience.

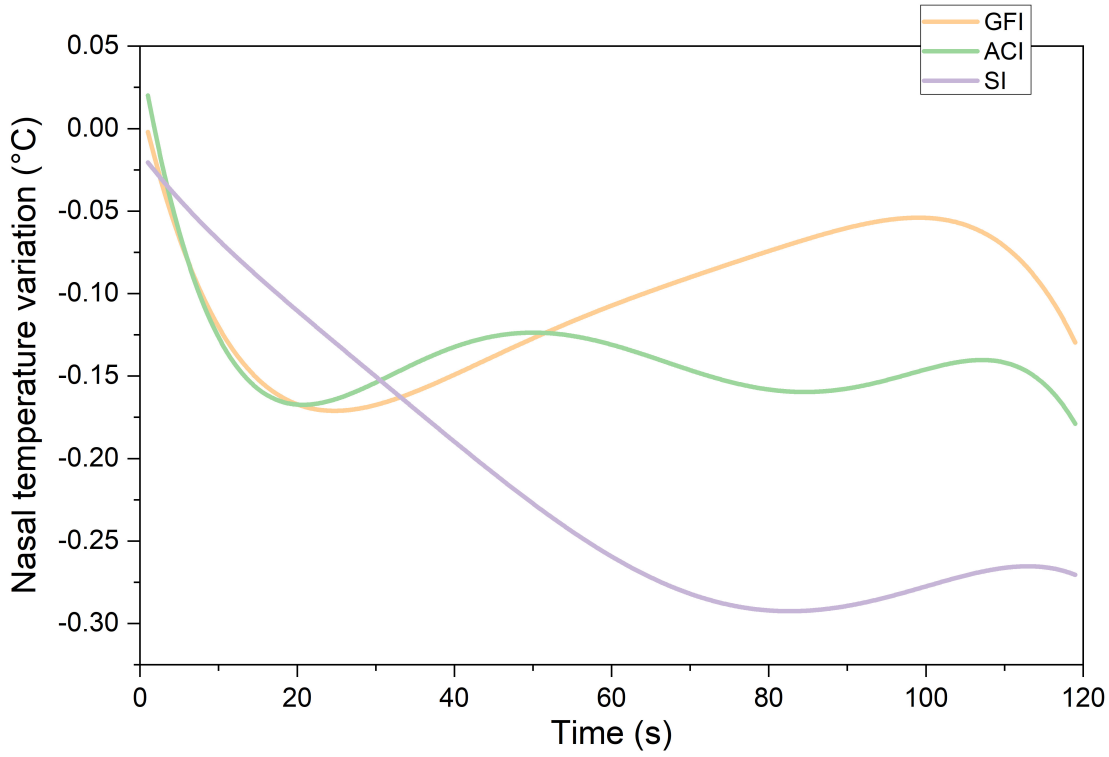


Figure 3.11: **Group-averaged nasal temperature changes.** Shows nasal temperature changes during presentations under the Good Friends Interviewer (GFI), Anime Character Interviewer (ACI), and Stranger Interviewer (SI) groups as a physiological indicator of anxiety, with lower temperatures correlated with heightened anxiety levels.

3.5 Discussion

With the mainstream adoption of video conferencing, an increasing number of studies have found that the anxiety experienced during remote interactions can be comparable to that of face-to-face interactions [163, 93]. This is especially true for individuals with high social anxiety, who may even experience heightened anxiety in video conferences [5]. As video conferencing continues to play a critical role in academic, professional, and interpersonal communication, developing strategies to reduce associated anxiety becomes increasingly important [88].

The findings show that participants who saw their interviewer as a familiar friend (GFI) reported lower anxiety and better self-evaluation than those who saw either a stranger (SI) or an anime character (ACI). Additionally, participants in the GFI condition exhibited more gaze behaviors toward the interviewer compared to those in the ACI condition and demonstrated a greater tendency to produce more verbal output than those in the SI condition. These patterns offer insight into how perceived social context and emotional cues shape user experience

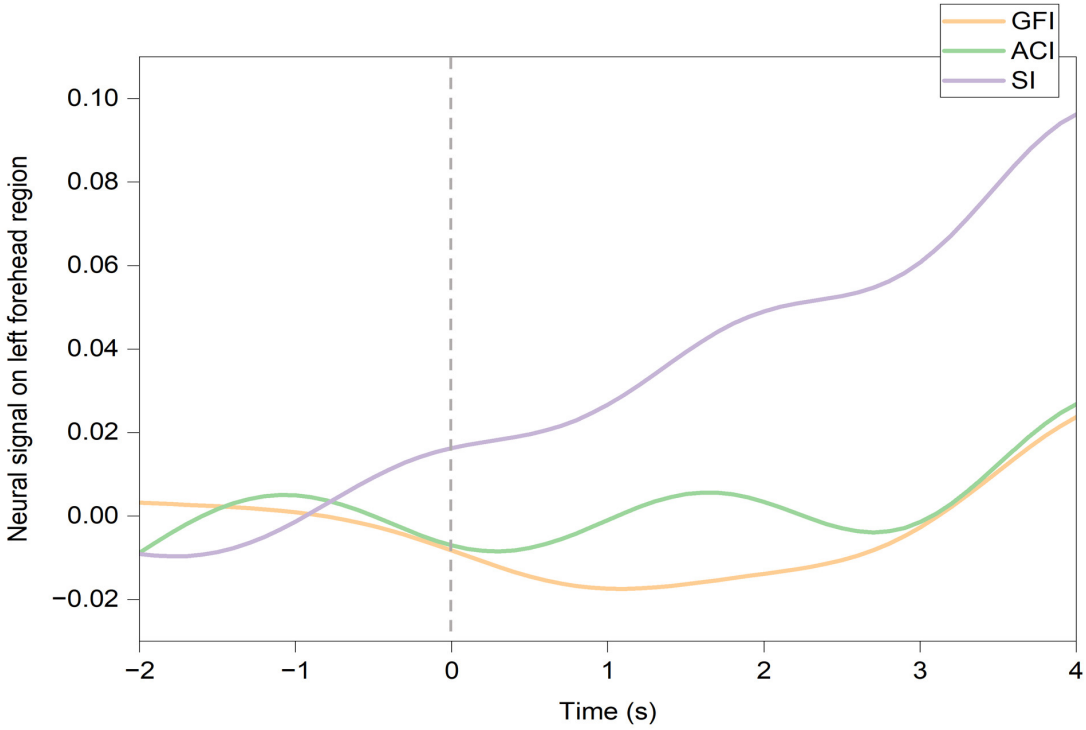


Figure 3.12: **Group-averaged changes in left prefrontal cerebral blood flow.** Shows changes in left prefrontal cortical blood flow at the time of interviewer appearance across the Good Friends Interviewer (GFI), Anime Character Interviewer (ACI), and Stranger Interviewer (SI) groups, reflecting potential anxiety-related neural responses.

in virtual settings, highlighting the potential of leveraging familiar and emotionally resonant visual representations to foster more comfortable and engaged interactions in video-mediated communication.

3.5.1 Familiarity promotes a sense of psychological safety

The GFI condition consistently showed more favorable outcomes across subjective measures. This is consistent with prior work suggesting that familiarity reduces social-evaluative threat by activating schemas of trust, acceptance, and prior positive experience [91, 85]. The friend's face may have served as a cognitive anchor, shifting the user's interpretation of the situation from formal evaluation to informal support. This may explain why participants felt more relaxed and rated their performance more positively—even though the facial expression itself was neutral.

Importantly, these findings extend previous work on face-to-face familiarity [92] into virtual environments where the familiarity is artificially induced. These results also extend our understanding of how social meanings are actively constructed and preserved during interactions.

According to Constructivist theory [154, 153], individuals interpret external stimuli through the lens of prior experiences and knowledge, dynamically shaping their perceptions in ways that align with their expectations and relational histories. In this study, when participants chose to believe they were interacting with a friend, they not only experienced reduced anxiety but also engaged in spontaneous meaning-making processes that preserved the social significance of facial identity. This suggests that the cognitive frameworks individuals bring into social interactions can, to some extent, shield them from the objective reality of the situation, allowing them to maintain emotional and relational continuity even when the interaction partner is known to be non-authentic. Although technical limitations during the experiment may have caused momentary confusion due to changes in facial angles, these disruptions did not fundamentally undermine participants' ability to preserve the social and emotional meanings attached to the interaction. Such findings are consistent with previous research demonstrating that cognition is inherently situated within social contexts and interactions [135], highlighting the flexibility of the mind in sustaining social connections through constructive processes. Moreover, these results imply that interventions leveraging perceived social presence could effectively reduce anxiety in contexts where in-person interaction is not feasible, such as online therapy or remote interviews.

3.5.2 Stylized avatars may disrupt social interpretation

In contrast, replacing the interviewer's face with an anime character failed to reduce anxiety or improve self-evaluation. The ACI condition elicited the least visual engagement (shorter and fewer eye fixations), despite being designed to be visually appealing. This suggests that aesthetic appeal alone is insufficient to foster comfort in evaluative contexts.

One possible explanation lies in the mismatch between the avatar's expressive capacity and participants' expectations. While anime characters are often associated with warmth and expressiveness [163], the avatar used in this study had a neutral or slightly stern expression, coupled with formal attire. This inconsistency may have produced a subtle dissonance or expectation violation [34], undermining the avatar's intended supportive function. Further, the simplified and stylized facial features likely limited participants' ability to interpret emotional nuance [2], reducing the effectiveness of social feedback. In high-anxiety situations, the ability to detect affirming or disapproving cues becomes critical [92], and the absence of such signals may increase uncertainty. The large eyes of the character, a common stylized feature, may have amplified the sense of being watched, contributing to gaze aversion—a typical anxiety response [122].

While the avatar was designed to provide a non-threatening social presence, its limited dy-

namic expressiveness may have constrained its capacity to modulate participants' anxiety effectively. Additionally, the static or limited gaze behavior of the avatar may have been interpreted as staring, which, in socially anxious individuals, can be perceived as threatening or evaluative, thereby inadvertently increasing tension rather than reducing it. Taken together, these factors highlight the complexity of using stylized avatars as social buffers, indicating that careful calibration of appearance, expressiveness, and context-specific behaviors is critical to achieving the desired psychological outcomes in high-stress environments.

3.5.3 Anxiety and behavioral correlates

Our correlational analysis revealed a strong positive association between self-reported state anxiety and poorer perceived performance, consistent with prior findings [79]. Individuals experiencing higher levels of state anxiety tend to interpret their performance more pessimistically, potentially due to heightened attentional focus on perceived mistakes or failures. This pessimistic interpretation may reduce their sense of accomplishment, which in turn can undermine their engagement and self-efficacy in future tasks.

Additionally, we observed a weak negative correlation between anxiety and word count, aligning with earlier findings that anxiety may reduce verbal output in public speaking situations [65, 121, 10]. Heightened anxiety is likely to increase cognitive load, which can restrict the cognitive resources available for speech planning and production, leading to reduced verbal output during presentation tasks.

Interestingly, although gaze behavior (fixation metrics) varied significantly between groups, it did not directly correlate with STAI-S scores. This may reflect the unique dynamics of video conferencing environments, where maintaining eye contact or gaze direction carries different social significance and cognitive demands compared to in-person interactions [66]. In virtual settings, factors such as screen layout, camera positioning, and the absence of direct physical presence may decouple gaze behaviors from subjective anxiety levels, suggesting that gaze metrics alone may not serve as reliable indicators of state anxiety in remote communication contexts.

3.5.4 Implications for design of virtual interactions

Our findings suggest that familiarity—even when artificially induced—can serve as a powerful psychological buffer in high-stressful video communication. The mere appearance of a trusted individual may trigger positive emotional associations and reduce anticipatory anxiety, thereby

facilitating a more adaptive emotional state during the task. These results highlight the importance of social context design in virtual platforms, where the interface can be customized to shape user experience and performance.

Conversely, the limited effect of the anime character on reducing anxiety suggests that anthropomorphic design alone is not universally effective. One possible explanation is that such stylized representations may not evoke consistent emotional resonance across individuals. Emotional comfort in social interactions often depends not only on visual aesthetics but also on personal associations and affective bonds [88]. For individuals lacking a strong connection to anime culture, the character may have felt unfamiliar or even distanced, thereby failing to provide a sense of psychological safety. This highlights the importance of allowing users to select or customize avatars that reflect their own preferences or emotional anchors.

3.5.5 Limitations and future directions

Several limitations of this study should be acknowledged.

First, the sample size of 45 participants, while consistent with standards in similar HCI studies [24], limits statistical power in a three-group between-subjects design. Future studies with larger and more diverse samples are needed to validate and generalize the current findings.

Second, our self-designed questionnaire was intended to assess distinct subjective dimensions—
such as familiarity, friendliness, and fear—rather than a single latent construct. Due to the conceptual heterogeneity of these items, we did not compute a composite score or perform psychometric analyses (e.g., Cronbach's alpha). While this approach allowed us to examine specific experiential components, analyzing each item separately increases the family-wise error rate and the possibility of Type I errors. Accordingly, these questionnaire results should be interpreted as exploratory and supplementary.

Third, although Bonferroni correction was applied within each set of planned comparisons, and the adjusted p values under SPSS tend to be more stringent, the absence of a global correction across all statistical tests in the study remains a limitation.

Fourth, participants were given the option to discuss either their strengths, weaknesses, or both. While this was intended to promote comfort and reduce presentation avoidance, it may have inadvertently introduced a confounding variable. Future studies should consider controlling topic content to ensure more uniform cognitive and emotional demands across participants.

Fifth, although Japanese language proficiency was controlled, other factors such as topic

familiarity, prior interview experience, and trait anxiety were not measured. These individual differences may interact with anxiety in complex ways [79, 91, 33] and merit closer attention in future work. Additionally, the gender of the interviewer as presented via the facial filters was not standardized across conditions. Participants in the GFI group may have seen male or female interviewers, whereas SI and ACI were limited to male-presenting avatars. Gender composition can influence social dynamics and anxiety [23], and should be balanced in subsequent designs.

Sixth, physiological data (infrared thermography and fNIRS) did not reveal significant group differences. This may be due to single-trial design, equipment limitations, or insufficient signal sensitivity. More robust physiological measures and multi-trial designs may yield stronger effects [50]. Additionally, because of practical constraints, the study did not include evaluations of presentation performance by independent third-party professionals.

Finally, due to technical constraints, real-time audience masking in this study was implemented by the interviewer rather than being controlled by the participants themselves. While this ensured standardization across sessions, it limited the autonomy of participants in managing their own anxiety during evaluative interactions.

In the future, under strict adherence to robust ethical principles [163], it would be valuable to develop systems that allow individuals who are more prone to experiencing anxiety to apply such facial filters during interactions without requiring disclosure to the other party. For example, when designing software, users can be prompted upon first entry to choose whether they consent to having face filters applied to them by others during interactions. Once the user has given consent, they will not receive further notifications when others apply these filters.

Such systems could empower users to engage in social, educational, or professional interactions with greater confidence, particularly in high-stakes situations such as interviews or public speaking, where visual self-presentation can be a significant source of stress. By giving users the ability to selectively manage the visual cues they receive, these tools could reduce cognitive load associated with self-monitoring and fear of negative evaluation, thereby supporting smoother and more effective communication.

Additionally, from a design perspective, the integration of customizable audience masking should be accompanied by clear, transparent consent protocols and opt-out options to ensure that all parties retain agency over their interaction experience. Such protocols not only uphold individual autonomy but also help set clear expectations within social and professional interactions, reducing potential misunderstandings or discomfort. Future research should systematically explore the psychological and interpersonal impacts of these user-controlled filtering systems, including their potential effects on trust, authenticity, relational dynamics, and perceptions of social presence. Understanding these impacts is crucial for identifying both the

benefits and unintended consequences of audience masking in diverse contexts. Insights from this research could inform the development of ethically grounded, user-centered technologies that address emotional and accessibility needs while preserving the integrity of interpersonal communication in both remote and hybrid settings.

3.6 Conclusion

This study provides initial empirical evidence that modifying the visual representation of an interviewer can affect users' subjective experience in video calls. Specifically, presenting the audience as a good friend was associated with lower anxiety and higher self-rated performance. In contrast, the anime character filter, though theoretically promising, failed to produce the same benefit—possibly due to reduced emotional expressiveness and mismatched expectations. These findings highlight the potential of visual customization for anxiety reduction, while also underscoring the need for careful design considerations. Future work should further explore adaptive, personalized, and real-time interventions for supporting user comfort in digital communication.

Chapter 4

Study 2: Enhancing User Engagement in Physical Exercise

4.1 Overview

Chapter 4 presents Study 2, which investigates how augmented-reality (AR) companions can support emotional regulation and motivation through designed forms of social presence. Building upon the findings of Study 1, which demonstrated how intimacy and detachment modulate anxiety in cognitive tasks, Study 2 examines whether similar mechanisms operate in embodied, action-oriented contexts. Three sequential experiments (shown in Fig 4.1) were conducted using an AR treadmill system to analyze how variations in social intimacy, aesthetic distance of stylized figure shape arousal and engagement. Experiment 1 compared real, video, AR, and no audiences to validate the potential of an AR audience in producing social facilitation effects within mediated environments. Experiment 2 manipulated relational and aesthetic qualities of AR companions—friend-like, liked-anime, neutral stylized, or none—to assess how emotional affinity influences motivation. The final validation experiment contrasted With_system and Without_system conditions and incorporated post-exercise interviews to interpret user experience. Quantitative results showed that the AR companion improved performance while reducing evaluative pressure and enhancing enjoyment, particularly among high-anxiety participants. Interview data further revealed that participants preferred the companion for its emotional safety and nonjudgmental presence, reporting that it helped them remain focused and emotionally supported. Collectively, these findings confirm that mediated companionship can regulate affect through emotional reassurance and attentional redirection. However, individuals' subjective liking for anime characters strengthens affective bonding; under this condition, anime figures promote task engagement and immersion through a combination of social dis-

tancing and personal affinity, making them the preferred choice for many participants.

Experiment	Focus	Key Variable	Type
Exp 1	Audience Modality	Live / Video-call / AR / None	Passive social presence (no feedback)
Exp 2	Emotional Affinity	Friend / Liked / Disliked / None	Static affective feedback (video-based)
Validation	Feedback Selection	System vs. None Preference for characters	Dynamic feedback (real-time physiological loop)

Figure 4.1: **Comparative Structure of Experiments in Study 2.**

4.2 Preliminary Experiments

4.2.1 Introduction

Sustaining self-regulated physical activity over time remains a persistent and multifaceted challenge in both clinical rehabilitation settings and everyday life [1, 39, 61, 94]. Although the physical and psychological benefits of regular exercise are well documented, ranging from improved cardiovascular health to reduced symptoms of anxiety and depression [16, 44, 120, 56, 126, 32, 124], many individuals struggle to maintain consistent motivation, particularly when engaging in physical activity alone [60]. In solitary environments where social cues, encouragement, and accountability are absent, adherence to exercise routines tends to decline [35, 52].

Social facilitation is a well-established psychological phenomenon wherein the mere presence of others enhances an individual's performance on tasks that are simple, habitual, or well-practiced. This concept, rooted in the foundational work of social psychologist Robert Zajonc [166], proposes that the presence of others increases physiological arousal, thereby enhancing performance on dominant (i.e., easier or familiar) tasks while potentially hindering performance on non-dominant (i.e., complex or unfamiliar) tasks. Based on this theoretical framework, an expanding body of interdisciplinary research, spanning psychology, human-computer interaction, and exercise science, has explored how emerging technologies can simulate or amplify the effects of social presence. In particular, immersive technologies, such as virtual reality (VR) and augmented reality (AR), have opened new possibilities for integrating various virtual agents into exercise contexts [78, 150], each of which embodies different social roles to

increase motivation [110]. For example, virtual rivals in VR-enhanced tasks boost effort and exercise engagement, especially among older adults and during recovery processes [3, 1]. The virtual coaches improved user confidence, perceived competence, and motivation through verbal feedback and demonstration-based learning [108, 89]. Other studies have explored companionship and social connectedness as motivating factors. For instance, an AR dog that mirrored the user's actions created a sense of mutual presence and emotional bonding during exercise [107], whereas smart glasses-based jacket partners encouraged sustained effort by mimicking the experience of running with a peer [60]. At the group level, immersive rowing simulations that require coordinated action with virtual teammates increase task commitment and performance through perceived group cohesion and shared effort [106]. These digital interventions aim to replicate the motivation and performance associated with social interaction and offer novel strategies to support sustained engagement in exercise.

However, a crucial component of social facilitation, the audience effect, has received minimal attention in the context of immersive technologies. In real-world settings, audience presence is typically categorized into two types [142]: passive audiences, whose mere observation can improve running performance [144, 159], and active audience members who provide feedback (positive or negative) and deliver visual or auditory stimuli that positively influence athletic performance [141, 146, 59].

A mini-review of passive audiences in virtual environments by Sterna et al. [139] highlights the limited and inconsistent findings in this area. Only one study reported both social facilitation for simple tasks and social inhibition for complex ones [119]. Additionally, three studies have reported that the mere presence of a virtual audience causes social inhibition during cognitively demanding or complex tasks [68, 167, 57]. Other studies have observed social facilitation effects; however, they involved interactions between the audience and participants, hindering isolation of the impact of mere presence [118, 106, 45]. Several studies found no significant effects of virtual audience presence [143, 64, 8, 41, 42]. One possible reason for these inconsistent findings is that the physical presence of the experimenter may have confounded the results, obscuring whether the participants' performance was influenced by real-world social presence or by the virtual manipulation [139]. As Bond and Titus [14] stated, the presence of others within a participant's visual field is not a necessary condition for audience effects to occur. Therefore, it is essential to control unintended real-world social influences when investigating audience effects in virtual settings. Although the above findings suggest that the mere presence of virtual observers may not reliably produce social facilitation or inhibition effects, this conclusion stems primarily from studies of cognitive task contexts. In physical activity settings where tasks are more automatic, the impact of virtual passive audiences may differ and remain underexplored. According to both Activation Theory [166] and Evaluation Apprehen-

sion Theory [30], the presence of others or fear of being evaluated can increase arousal, thereby facilitating performance in simple tasks. However, whether this arousal leads to a positive or negative experience depends on the individual characteristics. For example, individuals with high social anxiety may experience reduced intrinsic motivation. Individual differences and perceived pressures are critical factors that should not be overlooked in social facilitation studies. In Experiment 1, we accounted for these two factors to investigate how passive virtual audiences affect individuals' physical performance and subjective experience during self-regulated treadmill running in comparison with the presence of real spectators (in-person and via video-call). Additionally, to ensure safety during physical activity and avoid confounding effects of the experimenter's physical presence, we presented the virtual audience on large screens positioned close to each other.

Virtually active audiences, particularly those providing encouragement or applause, have consistently demonstrated positive effects on performances. For example, in a study by Xu et al. [160], gesture-based VR exergames featuring virtual audiences (providing both positive and negative feedback through cheering and clapping sounds) led to improved players' overall gameplay performance and experience compared with passive audiences. Additionally, synchronized virtual crowd applause with rhythmic beats in a VR cycling game significantly increased speed and heart rate during workouts, although it did not affect the subjective experience [59]. These findings underscore the synergistic effect of audience behavior and auditory cues in enhancing performance. The independent effects of auditory cues have been extensively explored; for instance, tempo has been shown to significantly enhance pacing, physical effort, and overall psychological experience during exercise [15, 40, 140, 60]. However, the individual effects of these visual stimuli remain largely unknown. In addition to audience behavior, the appearance of the audience, especially when emotional ties are involved, has been identified as an important factor that influences performance [86, 142, 41]. Previous studies have demonstrated that familiar audiences in VR can affect the speakers' anxiety levels during presentations [101, 10]. However, in the context of exercise, it remains unknown whether the presence of familiar-looking spectators has a measurable impact on physical performance [59, 164, 160]. While some empirical evidence supports the benefits of emotionally close audiences [142], other findings suggest that such familiarity may increase pressure, especially when the supporter has higher expectations, potentially leading to performance impairment because of choking under pressure [91, 48, 18]. This evidence suggests that emotional cues from familiar audiences may have inconsistent effects on performance. Therefore, in Experiment 2, we focused on visual stimuli and used an AR character resembling a close friend to examine whether visual encouragement from a familiar figure enhances performance. Prior research has shown that emotional engagement with anthropomorphic characters can evoke feelings of warmth and increase users' senses of realism and presence [163, 88]. The perceived realism of virtual spec-

tators has been identified as a key factor influencing individual performance [139, 22, 143]. Therefore, the preferred character chosen by the participants was used as the second condition, allowing a comparison with a neutral virtual character to evaluate the impact of character likability on performance. This experiment examined how identity-based personalization of virtual spectators, combined with visual encouragement, influenced users' motivation, exercise experience, and perceived social support.

Two experiments were conducted to advance theoretical and practical understanding of audience-based social facilitation in immersive exercise environments. By integrating both passive and active audience forms and considering individual psychological traits and audience-user relationships, this study offers a more nuanced perspective on how virtual spectators can support and enhance self-regulated exercise.

This study was approved (Approved No. 05-005) by the Life Science Committee of the Japan Advanced Institute of Science and Technology on June 20, 2023. All procedures were performed in compliance with relevant laws and institutional guidelines. The privacy rights of human subjects have been respected, and informed consent was obtained for experimentation with human subjects.

4.2.2 Experiment 1: impact of audience presence on treadmill running experiences

To investigate how virtual passive audiences affect the experience of self-regulated treadmill running, we conducted an experiment using a within-subjects design, where each participant completed four experimental conditions in a randomized order. These conditions, ranked by descending order of social presence, were live spectators (Live), video-call spectators (Video), AR spectators (AR), and no spectators (Absence). The AR character was generated using the ReadyPlayer Me software (https://readyplayer.me/) based on a photograph of the experimenter, as shown in Fig 4.2. It was animated in Unity (version 2022.3.22f1) to simulate natural standing body movements, such as alternating the supporting foot. The AR scene was developed in Unity, built using the Universal Windows Platform (UWP), and deployed via Visual Studio. The final application was executed on a Surface Pro Windows. The experimenter acted as a spectator in both the Live and Video conditions.

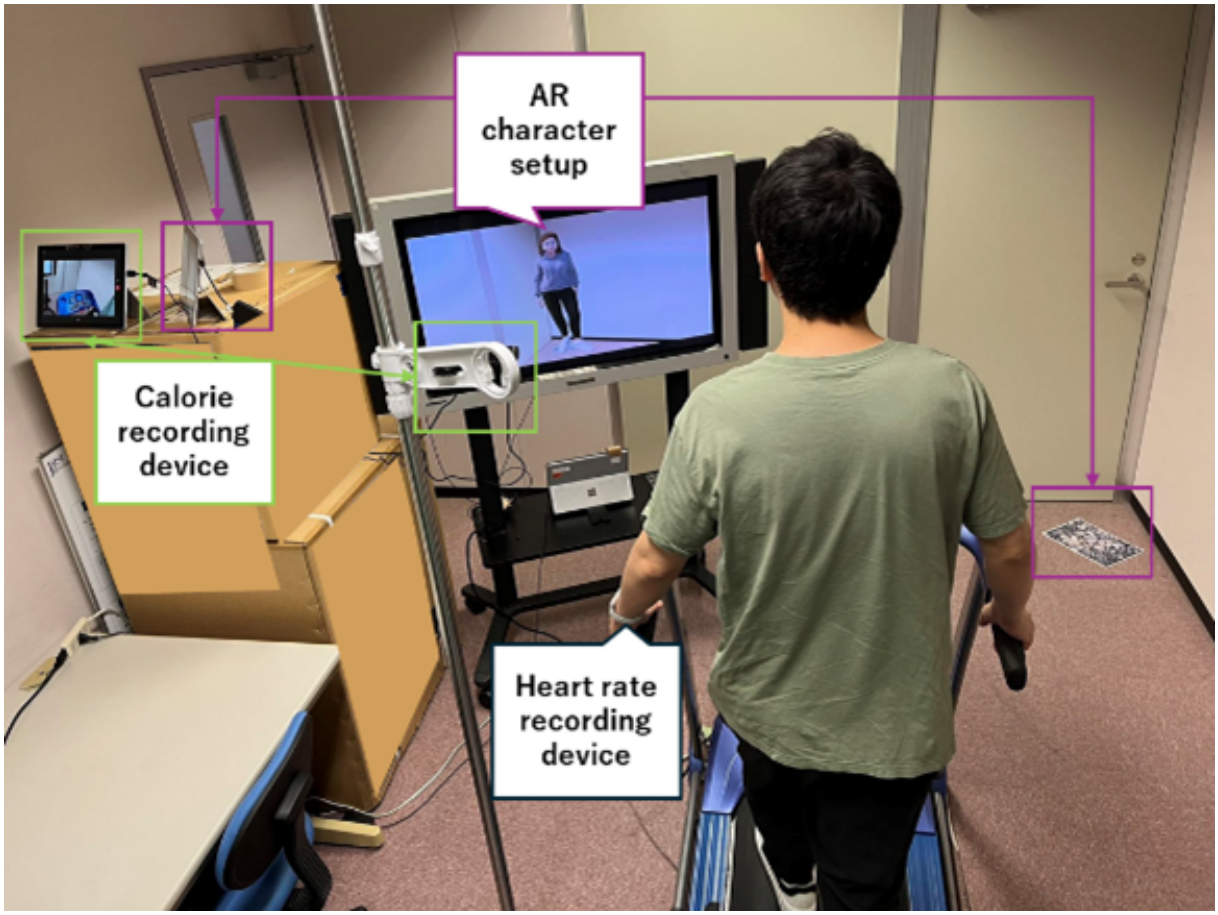


Figure 4.2: **Experimental room layout.** Displays the monitoring devices and the AR character's position in the room and on-screen.

Measures

The following measures were used to assess both subjective experiences and objective performance.

The State-Trait Anxiety Inventory (STAI) [137] consists of two scales: the State Anxiety Scale (STAI-S) and Trait Anxiety Scale (STAI-T). The STAI-S assesses an individual's feelings and emotions at present, whereas the STAI-T measures both general and enduring anxiety tendencies. For our experiment, we utilized the STAI-T (Form JYZ), which comprises 20 items (e.g., "I feel pleasant") rated on a 4-point Likert scale. The total trait anxiety score ranges from 20 to 80, with lower scores (20–44) indicating low anxiety, moderate scores (45–54) indicating moderate anxiety, and higher scores (55–80) indicating high anxiety. In this experiment, the STAI-T demonstrated excellent internal consistency, with Cronbach's α = 0.90. Participants' scores ranged from 26 to 66 (M = 44, SD = 9.49).

The physical burden of the running task was assessed using the Rating of Perceived Exertion (RPE) scale [17], which ranged from 6 (no exertion) to 20 (maximal exertion).

Intrinsic motivation was evaluated using the Intrinsic Motivation Inventory (IMI) [130], which is widely used in sports and exergaming research [29, 97, 168]. Considering the nature of our experimental conditions, we selected the pressure/tension subscale of the IMI, which serves as a negative predictor of intrinsic motivation [60]. Participants rated five items (e.g., "I did not feel nervous at all while doing this") on a 7-point Likert scale ranging from 1 (not at all true) to 7 (very true). This subscale exhibited good internal consistency, with Cronbach's α = 0.80.

To compare attentional focus across different experimental conditions, we employed the Focus of Attention Questionnaire (FAQ) [158], which includes two five-item subscales: self-focus (e.g., "I was focusing on what I thought of the other person"; Cronbach's α = 0.57) and external-focus (e.g., "I was focusing on the other person's appearance or dress"; Cronbach's α = 0.74). These subscales measure attention directed toward internal bodily reactions and external environmental cues. Participants responded using a 5-point scale ranging from 1 (not at all) to 5 (in total), selecting the option that best represented their experience during the running task.

The participants' physical performance was assessed using a treadmill's built-in calorie consumption display. In addition, real-time heart rate data were collected during the running task using an Apple Watch SE (2nd generation) paired with the HypeRate API (https://www.hyperate.io/). Fig 4.2 illustrates the layout of the room used in the experiment.

Procedure

Recruitment emails were sent to students inviting them to participate in the treadmill exercise experiment. The email included information, such as the experiment's duration, monetary compensation, a requirement that participants be healthy adults, and a note indicating that the experiment may involve a considerable physical demand.

A total of 24 students (6 females and 18 males, aged 22–33 years) expressed their willingness to participate during the designated recruitment period (May 7–27, 2024). On the day of the experiment, upon arrival at the room, participants were provided with a verbal explanation of the study's purpose, procedures, data handling (including anonymization), and compensation based on written materials approved by the Life Science Committee. Furthermore, it was emphasized that the participants could discontinue the experiment at any time in case of physical discomfort without any negative consequences.

The procedures for obtaining the informed consent form and its withdrawal documents are

explained. After confirming that they had no questions, the participants voluntarily signed a consent form and retained the withdrawal document. In addition, they completed two surveys to assess exercise frequency and trait anxiety (STAI-T). Under the supervision of the experimenter, the participants were equipped with an Apple Watch, and a heart rate monitoring application was initiated. They completed a 5-minute pre-running stretching routine by following a guided video. The experimenter instructed participants on treadmill use and allowed sufficient time for familiarization and adaptation. Once the participants began the task, the experimenter left the room. The task consisted of a 1-minute warm-up at 3.5 km/h, followed by 7 min of running at a self-selected pace. After the run, participants stopped the treadmill and performed a small post-run stretching routine. The experimenter returned to the room and instructed the participants to complete the RPE, FAQ, and IMI. Following a rest period of at least 10 minutes, the participants repeated the Running Task under different conditions and completed the same questionnaires. The entire session lasted approximately one hour.

The procedure varied slightly depending on the spectator's condition. In the absence condition, the experimenter left the room immediately after the participants began running. In the AR spectator condition, the experimenter launched the AR application in advance on a Surface Pro device and projected the AR character onto a large screen before leaving the room. In the video spectator condition, a pre-opened video call interface was projected onto a large screen. The experimenter left the room and appeared on the screen when the participant began to run. In the live spectator condition, the experimenter stood in front of a large screen and observed the participant directly throughout the running task, occasionally shifting the body weight between the feet to simulate the movement pattern of the AR character. For safety monitoring, the experimenter monitored the participants' safety during the AR spectator and absence conditions using a concealed camera (unknown to the participants) and real-time heart rate data. Under live and video spectator conditions, the experimenter ensured the participants' safety by directly observing their running status.

Across all the spectator conditions, the experimenter or virtual representation maintained a neutral expression without overt emotional reactions to standardize the spectator role.

After at least two days, the participants returned for a second session, during which the remaining experimental conditions were tested using the same procedure. Each participant received a compensation of 2,000 JPY for their participation.

Results

We performed linear mixed-effects modeling on the data from the self-regulated running phase using the General Analyses for Linear Models module (GAMLj3, version 3.5.1) within the

Jamovi software (version 2.6.44; https://www.jamovi.org/).

Linear mixed-effects models extend the general linear model by incorporating fixed and random effects. Contrarily to traditional repeated-measures ANOVA, these models can handle hierarchical data structures and are sufficiently flexible to accommodate unbalanced data, missing values, and irregular time intervals [100]. This flexibility allows researchers to model individual differences effectively and derive accurate estimates of population-level effects [6, 81, 83]. However, the flexibility of linear mixed-effects models increases the risk of overfitting [100].

To minimize this risk, we conducted the following procedure to select a suitable model. First, we developed a null baseline model containing a fixed intercept and random participant effects. Next, we added different combinations of fixed effects (condition, STAI-T) to the baseline model and evaluated each using the maximum likelihood (ML) estimation. By comparing these models, we determined the optimal set of fixed effects, which were included in the final model. Subsequently, we re-estimated the model parameters using restricted maximum likelihood (REML) to obtain the final fixed-effect coefficients. Table 4.1 summarizes the results for each dependent variable in the final model, including the fixed coefficients that showed significant differences or trends between the two groups ($p < 0.1$).

An a priori power analysis was conducted to determine the required sample size for detecting audience effects. Prior literature [84] and subsequent meta-analyses [14] have shown that such effects account for only 1–3% of the variance in motor performance. A similar study [128] reported a small audience effect with an effect size of Cohen's $d = 0.29$, which was used as the basis for our estimate. In a within-subjects experimental design with 4 conditions, $\alpha = 0.05$, and a desired power of 0.90, the required sample size was calculated to be 14 participants.

Table 4.1: **Significant fixed coefficients for the measures.**

				95% Confidence Interval				
Name	Effect	Estimate	SE	Lower	Upper	df	t	p
IMI_Pressure	Live - AR	0.81	0.26	0.28	1.33	66.0	3.07	.003**
	Live - Absence	0.96	0.26	0.43	1.48	66.0	3.64	.001**
	Video - AR	0.53	0.26	0.00	1.05	66.0	2.00	.050*
	Video - Absence	0.68	0.26	0.15	1.20	66.0	2.56	.013*
External_Focus	Live - AR	2.08	1.00	0.10	4.07	69.0	2.09	.041*
	Live - Absence	3.04	1.00	1.06	5.03	69.0	3.05	.003**
	Video - AR	1.88	1.00	-0.11	3.86	69.0	1.88	.065
	Video - Absence	2.83	1.00	0.85	4.82	69.0	2.84	.006**
Heart_Rate	Live - AR	6.86	8.54	1.80	11.92	62.9	2.70	.009*
	Live - Absence	7.48	2.50	2.51	12.46	62.8	2.99	.004**
	Video - AR	5.53	2.55	0.46	10.61	62.3	2.17	.034*
	Video - Absence	6.16	2.51	1.16	11.15	62.1	2.45	.017*
Calorie_consumption	Live - AR	1.84	0.99	-0.12	3.80	67.1	1.87	.066
	Live - Absence	2.53	1.00	0.54	4.52	67.1	2.53	.014*
	Video - Absence	1.69	0.99	-0.27	3.65	67.1	1.72	.091

*$p \leq 0.05$, **$p < 0.01$

Perceived exertion - RPE The experimental condition had no significant effect on the participants' perceived exertion during the running task, as indicated by similar RPE scores across groups: AR (M = 11.5, SD = 2.23), live (M = 12.1, SD = 2.39), video (M = 11.6, SD = 2.76), and absence (M = 11.4, SD = 2.12). However, the STAI-T showed a marginal effect on RPE scores (F = 4.17, p = .053), suggesting a trend in which higher trait anxiety was associated with greater perceived exertion.

Intrinsic motivation - IMI Pressure/Tension Higher IMI Pressure/Tension scores indicate greater subjective pressure. The experimental condition significantly affected the participants' perceived pressure during the self-controlled running task (F = 5.78, p = .001). As shown in Table 4.1, the participants reported significantly higher perceived pressure in the live and video conditions than in the AR and absence conditions. The STAI-T had a significant main effect on pressure/tension scores (F = 9.72, p = .005), indicating that individuals with higher trait anxiety reported experiencing greater pressure. Additionally, there was a marginal interaction between condition anxiety and trait anxiety (F = 2.58, p = .061). As illustrated in Fig 4.3, the perceived pressure of participants in the live condition was more strongly influenced by their levels of trait anxiety than those in the AR condition (b = 0.08, p = .009), with a similar trend observed in the video condition than in the AR condition(b = 0.05, p = .059).

Attention - external focus The experimental condition significantly influenced the participants' External Focus in the self-controlled running task (F = 4.37, p = .007). As shown in Table 4.1, the participants reported significantly higher scores for external focus in the live condition than in the AR and absence conditions. The participants in the video condition exhibited a trend toward greater external focus than in the AR condition and significantly more focus than in the absence condition. The STAI-T showed a significant main effect (F = 6.26, p = .020), suggesting that individuals with higher trait anxiety tended to focus more on external factors, such as the presence and appearance of others, during the task.

Physiological index - mean heart rate The experimental condition significantly affected participants' heart rate during the self-controlled running task (F = 4.48, p = .007). As shown in Table 4.1, the participants exhibited significantly higher heart rates during the self-paced running phase under the live and video conditions than under the AR and absence conditions.

Performance index - calorie consumption The experimental condition had a marginal effect on participants' calorie consumption during the self-controlled running task (F = 2.49, p = .067). As shown in Table 4.1, the participants burned significantly more calories in the live

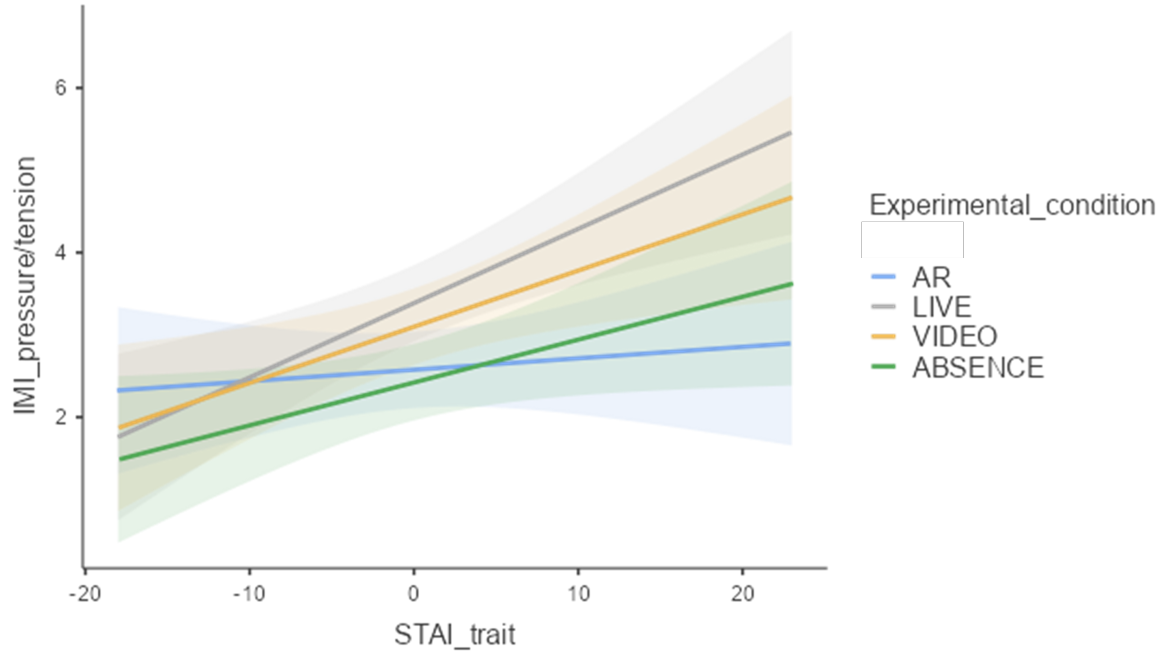


Figure 4.3: **Interaction between audience condition and trait anxiety on perceived pressure.** Perceived pressure varies with trait anxiety across audience conditions: AR, live, video-call, and no spectators. IMI: Intrinsic Motivation Inventory; STAI: State-Trait Anxiety Inventory.

condition than in the absence condition (p = .014). There was a trend toward higher calorie consumption in the live condition than in the AR condition (p = .066), and in the video condition than in the Absence condition (p = .091).

Correlation We conducted a Spearman correlation analysis using IBM SPSS Statistics 27 to examine the pairwise relationships among the measures. The results (see Table 4.2) showed that RPE, IMI_pressure, External_focus, and calorie consumption were positively correlated with one another.

Table 4.2: **Correlations between measures.**

	RPE	IMI_pressure	External_focus	Heart_rate	Calorie
RPE	1	.269**	.206*	-.015	.280**
IMI_pressure	.269**	1	.378**	-.073	.262*
External_focus	.206*	.378**	1	.144	.087
Heart_rate	-.015	-.073	.144	1	.141
Calorie	.280**	.262*	.087	.141	1

** Correlation is significant at the 0.01 level: 2-tailed
* Correlation is significant at the 0.05 level: 2-tailed

4.2.3 Experiment 2: impact of different types of supportive AR characters

To investigate how familiarity or likability of virtual characters influences the effectiveness of their encouragement, we experimented with a within-subject design. Each participant experienced four conditions: encouragement from an AR character resembling a good friend (FC), encouragement from a liked AR character (LC), encouragement from a neutral AR character (DC), and a control condition with no encouragement (NC).

Each participant used Ready Player Me software to create an AR character resembling the appearance of a close friend for use in the FC condition. The participants were instructed to upload a photo or use a random character and adjust their body shape, appearance details, and hairstyle or color to closely resemble their good friends. For the LC condition, we selected five different styles of female characters from a Unity Asset Store. Each participant chose their favorite character from these options, which was used as a virtual character in the LC experimental condition. The neutral AR character was a generic female model from Mixamo software (https://www.mixamo.com/) without facial details, as shown in Fig 4.4. The encouragement animation was created by combining the waving animation and cheering animation from Mixamo and was applied uniformly across all characters. The configuration of the AR application followed the same procedure as in the previous experiment.

(a) Waving (b) Cheering

Figure 4.4: **Encouragement animations under the DC (disliked or natural AR character) condition.** A generic female performs (a) waving and (b) cheering animations.

To investigate the effect of a supportive AR character, we designed a running task, as shown in Fig 4.5, which was a prerecorded video composed of different segments that changed over time. Based on preliminary trials and insights from our previous experiment, we set the warm-up speed at 4 km/h and established the following maximum speed guidelines: For participants who regularly engaged in exercise, the maximum speed was set at 7 km/h for women and 8

km/h for men. With a lower exercise frequency, the speed was limited to 6 km/h in women and 7 km/h in men. A 3-minute walking segment at 3 km/h was included after a period of higher-intensity running, with the dual purpose of facilitating cardiovascular recovery and enhancing the subjective contrast between exertion and rest. This contrast was intended to amplify the perceived sense of fatigue and increase the psychological urge to stop, which effectively simulated real-world scenarios in which individuals must decide whether to continue exercising despite fatigue. In this experiment, the 3-minute self-paced phase began immediately after the participants viewed a 12-second AR encouragement animation. The visual stimuli depicted in Fig 4.5 show one of the AR friend characters used in this experiment. As shown in Fig 4.6, the video content was displayed on a monitor positioned in front of the participants while running.

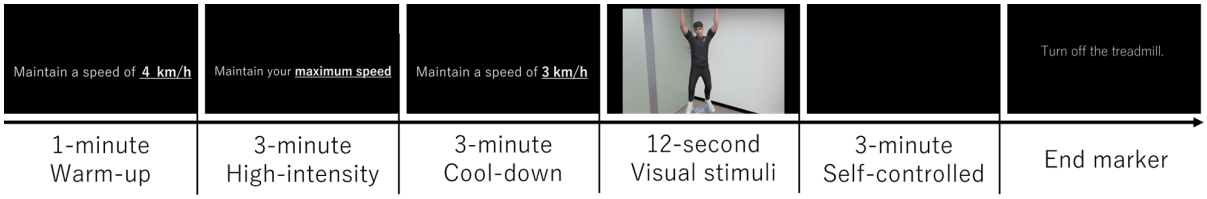


Figure 4.5: **Running task procedure in Experiment 2.** Sequences of screen content presented during the running task.

Procedure

The recruitment procedures conducted from September 17 to September 30, 2024, and from November 11 to November 17, 2024, were identical to those used in previous experiments. Twenty-one subjects (11 females and 10 males, aged 21-32 years) participated in this experiment.

Fig 4.7 illustrates the experimental procedure. Consistent with the previous experiment, the participants provided written informed consent following a detailed explanation of the study procedure. Subsequently, participants completed questionnaires to assess exercise frequency and trait anxiety (STAI-T). They were informed that they could stop the experiment at any point if they experienced discomfort, without any negative consequences. The participants wore an Apple Watch to begin heart rate monitoring. Next, they completed a questionnaire ranking the likability of various characters. Under the experimenter's guidance, the participants created a character that resembled a close friend. Subsequently, they completed a questionnaire evaluating the similarity of the character's appearance and ambiance with that of their friends. The experimenter then introduced the running task, explaining the treadmill speed display, the method for determining maximum speed, and the purpose of the all-black screen phase. Participants were informed that the suggested maximum speed was only a guideline and they were

Figure 4.6: **Experimental room layout in Experiment 2.** Shows the position of the large screen within the room and the way content appears on the screen.

encouraged to reduce their speed if it became too strenuous. The participants performed pre-run stretching exercises and familiarized themselves with the treadmill. When the participants were ready, the experimenter began recording the treadmill data, started the running task video, and exited the room. During the running task, the experimenter used a hidden camera and monitored the participants' heart rate to ensure their safety. After the treadmill was stopped, the experimenter returned to the room and administered the RPE scale, IMI-Pressure/Tension subscale, and a custom questionnaire evaluating the participants' feelings toward supportive characters. After a rest period of at least 10 minutes, the participants completed task B under different conditions and answered the same questionnaires. The entire session lasted approximately one hour.

Participants returned for a second session after two days. After a 5-minute resting heart rate measurement, they proceeded with the same procedure for the remaining experimental conditions. Each participant received a compensation of 2,000 yen for their participation. The LC, DC, and NC conditions were presented in randomized order across Tasks A, B, and C or D, as shown in Fig 4.7. Because the characters in the FC condition were created on the day of the experiment and recording the AR as well as producing the video required additional time,

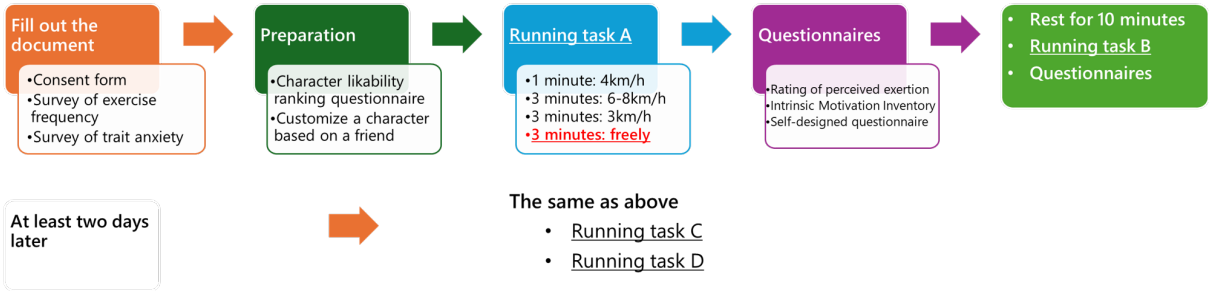


Figure 4.7: **Experimental procedure.** Outlines the participant flow after entering the room, including questionnaires and the running task.

the FC condition was included only in the second session and was randomly assigned to either Task C or D.

Measures

The Karvonen formula, also known as the heart rate reserve (HRR) method, was used to calculate exercise intensity [148]. In line with the current guidelines for healthy populations, the HRmax was estimated using the following age-based formula: 220 – age [90, 47, 112]. Exercise intensity was calculated as: (HR – resting HR) / HRR, where HRR was defined as HR_{max} – resting HR. The participants' heart rates were continuously monitored in real-time during the experiment using an Apple Watch.

Perceived exertion was measured using the RPE scale; trait anxiety was assessed using the STAI-T (participants' scores ranged from 25 to 68; M = 44, SD = 11.4); and intrinsic motivation was evaluated using the pressure/tension subscale of the IMI. Exercise performance was quantified using the treadmill's built-in calorie consumption feature. These measures were identical to those used in the previous experiment.

Results

We analyzed the data from the three-minute free-running phase using the same method as in the previous experiment. Table 4.3 summarizes the results for each dependent variable under the best-fitting model, including the fixed coefficients that showed significant differences or trends between the two groups ($p < 0.1$).

RPE The experimental condition had no significant effect on the participants' perceived exertion during self-paced running, as indicated by the similar RPE scores across the groups: FC

Table 4.3: **Significant fixed coefficients for the measures in Experiment 2.**

				95% Confidence Interval				
Name	Effect	Estimate	SE	Lower	Upper	df	t	p
IMI_Pressure	FC-LC	-0.47	0.19	-0.84	-0.10	60.0	-2.50	.015*
	FC-DC	-0.50	0.19	-0.88	-0.13	60.0	-2.70	.009*
Exercise_Intensity	FC-NC	4.01	2.03	-0.03	8.05	53.0	1.98	.053
	LC-NC	5.38	2.03	1.34	9.42	53.0	2.66	.010*
	LC-DC	3.43	2.03	-0.61	7.47	53.0	1.70	.096
Calorie_consumption	FC-NC	1.86	0.53	0.81	2.90	55.9	3.54	.001**
	FC-DC	1.20	0.53	0.15	2.25	55.9	2.29	.026*
	LC-NC	1.66	0.54	0.59	2.73	56.0	3.09	.003**
	LC-DC	1.00	0.53	-0.07	2.07	56.0	1.87	.067

*$p \leq 0.05$, **$p < 0.01$

(M = 11.1, SD = 1.71), LC (M = 11.3, SD = 2.15), DC (M = 11.6, SD = 2.20), and NC (M = 11.4, SD = 1.75).

IMI Pressure/Tension The experimental condition significantly affected the participants' perceived pressure during self-paced running (F = 3.06, p = .035). As shown in Table 4.3, the participants reported significantly lower perceived pressure in the FC condition than in the LC (p = .015) and DC (p = .009) conditions.

Exercise intensity The experimental conditions significantly affected exercise intensity during self-paced running (F = 2.68, p = .056). As shown in Table 4.3, the participants exhibited significantly higher levels of exercise intensity in the LC condition than in the NC condition (p = .010). Additionally, a similar trend was observed when comparing the LC and DC conditions (p = .096), and a marginally significant difference was observed between the FC and NC conditions (p = .053). There was a significant interaction between condition and trait anxiety (F = 7.32, p = .001). As illustrated in Fig 4.8, the exercise intensity of participants in the NC condition was more strongly affected by variations in trait anxiety than those in the FC (b = 0.85, p = .001), LC (b = 0.56, p = .006), and DC (b = 0.63, p = .001) conditions.

Calorie consumption The experimental conditions significantly affected the participants' calorie consumption during self-paced running (F = 5.44, p = .002). As shown in Table 4.3, the participants burned significantly more calories in the FC condition than in the NC (p = .001) and DC (p = .026) conditions. Similarly, calorie expenditure was significantly higher in the LC condition than in the NC condition (p = .003), with a trend toward higher calorie consumption than in the DC condition (p = .067). There was a significant interaction between condition and

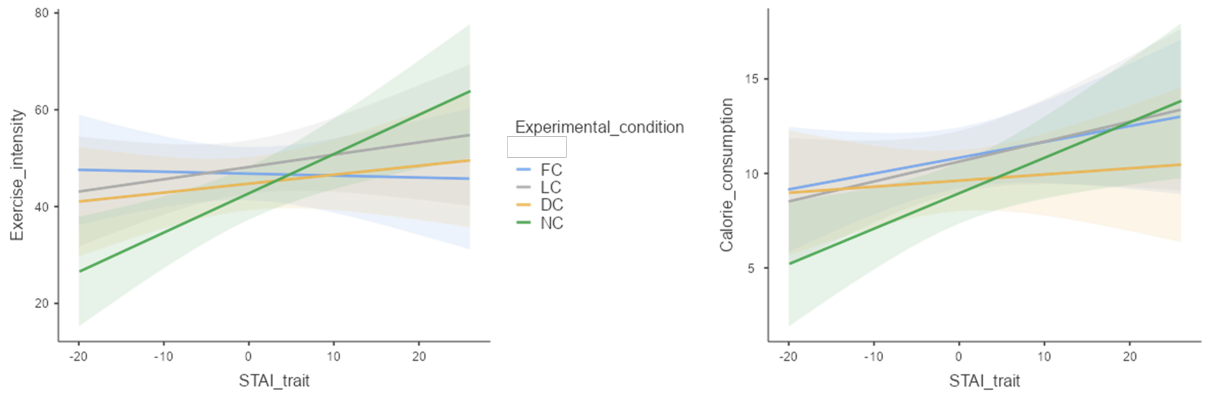


Figure 4.8: **Interaction between audience condition and trait anxiety on exercise intensity and calorie consumption.** Changes in exercise intensity (left) and calorie consumption (right) with increasing trait anxiety under four audience conditions: FC (AR character resembling a good friend), LC (liked AR character), DC (disliked AR character), and NC (no encouragement).

trait anxiety (F = 3.77, p = .015). As illustrated in Fig 4.8, the calories burned by participants in the NC condition were more strongly influenced by their levels of trait anxiety than those in the FC (b = 0.10, p = .031) and DC conditions (b = 0.16, p = .002).

Designated questionnaire To explore participants' feelings about supportive characteristics, we administered a self-designed questionnaire. The items are listed in Table 4.4. Item 3 used a scale ranging from 1 (very negative) to 7 (very positive), whereas the other items used a scale ranging from 1 (not at all) to 7 (very much). As shown in Table 4.4, participants reported lower familiarity and attractiveness of the character in the DC condition than in the FC and LC conditions. In addition, they perceived that characters in the DC condition elicited higher levels of anxiety and negative emotions. There were no significant differences in participants' feelings about the characters between the LC and FC conditions, although small trends were observed in terms of familiarity and attractiveness.

Table 4.4: **Differences in self-designed questionnaire items between conditions.**

	FC-DC		LC-DC		FC-LC	
	Estimate	p	Estimate	p	Estimate	p
1. How familiar do you feel with this character	2.57	.001**	1.67	.001**	0.91	.057
2. How attractive do you find this character	2.52	.001**	1.86	.001**	0.67	.092
3. What emotions do you have towards this character	1.90	.001**	1.81	.001**	0.10	.777
4. How natural do you find the character's movements	0.91	.052	1.33	.005**	-0.43	.349
5. Do the character's movements provoke excitement	1.19	.015*	0.81	.092	0.38	.422
6. Do the character's movements evoke feelings of anxiety	-1.48	.002**	-1.57	.001**	0.10	.829

*$p \leq 0.05$, **$p < 0.01$.

Table 4.5: **Correlations between measures in Experiment 2.**

	RPE	IMI_pressure	Exercise_intensity	Calorie
RPE	1	.222*	.020	-.065
IMI_pressure	.222*	1	.000	-.153
Exercise_intensity	.020	.000	1	.643**
Calorie	-.065	-.153	.643**	1

** Correlation is significant at the 0.01 level: 2-tailed
* Correlation is significant at the 0.05 level: 2-tailed

4.2.4 Discussion

This study investigated the effects of passive and active virtual audiences on self-regulated physical activity, focusing on perceived exertion, psychological experience, and physiological performance. Across two experiments, we found that while the presence of a virtual audience did not significantly influence perceived physical exertion, it did affect participants' psychological pressure and exercise outcomes—particularly when audience characteristics (appearance) and individual differences (trait anxiety), were considered.

Audience modality and social presence Our first experiment confirmed that real spectators significantly increased both psychological pressure and calorie consumption, consistent with the Yerkes–Dodson law [161] and classical findings that social presence enhances performance on effort-based tasks [14]. This enhancement appears to be driven by heightened evaluation apprehension [30] or attentional conflict [9]. The finding that video-call spectators elicited levels of perceived pressure and attentional focus comparable to those induced by in-person audiences suggests that mediated communication can effectively convey social presence, consistent with the social presence theory proposed by Short et al. [132], which emphasizes the role of perceived interpersonal immediacy in mediated interactions.

In contrast, the mere presence of AR virtual characters did not lead to significant changes in either perceived pressure or performance. One possible explanation lies in the direction of participants' attention: participants paid significantly less attention to the AR characters than to real spectators. This may be because virtual characters are unable to provide immediate feedback or social evaluation, thereby reducing their perceived presence and authenticity. Consequently, the lower sense of being evaluated could explain the reduced pressure experienced under virtual spectators compared to real ones. This finding is consistent with previous studies showing that the presence of virtual agents does not necessarily enhance task performance [42, 160].

Moreover, in the presence of virtual spectators, pressure levels were largely unaffected by

participants' trait anxiety. While individuals with higher trait anxiety reliably reported greater pressure when facing real audiences, a similar trend was observed even in the no-audience condition. This may be attributed to their heightened sensitivity to potential monitoring, given their awareness of being part of a lab-based experiment. As suggested by the Hawthorne effect [98], participants may feel observed regardless of the experimental condition [42]. Notably, the visible but clearly non-evaluative nature of AR characters may have buffered this imagined pressure, offering a psychologically safe alternative to real audiences for anxious exercisers.

Audience familiarity and visual encouragement Experiment 2 extended previous research by examining how audience identity and visual feedback influence psychological pressure and performance outcomes. Overall, encouragement from virtual characters—regardless of type—did not significantly increase perceived pressure. This may be due to the relatively simple and low-stakes nature of the treadmill running task, which likely failed to trigger strong evaluative concerns. However, significant differences emerged across character types. Participants exposed to familiar, friend-resembling virtual characters reported significantly lower pressure compared to those exposed to merely liked characters. This suggests that the customization process not only enhanced character realism but also increased psychological relevance, fostering a stronger emotional connection. In contrast, static "liked" characters lacked the personal context necessary to elicit similar effects. These findings highlight the importance of identity-based personalization in enhancing the motivational impact of virtual agents. Rather than visual appeal alone, it is the emotional resonance and personal relevance of the character that shape users' interpretations of and responses to virtual encouragement. In our custom questionnaire, responses to the familiarity item revealed a near-significant difference between the two conditions, suggesting subtle but meaningful variations in emotional connection. This aligns with prior research showing that user-customized agents can influence users' emotional states [157, 3].

Importantly, both conditions enhanced performance, indicating that high familiarity or emotional affinity with a virtual audience can significantly boost task outcomes. This supports previous findings that both the perceived relationship between the individual and the audience [142, 160, 59, 164] and audience characteristics such as attractiveness [60, 41] play key roles in modulating performance. The lack of a performance benefit from a neutral virtual character reinforces that encouragement alone is not universally effective—it depends on who delivers it and how personally meaningful they are perceived to be. While many previous studies have discussed the theoretical potential of personalized characters [3, 95, 80, 136], few have empirically demonstrated their practical benefits. Our findings offer empirical support for their effectiveness, aligning with prior evidence that emotionally meaningful digital companions can enhance task engagement [49].

Finally, the interaction between audience condition and trait anxiety revealed that, across all virtual audience types, performance was less affected by participants' trait anxiety compared to the no-audience condition. This suggests that virtual audiences may buffer the negative impact of individual anxiety, providing a psychologically safer and more supportive exercise environment.

Collectively, our findings highlight a nuanced tradeoff between performance enhancement and psychological safety. While real audiences or competitive settings can boost effort, they also risk inducing pressure, particularly for socially anxious individuals. Passive AR spectators, though less effective at enhancing performance, may offer a low-pressure social presence more conducive to long-term motivation. Meanwhile, active and emotionally resonant virtual characters—especially those perceived as close companions—appear capable of enhancing both performance and subjective experience without inducing stress.

Theoretical and practical implications These findings offer several key implications for the design and application of virtual audience systems in exercise settings:

- Emotionally Safer Alternatives to Real Audiences: Passive or semi-interactive virtual audiences can offer a less threatening form of social presence compared to real spectators. This moderated presence helped buffer the negative effects of trait anxiety while still providing motivational benefits. Such virtual audiences are particularly well-suited for individuals who may be sensitive to social evaluation, offering a psychologically safer environment that supports engagement without inducing performance pressure.
- Importance of Audience Appearance and Emotional Affinity: The appearance and identity of virtual agents significantly affect their behavioral impact. Characters resembling familiar or personally meaningful figures are more effective in enhancing user performance. These findings highlight the role of emotional resonance and familiarity in shaping users' responses to virtual encouragement.
- Advantages of Customization: User-involved customization processes—not just character aesthetics—are central to maximizing the motivational potential of virtual agents. Personalized agents foster greater emotional connection and perceived realism, leading to improved task engagement and performance. Systems that allow users to shape or co-create their digital companions may yield long-term adherence benefits in fitness and wellness applications.

Limitations and future directions This study has several limitations that should be acknowledged. First, to ensure participant safety during treadmill running, virtual characters were displayed on a standard computer monitor rather than through a head-mounted display (HMD). While this setup was practical, it may have reduced the sense of immersion and social presence. Prior studies suggest that social facilitation or inhibition effects are often amplified in immersive HMD environments compared to non-immersive displays [164, 41].

Second, the characteristics of spectators—such as their number, visual size, and gender—are known to influence the experience of performers [160, 41, 57]. In our study, the virtual audience was represented by a single character, whose on-screen presence was considerably smaller than that of a real human observer. This difference may have attenuated the perceived salience of the virtual audience. Moreover, while prior research has examined the impact of audience quantity, valence, and interactivity, these variables are often intertwined, making it difficult to isolate the effect of each factor [160].

Third, gender dynamics may have played an unintentional role in influencing outcomes. For instance, the presence of opposite-gender spectators has been associated with enhanced social facilitation effects in previous work [39]. In Experiment 1, the real and AR spectators were female, whereas in Experiment 2, the gender of the virtual characters in the friend-like condition was not standardized or matched to the participant. This lack of gender control may have introduced additional variability.

Finally, this study focused on short-term, single-session effects. While the immediate impact of virtual audiences on motivation and performance is encouraging, long-term engagement may depend more on emotional safety, personalization, and sustained intrinsic motivation. Future research should explore the longitudinal effectiveness of virtual audience systems, including potential habituation or novelty effects over time.

Emerging techniques such as adaptive feedback—e.g., dynamically triggered encouragement based on physiological metrics like heart rate—could further optimize both motivational and emotional outcomes. The integration of emotionally intelligent, user-customized virtual agents into AR/VR exercise platforms offers a promising avenue for supporting long-term adherence and behavioral change in physical activity interventions.

4.2.5 Conclusion

This study explores the psychological and performance effects of different types of audience presence during treadmill exercise, focusing on both passive and active virtual audiences. In Experiment 1, real spectators elicited classical social facilitation effects, increasing both per-

ceived pressure and physical performance. In contrast, passive AR spectators failed to produce such effects; however, pressure levels under AR conditions remained stable across different levels of trait anxiety, indicating a less threatening experience for anxiety-prone individuals. Experiment 2 examined the impact of active virtual encouragement. While neutral characters had minimal effect, both liked and friend-resembling characters significantly enhanced performance, with highly customized, familiar characters also reducing perceived pressure.

These findings address the existing gap in research on non-interactive virtual audiences and highlight the motivational value of emotionally resonant, personalized digital agents. Such agents can serve as low-pressure alternatives to real spectators, offering promising design directions for future immersive and adaptive exercise technologies.

4.3 Validation Experiment

To extend these preliminary findings, a validation experiment was conducted to examine whether the observed effects of virtual audiences could be integrated into an adaptive feedback AR exercise system. While Experiments 1 and 2 focused on understanding how different types of audiences (passive vs. active) influence perceived pressure and motivation, the validation experiment aimed to translate these insights into an applied design. Specifically, the study tested whether a virtual companion that adjusts its behavior based on users' physiological feedback (e.g., heart rate) could maintain motivation and comfort during exercise. This final stage served to bridge theoretical understanding and practical implementation, offering guidance for the development of emotionally adaptive virtual audiences in future fitness technologies.

4.3.1 Introduction

Regular physical activity provides substantial physical and psychological benefits, yet sustaining motivation during solitary exercise remains difficult for many people [1, 39, 61, 94, 60]. In the absence of social cues or companionship, self-paced workouts often feel less engaging and more effortful, which can undermine adherence over time [35, 52]. This challenge has motivated growing interest in technologies that introduce supportive forms of social presence into otherwise solitary exercise routines.

Augmented reality (AR) offers a unique opportunity to embed virtual agents directly into users' physical environments. Prior research has shown that virtual coaches, teammates, and companion figures can influence motivation, enjoyment, and perceived support during exercise [108, 89, 106, 107, 60]. In digital settings, emotionally expressive or anthropomorphic char-

acters can evoke warmth, social connectedness, and engagement [163, 88]. However, most existing work has focused on instructional or competitive roles, and relatively little is known about how emotionally expressive AR companions affect users during self-regulated physical activity.

Another key but underexplored factor is avatar appearance. Visual style, realism, and familiarity can strongly shape comfort, trust, and emotional responses [163, 143, 142]. While stylized characters may offer approachability and emotional distance, human-like or familiar avatars can evoke more complex reactions, sometimes increasing connection but also potentially introducing social pressure. Understanding these nuances is essential for designing companions that feel supportive rather than intrusive.

The present study investigates how an emotionally expressive AR companion shapes users' affective experiences, perceptions, and exercise behavior during self-paced treadmill running. Using a mixed-methods approach, we examine how participants interpret different companion designs, how they experience social and emotional aspects of the interaction, and how individual traits such as anxiety may relate to these perceptions.

This research contributes to a growing understanding of how virtual companionship can be designed for physical activity and identifies factors that influence comfort, engagement, and the perceived social presence of AR exercise partners.

This study was approved (Approved No. 05-005) by the Life Science Committee of the Japan Advanced Institute of Science and Technology on June 20, 2023. All procedures were performed in compliance with relevant laws and institutional guidelines. The privacy rights of human subjects have been respected, and informed consent was obtained for experimentation with human subjects. All procedures were performed in compliance with relevant laws and institutional guidelines. The privacy rights of human subjects have been respected, and informed consent was obtained for experimentation with human subjects.

4.3.2 System design

We developed an exercise companion system in which an AR character dynamically responds to the user's real-time exercise intensity through corresponding animations. Based on prior research indicating that 40% intensity is the minimum threshold for improving aerobic capacity [112], this value served as the system's baseline.

Upon launch, the software retrieves the user's resting heart rate and age from the input (in Fig 4.9) and continuously calculates exercise intensity using real-time heart rate data. When intensity rises above 40%, the AR character displays a standing animation, functioning as a

passive observer or presence. When intensity falls below 40%, the character switches to a cheering animation. The character returns to the standing animation only when the user's effort increases sufficiently to reach 50% intensity. The 50% threshold was selected as it represents the moderate-intensity exercise and, following vigorous activity, is considered to impose a moderate physiological load [148].

There are five different types of AR characters that can be selected for the feedback system:

1. a custom-made avatar resembling a close friend,
2. a personally preferred anime character,
3. a human-like avatar of the opposite gender to their friend,
4. a simplified blue skeletal male model, or
5. a simplified pink skeletal female model.

The human-like characters were created using ReadyPlayerMe, appealing anime-style characters were sourced from Unity assets, and skeletal models were obtained from Mixamo. All characters were animated in Unity (version 2022.3.22f1) to simulate natural standing body movements, such as shifting weight between feet. A cheering animation was implemented by animating both arms swinging above the head in a forward–backward motion. The AR scene was developed in Unity, built using the Universal Windows Platform (UWP), and deployed through Visual Studio. The final application was run on a Surface Pro device.

The layout of the experimental room is shown in Fig 4.9 and Fig 4.10.

Figure 4.9: **Experimental room.**

4.3.3 Measures

The study incorporated several measures to evaluate both subjective experiences and objective performance outcomes.

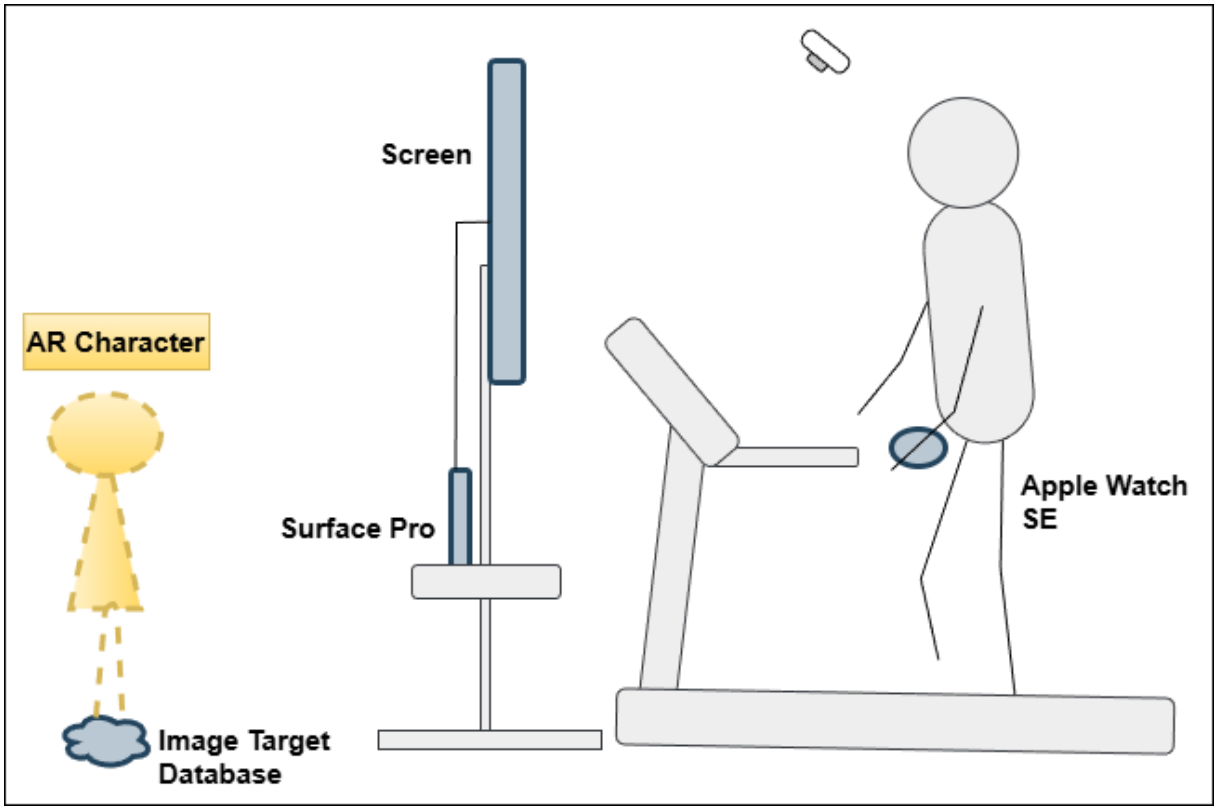


Figure 4.10: **A simple diagram of the experiment room.**

To assess anxiety, we used the State–Trait Anxiety Inventory (STAI; [137]), which includes the State Anxiety Scale (STAI-S) and the Trait Anxiety Scale (STAI-T). The present study employed the STAI-T (Form JYZ), a 20-item measure (e.g., "I feel pleasant") rated on a 4-point Likert scale. Total scores range from 20 to 80, with 20–44 indicating low anxiety, 45–54 indicating moderate anxiety, and 55–80 indicating high anxiety. Internal consistency in the experiment was excellent (Cronbach's α = .94). Participants' trait anxiety scores ranged from 27 to 73 (M = 46.75, SD = 12.39).

Perceived physical exertion during the self-regulated phase was measured using the Rating of Perceived Exertion (RPE) scale [17], which ranges from 6 (no exertion) to 20 (maximal exertion).

Intrinsic motivation was evaluated with the Intrinsic Motivation Inventory (IMI; 130), a widely used tool in sport and exergaming research [29, 97, 168]. We administered the pressure/tension subscale, a negative predictor of intrinsic motivation [60], consisting of five items rated on a 7-point Likert scale from 1 (not at all true) to 7 (very true). This subscale demonstrated moderate internal consistency (Cronbach's α = .64). Intrinsic motivation was also assessed using the Interest/Enjoyment subscale (seven items; e.g., "I enjoyed doing this activity

very much"), which showed excellent reliability (Cronbach's α = .90).

Attentional focus across conditions was examined with the Focus of Attention Questionnaire (FAQ; 158), comprising two five-item subscales: self-focus (e.g., "I was focusing on what I thought of the other person"; Cronbach's α = .63) and external-focus (e.g., "I was focusing on the other person's appearance or dress"; α = .78). These subscales capture attention directed toward internal reactions versus external cues. Items were rated on a 5-point scale from 1 (not at all) to 5 (in total).

Objective physical performance was captured using the treadmill's built-in calorie expenditure display. Additionally, real-time heart rate data were recorded throughout the running task using an Apple Watch SE (2nd generation) connected to the HypeRate API (`https://www.hyperate.io/`).

Finally, participants took part in a semi-structured interview lasting approximately 25 minutes. They were encouraged to reflect on their feelings during the experiment and to draw on their recent experiences when responding. The aim of the interview was to explore how the appearance of different types of virtual characters influenced participants' emotional responses. A set of pre-prepared questions provided a loose structure, addressing themes such as attentional focus, goals, reasoning behind character selection, emotional reactions, perceived pressure, exertion, and customization preferences. Follow-up questions were used when necessary to clarify participants' statements and to elicit more nuanced explanations of the ideas and concepts they raised.

4.3.4 Procedure

Twenty participants (9 females, 11 males; aged 22–36 years) took part in the study. The experiment comprised two sessions.

Session 1: Preparation Phase Participants first visited the laboratory, received a comprehensive explanation of the study procedures, and provided written informed consent. They completed questionnaires assessing exercise frequency, age, and trait anxiety (STAI-T), and rated the likability of several appealing anime characters. Resting heart rate was then recorded while participants wore an Apple Watch.

Next, each participant created an AR character modeled after a close friend, followed by a similarity questionnaire evaluating how closely the character's appearance and overall atmosphere matched that friend.

Participants' treadmill running speeds were subsequently calibrated. This involved determining (a) the running pace necessary to maintain a heart rate corresponding to 60–70% intensity for 2–3 minutes, and (b) the time required for heart rate to drop below 40% after transitioning to a reduced pace. Based on these assessments, the experimenter confirmed each participant's individualized "high-speed" and "low-speed" settings.

Session 2: System Experience Phase After a delay of 2–7 days, participants returned for the second session. At the beginning of the session, the experimenter explained the system's animation-triggering principles and presented 3–5-second preview animations of each character's cheering and standing motions within the experimental room. Participants viewed the previews and selected one companion from five options.

Participants then equipped the heart-rate monitor, completed warm-up stretching, and jogged lightly on the treadmill. When ready, they signaled to the experimenter, who announced the start of the trial and exited the room. Participants ran at their predetermined high speed for 4–5 minutes. After verifying that the required intensity and duration were achieved, the experimenter re-entered the room to cue the transition to the low-speed setting without interaction.

In the system condition, the experimenter activated the AR companion software from behind the large display, ensuring that the visual output was correctly mirrored before leaving the room. Participants then maintained the low speed for an additional 4–5 minutes—timed using individual calibration data to ensure that heart rate fell below the 40% intensity threshold once the experimenter had exited.

In the system condition, the experimenter, positioned behind the large display, activated the AR companion software and entered the participant's resting heart rate and age into the system. During the next 4–5 minutes, participants maintained the lower speed. This duration was calibrated from the first session's data to ensure that participants' heart rates fell below the 40% intensity threshold.

Following the experimenter's departure, participants were allowed to adjust their running speed freely, with an upper limit set at 85% intensity for safety. After 7 minutes, the experimenter re-entered and announced the end of the running task. Participants then completed a post-exercise questionnaire and rested for 15–20 minutes before performing the same running task under the opposite condition (With_system vs. Without_system). The order of conditions was counterbalanced across participants.

After completing both runs, each participant took part in a semi-structured interview.

The first session lasted approximately 30 minutes, and the second session approximately 90

minutes. Participants received compensation in accordance with institutional guidelines.

4.3.5 Quantitative evidence

We conducted linear mixed-effects modeling on data from the self-regulated running phase using the General Analyses for Linear Models module (GAMLj3, version 3.5.1) in Jamovi (version 2.6.44; `https://www.jamovi.org/`).

Linear mixed-effects models extend the general linear model by incorporating both fixed and random effects. Unlike traditional repeated-measures ANOVA, these models are well suited for hierarchical data structures and can flexibly accommodate unbalanced designs, missing observations, and irregular time intervals [100]. This flexibility enables more accurate estimation of population-level effects while appropriately accounting for individual differences [6, 81, 83]. However, the increased modeling flexibility also heightens the risk of overfitting [100].

To mitigate this risk, we followed a stepwise model-building procedure. We first constructed a baseline null model that included only a fixed intercept and random participant effects. Next, we added different combinations of fixed effects (condition, STAI-T) to this baseline and evaluated each candidate model using maximum likelihood (ML) estimation. Model comparisons were used to identify the optimal fixed-effects structure, which was then incorporated into the final model. Finally, we refitted the selected model using restricted maximum likelihood (REML) to obtain unbiased estimates of the fixed-effect coefficients.

Perceived exertion - RPE The experimental condition had no significant effect on the participants' perceived exertion during self-paced running, as indicated by the similar RPE scores across the groups: With_system (mean = 9.4, SD = 2.23) and Without_system (M = 9.3, SD = 2.38).

Attention - FAQ The experimental condition had no significant effect on the participants' external focus or internal focus during self-paced running, as indicated by the similar scores across the groups: With_system (mean = 11.7, SD = 5.00) and Without_system (M = 11.0, SD = 5.59) in external focus, With_system (mean = 12.0, SD = 3.57) and Without_system (M = 12.3, SD = 4.42) in internal focus.

Intrinsic motivation - IMI The experimental condition significantly affected the participants' felt pressure/tension (F = 6.98, p = .016), and interest/enjoyment during self-paced running (F

= 7.87, p = .011). The participants reported significantly lower perceived pressure and more interest/enjoyment in the With_system condition than in the Without_system condition.

Performance index - Calorie consumption The experimental condition significantly affected the participants' calorie consumption during self-paced running (F = 7.96, p = .011). The participants burned significantly more calories in the With_system condition than in the Without_system condition. There was a slight interaction between condition and trait anxiety (F = 3.55, p = .076). As illustrated in Fig 4.11, the calories burned by participants with moderate (F = 7.96, p = .011) and higher (F = 11.04, p = .004) trait anxiety were more in the With_system condition than in the Without_system condition. Participants' STAI-T scores ranged from 27 to 73 (M = 46.8, SD = 12.2).

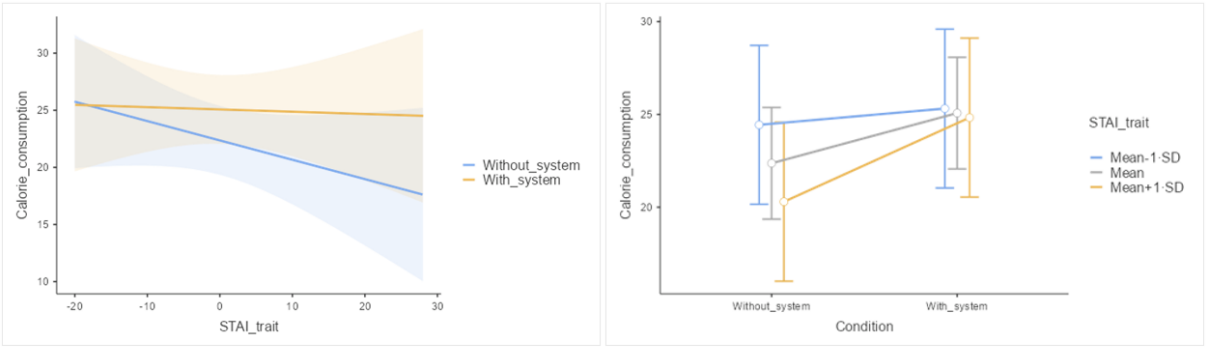


Figure 4.11: **Interaction between condition and trait anxiety on calorie consumption.** Participants with higher trait anxiety burned more calories in the With_system condition compared to the Without_system condition. STAI: State-Trait Anxiety Inventory.

Correlation We conducted a Spearman correlation analysis using IBM SPSS Statistics 27 to examine the pairwise relationships among the measures. The results (see Table 4.6) showed that RPE and External_focus, IMI_interest and calorie consumption, External_focus and Internal_focus, were positively correlated.

4.3.6 Qualitative evidence

This subsection presents participants' overall preference rankings for the different avatar types, followed by descriptive interpretations of why they favored or rejected certain designs. These patterns provide contextual grounding for the thematic analysis that follows.

Table 4.6: **Correlations between measures in validation experiment.**

	RPE	IMI_interest	IMI_pressure	External_focus	Internal_focus	Calorie
RPE	1	-.209	.293	.392*	.201	.167
IMI_interest	-.209	1	-.299	.019	.026	.328*
IMI_pressure	.293	-.299	1	.210	.236	-.192
External_focus	.392*	.019	.210	.1	.689**	-.140
Internal_focus	.201	.026	.236	.689**	1	.029
Calorie	.167	.328*	-.192	-.140	.029	1

** Correlation is significant at the 0.01 level: 2-tailed
* Correlation is significant at the 0.05 level: 2-tailed

Preference patterns across avatar types

When asked to rank their willingness to use each character, most participants (n = 13) selected the preferred animated character as their first choice, describing it as visually appealing, emotionally positive, and "comfortable to look at." Four participants favored the friend-like custom character, and two preferred the opposite-gender human-like avatar, often citing realism and familiarity as motivating factors. Only one participant selected the blue skeletal male model as his most preferred option.

In contrast, the skeletal models were overwhelmingly ranked as the least desirable by fifteen participants, who described them as "emotionless," "unnatural," or "unsettling." Three participants least preferred the opposite-gender human-like avatar, and two expressed low willingness to use their friend-like custom character.

These patterns suggest that participants gravitated toward avatars that conveyed emotional warmth, aesthetic appeal, or a sense of companionship, while avoiding designs perceived as abstract, mechanical, or uncanny.

Theme 1: How Avatars Feel to Users (Affective and Social Responses)

Participants described rich emotional and social reactions to avatar appearance. Many sought avatars that produced comfort, positivity, and warmth—qualities expressed as "harmless," "friendly," and "energetic" (P01–P03, P24). P01 emphasized that comfort was essential: the avatar "looked non-aggressive, and the appearance is what I like. Comfort is my main criterion—only when I feel comfortable can I interact with it." P24 also valued the lively expression: "the anime character's colors and expression are very lively and full of energy. When I feel a bit tired, seeing her cheering makes me want to persist."

Familiarity shaped social feelings in both positive and negative ways. P23 felt supported—"I

felt relaxed and supported, like when we used to work out together"—and P16 said familiarity "made me feel lighter and more natural." P15 appreciated reduced loneliness: "The friend-like character gives me a stronger sense of connection because it's someone from my real life—it reduces loneliness during exercise."

But familiarity also generated invisible expectations. P21 said, "Seeing a familiar face brings invisible social pressure, because you relate it to real-life interactions and feel you must behave accordingly," and P12 noted, "The more similar the character is to someone I know, the less relaxed I feel."

Participants reacted strongly to the gaze. P13 disliked being observed: "I don't like being stared at—it's awkward. The faceless model was my favorite because it doesn't have a gaze." P21 felt reassured: "I like being looked at—it feels kind and makes me feel noticed by the character, which relaxes me."

Visual complexity influenced attention and comfort. P12 said, "I don't need interaction or gaze—I just want something bright and strange that keeps my attention off tiredness," while P20 noted, "The anime character is my favorite, but she doesn't match the exercise setting—she's too vivid and distracts my attention."

When avatars appeared strict or unrealistic, participants felt pressure or discomfort. P08 said, "The realistic model felt like a strict coach—the emotional support was not gentle, more like supervision." Idealized avatars caused anxiety for some: P22 said, "The anime character's body and facial features are too perfect, too thin—it makes me anxious," and P13 said, "Her body proportions are strange and unrealistic; it feels creepy and distant." Mismatched realism also triggered discomfort, as P07 and P19 noted that "when the body shape didn't match my real friend's, it felt uncomfortable."

Theme 2: What Avatars Do for Users (Motivational and Functional Support)

Participants also evaluated avatars with respect to how they support exercise goals. Many selected avatars that appeared strong, capable, or similar to their gender identity. P03 said that "He looks strong and is the same gender, which gives me confidence—like exercising with a reliable partner." P22 similarly said, "her body and skin color make me feel secure; the same gender gives a sense of closeness and motivates me to reach that kind of body shape."

Participants also described role models as strong motivators. P10 said, "if a professional athlete were the companion, I would feel more driven to exercise." P07 and P23 added that "fitness influencers could motivate me more because their body is my goal," and "knowing their effort makes me want to move closer to that target."

Users also expected functional support from avatars, especially for safety. P19 said, “It would be better if the avatar reminded me when I overdo it to prevent injury.” Participants also valued audio cues, noting that “a bit of cheering or voice feedback would make the companion feel more alive.”

Gamification was widely supported. P22 said, “If my running time determined how long my virtual dog could walk, I’d definitely keep running longer,” and others (P02, P03, P05) mentioned coin rewards as engaging features.

4.3.7 Discussion

The validation experiment served as the final stage of this research, extending the two preliminary studies by translating their findings into a feedback system context. Whereas Experiments 1 and 2 investigated the psychological mechanisms underlying virtual audience effects—specifically, how the presence and emotional qualities of virtual spectators influence performance, perceived pressure, and motivation—the validation experiment aimed to determine whether these insights could be operationalized into a company AR system that responds to users’ physiological. This experiment therefore functioned as a bridge between theory and application, exploring how emotional companions might sustain motivation during self-regulated exercise.

This study demonstrates that emotionally expressive AR companions can enhance users’ affective experiences and exercise outcomes during self-paced running. Compared with running alone, exercising with the companion reduced perceived pressure, increased enjoyment, and led to higher calorie expenditure. These benefits were especially notable among participants with moderate to high trait anxiety. The qualitative findings further clarify why these effects occurred, revealing that users’ emotional responses and comfort levels were strongly shaped by the companion’s appearance, expressiveness, and perceived social role.

Emotional Safety as a Mechanism for Improved Experience The reduction in perceived pressure suggests that the AR companion created a psychologically safer environment for exercise. Running alone in a laboratory can heighten self-awareness and subtle evaluative stress[42, 20], particularly for individuals who are sensitive to being observed. In contrast, the presence of an emotionally positive avatar appeared to redirect attention away from the evaluative context and toward a more comforting focal point.

Participants described warm, expressive characters as relaxing, encouraging, and less socially demanding. This emotional cushioning likely benefited participants with higher trait anx-

iety the most: they showed the greatest improvement in calorie expenditure when accompanied by the system. The companion may have reduced internal tension, allowing them to maintain effort more comfortably. This aligns with prior work showing that virtual social presence can reproduce motivational aspects of social facilitation while reducing the pressure reducing associated with real or imagined social evaluation [131, 88].

Avatar Appearance Shapes Comfort and Engagement The qualitative results reveal that users' acceptance of the companion depended heavily on its visual and social characteristics. Animated and stylized avatars were consistently preferred; they were seen as friendly, non-threatening, and energizing. In contrast, human-like or skeletal avatars were judged as unsettling, overly strict, or emotionally flat. Even minor discrepancies in realism—such as body proportions that did not match the expected appearance—reduced trust and comfort, consistent with perceptual-mismatch accounts of the uncanny valley [76, 103].

Familiarity produced mixed effects. While friend-like avatars strengthened emotional connection for some participants, they also introduced subtle social pressure for others, who associated the avatars with real interpersonal expectations. This indicates that social distance is a critical design factor, determined by whether users perceive the character as providing intimate support or as being evaluative.

Motivational and Functional Expectations of AR Companions Beyond emotional comfort, participants evaluated avatars in terms of how they supported their exercise goals. Expressive, stylized avatars were perceived as effective emotional companions, whereas more realistic or athletic avatars were viewed as better for performance modeling or credible feedback. This distinction suggests that users implicitly assign different functional roles to different avatar types.

Participants also expressed clear expectations for future capabilities, including adaptive encouragement, safety alerts, and multimodal feedback such as voice or sound. Many viewed the companion not as an observer but as a potential training partner. Gamified interactions—such as earning coins or caring for a virtual pet—were also seen as powerful motivators, indicating that playful goal structures can complement emotional support.

Design implications for Real-Time Feedback Companions Taken together, the findings show that AR exercise companions must balance social presence, emotional safety, and motivational support. Users react not only to what companions do, but to who they appear to be. Emotional expressiveness and warmth promote comfort and enjoyment, while mismatched or

overly realistic designs risk inducing pressure or discomfort. Familiarity must be used carefully, as it can both strengthen and undermine engagement.

Designers should therefore aim for avatars that provide encouragement without surveillance, presence without pressure, and support without imposing real-world social expectations. As AR fitness applications continue to evolve, understanding how appearance, emotional tone, and adaptive behavior interact will be essential for creating companions that enhance—not hinder—users' physical and psychological experiences.

Based on both behavioral outcomes and participant feedback, three key design directions emerge:

- Role-aligned aesthetics – Match avatar visual style to function: stylized or cute characters for emotional companionship and attention support; realistic, physically fit models for instructional or goal-oriented feedback.
- Adaptive feedback intensity – Expand beyond low-intensity triggers toward context-aware responses that balance motivation with relaxation.
- Personalization and emotional distance – Allow users to adjust familiarity, realism, and gaze behavior to align with their comfort level. Idol-like or aspirational avatars may provide an optimal middle ground—motivational but emotionally safe.

Limitations and Future Directions In addition to the limitations of the preliminary experiments, the validation experiment also presents several constraints that should be considered. First, although the company AR system demonstrated promising results in improving user motivation and comfort, it was tested within a short, single-session context. Longer-term trials are required to evaluate whether users maintain engagement with virtual companions over extended use. Second, the feedback mechanism in this version of the system relied solely on heart rate data, which captures physical intensity but not psychological or emotional states. Future iterations should integrate additional biosignals such as facial expression, gaze, or affect recognition to enable richer emotional adaptation. Third, the visual and behavioral complexity of the AR companion was limited to simple animations (e.g., cheering or standing), which may not fully represent realistic social feedback. Incorporating more naturalistic gestures, speech, or contextual reactions could improve immersion and user identification.

The validation experiment contributes to both theoretical understanding and applied development of virtual audience research. Theoretically, it demonstrates that even minimal, reactive feedback grounded in physiological data can reproduce the psychological benefits of social facilitation while maintaining comfort. Practically, it offers a foundation for the gradual evolution

of real-time feedback companions into more adaptive, emotionally aware systems capable of interpreting users' internal states with greater nuance.

Future research should explore several key directions. First, expanding beyond heart rate as the sole feedback channel would provide a richer understanding of user states. Second, longitudinal studies are needed to assess whether feedback-based systems can sustain motivation and adherence to exercise routines over longer periods. Third, testing the system in more immersive AR or VR environments—and under varying social or environmental conditions—could clarify how context influences the perceived realism and effectiveness of feedback. Finally, collaborative or multi-user configurations may allow feedback companions to support cooperative or competitive virtual exercise experiences, broadening the social and motivational dimensions of the system.

Conclusion This study demonstrates that emotionally expressive AR companions can meaningfully enhance users' affective experiences and exercise performance during treadmill running. Exercising with the companion reduced perceived pressure, increased enjoyment, and led to higher calorie expenditure, with particularly strong benefits for individuals with moderate to high trait anxiety. These findings indicate that virtual companions can provide a motivating and emotionally safe form of social presence that supports engagement without creating the evaluative pressure often associated with real human audiences.

The qualitative results further show that users' comfort and acceptance depend strongly on avatar appearance and perceived social role. Warm, stylized characters fostered approachability and emotional support, whereas overly realistic or mismatched designs generated discomfort or social pressure. Participants also expressed a desire for companions that combine emotional expressiveness with adaptive feedback and playful goal structures.

Taken together, the findings highlight the importance of aligning a companion's visual design, emotional tone, and behavioral capabilities to create a supportive and enjoyable exercise experience. As AR and mixed-reality fitness applications become more widespread, designing companions that balance social presence with emotional safety will be essential for promoting sustained motivation and positive user experiences.

4.4 Summary

This chapter examined how designed social and emotional cues can enhance motivation and engagement during solitary physical activity. Building upon the theoretical foundation of social facilitation and the Yerkes–Dodson law of optimal arousal, the study explored whether vir-

tual representations of social presence—delivered through augmented reality (AR)—can sustain users' effort and enjoyment in exercise contexts where motivation typically declines over time. We selected a treadmill as the exercise apparatus for this study for the following two reasons. First, as one of the most common and accessible exercise devices, the treadmill can be used easily by participants after a brief familiarization period, which aligns well with our intention to employ a simple physical task. Second, running is a whole-body activity, and a relatively high level of exercise intensity can be achieved within a short period by adjusting the speed. This characteristic allowed us to better control and reduce the overall experimental duration.

Two preliminary experiments were conducted to isolate the effects of distinct forms of virtual social presence. **Experiment 1 (Audience Presence)** implemented a minimally interactive virtual audience with subtle standing animations to examine the influence of passive observation, comparing with other forms of social presence. Results indicated that limited social cues significantly triggered lower pressure. **Experiment 2 (Supportive AR Characters)** compared the influence of avatars with varying levels of familiarity and aesthetic appeal serving as cheering companions after users' high-intensity exercise.

To consolidate these insights, a **Validation Experiment** was developed as an integrated AR system combining the two prior conditions—the standing audience and cheering character animations. This system aimed to simulate a comprehensive affective-social environment, providing both subtle observation and active encouragement. Empirical findings revealed that the combined system effectively maintained participants' physiological arousal and engagement across exercise sessions, leading to higher running speed and self-reported enjoyment relative to control conditions. The qualitative results further show that users' comfort and acceptance depend strongly on avatar appearance and perceived social role. Warm, stylized characters fostered approachability and emotional support, whereas overly realistic or mismatched designs generated discomfort or social pressure.

Taken together, the findings highlight the importance of aligning a companion's visual design, emotional tone, and behavioral capabilities to create a supportive and enjoyable exercise experience. As AR and mixed-reality fitness applications become more widespread, designing companions that balance social presence with emotional safety will be essential for promoting sustained motivation and positive user experiences.

Appearance in two studies Across both studies, appearance is defined as the visual identity of the character (e.g., a familiar close friend or a highly preferred animated character). Appearance is not treated as a performance-related variable per se, but rather as a design mechanism through which the level of engagement with affective channels—namely familiarity and preference—is manipulated.

In Study 2, the research objective was to improve exercise performance. To support this goal, cheering behavior was introduced as a task-relevant motivational action. Although cheering involves movement, it was implemented in an identical manner across all experimental conditions. Thus, behavioral differences were not manipulated; instead, the experimental manipulation concerned who performed the cheering, rather than how the cheering was performed. Consequently, the character's appearance remained the key independent variable.

The two studies validate the same underlying design proposal. In Study 1, the appearance of a familiar friend acting as an interviewer effectively reduced participants' anxiety in interview and oral presentation contexts. In Study 2, the same appearance types—familiar friends and highly favored animated characters—were shown to enhance exercise performance when they functioned as cheering companions.

Chapter 5

Discussion

5.1 Overview

This chapter integrates findings (Table 5.1) from two empirical studies to examine how the appearance of others regulates emotional experience and behavioral performance in mediated interaction. Although the studies focus on different task domains—evaluative cognitive communication and motivational physical activity—they are unified by a common mechanism: appearance-based modulation of perceived social presence. By comparing familiarity and liking as affective channels across contexts, this discussion synthesizes how appearance cues calibrate psychological distance, influence appraisal of social evaluation or support, and ultimately shape affective and performance-related outcomes. The chapter first interprets each study in its respective task context, then develops a cross-study synthesis that identifies shared mechanisms, contextual divergences, and implications for theory and design.

Table 5.1: **Comparison of Interaction Dimensions Across Two Studies**

Dimension	**Study 1: Evaluative Interaction**	**Study 2: Action-Oriented Interaction**
Interaction Context	Evaluative, cognitive, socially judgmental	Action-oriented, physical, self-regulated
Role of Appearance	Regulating perceived evaluation	Shaping supportive companionship
Targeted Social Presence	Attenuated, non-threatening presence	Emotionally supportive presence
Primary Psychological Effect	Reduced state anxiety	Increased enjoyment & motivation
Performance Outcome	Improved communicative engagement	Enhanced physical engagement
Optimal Appearance Strategy	Familiar, emotionally safe faces	Liked or emotionally resonant avatars
Key Risk if Mismatched	Stylization $\rightarrow$ emotional ambiguity	Real spectators $\rightarrow$ pressure & stress

5.2 Study 1: Appearance-Based Regulation in Evaluative Cognitive Tasks

Study 1 examined how manipulating the appearance of an interaction partner regulates emotional experience during a cognitively demanding and socially evaluative task. By varying the interviewer's appearance (familiar friend, unfamiliar stranger, anime-stylized), the study isolated two competing regulatory pathways—affective safety through familiarity and cognitive distancing through abstraction—and tested their effectiveness in a controlled video-interview context.

The results demonstrate that familiar appearance reliably regulates anxiety in evaluative settings. Participants interacting with a friend-like interviewer consistently reported lower state anxiety, higher perceived performance, and greater verbal fluency. Self-reported experiences characterized these interactions as comforting and supportive, suggesting that familiarity activates schemas of acceptance and reduced judgment. In this sense, familiar appearance functions as an external affective scaffold—lowering perceived evaluation threat and enabling participants to allocate cognitive resources toward task execution rather than self-monitoring. Notably, these positive effects persisted even when friend-like avatars introduced slight uncanny valley effects, suggesting that familiarity-based emotional safety remains the dominant factor in reducing evaluation anxiety.

By contrast, stylized abstraction produced mixed and unstable effects. The anime condition reduced gaze engagement, indicating partial cognitive distancing from the evaluative situation. For some participants, this abstraction enabled reappraisal—reframing the interaction as less real or less consequential. However, for others, reduced realism impaired social interpretability and introduced ambiguity, which in turn heightened uncertainty and anxiety. Rather than functioning as a consistent buffer, abstraction became a double-edged mechanism: capable of dampening evaluation for some users while undermining emotional regulation for others.

Taken together, Study 1 indicates that in cognitively evaluative contexts, emotional regulation benefits more from affective presence than from perceptual distance. Familiarity provides informational richness and emotional predictability, whereas abstraction risks disrupting meaning-making when users must interpret intent, judgment, or feedback. In conclusion, the overall pattern suggests that realism paired with emotional safety is more effective than symbolic distancing when composure and clarity are required.

5.3 Study 2: Appearance-Based Regulation in Motivational Physical Tasks

Study 2 extended the investigation from anxiety down-regulation to motivation up-regulation in embodied physical activity. Across multiple experiments, participants exercised under varying forms of social presence, including real audiences, mediated spectators, and AR companions with different appearances. The findings demonstrate that appearance-driven modulation of social presence plays a critical role in sustaining motivation while managing perceived pressure.

Compared to real or video audiences, AR-mediated presence consistently reduced pressure without eliminating social engagement. In subsequent experiments, both friend-like companions and liked anime characters enhanced effort and performance, but through partially distinct pathways. Friend-like companions operated through relational warmth and emotional reassurance, supporting sustained effort, especially among experienced exercisers or individuals with lower social anxiety. Liked anime characters, in contrast, leveraged aesthetic affinity and playfulness, transforming repetitive exercise into an engaging interaction with a nonjudgmental companion.

Preference data and interviews clarified this distinction. Many participants favored stylized, expressive characters because these representations reduced evaluation apprehension while maintaining motivational presence. The benefits of stylization emerged not from abstraction alone, but from personal liking—a factor that narrowed affective distance even as perceptual distance remained. In this context, abstraction was energizing rather than alienating, supporting engagement in tasks that demand activation rather than composure.

Overall, Study 2 reveals that motivational contexts invert the regulatory logic observed in Study 1. In evaluative tasks, familiarity and realism provide emotional safety that enables composure. In motivational tasks, stylization and aesthetic affinity provide energizing presence without triggering evaluation apprehension. Effective regulation therefore depends on aligning appearance with task-specific emotional demands.

5.4 Cross-Study Synthesis: Appearance as a Regulator of Social Presence

5.4.1 Familiarity and Emotional Safety

Across both studies, familiarity emerged as a powerful but context-sensitive regulator. Familiar appearance cues activate trust, acceptance, and prior positive interaction schemas, often reducing perceived threat and enhancing comfort. However, familiarity can also amplify evaluation when it evokes real-world expectations or interpersonal standards. Thus, familiarity operates not as a uniformly positive factor, but as a high-impact amplifier whose effect depends on whether the interaction is framed as supportive or evaluative.

- **Study 1 (video interviews):** Seeing the interviewer as a familiar friend reduced state anxiety, improved self-evaluation, and encouraged more verbal output.
- **Study 2 (exercise with AR companions):** Friend-like characters reduced perceived pressure and supported sustained physical performance.

These findings align with social-cognitive accounts suggesting that familiar identities activate schemas of trust, acceptance, and prior positive experiences. In mediated contexts—where nonverbal cues may be limited or distorted—familiar appearance can function as perceptual reassurance, transforming an evaluative or demanding situation into one that feels more supportive.

At the same time, familiarity is not uniformly benign. For some participants, friend-like characters also intensified feelings of being evaluated, because they implicitly carried real-world expectations. This dual role suggests that familiarity:

- Buffers social-evaluative threat when interpreted as supportive presence
- Amplify evaluation pressure when interpreted as representing real interpersonal standards.

Thus, familiarity is best understood as a powerful but context-sensitive regulator of both emotional safety and perceived evaluation.

5.4.2 Stylization as Calibrated Distance

Stylization functioned as a mechanism for regulating psychological distance, but its effectiveness depended on task seriousness, user preference, and interpretability. In formal evaluative contexts, stylization risked ambiguity and reduced emotional clarity. In contrast, in physical and motivational contexts, stylization—when aligned with user liking—reduced pressure and increased enjoyment. These findings position stylization not as an aesthetic choice, but as a context-dependent moderator of social meaning.

Divergence from Prior Work and the Role of Interaction Authenticity:

A critical finding is that anime-style face filters did not reduce anxiety in Study 1's high-pressure video interviews and significantly reduced gaze engagement—participants looked less at stylized interviewers. This contradicts prior research suggesting stylization's potential for anxiety buffering [88, 163].

This divergence reveals an important methodological insight: prior studies examined stylization in practice scenarios, casual conversations, or pre-recorded interactions where evaluation stakes were low or absent. In contrast, Study 1 employed real-time, consequential video interviews where participants faced genuine evaluation. Under these authentic high-stakes conditions, stylization introduced interpretive ambiguity and disrupted social meaning-making rather than providing psychological distance.

The gaze behavior finding is particularly revealing. Reduced visual engagement with anime-styled interviewers suggests that stylization not only failed to comfort but actively disrupted normal interaction patterns in formal contexts—a behavioral marker absent in prior low-stakes studies.

Context-Contingent Effectiveness:

However, Study 2 demonstrated that stylization succeeded in motivational contexts. Liked anime companions enhanced enjoyment, reduced pressure, and increased engagement during exercise. The key difference: exercise contexts are playful and less interpretively demanding, allowing stylization to function as intended—creating presence without judgment.

Implication: Stylization's effectiveness is task-contingent and stakes-dependent. It succeeds when:

- Tasks are motivational rather than evaluative
- Interpretive clarity is less critical
- Aesthetic affinity can compensate for perceptual distance

It fails when:

- Evaluation stakes are high and genuine
- Social interpretation and feedback reading are essential
- Ambiguity introduces uncertainty rather than comfort

This highlights the importance of testing appearance manipulations under ecologically valid, high-pressure conditions rather than extrapolating from simulated or practice contexts.

5.4.3 Effects of User Goals, and Individual Characteristics

User Goals Within the Same Context:

Even within motivational exercise contexts, different user goals shaped appearance preferences.

- Participants prioritizing emotional regulation and enjoyment preferred cute, playful, or aesthetically appealing companions that reduced pressure and made exercise feel pleasant
- Participants prioritizing performance outcomes were more drawn to companions with athletic physiques that provided aspirational modeling

This indicates that appearance preferences depend not only on task type (evaluative vs. motivational) but also on what users aim to achieve within that task.

Trait Anxiety as Moderator:

High trait anxiety participants benefited more from supportive appearance designs, experiencing stronger motivational and emotional regulation effects compared to low-anxiety participants.

5.4.4 Emotion–Performance Coupling

A central insight across both studies is the tight coupling between affect and performance. Reduced anxiety, increased comfort, and feelings of support consistently aligned with better cognitive fluency and sustained physical effort. Emotional regulation is therefore not peripheral to task performance in mediated environments; it is a precondition for effective engagement.

- In Study 1, higher state anxiety correlated with poorer perceived performance and tended to be associated with reduced verbal output. Participants who felt more at ease (especially in the friend condition) spoke more and judged their performance more positively. Familiar mediated presence therefore not only improved subjective comfort but also supported more fluent and confident task behavior.
- In Study 2, conditions that reduced pressure and increased enjoyment (e.g., emotionally resonant AR companions) were also associated with higher calorie expenditure and better performance. Participants described feeling "encouraged but not judged," which allowed them to sustain effort without the mental cost of intense evaluation pressure.

5.5 Theoretical Contribution

Extending Avatar Research Beyond Self-Representation

This dissertation addresses a critical gap in avatar and embodiment research by demonstrating that others' appearance—not only self-representation—systematically shapes users' emotional states and behavioral performance in mediated environments. While prior work has predominantly examined how one's own avatar influences behavior (e.g., the Proteus effect), this research reveals that interaction partners' visual representations constitute an equally powerful, yet understudied, mechanism of social influence.

Mechanism: Appearance as a Tunable Mechanism for Regulating Social Presence

Moving beyond treating appearance as merely an aesthetic attribute, this work establishes appearance as a regulatory mechanism that systematically modulates perceived social presence. The empirical findings demonstrate that familiarity and stylization can either amplify or attenuate social presence depending on context, task demands, and individual differences. This reframes social presence not as a uniformly beneficial property to be maximized, but as a context-

dependent, tunable design variable that must be calibrated to support different emotional and performance goals.

Methodological Contribution: Ecological Validity as a Boundary Condition

This dissertation reveals that appearance effects are shaped by interaction authenticity. Appearance manipulations that succeed in simulated or low-stakes contexts may fail—or even backfire—when evaluation stakes are genuine and social interpretability is critical. Study 1's finding that stylization disrupted gaze behavior and failed to reduce anxiety in video interviews with real evaluation stakes, despite prior evidence of its buffering potential in practice scenarios, demonstrates that ecological validity functions as a boundary condition for appearance-based regulation.

This challenges assumptions that effects observed in laboratory simulations or casual interactions generalize straightforwardly to high-pressure real-world contexts. The divergence suggests that when users face authentic evaluation, their need for interpretive clarity may override the potential comfort of visual abstraction. This insight highlights the importance of testing appearance interventions under conditions that authentically replicate deployment contexts—including genuine evaluation stakes and real-time interaction—to more accurately assess intervention effectiveness and identify when design strategies are more or less likely to succeed.

Establishing an Appearance-Based Affective Co-Regulation Framework

Finally, this work introduces an appearance-based affective co-regulation framework showing that digital representations function as emotional partners that scaffold users' affective states and performance. By integrating emotion regulation theory, social facilitation, and social presence research, the dissertation demonstrates that:

- Affective regulation is not peripheral but foundational to performance in mediated interaction
- Appearance cues (familiarity, stylization, aesthetic affinity) operate through appraisal pathways that shape perceived social distance, evaluation threat, and motivational support
- Individual differences (particularly trait anxiety) moderate these effects, with high-anxiety users benefiting disproportionately from supportive, familiar, or stylized companions

This framework provides a theoretical foundation for designing psychologically adaptive virtual and augmented environments that support not only task performance but also emotional well-being. The regulatory pathway is shown in Fig 5.1.

Regulatory pathway:

```
Appearance Cues (familiarity, stylization, aesthetic affinity)
    ↓
Modulation of Perceived Social Distance & Evaluation Threat
    ↓
Emotional Regulation (anxiety down-regulation OR motivation up-regulation)
    ↓
Performance Outcomes (cognitive fluency, sustained effort)
```

Figure 5.1: **Regulatory Pathway in Appearance-Based Affective Co-Regulation Framework**

5.6 Practical Implications

The empirical findings reveal how appearance cues—familiarity, stylization, and aesthetic affinity—regulate users' emotional responses and task performance across mediated interaction contexts. These insights translate into actionable design guidance for developing emotionally intelligent AR and virtual systems. Below, we outline context-specific recommendations and general principles, emphasizing the importance of emotional congruence, adaptive presence, and task-aligned appearance design.

AR Communication Systems: Supporting Emotional Safety in Evaluative Contexts

1. Problem Identified: Users experience heightened anxiety and evaluation apprehension in video-mediated communication, particularly during high-stakes interactions (interviews, presentations, performance reviews).
2. Mechanism-Based Solution: Familiar appearance cues increase emotional safety by activating schemas of trust and acceptance.
3. Design Recommendations:
    - Prioritize human-like, familiar appearances in real-time communication tools where users value immediate feedback and emotional reassurance

- Design for emotional fit over photorealism: prioritize appearance cues that signal warmth, acceptance, and nonjudgmental presence

AR Fitness and Motivational Companions: Enhancing Engagement Without Pressure

1. Problem Identified: Users struggle with low motivation in self-regulated exercise; real audiences increase performance but often at the cost of excessive pressure.
2. Mechanism-Based Solution: Stylized, aesthetically liked companions provide motivational presence while reducing evaluation apprehension through calibrated psychological distance.
3. Design Recommendations:
    - Create emotionally safe motivational presence through expressive, stylized, or aesthetically appealing characters (e.g., liked anime companions) that reduce feelings of judgment while maintaining engagement
    - Support identity alignment and long-term engagement by offering customization options that allow users to select companions matching their aesthetic preferences, relational expectations, and motivational needs
    - For high-anxiety users, emphasize nonjudgmental, playful companions that transform exercise from a performance task into an enjoyable interaction

General Principles for Designing Mediated Social Presence

Effective mediated social presence requires aligning appearance design with task demands, emotional goals, and individual user characteristics.

Core Design Principles:

- Emotional congruence over realism: Prioritize how a character feels (warm, supportive, nonjudgmental) over how realistic it appears; emotional fit matters more than visual fidelity
- Familiarity and aesthetic affinity drive comfort: Users feel safer and more engaged with representations they already like, relate to, or find aesthetically appealing—leverage personal preferences, not just demographic matching

- Adaptive presence calibration: Design presence that adjusts to users' emotional needs (supportive and visible when sought, unobtrusive and low-pressure when autonomy is preferred), avoiding the feeling of constant surveillance
- Task-contingent appearance design: Match appearance to task context:
  - Evaluative tasks (interviews, presentations): Favor familiar, realistic, warm appearances that reduce anxiety
  - Motivational tasks (exercise, learning): Favor stylized, aesthetically liked appearances that enhance engagement without judgment
- Design for perceived identity and relational role: Users respond not only to a companion's features or functions but to its implied personality, intentions, and relational role ("who it is")—craft appearance that signals the desired social relationship (friend, coach, companion)

Ethical Considerations

Core principle: Co-regulative systems must respect transparency and consent.

- Prioritize privacy, dignity, and trust as foundational ethical principles

This framework advocates for human-centered character design where empathy, emotional support, and psychological well-being form the foundation of technological intervention, ensuring that virtual and augmented environments are not only engaging but also emotionally sustainable.

5.7 Revisiting the Research Objectives

To explicitly answer the Research Objectives stated in Chapter 2, this section summarizes how each objective was addressed by the two studies.

This dissertation set out to validate an appearance-based social presence regulation mechanism for addressing real-world challenges in mediated interaction. The research objectives outlined in Chapter 2 are addressed through two complementary empirical studies examining distinct yet theoretically connected contexts.

First, the dissertation aimed to examine how variations in others' appearance shape perceived social presence and evaluation in mediated interactions (Objective 1). Study 1 demonstrated that familiar appearances significantly increased perceived emotional support and reduced evaluation threat in video-mediated interviews, whereas stylized (anime) appearances disrupted social interpretation and introduced perceptual ambiguity. Study 2 further revealed that AR companions' visual identity (particularly aesthetic affinity toward liked characters) influenced how participants appraised their presence as supportive versus pressuring, demonstrating that personal liking can narrow affective distance even when perceptual distance remains high.

Second, the research sought to understand how appearance cues regulate users' emotional states across contexts (Objective 2). Study 1 showed that familiar faces effectively down-regulated anxiety in high-stakes evaluative tasks, while stylization alone did not reliably alleviate evaluative stress and sometimes introduced uncertainty. In contrast, Study 2 revealed that liked stylized companions—particularly aesthetically appealing characters—up-regulated motivation and enjoyment while reducing perceived pressure during physical exercise. This demonstrates context-contingent regulatory pathways: familiarity provides affective safety under evaluation, whereas stylization paired with aesthetic affinity enhances engagement in motivational tasks.

Third, the dissertation investigated how these perceptual and emotional effects translate into behavioral and performance outcomes (Objective 3). In Study 1, reduced anxiety and increased emotional comfort were associated with greater verbal fluency and more positive self-evaluation. Participants who felt emotionally safe spoke more and judged their performance more favorably. In Study 2, conditions that reduced pressure while increasing enjoyment (e.g., liked AR companions) supported sustained physical effort and higher calorie expenditure. Across both studies, emotional regulation emerged not as peripheral but as foundational to effective task performance, with affective comfort enabling cognitive resources to be allocated toward task execution rather than self-monitoring.

Finally, by integrating findings across evaluative cognitive and motivational physical tasks, the dissertation aimed to develop a generalizable framework of appearance-mediated social presence and derive design implications for adaptive systems (Objective 4). The cross-study synthesis reveals that appearance operates as a tunable mechanism for calibrating social presence to match task demands and individual needs. Three key principles emerged: (1) appearance effectiveness is task-contingent—stylization failed in evaluative contexts but succeeded in motivational contexts; (2) individual differences, particularly trait anxiety, systematically moderate effects—high-anxiety users benefit disproportionately from supportive representations; (3) the appearance $\rightarrow$ social presence appraisal $\rightarrow$ emotional regulation $\rightarrow$ performance pathway

operates consistently but requires context-appropriate calibration. These findings provide actionable design guidance: prioritize familiar, warm appearances in evaluative communication to reduce anxiety; use stylized, aesthetically liked companions in motivational contexts to enhance engagement without pressure.

Together, these studies demonstrate that manipulating others' appearance provides a systematic means of regulating social presence, emotion, and performance in mediated environments. Appearance cues shape how safe, supported, or scrutinized users feel, and these appraisals directly influence cognitive fluency, motivation, and sustained engagement. By revealing how familiarity, stylization, and aesthetic affinity interact with task demands and individual differences, this dissertation establishes appearance as a foundational design lever for creating empathetic, adaptive, and psychologically supportive virtual and augmented interaction systems.

5.8 Summary

In summary, this chapter consolidates findings from two studies into a unified framework showing that the appearance of others functions as a systematic mechanism for regulating perceived social presence, which in turn shapes emotional and performance-related outcomes in mediated interaction. Rather than treating social presence as a fixed property of a medium, the results demonstrate that it can be deliberately calibrated through appearance design to meet the psychological demands of different tasks.

Across cognitive-evaluative and embodied-motivational contexts, the studies reveal that appearance influences interaction not directly, but through its impact on how users interpret social distance, evaluation, and support. Familiarity, stylization, and aesthetic affinity operate as appearance cues that modulate the intensity and meaning of social presence, enabling either affective down-regulation or motivational up-regulation depending on context. This supports a relational view of appearance in which others' visual representations shape users' appraisal processes rather than merely serving as decorative or identity-related elements.

Theoretically, these findings extend existing work on digital embodiment by shifting the focus from self-avatar effects to appearance-mediated interpersonal regulation, addressing a critical gap in prior research. By empirically linking appearance design to social presence calibration and downstream emotional and behavioral outcomes, this dissertation contributes a generalizable framework for understanding how mediated representations can support psychological well-being and task engagement. This framework provides a foundation for future research and design of virtual and augmented social systems that are not only immersive but also emotionally adaptive and socially intelligent.

Beyond its immediate empirical contributions, this framework carries broader implications for the design and sustainability of future virtual and augmented communities. As social interaction increasingly takes place in persistent digital spaces, individuals will be exposed to a growing volume of mediated social presence. By conceptualizing others' appearance as a mechanism for regulating social presence rather than merely increasing it, this work highlights a pathway toward more harmonious and psychologically supportive virtual communities. The findings suggest that virtual environments need not rely on maximal realism or constant social intensity to foster connection. Instead, adaptive appearance design—tuning familiarity, stylization, and emotional resonance—can help balance social connection with psychological safety. The framework supports a vision of future virtual communities in which interaction is not only engaging and effective, but also emotionally sustainable—promoting inclusion, well-being, and long-term social harmony.

Chapter 6

Conclusion

6.1 Overview

This dissertation investigates how social presence can be deliberately regulated through the appearance of others in mediated and augmented interaction. Motivated by the growing prevalence of virtual communication and AR-based companionship, the research addresses a critical gap in existing literature: while prior work has largely focused on self-avatar representation or behavioral cues, far less is known about how the visual appearance of interaction partners shapes users' emotional experience and performance. Through two complementary studies—one examining anxiety regulation in evaluative communication, and the other examining motivation and performance in self-regulated physical activity—this work adopts a controlled, appearance-centered approach to reveal how familiarity, stylization, and affective affinity operate as social presence cues. Together, the studies establish appearance as a central design lever for shaping psychologically supportive mediated interactions. The research framework is shown in Fig. 6.1, with its explanation presented in Fig. 6.2.

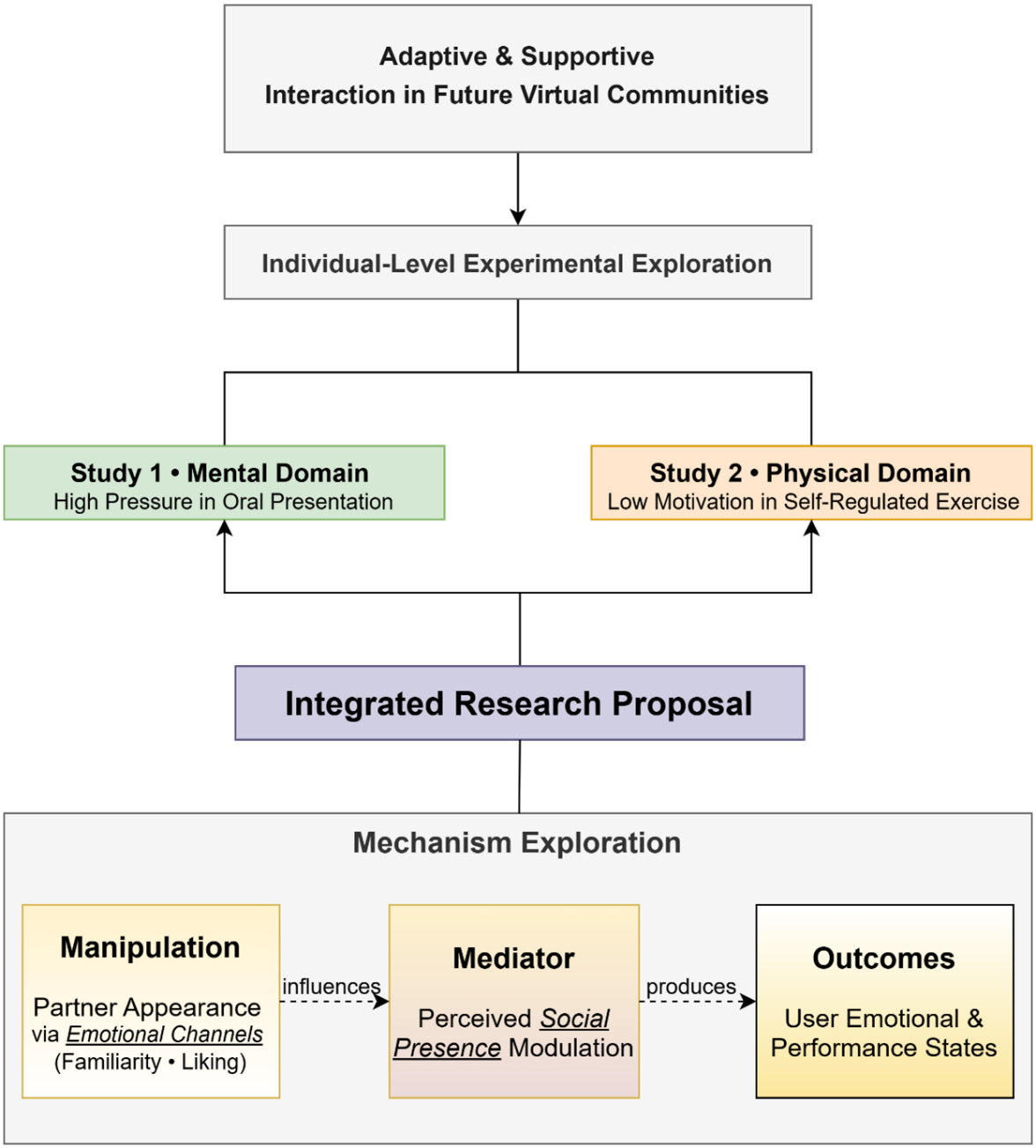


Figure 6.1: **Overview of the Research Framework**

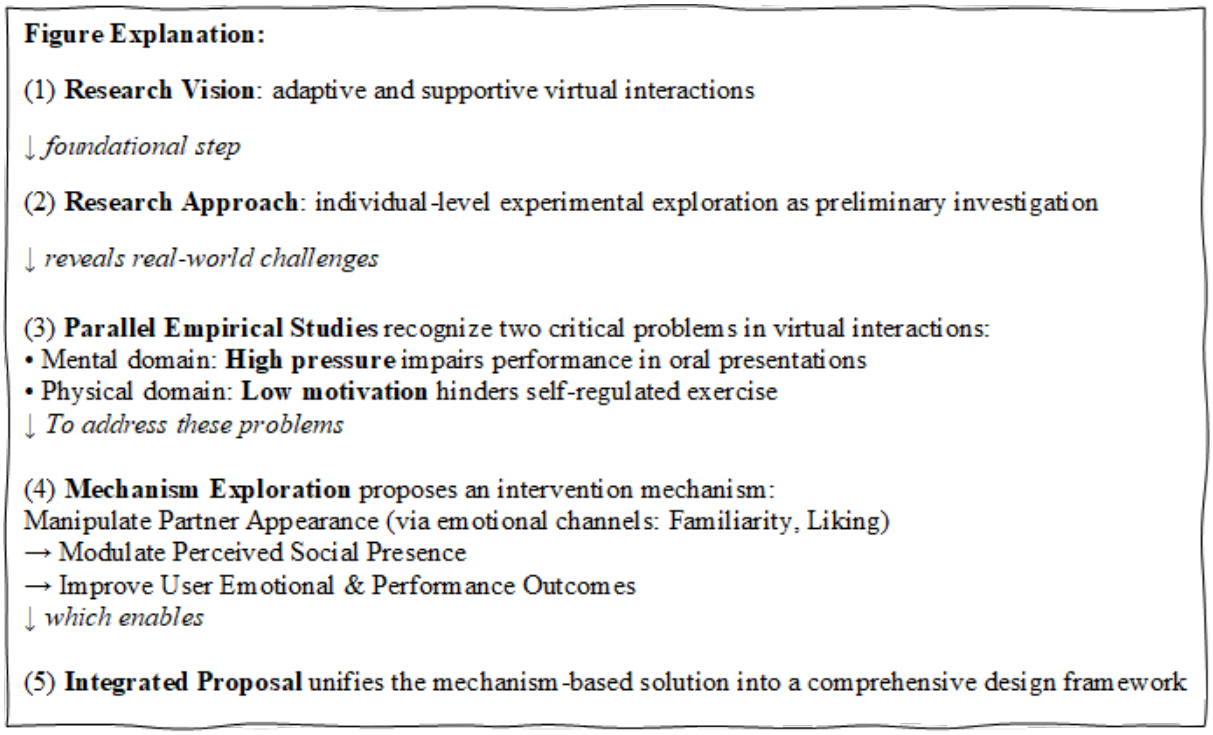


Figure 6.2: **Explanation of the Research Framework**

6.2 Purpose and Scope of This Research

The overarching purpose of this dissertation is to understand how the appearance of others in mediated and augmented interaction can be deliberately designed to regulate social presence, and how such regulation shapes users' emotional states, motivation, and performance.

This research is motivated by a broader challenge facing contemporary virtual and augmented systems: as social interaction increasingly occurs through digital mediation, users are often exposed to heightened social evaluation, emotional strain, or motivational fatigue. While prior work has focused extensively on system functionality or on users' self-representation, far less attention has been given to how others' appearance—interviewers, audiences, or virtual companions—can be systematically leveraged to support users psychologically.

To address this gap, this dissertation investigates appearance not as a decorative feature, but as a designable social variable that influences how social presence is perceived, interpreted, and experienced.

6.3 Research Structure and Methodological Strategy

The dissertation adopts a two-study structure, intentionally spanning distinct interaction contexts to test the generality of appearance-based social regulation.

- Study 1 examines a cognitively demanding and socially evaluative scenario—video-mediated interviews. By manipulating interviewer appearance (familiar, stranger, stylized anime) while holding behavior constant, the study isolates appearance as the sole variable influencing emotional response. This design allows a controlled investigation of whether and how appearance alone can modulate anxiety and perceived performance.
- Study 2 extends this inquiry into an embodied, self-regulated activity—physical exercise with AR companions and audiences. Through a series of experiments, it systematically varies the form of social presence (real, video, AR, none) and the appearance of AR companions (friend-like, liked-anime, neutral), examining their effects on motivation, pressure, and performance.

This cross-contextual structure is deliberate. By pairing a high-stakes evaluative task with a motivation-driven physical task, the dissertation demonstrates that appearance-based regulation operates across domains, while remaining sensitive to task demands and individual differences.

6.4 Rationale for the Design Approach

A defining methodological choice in this work is the focus on appearance manipulation without changing agent behavior, authority, or functional role. This approach was adopted for three reasons.

First, it allows the research to isolate perceptual and affective mechanisms, showing that emotional and behavioral outcomes can be altered even when interaction dynamics remain unchanged.

Second, appearance is uniquely suited to subtle and user-respectful intervention. Unlike explicit feedback, instruction, or persuasion, visual social cues can reshape experience without imposing cognitive load or overt control.

Third, appearance-based design is highly compatible with emerging AR systems, which can modify how others are perceived while preserving grounding in the real world. This makes the findings directly relevant to future mediated communication and companionship technologies.

6.5 Research Motivation and Originality

The overarching aim of my research is to promote more harmonious and effective communication and interaction in future virtual environments. As a foundational step toward this aim, the present work addresses common challenges in contemporary technology-mediated interactions, including heightened stress in video-based communication and reduced motivation in self-regulated exercise contexts. To tackle these issues, this research advances a shared core proposition: that social presence, when mediated through visual appearance and affective channels such as familiarity and affinity, can meaningfully shape users' emotional experiences and behavioral outcomes during interactions with avatars or virtual agents.

Study 1 focuses on stressful interaction contexts and examines whether modifying an avatar's appearance can reduce users' nervousness and create a more emotionally comfortable interaction atmosphere. The originality of this study lies in demonstrating that interpersonal impressions and emotional responses can be altered solely through appearance-mediated social presence, without changing the behavior, role, or intrinsic characteristics of the interaction partner. By isolating appearance as the key variable, Study 1 provides evidence that emotional responses in mediated interactions can be regulated at a perceptual level, highlighting the importance of visual social cues in shaping users' psychological comfort.

Study 2 extends this proposal to a self-regulated exercise context, where maintaining motivation and performance is inherently challenging. The purpose of Study 2 is to examine whether interaction with a virtual agent can enhance exercise motivation and performance by providing a sense of social presence that is emotionally supportive rather than pressure. In this study, the avatar is designed to function as a virtual presence that encourages continued participation while respecting the user's autonomy and self-regulation. The originality of Study 2 lies in its exploration of avatar-mediated support as a means of facilitating motivation and performance through affective channels, rather than through explicit instruction, authority, or external pressure. By emphasizing familiarity and affinity in the avatar's appearance and presence, the study demonstrates how social presence can positively influence exercise engagement and outcomes.

Together, these studies illustrate how a shared design framework—centered on social presence, appearance, and emotional channels—can be applied across different contexts to support users' emotional well-being, motivation, and performance. They highlight the potential of avatar-mediated interaction to function not merely as a substitute for human interaction, but as a distinct and, in some cases, advantageous form of support.

6.6 Limitations and Future Work

This dissertation provides preliminary evidence that others' appearance can be used to regulate social presence and influence emotional and behavioral outcomes. Several limitations define the current scope of its contributions and point toward important directions for future research.

6.6.1 Scope of Participants and Cultural Specificity

First, the participant samples were relatively homogeneous, consisting primarily of young adults with prior exposure to digital media and, in some cases, familiarity with anime aesthetics or AR interfaces. As a result, the observed effects of familiarity and stylization may be shaped by cultural norms, media literacy, or generational preferences.

Future research should examine whether appearance-based social presence regulation operates similarly across broader demographic groups, including older adults, individuals with limited exposure to virtual characters, and participants from different cultural backgrounds. Such work would help determine whether familiarity and affinity function as universal affective channels or culturally situated ones.

6.6.2 Temporal Scope and Longitudinal Effects

Second, both studies focused on short-term interactions and immediate affective and performance outcomes. While this approach was appropriate for isolating causal effects, it does not capture longer-term dynamics such as habituation, emotional adaptation, or shifts in user–agent relationships over time.

Future work should adopt longitudinal designs to investigate how repeated exposure to familiar or stylized avatars influences emotional regulation, motivation, and trust. In particular, it remains unclear whether the benefits of supportive appearance persist, diminish, or transform as users develop expectations or relational histories with mediated agents.

6.6.3 Limited Expressive and Relational Complexity

Third, the avatars and AR companions used in this research were intentionally constrained in their expressive range and behavioral variability. This design choice enabled isolation of appearance effects but limited exploration of richer social dynamics, such as reciprocal emotion expression, adaptive relational roles, or conversational depth.

Future systems could integrate more expressive behaviors, dynamic facial cues, or adaptive relational framing to examine how appearance interacts with behavior over time. This would allow researchers to explore not only appearance-mediated regulation, but also co-evolution of appearance and interaction as a form of sustained social presence.

6.6.4 Contextual Boundaries of the Framework

Finally, the appearance-based regulation framework proposed in this dissertation should be understood as context-sensitive rather than universally optimal. The effectiveness of familiarity or stylization depends on task demands, emotional goals, and individual traits. Appearance that is supportive in one context (e.g., exercise) may be distracting or inappropriate in another (e.g., formal evaluation).

Future research should therefore explore additional contexts—such as collaborative work, education, healthcare, and group-based interaction—to refine when and how appearance-based social presence regulation is beneficial. Comparative studies across task types would help formalize design principles that specify not only what works, but under what conditions it works.

6.6.5 Toward Adaptive and Ethical Appearance Design

Looking forward, a promising direction lies in developing adaptive appearance systems that respond to users' emotional and cognitive states in real time. Advances in physiological sensing, affect recognition, and AR rendering could enable avatars to dynamically adjust visual style, expressiveness, or familiarity cues to better support users.

At the same time, such systems raise ethical questions regarding transparency, autonomy, and emotional influence. Future work must therefore integrate ethical design principles, ensuring that appearance-based regulation remains interpretable and respectful of psychological boundaries.

Summary Together, these limitations delineate the current boundaries of the dissertation's contributions while outlining a clear agenda for future research. By extending the framework across populations, time scales, interaction richness, and application domains, future work can build on this foundation to develop more adaptive, ethical, and psychologically informed mediated social systems.

6.6.6 Extending to Face-to-Face Interaction

When AR devices become more portable and technically mature—such that AR lenses can effectively replace computer screens as carriers of virtual avatars—it may become possible to manipulate the perceived appearance of others during face-to-face communication through AR glasses. In such scenarios, users would engage in in-person interaction while visually perceiving an altered appearance of their communication partner. Compared to screen-based interaction, physical or hybrid environments provide richer contextual information, including bodily movement and spatial cues. In this context, attentional focus becomes particularly important. Users may continue to direct their attention toward the face to seek reassurance from a familiar and comforting appearance. Alternatively, mismatches between facial appearance and bodily movement could increase the risk of an uncanny valley effect. However, when users actively choose to apply such face filters, they are explicitly aware that the perceived appearance does not represent the real person. This awareness may help mitigate choking under pressure and allow users to anticipate and tolerate potential mismatches between appearance and behavior. For these reasons, I believe that the core findings of Study 1 are likely to extend to real-world or hybrid interaction contexts. However, additional factors—such as the congruence between visual appearance and voice—would need to be carefully considered in physical interactions, and these issues represent important directions for future research.

6.7 Contribution to Knowledge Science

This dissertation contributes to knowledge science by advancing theoretical understanding of how social information is perceived, processed, and acted upon in mediated environments.

Conceptual Contribution: Appearance as a Knowledge-Bearing Social Signal

The research establishes appearance as a knowledge-bearing signal that conveys information about safety, evaluation, and relational intent. It demonstrates that users draw inferences about social meaning not only from behavior or language, but from visual representation alone. This extends knowledge-science perspectives on social cognition by highlighting appearance as an input channel through which social meaning is constructed.

Theoretical Contribution: Regulating Social Presence as a Mechanism

The dissertation contributes a structured account of how appearance regulates perceived social presence, which in turn shapes emotional and behavioral outcomes. By empirically linking appearance $\rightarrow$ social presence $\rightarrow$ emotion $\rightarrow$ performance across contexts, the work clarifies a mechanism that has previously been treated implicitly or descriptively in the literature.

Individual-Difference Contribution: Anxiety-Sensitive Knowledge Design

The findings show that social information is not processed uniformly: trait anxiety systematically alters how appearance cues are interpreted. This contributes to knowledge science by reinforcing the importance of user-sensitive models of social perception, in which the same stimulus produces different cognitive and emotional outcomes across individuals.

6.8 Significance and Broader Implications

The significance of this research lies in its demonstration that mediated social environments can be designed to support psychological functioning, not merely communication efficiency. By revealing how appearance-based cues shape emotional safety, evaluation, and motivation, the dissertation provides foundational knowledge for designing future systems that are empathetic, adaptive, and inclusive.

Beyond specific applications, the work contributes to a broader understanding of how humans construct social meaning in hybrid physical–digital environments. As AR, AI agents, and virtual communities become more pervasive, the ability to regulate social presence through appearance will be central to sustaining healthy interaction at scale.

6.9 Summary

This dissertation shows that the appearance of others is a powerful yet underexplored dimension of mediated interaction. By systematically investigating how appearance regulates social presence across evaluative and motivational contexts, the work advances both theoretical understanding and practical knowledge for designing psychologically supportive digital environments. It argues that future virtual and augmented systems should treat appearance not as an afterthought, but as a core component of social experience—one that shapes how people feel, perform, and relate to one another.

This dissertation demonstrates that the appearance of others in mediated environments is not a peripheral aesthetic choice but a core mechanism through which social presence, emotion, and performance are intertwined. By systematically showing that visual representations alone—without changes in behavior or authority—can alter users' sense of safety, pressure, and engagement, the work advances understanding of how social meaning is constructed in hybrid physical–digital contexts. The findings highlight the importance of designing mediated presence that is emotionally congruent, adaptable, and sensitive to individual differences such as trait anxiety. As virtual and augmented systems continue to integrate into everyday life, the insights from this research provide a foundational framework for creating interaction environments that support not only task performance, but also psychological well-being and sustainable engagement.

Looking to the Future As virtual and augmented environments become embedded in daily life, the question of how people relate to the appearance of others in mediated spaces will grow in both urgency and complexity. The next decade will likely see transformations that make today's mediated interactions feel rudimentary by comparison. AR glasses may replace smartphones; AI agents may become everyday social partners; and virtual presence may blur with physical intimacy. In such a landscape, designing the appearance layer of digital interaction will not be a peripheral aesthetic choice—it will be a core infrastructure of social experience.

One promising direction is the development of adaptive social interfaces that respond to users' emotional states in real time. Emerging physiological sensing technologies—such as wearable biosignals, eye-tracking, and ambient environmental sensors—will make it possible for AR companions or communication systems to modify appearance in ways that regulate emotion, redirect attention, or reduce stress. Instead of static avatars, future systems may generate context-sensitive representations that shift between supportive, neutral, or motivational forms depending on whether a user is anxious, fatigued, or seeking focus.

Another frontier lies in the domain of personalized social ecosystems. As individuals curate aesthetic identities and emotional preferences online, the appearance of others—friends, mentors, AI agents—may be customizable to a degree not possible today. Users might choose stylized presences for low-pressure engagement, realistic presences for intimacy or trust, and familiar presences for emotional grounding. Designing systems that support such diversity while ensuring ethical and psychological safety will be a major challenge for the coming years.

Future mediated environments will also need to navigate the tension between social closeness and social evaluation. As this dissertation shows, familiarity can both alleviate and intensify pressure; stylization can comfort or confuse. Understanding these dualities will be essential for building AR spaces that support individuals with diverse emotional needs, personality traits,

and cultural backgrounds. Particularly for users with high trait anxiety, future systems must prioritize nonjudgmental presence and control over social intensity, allowing users to tune the emotional weight of interactions.

From a societal standpoint, the increasing presence of AI companions and augmented human figures raises profound questions. How should emotionally supportive agents look? Who gets to design the default "faces" of future social technologies? How do we ensure that mediated appearances do not reinforce biases or narrow social norms? As mediated social presence becomes ubiquitous, issues of representation, diversity, and psychological inclusivity will need to be addressed with care.

Technologically, as AR hardware becomes lighter, more immersive, and more integrated into everyday life, the role of mediated appearance will extend beyond communication or exercise into education, healthcare, workplace collaboration, and personal well-being. Future research must investigate how appearance-based cues influence group dynamics, long-term mental health, and emergent community norms in hybrid virtual–physical worlds.

Looking forward, the insights from this dissertation provide an early foundation for shaping these futures. By demonstrating how the appearance of others affects emotion, attention, and performance, the work highlights the need for human-centered, emotion-aware design in the next generation of mediated systems. Ultimately, the future of AR and virtual communication will depend not only on technological advances, but on our ability to build systems that respect human vulnerability, support psychological safety, and create spaces where people—regardless of their emotional profile—can feel confident, motivated, and authentically connected.

Bibliography

[1] Alhirsan, S. M., Capó-Lugo, C. E., and Brown, D. A. (2021). Effects of different types of augmented feedback on intrinsic motivation and walking speed performance in post-stroke: A study protocol. *Contemporary Clinical Trials Communications*, 24:100863.

[2] Ambadar, Z., Cohn, J. F., and Reed, L. I. (2009). All smiles are not created equal: Morphology and timing of smiles perceived as amused, polite, and embarrassed/nervous. *Journal of nonverbal behavior*, 33:17–34.

[3] Anderson-Hanley, C., Snyder, A. L., Nimon, J. P., and Arciero, P. J. (2011). Social facilitation in virtual reality-enhanced exercise: competitiveness moderates exercise effort of older adults. *Clinical interventions in aging*, pages 275–280.

[4] Andrist, S., Mutlu, B., and Gleicher, M. (2013). Conversational gaze aversion for virtual agents. In *Intelligent Virtual Agents: 13th International Conference, IVA 2013, Edinburgh, UK, August 29-31, 2013. Proceedings 13*, pages 249–262. Springer.

[5] Azriel, O., Lazarov, A., Segal, A., and Bar-Haim, Y. (2020). Visual attention patterns during online video-mediated interaction in socially anxious individuals. *Journal of Behavior Therapy and Experimental Psychiatry*, 69:101595.

[6] Baayen, R. H., Davidson, D. J., and Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of memory and language*, 59(4):390–412.

[7] Bailenson, J. N., Blascovich, J., Beall, A. C., and Loomis, J. M. (2003). Interpersonal distance in immersive virtual environments. *Personality and social psychology bulletin*, 29(7):819–833.

[8] Baldwin, N., Branyon, J., Sethumadhavan, A., and Pak, R. (2015). In search of virtual social facilitation effects. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, volume 59, pages 90–94. SAGE Publications Sage CA: Los Angeles, CA.

[9] Baron, R. S. (1986). Distraction-conflict theory: Progress and problems. *Advances in experimental social psychology*, 19:1–40.

[10] Barrett, A., Pack, A., Monteiro, D., and Liang, H.-N. (2024). Exploring the influence of audience familiarity on speaker anxiety and performance in virtual reality and real-life presentation contexts. *Behaviour & Information Technology*, 43(4):787–799.

[11] Beck, A. T., Emery, G., and Greenberg, R. L. (1985). *Anxiety disorders and phobias: A cognitive perspective.* Basic Books/Hachette Book Group.

[12] Bodie, G. D. (2010). A racing heart, rattling knees, and ruminative thoughts: Defining, explaining, and treating public speaking anxiety. *Communication education*, 59(1):70–105.

[13] Boetje, J. and van Ginkel, S. (2021). The added benefit of an extra practice session in virtual reality on the development of presentation skills: A randomized control trial. *Journal of Computer Assisted Learning*, 37(1):253–264.

[14] Bond, C. F. and Titus, L. J. (1983). Social facilitation: A meta-analysis of 241 studies. *Psychological bulletin*, 94(2):265.

[15] Bood, R. J., Nijssen, M., Van Der Kamp, J., and Roerdink, M. (2013). The power of auditory-motor synchronization in sports: enhancing running performance by coupling cadence with the right beats. *PloS one*, 8(8):e70758.

[16] Booth, F. W., Roberts, C. K., and Laye, M. J. (2012). Lack of exercise is a major cause of chronic diseases. *Comprehensive physiology*, 2(2):1143.

[17] Borg, G. A. (1982). Psychophysical bases of perceived exertion. *Medicine and science in sports and exercise*, 14(5):377–381.

[18] Brown, B. R. and Garland, H. (1971). The effects of incompetency, audience acquaintanceship, and anticipated evaluative feedback on face-saving behavior. *Journal of experimental social psychology*, 7(5):490–502.

[19] Brubaker, J. R., Venolia, G., and Tang, J. C. (2012). Focusing on shared experiences: moving beyond the camera in video communication. In *Proceedings of the Designing Interactive Systems Conference*, pages 96–105.

[20] Burr, D. A., Pizzie, R. G., and Kraemer, D. J. (2021). Anxiety, not regulation tendency, predicts how individuals regulate in the laboratory: An exploratory comparison of self-report and psychophysiology. *PloS one*, 16(3):e0247246.

[21] Buss, A. H. (1980). Self-consciousness and social anxiety. *(No Title)*.

[22] Bystrom, K.-E., Barfield, W., and Hendrix, C. (1999). A conceptual model of the sense of presence in virtual environments. *Presence: Teleoperators & Virtual Environments*, 8(2):241–244.

[23] Caballo, V. E., Arias, B., Salazar, I. C., Irurtia, M. J., and Hofmann, S. G. (2015). Psychometric properties of an innovative self-report measure: The social anxiety questionnaire for adults. *Psychological assessment*, 27(3):997.

[24] Caine, K. (2016). Local standards for sample size at chi. In *Proceedings of the 2016 CHI conference on human factors in computing systems*, pages 981–992.

[25] CD, S. (1983). State-trait anxiety inventory. a comprehensive bibliography.

[26] Cho, Y. and Bianchi-Berthouze, N. (2019). Physiological and affective computing through thermal imaging: A survey. *arXiv preprint arXiv:1908.10307*.

[27] Chollet, M. and Scherer, S. (2017). Perception of virtual audiences. *IEEE computer graphics and applications*, 37(4):50–59.

[28] Chua, S. L., Chen, D.-T., and Wong, A. F. (1999). Computer anxiety and its correlates: a meta-analysis. *Computers in human behavior*, 15(5):609–623.

[29] Clancy, R. B., Herring, M. P., and Campbell, M. J. (2017). Motivation measures in sport: A critical review and bibliometric analysis. *Frontiers in psychology*, 8:348.

[30] Cottrell, N. B., Wack, D. L., Sekerak, G. J., and Rittle, R. H. (1968). Social facilitation of dominant responses by the presence of an audience and the mere presence of others. *Journal of personality and social psychology*, 9(3):245.

[31] Craven, M. (2015). *Effects of self-focused attention and fear of evaluation on anxiety and perception of speech performance*. PhD thesis, Ohio university.

[32] Da Silveira, M. P., da Silva Fagundes, K. K., Bizuti, M. R., Starck, É., Rossi, R. C., and de Resende e Silva, D. T. (2021). Physical exercise as a tool to help the immune system against covid-19: an integrative review of the current literature. *Clinical and experimental medicine*, 21(1):15–28.

[33] Daly, J. A., Stafford, L., et al. (1984). Correlates and consequences of social-communicative anxiety. *Avoiding communication: Shyness, reticence, and communication apprehension*, pages 125–143.

[34] Davis, A., Linvill, D. L., Hodges, L. F., Da Costa, A. F., and Lee, A. (2020). Virtual reality versus face-to-face practice: A study into situational apprehension and performance. *Communication Education*, 69(1):70–84.

[35] Day, M. L., McGuigan, M. R., Brice, G., and Foster, C. (2004). Monitoring exercise intensity during resistance training using the session rpe scale. *The Journal of Strength & Conditioning Research*, 18(2):353–358.

[36] Dictionary, O. (2010). Oxford dictionary of english. *Oxford Dictionary of English, 3rd edn. Oxford University Press. China Translation & Printing Services Ltd, China*.

[37] Dunbar, N. E. and Burgoon, J. K. (2005). Perceptions of power and interactional dominance in interpersonal relationships. *Journal of Social and Personal Relationships*, 22(2):207–233.

[38] Dwyer, K. K. and Davidson, M. M. (2012). Is public speaking really more feared than death? *Communication Research Reports*, 29(2):99–107.

[39] Edwards, A. M., Dutton-Challis, L., Cottrell, D., Guy, J. H., and Hettinga, F. J. (2018). Impact of active and passive social facilitation on self-paced endurance and sprint exercise: encouragement augments performance and motivation to exercise. *BMJ open sport & exercise medicine*, 4(1):e000368.

[40] Edworthy, J. and Waring, H. (2006). The effects of music tempo and loudness level on treadmill exercise. *Ergonomics*, 49(15):1597–1610.

[41] Emmerich, K. and Masuch, M. (2016). The influence of virtual agents on player experience and performance. In *Proceedings of the 2016 Annual Symposium on Computer-Human Interaction in Play*, pages 10–21.

[42] Emmerich, K. and Masuch, M. (2018). Watch me play: does social facilitation apply to digital games? In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, pages 1–12.

[43] Engert, V., Merla, A., Grant, J. A., Cardone, D., Tusche, A., and Singer, T. (2014). Exploring the use of thermal infrared imaging in human stress research. *PloS one*, 9(3):e90782.

[44] Ensari, I., Sandroff, B. M., and Motl, R. W. (2016). Effects of single bouts of walking exercise and yoga on acute mood symptoms in people with multiple sclerosis. *International journal of MS care*, 18(1):1–8.

[45] Far, I. K., Ferron, M., Ibarra, F., Baez, M., Tranquillini, S., Casati, F., and Doppio, N. (2015). The interplay of physical and social wellbeing in older adults: investigating the relationship between physical training and social interactions with virtual social environments. *PeerJ Computer Science*, 1:e30.

[46] Felnhofer, A., Hlavacs, H., Beutl, L., Kryspin-Exner, I., and Kothgassner, O. D. (2019). Physical presence, social presence, and anxiety in participants with social anxiety disorder during virtual cue exposure. *Cyberpsychology, Behavior, and Social Networking*, 22(1):46–50.

[47] FOX III, S. (1971). Physical activity and the prevention of coronary heart disease. *Ann Clin Res*, 3:404–432.

[48] Froming, W. J., Corley, E. B., and Rinker, L. (1990). The influence of public self-consciousness and the audience's characteristics on withdrawal from embarrassing situations. *Journal of Personality*, 58(4):603–622.

[49] Gardner, W. L. and Knowles, M. L. (2008). Love makes you real: Favorite television characters are perceived as "real" in a social facilitation paradigm. *Social Cognition*, 26(2):156–168.

[50] Giannakakis, G., Grigoriadis, D., Giannakaki, K., Simantiraki, O., Roniotis, A., and Tsiknakis, M. (2019). Review on psychological stress detection using biosignals. *IEEE Transactions on Affective Computing*, 13(1):440–460.

[51] Girondini, M., Stefanova, M., Pillan, M., and Gallace, A. (2023). Speaking in front of cartoon avatars: A behavioral and psychophysiological study on how audience design impacts on public speaking anxiety in virtual environments. *International Journal of Human-Computer Studies*, 179:103106.

[52] Glass, S. C. and Stanton, D. R. (2004). Self-selected resistance training intensity in novice weightlifters. *The Journal of Strength & Conditioning Research*, 18(2):324–327.

[53] Gong, Z. and Kanai, H. (2023). Comparison of two methods for altering the appearance of interviewers: Analysis of multiple biosignals. In *International Conference on Human-Computer Interaction*, pages 53–64. Springer.

[54] Gray, E., Beierl, E. T., and Clark, D. M. (2019). Sub-types of safety behaviours and their effects on social anxiety disorder. *Plos one*, 14(10):e0223165.

[55] Grillon, H., Riquier, F., Herbelin, B., and Thalmann, D. (2006). Virtual reality as a therapeutic tool in the confines of social anxiety disorder treatment. *International journal on disability and human development*, 5(3):243–250.

[56] Guthold, R., Stevens, G. A., Riley, L. M., and Bull, F. C. (2018). Worldwide trends in insufficient physical activity from 2001 to 2016: a pooled analysis of 358 population-based surveys with 1· 9 million participants. *The lancet global health*, 6(10):e1077–e1086.

[57] Hall, B. and Henningsen, D. D. (2008). Social facilitation and human–computer interaction. *Computers in human behavior*, 24(6):2965–2971.

[58] Hall, E. T. and Hall, E. T. (1966). *The hidden dimension*, volume 609. Anchor.

[59] Haller, J. C., Jang, Y. H., Haller, J., Shaw, L., and Wünsche, B. C. (2019). Hiit the road: Using virtual spectator feedback in hiit-based exergaming. In *Proceedings of the Australasian Computer Science Week Multiconference*, pages 1–9.

[60] Hamada, T., Hautasaari, A., Kitazaki, M., and Koshizuka, N. (2022). Solitary jogging with a virtual runner using smartglasses. In *2022 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*, pages 644–654. IEEE.

[61] Hamine, S., Gerth-Guyette, E., Faulx, D., Green, B. B., and Ginsburg, A. S. (2015). Impact of mhealth chronic disease management on treatment adherence and patient outcomes: a systematic review. *Journal of medical Internet research*, 17(2):e52.

[62] Hanna, J. E. and Brennan, S. E. (2007). Speakers' eye gaze disambiguates referring expressions early during face-to-face conversation. *Journal of Memory and Language*, 57(4):596–615.

[63] Harris, S. R., Kemmerling, R. L., and North, M. M. (2002). Brief virtual reality therapy for public speaking anxiety. *Cyberpsychology & behavior*, 5(6):543–550.

[64] Hayes, A. L., Ulinski, A. C., and Hodges, L. F. (2010). That avatar is looking at me! social inhibition in virtual worlds. In *Intelligent Virtual Agents: 10th International Conference, IVA 2010, Philadelphia, PA, USA, September 20-22, 2010. Proceedings 10*, pages 454–467. Springer.

[65] Hewitt, E. and Stephenson, J. (2012). Foreign language anxiety and oral exam performance: A replication of phillips's mlj study. *The Modern Language Journal*, 96(2):170–189.

[66] Hietanen, J. O., Peltola, M. J., and Hietanen, J. K. (2020). Psychophysiological responses to eye contact in a live interaction and in video call. *Psychophysiology*, 57(6):e13587.

[67] Hirsch, C., Meynen, T., and Clark, D. (2004). Negative self-imagery in social anxiety contaminates social interactions. *Memory*, 12(4):496–506.

[68] Hoyt, C. L., Blascovich, J., and Swinth, K. R. (2003). Social inhibition in immersive virtual environments. *Presence*, 12(2):183–195.

[69] Hwang, A. H.-C., Wang, C. Y., Yang, Y.-Y., and Won, A. S. (2021). Hide and seek: Choices of virtual backgrounds in video chats and their effects on perception. *Proceedings of the ACM on Human-Computer Interaction*, 5(CSCW2):1–30.

[70] Iyamuremye, E., Ndayambaje, I., and Muwonge, C. M. (2023). Relationships of mathematics achievement with self-determined motivation and mathematics anxiety among senior two students in northern rwanda. *Heliyon*, 9(4).

[71] Kalra, R. and Siribud, S. (2020). Public speaking anxiety in the thai efl context. *LEARN Journal: Language Education and Acquisition Research Network*, 13(1):195–209.

[72] Kana Naruse, Shigeo Yoshida, e. a. (2018). Preliminary investigation of the relaxation effects through real-time transformed auditory feedback (japanese). *Entertainment Computing(EC)*, 2018(17):1–8.

[73] Kang, N., Brinkman, W.-P., Van Riemsdijk, M. B., and Neerincx, M. (2016). The design of virtual audiences: noticeable and recognizable behavioral styles. *Computers in Human Behavior*, 55:680–694.

[74] Kang, N., Brinkman, W.-P., van Riemsdijk, M. B., and Neerincx, M. A. (2013). An expressive virtual audiencewith flexible behavioral styles. *IEEE Transactions on Affective Computing*, 4(4):326–340.

[75] Kaptelinin, V., Danielsson, K., Kaiser, N., Kuenen, C., and Nordin, M. (2021). Understanding the interpersonal space of online meetings: An exploratory study of "we-ness". In *Companion Publication of the 2021 Conference on Computer Supported Cooperative Work and Social Computing*, pages 79–83.

[76] Kätsyri, J., Förger, K., Mäkäräinen, M., and Takala, T. (2015). A review of empirical evidence on different uncanny valley hypotheses: support for perceptual mismatch as one road to the valley of eeriness. *Frontiers in psychology*, 6:390.

[77] Kazdin, A. E., Association, A. P., et al. (2000). *Encyclopedia of psychology*, volume 8. American Psychological Association Washington, DC.

[78] Kim, J., Ka, J., Lee, Y., Lee, Y., Park, S., and Kim, W. (2022). Mixed reality-based outdoor training system to improve football player performance. In *2022 International Conference on Engineering and Emerging Technologies (ICEET)*, pages 1–3. IEEE.

[79] King, P. E. and Finn, A. N. (2017). A test of attention control theory in public speaking: Cognitive load influences the relationship between state anxiety and verbal production. *Communication Education*, 66(2):168–182.

[80] Kowatsch, T., Lohse, K.-M., Erb, V., Schittenhelm, L., Galliker, H., Lehner, R., and Huang, E. M. (2021). Hybrid ubiquitous coaching with a novel combination of mobile and

holographic conversational agents targeting adherence to home exercises: four design and evaluation studies. *Journal of medical Internet research*, 23(2):e23612.

[81] Krueger, C. and Tian, L. (2004). A comparison of the general linear mixed model and repeated measures anova using a dataset with multiple missing data points. *Biological research for nursing*, 6(2):151–157.

[82] Lai, K., Liu, Y., Iino, T., and Fujinami, T. (2022). Analysis on mechanisms of idea generation: evidences from fnirs hyperscanning. In *Proceedings of the 6th International Conference on Medical and Health Informatics*, pages 176–182.

[83] Laird, N. M. and Ware, J. H. (1982). Random-effects models for longitudinal data. *Biometrics*, pages 963–974.

[84] Landers, D. M., Snyder Bauer, R., and Feltz, D. L. (1978). Social facilitation during the initial stage of motor learning: a re-examination of martens' audience study. *Journal of Motor Behavior*, 10(4):325–337.

[85] Langer, J. K., Lim, M. H., Fernandez, K. C., and Rodebaugh, T. L. (2017). Social anxiety disorder is associated with reduced eye contact during conversation primed for conflict. *Cognitive therapy and research*, 41:220–229.

[86] Latané, B. (1981). The psychology of social impact. *American psychologist*, 36(4):343.

[87] Leeds, D. (2003). *Power speak: Engage, inspire, and stimulate your audience*. Red Wheel/Weiser.

[88] Leong, J., Perteneder, F., Rajvee, M. R., and Maes, P. (2023). "picture the audience...": Exploring private ar face filters for online public speaking. In *Proceedings of the 2023 CHI conference on human factors in computing systems*, pages 1–13.

[89] Li, B. J., Ratan, R., and Lwin, M. O. (2022). Virtual game changers: How avatars and virtual coaches influence exergame outcomes through enactive and vicarious learning. *Behaviour & Information Technology*, 41(7):1529–1543.

[90] Liguori, G., of Sports Medicine, A. C., et al. (2020). *ACSM's guidelines for exercise testing and prescription*. Lippincott Williams & Wilkins.

[91] MacIntyre, P. D. and Thivierge, K. A. (1995). The effects of audience pleasantness, audience familiarity, and speaking contexts on public speaking anxiety and willingness to speak. *Communication Quarterly*, 43(4):456–466.

[92] MacIntyre, P. D., Thivierge, K. A., and MacDonald, J. R. (1997). The effects of audience interest, responsiveness, and evaluation on public speaking anxiety and related variables. *Communication research reports*, 14(2):157–168.

[93] Maeda, S. (2023). No differential responsiveness to face-to-face communication and video call in individuals with elevated social anxiety. *Journal of Affective Disorders Reports*, 11:100467.

[94] Marcolino, M. S., Oliveira, J. A. Q., D'Agostino, M., Ribeiro, A. L., Alkmim, M. B. M., and Novillo-Ortiz, D. (2018). The impact of mhealth interventions: systematic review of systematic reviews. *JMIR mHealth and uHealth*, 6(1):e8873.

[95] Marín-Lora, C., Chover, M., Martín, M. Y., and García-Rytman, L. (2023). Comparative study of interaction methods for mobile gaming while running on a treadmill. *Computers & Graphics*, 117:164–171.

[96] Marinho, A. C. F., Medeiros, A. M. d., Lima, E. d. P., Pantuza, J. J., and Teixeira, L. C. (2019). Prevalence and factors associated with fear of public speaking. In *CoDAS*, volume 31, page e20180266. SciELO Brasil.

[97] McAuley, E., Duncan, T., and Tammen, V. V. (1989). Psychometric properties of the intrinsic motivation inventory in a competitive sport setting: A confirmatory factor analysis. *Research quarterly for exercise and sport*, 60(1):48–58.

[98] McCambridge, J., Witton, J., and Elbourne, D. R. (2014). Systematic review of the hawthorne effect: new concepts are needed to study research participation effects. *Journal of clinical epidemiology*, 67(3):267–277.

[99] McCroskey, J. C. (1984). The communication apprehension perspective. *Avoiding communication: Shyness, reticence, and communication apprehension*, pages 13–38.

[100] McKendrick, R. D. and Cherry, E. (2018). A deeper look at the nasa tlx and where it falls short. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, volume 62, pages 44–48. SAGE Publications Sage CA: Los Angeles, CA.

[101] Monteiro, D., Liang, H.-N., Li, H., Fu, Y., and Wang, X. (2020). Evaluating the need and effect of an audience in a virtual reality presentation training tool. In *Computer Animation and Social Agents: 33rd International Conference on Computer Animation and Social Agents, CASA 2020, Bournemouth, UK, October 13-15, 2020, Proceedings 33*, pages 62–70. Springer.

[102] Monteiro, D., Wang, A., Wang, L., Li, H., Barrett, A., Pack, A., and Liang, H.-N. (2023). Effects of audience familiarity on anxiety in a virtual reality public speaking training tool. *Universal Access in the Information Society*, pages 1–12.

[103] Mori, M. (1970). Bukimi no tani [the uncanny valley]. *Energy*, 7:33.

[104] Mostajeran, F., Balci, M. B., Steinicke, F., Kühn, S., and Gallinat, J. (2020). The effects of virtual audience size on social anxiety during public speaking. In *2020 IEEE conference on virtual reality and 3D user interfaces (VR)*, pages 303–312. IEEE.

[105] Murali, P., Trinh, H., Ring, L., and Bickmore, T. (2021). A friendly face in the crowd: Reducing public speaking anxiety with an emotional support agent in the audience. In *Proceedings of the 21st ACM International Conference on Intelligent Virtual Agents*, pages 156–163.

[106] Murray, E. G., Neumann, D. L., Moffitt, R. L., and Thomas, P. R. (2016). The effects of the presence of others during a rowing exercise in a virtual reality environment. *Psychology of Sport and Exercise*, 22:328–336.

[107] Nam, H., Lee, K., Park, J.-I., and Kim, K. (2025). Ar fitness dog: Effects of a user-mimicking interactive virtual pet on user experience and social presence in physical exercise. *IEEE Transactions on Visualization and Computer Graphics*.

[108] Nekar, D. M., Kang, H. Y., and Yu, J. H. (2022). Improvements of physical activity performance and motivation in adult men through augmented reality approach: a randomized controlled trial. *Journal of Environmental and Public Health*, 2022(1):3050424.

[109] Nesse, R. M. et al. (1994). Fear and fitness: An evolutionary analysis of anxiety disorders. *Ethology and sociobiology*, 15(5-6):247–261.

[110] Neumann, D. L., Moffitt, R. L., Thomas, P. R., Loveday, K., Watling, D. P., Lombard, C. L., Antonova, S., and Tremeer, M. A. (2018). A systematic review of the application of interactive virtual reality to sport. *Virtual Reality*, 22:183–198.

[111] Nhan, B. R. and Chau, T. (2009). Classifying affective states using thermal infrared imaging of the human face. *IEEE Transactions on Biomedical Engineering*, 57(4):979–987.

[112] Norozi, K., Wessel, A., Alpers, V., Arnhold, J. O., Binder, L., Geyer, S., Zoege, M., and Buchhorn, R. (2007). Chronotropic incompetence in adolescents and adults with congenital heart disease after cardiac surgery. *Journal of cardiac failure*, 13(4):263–268.

[113] Nozawa, A. and Tacano, M. (2009). Correlation analysis on alpha attenuation and nasal skin temperature. *Journal of Statistical Mechanics: Theory and Experiment*, 2009(01):P01007.

[114] Nozawa, T., Sasaki, Y., Sakaki, K., Yokoyama, R., and Kawashima, R. (2016). Interpersonal frontopolar neural synchronization in group communication: an exploration toward fnirs hyperscanning of natural interactions. *Neuroimage*, 133:484–497.

[115] Oh, C. S., Bailenson, J. N., and Welch, G. F. (2018). A systematic review of social presence: Definition, antecedents, and implications. *Frontiers in Robotics and AI*, 5:114.

[116] O'Malley, C., Langton, S., Anderson, A., Doherty-Sneddon, G., and Bruce, V. (1996). Comparison of face-to-face and video-mediated interaction. *Interacting with computers*, 8(2):177–192.

[117] Orr, E. M. and Moscovitch, D. A. (2010). Learning to re-appraise the self during video feedback for social anxiety: Does depth of processing matter? *Behaviour Research and Therapy*, 48(8):728–737.

[118] Pan, X. and Hamilton, A. F. d. C. (2015). Automatic imitation in a rich social context with virtual characters. *Frontiers in psychology*, 6:790.

[119] Park, S. and Catrambone, R. (2007). Social facilitation effects of virtual humans. *Human factors*, 49(6):1054–1060.

[120] Pedrinolla, A., Schena, F., and Venturelli, M. (2017). Resilience to alzheimer's disease: the role of physical activity. *Current Alzheimer Research*, 14(5):546–553.

[121] Pérez Castillejo, S. (2019). The role of foreign language anxiety on l2 utterance fluency during a final exam. *Language Testing*, 36(3):327–345.

[122] Piccirillo, M. L., Dryman, M. T., and Heimberg, R. G. (2016). Safety behaviors in adults with social anxiety: Review and future directions. *Behavior Therapy*, 47(5):675–687.

[123] Poeschl, S. (2017). Virtual reality training for public speaking—a quest-vr framework validation. *Frontiers in ICT*, 4:13.

[124] Ranasinghe, C., Ozemek, C., and Arena, R. (2020). Exercise and well-being during covid 19–time to boost your immunity. *Expert review of anti-infective therapy*, 18(12):1195–1200.

[125] Rapee, R. M. and Lim, L. (1992). Discrepancy between self-and observer ratings of performance in social phobics. *Journal of abnormal psychology*, 101(4):728.

[126] Reiner, M., Niermann, C., Jekauc, D., and Woll, A. (2013). Long-term health benefits of physical activity–a systematic review of longitudinal studies. *BMC public health*, 13:1–9.

[127] Reiss, S. (1985). The expectancy model of fear. *Theoretical issue in behavior therapy*, pages 107–121.

[128] Rhea, M. R., Landers, D. M., Alvar, B. A., and Arent, S. M. (2003). The effects of competition and the presence of an audience on weight lifting performance. *The Journal of Strength & Conditioning Research*, 17(2):303–306.

[129] Rodebaugh, T. L. (2004). I might look ok, but i'm still doubtful, anxious, and avoidant: The mixed effects of enhanced video feedback on social anxiety symptoms. *Behaviour Research and Therapy*, 42(12):1435–1451.

[130] Ryan, R. M. (1982). Control and information in the intrapersonal sphere: An extension of cognitive evaluation theory. *Journal of personality and social psychology*, 43(3):450.

[131] Schneider, S. and Kümmert, F. (2016). Exercising with a humanoid companion is more effective than exercising alone. In *2016 IEEE-RAS 16th International Conference on Humanoid Robots (Humanoids)*, pages 495–501. IEEE.

[132] Short, J., Williams, E., and Christie, B. (1976). The social psychology of telecommunications. *(No Title)*.

[133] Slater, M., Pertaub, D.-P., Barker, C., and Clark, D. M. (2006). An experimental study on fear of public speaking using a virtual environment. *CyberPsychology & Behavior*, 9(5):627–633.

[134] Slater, M., Pertaub, D.-P., and Steed, A. (1999). Public speaking in virtual reality: Facing an audience of avatars. *IEEE Computer Graphics and Applications*, 19(2):6–9.

[135] Smith, E. R., Semin, G. R., et al. (2004). Socially situated cognition: Cognition in its social context. *Advances in experimental social psychology*, 36:57–121.

[136] Soltani, P. and Morice, A. H. (2020). Augmented reality tools for sports education and training. *Computers & Education*, 155:103923.

[137] Spielberger, C. D., Gonzalez-Reigosa, F., Martinez-Urrutia, A., Natalicio, L. F., and Natalicio, D. S. (1971). The state-trait anxiety inventory. *Revista Interamericana de Psicologia/Interamerican journal of psychology*, 5(3 & 4).

[138] Stein, M. B., Baird, A., and Walker, J. R. (1996). Social phobia in adults with stuttering. *The American journal of psychiatry*, 153(2):278–280.

[139] Sterna, R., Strojny, P., and Rębilas, K. (2019). Can virtual observers affect our behavior?: social facilitation in virtual environments: a mini-review. *Placeholder Journal*.

[140] Stork, M. J., Kwan, M. Y., Gibala, M. J., and Ginis, K. A. M. (2015). Music enhances performance and perceived enjoyment of sprint interval exercise. *Medicine & Science in Sports & Exercise*, 47(5):1052–1060.

[141] Strauss, B. (2002). The impact of supportive spectator behavior on performance in team sports. *International Journal of Sport Psychology*.

[142] Strauss, B., Staufenbiel, K., van Meurs, E., and MacMahon, C. (2023). Social influence of sport spectators. In *Sport and Exercise Psychology: Theory and Application*, pages 425–444. Springer.

[143] Strojny, P. M., Dużmańska-Misiarczyk, N., Lipp, N., and Strojny, A. (2020). Moderators of social facilitation effect in virtual reality: Co-presence and realism of virtual agents. *Frontiers in psychology*, 11:1252.

[144] Strube, M. J., Miles, M. E., and Finch, W. H. (1981). The social facilitation of a simple task: Field tests of alternative explanations. *Personality and Social Psychology Bulletin*, 7(4):701–707.

[145] Stupar-Rutenfrans, S., Ketelaars, L. E., and van Gisbergen, M. S. (2017). Beat the fear of public speaking: Mobile 360 video virtual reality exposure training in home environment reduces public speaking anxiety. *Cyberpsychology, Behavior, and Social Networking*, 20(10):624–633.

[146] Thornby, M. A., Haas, F., and Axen, K. (1995). Effect of distractive auditory stimuli on exercise tolerance in patients with copd. *Chest*, 107(5):1213–1217.

[147] Tomić, B., Stojanović, T., and Antović, I. (2022). Examining students'test anxiety and pre-university programming education in an undergraduate introductory programming course. In *EDULEARN22 Proceedings*, pages 3848–3857. IATED.

[148] Tran, D. L., Kamaladasa, Y., Munoz, P. A., Kotchetkova, I., D'Souza, M., Celermajer, D. S., Maiorana, A., and Cordina, R. (2022). Estimating exercise intensity using heart rate in adolescents and adults with congenital heart disease: Are established methods valid? *International Journal of Cardiology Congenital Heart Disease*, 8:100362.

[149] Tudor, A.-D., Poeschl, S., and Doering, N. (2013). Virtual audience customization for public speaking training procedures. In *2013 IEEE Virtual Reality (VR)*, pages 61–62. IEEE.

[150] Tzetzis, G., Votsis, E., and Kourtessis, T. (2008). The effect of different corrective feedback methods on the outcome and self confidence of young athletes. *Journal of sports science & medicine*, 7(3):371.

[151] Villavicencio, F. T. and Bernardo, A. B. (2016). Beyond math anxiety: Positive emotions predict mathematics achievement, self-regulation, and self-efficacy. *The Asia-Pacific Education Researcher*, 25:415–422.

[152] Vinkers, C. H., Penning, R., Hellhammer, J., Verster, J. C., Klaessens, J. H., Olivier, B., and Kalkman, C. J. (2013). The effect of stress on core and peripheral body temperature in humans. *Stress*, 16(5):520–530.

[153] Von Glasersfeld, E. (2013). *Radical constructivism*, volume 6. Routledge.

[154] Vygotsky, L. S. and Cole, M. (1978). *Mind in society: Development of higher psychological processes*. Harvard university press.

[155] Webb, T. L., Ononaiye, M. S., Sheeran, P., Reidy, J. G., and Lavda, A. (2010). Using implementation intentions to overcome the effects of social anxiety on attention and appraisals of performance. *Personality and Social Psychology Bulletin*, 36(5):612–627.

[156] Westgarth, M. M., Hogan, C. A., Neumann, D. L., and Shum, D. H. (2021). A systematic review of studies that used nirs to measure neural activation during emotion processing in healthy individuals. *Social Cognitive and Affective Neuroscience*, 16(4):345–369.

[157] Wilson, K. S. and Spink, K. S. (2009). Social influence and physical activity in older females: Does activity preference matter? *Psychology of sport and exercise*, 10(4):481–488.

[158] Woody, S. R. (1996). Effects of focus of attention on anxiety levels and social performance of individuals with social phobia. *Journal of abnormal psychology*, 105(1):61.

[159] Worringham, C. J. and Messick, D. M. (1983). Social facilitation of running: An unobtrusive study. *The Journal of social psychology*, 121(1):23–29.

[160] Xu, W., Yu, K., Meng, X., Monteiro, D., Kao, D., and Liang, H.-N. (2023). Exploring the effect of the group size and feedback of non-player character spectators in virtual reality exergames. *Frontiers in Psychology*, 14:1079132.

[161] Yerkes, R. M., Dodson, J. D., et al. (1908). The relation of strength of stimulus to rapidity of habit-formation. *Journal of comparative neurology and psychology*, 18(5):459–482.

[162] Yoneyama, J., Fujimoto, Y., Okazaki, K., Sawabe, T., Kanbara, M., and Kato, H. (2023). Augmented reality visual effects for mitigating anxiety of in-person communication for individuals with social anxiety disorder. In *APMAR*.

[163] Yoneyama, J., Fujimoto, Y., Okazaki, K., Sawabe, T., Kanbara, M., and Kato, H. (2024). Augmented conversations: Ar face filters for facilitating comfortable in-person interactions. *Journal on Multimodal User Interfaces*, pages 1–18.

[164] Yu, K., Wen, S., Xu, W., Caon, M., Baghaei, N., and Liang, H.-N. (2023). Cheer for me: effect of non-player character audience feedback on older adult users of virtual reality exergames. *Virtual Reality*, 27(3):1887–1903.

[165] Yusuf, A. and Yusuf Muhammad, A. (2024). Exploring clusters of novice programmers' anxiety-induced behaviors during block-and text-based coding: A predictive and moderation analysis of programming quality and error debugging skills. *Journal of Educational Computing Research*, 62(7):1798–1836.

[166] Zajonc, R. B. (1965). Social facilitation. *Science*, 149(3681):269–274.

[167] Zanbaka, C. A., Ulinski, A. C., Goolkasian, P., and Hodges, L. F. (2007). Social responses to virtual humans: implications for future interface design. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 1561–1570.

[168] Zhao, Z., Arya, A., Whitehead, A. D., Chan, G., and Etemad, S. A. (2017). Keeping users engaged through feature updates: A long-term study of using wearable-based exergames. In *CHI*, pages 1053–1064.