

# Who's Watching Me?: Exploring the Impact of Audience Familiarity on Player Performance, Experience, and Exertion in Virtual Reality Exergames

Zixuan Guo   
Xi'an Jiaotong-Liverpool  
University

Wenge Xu   
Birmingham City University

Jialin Zhang   
Xi'an Jiaotong-Liverpool  
University

Hongyu Wang   
Xi'an Jiaotong-Liverpool  
University

Cheng-Hung Lo   
Xi'an Jiaotong-Liverpool  
University

Hai-Ning Liang   
Xi'an Jiaotong-Liverpool  
University

## ABSTRACT

Familiarity with audiences plays a significant role in shaping individual performance and experience across various activities in everyday life. This study delves into the impact of familiarity with non-playable character (NPC) audiences on player performance and experience in virtual reality (VR) exergames. By manipulating of NPC appearance (face and body shape) and voice familiarity, we explored their effect on game performance, experience, and exertion. The findings reveal that familiar NPC audiences have a positive impact on performance, creating a more enjoyable gaming experience, and leading players to perceive less exertion. Moreover, individuals with higher levels of self-consciousness exhibit heightened sensitivity to the familiarity with NPC audiences. Our results shed light on the role of familiar NPC audiences in enhancing player experiences and provide insights for designing more engaging and personalized VR exergame environments.

**Index Terms:** Human-centered computing—Human computer interaction (HCI)—Interaction paradigms—Virtual reality; Software and its engineering—Software organization and properties—Virtual worlds software—Interactive games

## 1 INTRODUCTION

Virtual Reality (VR) exergames have gained significant popularity as an innovative approach that combines physical exercise with immersive gaming experiences [43,45]. As sedentary lifestyles have become increasingly prevalent, with prolonged sitting leading to various health concerns, VR exergames offer a promising solution by promoting physical movement and mitigating the negative effects associated with excessive sitting [51]. These interactive virtual environments allow users to engage in physical activities while being transported to fun and captivating virtual worlds [44,46].

Non-Player Characters (NPC) and their feedback have been used in games to enhance gameplay experiences [16,50]. However, there is little research on the mechanisms for establishing meaningful connections between these audiences and players, particularly in VR exergames. Tice et al. [38] argued that how individuals present themselves is influenced by their audience; specifically, individuals may behave differently when facing familiar and unfamiliar audiences [27,40]. Furthermore, several studies [6,32] have supported the positive influence of familiar audiences on individual performance, as encountering familiar individuals has been shown to evoke a sense of connection, comfort, and heightened engagement. Conversely, encountering unfamiliar individuals may elicit heightened nervousness or anxiety, but sometimes it can also stimulate individuals to have a stronger desire to perform and excel [22]. Research suggests that the extent of an audience's influence on users

depends on an individual's level of self-consciousness [2,7,39]. Individuals with high self-consciousness are more concerned about others' judgments and experience increased anxiety when being observed, which can affect their performance and emotions.

A familiar audience has been shown to enhance social presence and increase user motivation and enjoyment in various contexts, including virtual environments. For instance, recent studies [1,25] have demonstrated that when delivering a speech in a VR environment and facing familiar audiences, users tend to experience a greater sense of relaxation and exhibit improved performance. However, the impact of audience familiarity on player experiences in gaming contexts remains relatively unexplored.

Moreover, the role of voice familiarity has been largely overlooked. Notably, audience familiarity can be perceived through different patterns of signals originating not only from facial features but also from voices [49]. Familiar appearances and voices can convey identity information and are believed to create impressions of warmth, competence, and other social characteristics [10,34]. Understanding the influence of audience familiarity in VR exergames is essential for creating engaging and personalized experiences. Therefore, the primary objective of this work is to explore how familiarity with NPC audiences, encompassing both appearance (facial features and body shape) and voice, impacts players' performance, experience, and exertion within the context of VR exergames. Specifically, we seek to answer the following research questions:

- RQ1:** How would the familiar/unfamiliar appearances of NPC audiences affect players' performance, experience, and exertion in VR exergames?
- RQ2:** How would the familiar/unfamiliar voices of NPC audiences impact players' performance, experience, and exertion in VR exergames?
- RQ3:** How does the level of self-consciousness relate to the impact of audience familiarity on players' performance, experience, and exertion in VR exergames?

To answer these RQs, we conducted a  $2 \times 2$  within-subjects experiment (1. Audience Appearance—unfamiliar and familiar and 2. Audience Voice—unfamiliar and familiar) with 24 young adults participating in a VR exergame. Our findings point to the positive impact of both familiar audience appearance and voice on players' overall gaming experience, emotional state, and perceived exertion. Furthermore, the familiar voice enhances player performance in the game. Additionally, individuals with higher levels of self-consciousness demonstrate a heightened sensitivity to the familiarity with NPC audiences, leading to a greater sense of relaxation when facing familiar audiences. We provide two design recommendations incorporating audience familiarity to support designers and developers in shaping players' game performance and experience, ultimately contributing to a healthier lifestyle.

\*Corresponding author (e-mail: haining.liang@xjtlu.edu.cn)

## 2 RELATED WORK

### 2.1 NPC Audience and Its Impact on Users

NPC audiences can have a great impact on users' performance and experience. For instance, Pertaub et al. [29] found that negative virtual audiences triggered heightened anxiety levels among users during public speaking, in contrast to positive and neutral audiences. Yakura and Goto [48] discovered that when users are part of the concert audience, the behavior of other concert NPC audiences could significantly enhance their sense of co-presence and improve their overall concert experience. Strojny et al. [36] studied VR training for firefighter first response and found that social presence plays a crucial role in moderating the impact of virtual audiences on performance in such scenarios. Kabeko et al. [19] found NPCs' feedback can enhance the user's experience and create a sense of "unity" with the audiences in a VR live house. Raimbaud et al. [30] investigated the gaze behavior of virtual audiences in VR, revealing that participants demonstrated quicker detection and prolonged gaze duration in response to directed gaze, mirroring real-world tendencies.

Prior research has also explored the influence of NPC audiences on video games. For example, Haller et al. [16] suggested that feedback from a virtual audience could lead to better performance and higher heart rates in players. Xu et al. [47] found that the group size of an NPC audience and its feedback could positively affect a player's performance and experience while playing VR exergames. Focusing on older players, Yu et al. [50] found that virtual audiences with reactive feedback could enhance their performance and experience in VR exergames. Yet, no studies have focused on the investigation of audience familiarity specifically in the context of VR games, including VR exergame. This work fills this gap.

### 2.2 Users' Sense of Familiarity and Its Impact

From a psychological standpoint, familiarity can be defined as the level of knowledge or acquaintance that an individual possesses regarding specific stimuli, such as physical stimuli, environmental elements, and other individuals [15]. For other individuals, familiarity can be perceived on a continuum, ranging from complete strangers, and acquaintances, to close friends and family.

Existing research has demonstrated a link between sensory familiarity and the experience of positive affect [13, 14, 17]. For visual stimuli, Harmon-Jones and Allen [17] found that participants exhibited greater activity in facial muscles associated with positive affect after being exposed to familiar photos than unfamiliar photos. Garcia-Marques et al. [13] demonstrated that familiarity and positive affect influence each other's judgment latencies through a symbol judgment task. In the realm of auditory stimuli, Garcia-Marques et al. [14] delved into the evaluation of recorded statements and unveiled that the processing of familiar stimuli resulted in a pleasurable subjective experience.

When confronted with stimuli beyond simple sensory input, such as familiar people, individuals can also demonstrate more positive performance and psychological states. For example, MacIntyre and Thivierge [23] found that in public speaking, facing familiar audiences can make participants perceive stronger agreeableness and willingness to speak and lower anticipated anxiety than unfamiliar audiences. Burkitt et al. [6] investigated the effect of audience familiarity on children's expressive drawings of themselves and found that children showed more positive expressiveness when faced with familiar peers and adult audiences. Monteiro et al. [25] investigated the influence of audience familiarity in a VR public speaking training tool and found that participants with medium public speaking anxiety exhibited a significantly more relaxed demeanor when faced with familiar NPC audiences.

These studies provide evidence that familiar audiences have the potential to induce a more relaxed mood and facilitate participants' performance. However, there is currently limited research that incorporates the familiarity with the NPC audiences within virtual

gaming environments. Furthermore, our understanding of the influence of visual and auditory stimuli presented by audiences remains incomplete. To address these gaps, this study aims to investigate how familiarity with the NPC audience's appearance (including face and body shape) and voice influence players' performance and overall experience in VR exergames.

### 2.3 Self-consciousness

While the presence of an audience is often believed to enhance individual performance, it is essential to consider the role of self-consciousness, as it can significantly impact how individuals respond to audience interactions [3]. Self-consciousness refers to an individual's awareness and attention to their own thoughts, feelings, and behaviors in relation to others [12]. It involves a focus on how one is perceived by others and a concern for social evaluation.

Several studies [2, 7, 39] have explored the relationship between self-consciousness and the influence of audiences on personal performance. For example, Wang et al. [39] suggested that individuals with high levels of self-consciousness may experience heightened anxiety and decreased performance in front of an audience. This effect appears to be more pronounced when individuals are confronting unfamiliar audiences. Mesagno et al. [24] observed that basketball players with high self-consciousness showed significantly increased anxiety and diminished performance when facing unfamiliar audiences. In these instances, the presence of an audience may lead to self-doubt and increased self-monitoring, which can disrupt performance and hinder task execution [2]. Existing research [16, 20, 50] on audience familiarity in virtual environments has paid limited attention to users' self-consciousness. Considering that self-consciousness may be a crucial influencing factor, we aim to explore the relationship between players' self-consciousness and their exposure to NPC audience familiarity in this work.

## 3 EXERGAME APPLICATION

We have developed a VR exergame, which is inspired by Fruit Ninja, using the Unity3D engine. Expanding on the base game, our design includes a fog interference feature, encouraging players to squat and clear the fog, boosting exertion, and adding a challenging aspect. Furthermore, our game includes virtual audiences to enrich the player experience. All variables reported in this section were decided through several playtesting sessions with four game testers.

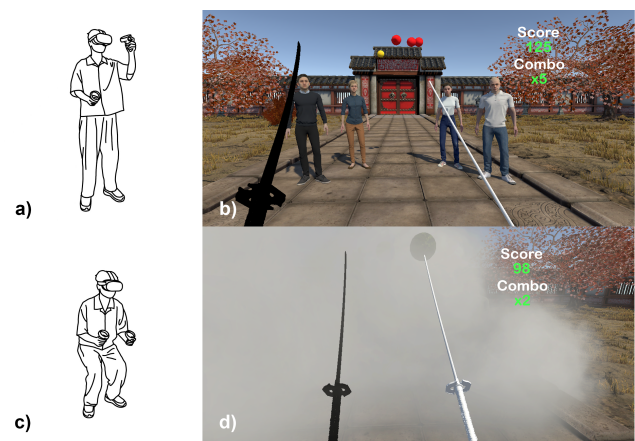


Figure 1: (a) Real and (b) Virtual perspectives of fruit slicing gameplay using controllers. (c) Real and (d) Virtual perspectives of squatting to clear fog interference.

### 3.1 Rules and Logic

During the game, players are required to use two hand-held controllers to wield virtual swords to slice as many fruits (watermelon, apple, and lemon) as possible and dodge bombs. Figure 1 shows that fruits and bombs are launched towards the left and right sides of the player, following a parabolic trajectory that enables them to land within the reachable area of the player’s virtual sword. To ensure trajectory reliability, we first adjusted the gravity setting based on feedback from the four game testers of different heights (160cm - 185cm), thereby extending the flight time of the targets. We also adjusted the initial velocities on the X, Y, and Z axes to restrict the angle range, allowing participants to hit the targets within a consistent area of approximately one square meter.

The game lasts 5 minutes and 12 seconds, composed of 5 one-minute game sequences with 3 seconds of rest time between each game sequence. Each one-minute sequence consists of 20 3-second rounds that generate a fixed number of fruits and bombs in random order. Each round contains 2-4 fruits, with a bomb appearing every 4-5 rounds. In total, there are 100 rounds (20 per minute), and the occurrences of each target are as follows: (1) watermelon: 90, (2) apple: 80, (3) lemon: 85, and (4) bomb: 20.

The game also includes a fog interference feature which obstructs the player’s view to elevate the challenge and variation of the gameplay. The player needs to squat to disperse the fog and regain their vision (Figure 1). We utilized the Steam effects from Unity’s Particle pack to create the fog, with a simulation speed of 1 and a resulting shape approximating a sphere with a diameter of 1.8 meters. We placed four white-gray smoke effects around the camera to completely obstruct the player’s view. The fog generation frequency was set at 6-10 seconds, resulting in the appearance of fog approximately 25-35 times per game, based on the player’s 3-second reaction and squatting duration. This ensured the fog interference mechanism was sufficiently challenging without compromising the players’ overall gaming experience.

### 3.2 Audience Setting

To ensure optimal visibility of each NPC audience’s faces, we positioned four NPCs as audiences in the player’s field of view. These NPCs provided both positive and negative feedback to the player, with each feedback consisting of two different animations and two voice lines (Figure 2). The two positive feedback animations include raising both hands and jumping while waving one hand upward when cheering for the player. The two negative feedback animations consist of covering the face with both hands and lowering the head while waving one hand downward. The positive voice lines are delivered with excitement and joy, while the negative voice lines convey a sense of disappointment and frustration. To enable participants to identify the voices of the NPCs, their voice feedback was set to first emit some exclamations (such as “Oh” and “Wow”) together, and then a random audience member played a voice line with content (such as “Great job” and “What a pity”). The four NPCs provided feedback simultaneously, but the specific animations and voice lines played are randomized to prevent predictability.

### 3.3 Scoring System And Feedback Activation

As shown in Table 1, players received three types of scores: total score, combo count, and rating score. The total score shows on the upper right corner of the gameplay screen and reflects the total points gained at the end of the gameplay. Combo count tracked the consecutive number of sliced fruits, which is presented below the total score display. The rating score with rating level served as a trigger for NPC feedback and was not visible to players. The rating score and rating level both start at 0. Each level is separated by a 10-point interval, which is determined by dividing the current rating score by 10 and taking the integer part. When players reach a higher level, it triggers positive feedback, while descending to a lower level

Table 1: Scoring Metrics for Fruits and Bombs - Total Score, Combo Count, and Rating Score

Item	State	Total Score	Combo Count	Rating Score
Watermelon	Sliced	+3	+1	+5
	Missed	-7	Reset to 0	-5
Apple	Sliced	+5	+1	+5
	Missed	-5	Reset to 0	-5
Lemon	Sliced	+7	+1	+5
	Missed	-3	Reset to 0	-5
Bomb	Hit	-10	Reset to 0	-10

triggers negative feedback. There is no upper limit to the rating level, but the lower limit is set at 0. If a player remains at level 0 for over 6 seconds, it also triggers negative feedback. To prevent excessive disruptions, a 4.5-second cooldown period followed each feedback event, during which no additional feedback was triggered.

## 4 EXPERIMENT

### 4.1 Experiment Design

The study employed a  $2 \times 2$  within-subjects design that incorporated two factors: Audience Appearance (unfamiliar and familiar) and Audience Voice (unfamiliar and familiar). In total, we have evaluated 4 conditions: (1) Unfamiliar Appearance with Unfamiliar Voice (UAUV), (2) Unfamiliar Appearance with Familiar Voice (UAFV), (3) Familiar Appearance with Unfamiliar Voice (FAUV), and (4) Familiar Appearance with Familiar Voice (FAFV). The order of these conditions was counterbalanced in the experiment.

When creating familiar audiences, the familiar face was obtained via scanned photographs of participants’ friends (using the Avatar Maker Pro plugin), and the body shape and height were generated by adjusting the size of the models. The same individuals also provided familiar voices through voice recordings. On the other hand, the unfamiliar audiences were created using realistic computer character models, and the unfamiliar voices were generated by Microsoft Azure, an AI voice generator that produces natural and smooth human speech based on text. Other factors such as the phrases used for the voices, physical characteristics of the NPC audiences (adjusted to suit the individuals known to the player), and the number of NPCs were controlled in all conditions.

### 4.2 Outcome Measures

- **Player Performance.** We collected the following performance measurements: (1) game score; (2) success rate of fruit hit and avoiding bomb; (3) count of positive and negative feedback; and (4) maximum combo count. See Section 3 for a detailed description.
- **Player Experience.** We employed the following three scales:
  - Player Experience of Need Satisfaction (PENS) scale [33]: a 21-item scale designed to measure game experience across five dimensions. (1) Competence: the players’ perception of their own capabilities. (2) Autonomy: the extent to which players felt free to make choices within the game. (3) Relatedness: the players’ sense of connection to “other players in the game,” (i.e., the NPC audience in our study). (4) Presence/Immersion: players’ emotional engagement in the game. (5) Intuitive Controls: the degree to which players feel that their choices translate smoothly and seamlessly into in-game actions.
  - Pick-A-Mood (PAM) tool [9]: a cartoon-based pictorial instrument for quick and intuitive reporting and expression of mood states. This tool categorized eight dis-

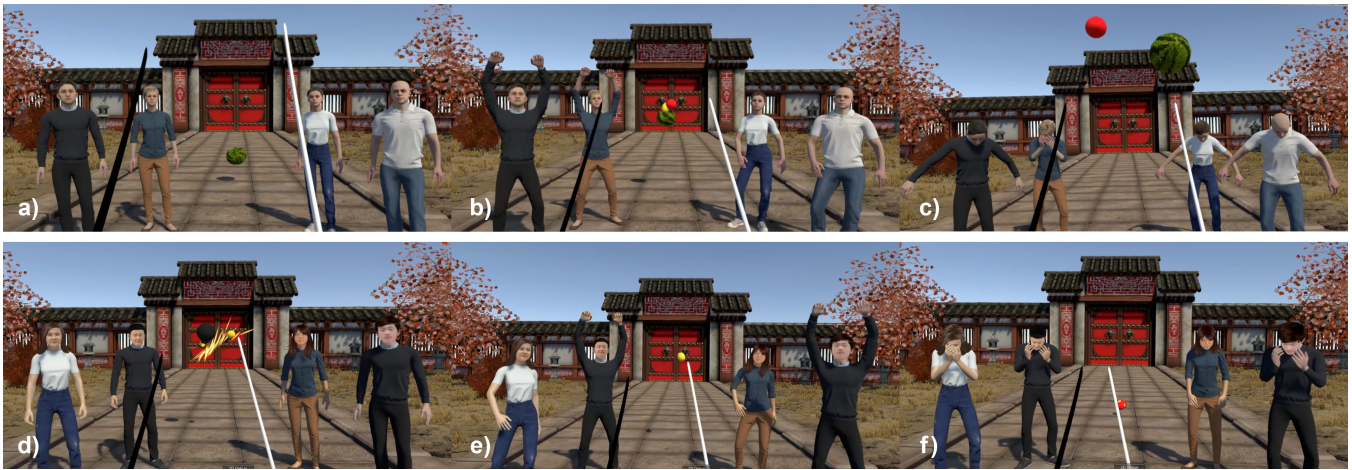


Figure 2: Top Row: (a) Unfamiliar NPC audiences, (b) Unfamiliar NPC audiences providing positive feedback, (c) Unfamiliar NPC audiences expressing negative feedback. Bottom Row: (d) Familiar NPC audiences, (e) Familiar NPC audiences providing positive feedback, (f) Familiar NPC audiences expressing negative feedback.

tinct mood states into four main categories: Energized-Pleasant (excited and cheerful), Calm-Pleasant (relaxed and calm), Energized-Unpleasant (irritated and tense), and Calm-Unpleasant (bored and sad).

- Simulator Sickness Questionnaire (SSQ) [21]: a commonly used 16-item self-report questionnaire that produces 3 measures of cybersickness (nausea, oculomotor, and disorientation). A total SSQ score ranging from 20 to 30 indicates mild to moderate simulator sickness, while a score exceeding 40 is indicative of a poor simulator experience [8].
- **Exertion.** The exertion was assessed by (1) the average heart rate (AvgHR%) using the formula the average heart rate / the participant’s estimated maximum heart rate (i.e.,  $211 - 0.64 \times \text{age}$  [26]); (2) calories burned; and (3) the Borg Rating of Perceived Exertion (RPE) 6-20 scale [4].
- **Self-Consciousness Scale (SCS).** The SCS [12] consisted of three subscales (private self-consciousness, public self-consciousness, and social anxiety), and the scores from the three subscales are summed to obtain an individual’s total self-consciousness score. Private self-consciousness refers to an individual’s tendency to introspect and reflect on their inner thoughts, feelings, and experiences. Public self-consciousness, on the other hand, refers to an individual’s awareness of how they are perceived by others in social situations. Social anxiety is the tendency to experience discomfort and anxiety in social situations.
- **Ranking and Interviews.** We asked participants to rank the evaluated conditions and provide reasons for their rankings. Then, we employed a semi-structured interview with the following open-ended questions: (1) “How do you think the NPCs’ familiarity affects your overall performance of the game?”; (2) “How do you think the NPCs’ familiarity affects your overall game experience?”; (3) “Which factor of the NPCs do you think has a greater impact on your game experience, the appearance or the voice?”; (4) “How do you think positive and negative feedback from the NPCs affect you?”; and (5) “Would having familiar NPCs in future VR exergames encourage you to continue playing in the long term?”. Interviews were recorded and transcribed for data analysis.

### 4.3 Participants

We recruited 24 participants (8 females; mean age = 23.00, SD = 2.02, range = 20 to 28) through an on-campus social media platform. Out of the 18 participants who reported having prior experience with VR HMDs, 9 were frequent weekly users. 11 of them had played exergames before, but none were regular players. Regarding how engaged they were with daily exercise, 5 participants exercised regularly (more than 3 hours per week), 13 participants were less regular (1-3 hours per week), and 6 were inactive (less than 1 hour per week).

### 4.4 Apparatus and Setup

The experiment used a Meta/Oculus Quest 2 as the VR device and a Polar OH1 for monitoring participants’ HR and calories burned. The experiment was conducted in a controlled laboratory environment that was well-illuminated and isolated from external disturbances. The room temperature was regulated by an air conditioner that maintained a constant temperature of 24°C during the experiment. The experiment was conducted under close supervision by an experimenter to prevent any potential risks. This experiment received approval from the University Ethics Committee at the host institution.

### 4.5 Procedures

Participants were provided with a comprehensive introduction of the experimental protocol to read and a consent form to sign off prior to participating in the study. Then, they were required to bring 4 friends in a separate appearance and voice collection session to generate NPC before the experiment. Their friends would also need to sign the consent form to agree that they were willing to form the NPCs in the game. During the appearance and voice collection session, participants’ friends were instructed (1) to maintain a neutral expression and to fully exhibit their facial features to generate the best facial expression, (2) to provide verbal feedback based on a list of prompts, which consisted of both exclamations (such as “Oh” and “Wow”) and complete sentences (such as “Great job”, “Well done”, “Oh no”, and “What a pity”) [11]. All photographs and voice recordings were collected using an iPhone 11. After creating the familiar NPCs, participants were asked to independently identify the models and voices of the NPCs to ensure their recognizability. No difficulties were reported during this process.

On the day of the experiment, participants needed to complete a pre-experiment questionnaire asking demographic questions includ-

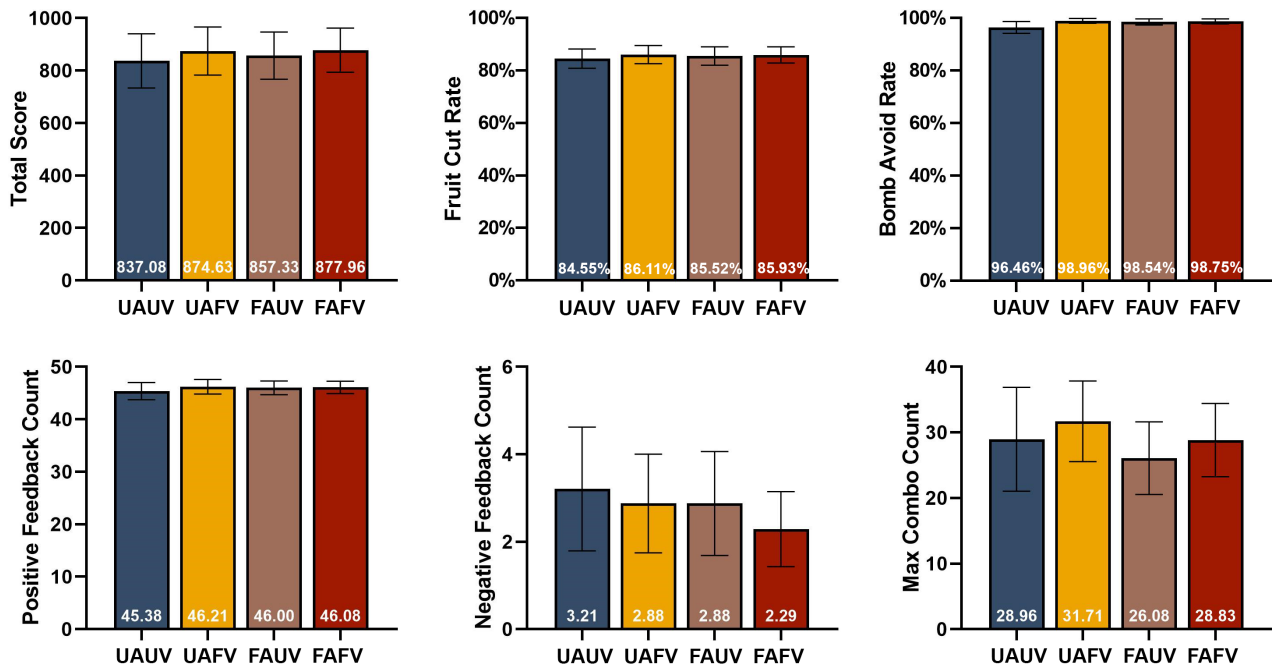


Figure 3: Mean total score, fruit cut rate, bomb avoidance rate, positive feedback count, negative feedback count, and maximum combo count for each condition. Error bars indicate 95% confidence intervals.

ing age, gender, frequency of exercise, experience with exergames and VR HMDs, as well as self-consciousness via SCS. Then, participants were required to enter their personal information (i.e., age, gender, height, and weight) into the Polar Beat mobile application. We collected the rest level HR using the Polar OH1 HR monitor by asking participants to relax and sit still for over 1 minute.

Before the formal gameplay sessions, a 3-minute training phase without any NPCs was administered to enable participants to practice moves and become more acquainted with the equipment. For each experimental condition, participants were assisted by the experimenter to wear the Polar OH1 and HMD. The experiment commenced only when the participant’s heart rate had reached the resting level. After each condition, participants needed to complete four questionnaires: PENS [33], PAM [9], SSQ [21], and Borg RPE 6-20 [4]. They then rested until they felt ready to begin the next experimental condition and had reached a resting heart rate level. At the end of the experiment, participants underwent a semi-structured interview to rank the experimental conditions and provide feedback on their experiences. The experiment lasted about 50 minutes for each participant.

## 5 RESULTS

We conducted statistical analysis on the data using a two-way repeated measures ANOVA with Audience Appearance (unfamiliar and familiar) and Audience Voice (unfamiliar and familiar) as within-subjects variables. The normal distribution of the data was assessed using Shapiro-Wilks tests and Q-Q plots. For non-normally distributed data, we applied the Aligned Rank Transform (ART) to transform the data before performing repeated-measures ANOVAs. We utilized Bonferroni correction for pairwise comparisons, and we reported effect sizes whenever feasible ( $\eta_p^2$ ).

To analyze correlations between the participants’ self-consciousness and their performance and experience, we conducted Pearson’s bivariate correlation analyses.

### 5.1 Player Performance

Figure 3 shows the performance measures for each condition. ANOVA tests yielded a significant main effect of Audience Voice on total score ( $F_{1,23} = 6.843, p = .015, \eta_p^2 = 0.229$ ), but not for Audience Appearance ( $F_{1,23} = 0.437, p = .515$ ) and their interaction effect ( $F_{1,23} = 0.209, p = .652$ ). Pairwise comparisons revealed that participants achieved higher scores when exposed to a familiar audience voice ( $M = 876.29, SE = 40.50$ ) than unfamiliar audience voice ( $M = 847.21, SE = 44.94$ ).

For fruit cut rate, there was a significant main effect of Audience Voice ( $F_{1,23} = 5.691, p = .026, \eta_p^2 = 0.198$ ), but no significant effects were observed for Audience Appearance ( $F_{1,23} = 0.346, p = .562$ ) and Audience Appearance  $\times$  Audience Voice ( $F_{1,23} = 0.825, p = .373$ ). Subsequent pairwise comparisons revealed that participants exhibited a higher fruit cut rate when exposed to familiar audience voice ( $M = 86\%, SE = 0.02$ ) in comparison to unfamiliar audience voice ( $M = 85\%, SE = 0.02$ ).

The data of maximum combo count was not normally distributed and underwent an ART prior to conducting the RM-ANOVA. Our analysis suggested Audience Voice ( $F_{1,23} = 8.214, p = .009, \eta_p^2 = 0.263$ ) had a significant effect on player’s maximum combo count, but not Audience Appearance ( $F_{1,23} = 1.849, p = .187$ ) and Audience Appearance  $\times$  Audience Voice ( $F_{1,23} = 1.047, p = .317$ ). Post-hoc pairwise comparisons revealed that participants achieved a significantly higher maximum combo count in familiar voice conditions ( $M = 30.27, SE = 2.63$ ) than in unfamiliar voice conditions ( $M = 27.52, SE = 2.68$ ). No significant effects were found for bomb avoidance rate, positive feedback count, and negative feedback count.

Pearson’s correlation unveiled a significant correlation between participants’ self-consciousness and (1) the number of negative feedback they triggered when facing an unfamiliar audience appearance ( $r = -.437, p = .033$ ); (2) their maximum combo count when confronted with a familiar audience appearance ( $r = .414, p = .044$ ) and (3) familiar audience voice ( $r = .422, p = .040$ ). No other significant

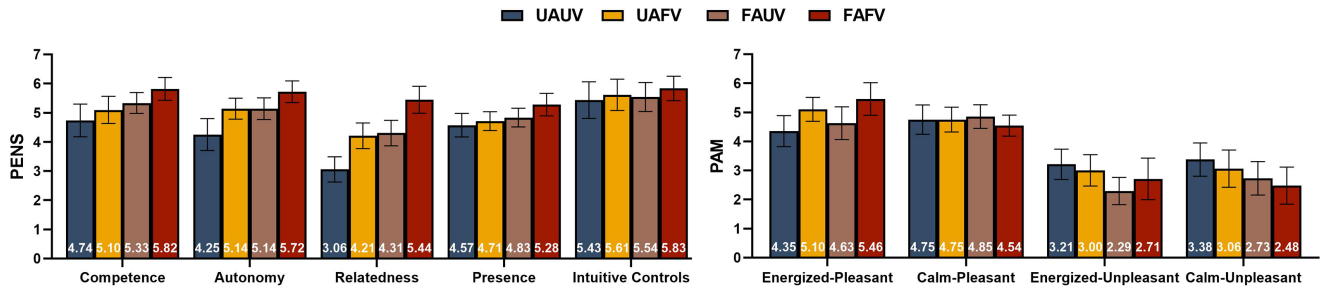


Figure 4: PEN and PAM ratings for each condition. Error bars indicate 95% confidence intervals.

correlations were found.

## 5.2 Player Experience

### 5.2.1 Game Experience

Values for each condition for PENS ratings can be found in Figure 4. Concerning Competence, ANOVA analysis revealed significant main effects for Audience Appearance ( $F_{1,23} = 13.327, p = .001, \eta_p^2 = 0.367$ ) and Audience Voice ( $F_{1,23} = 14.266, p = .001, \eta_p^2 = 0.383$ ), while no significant interaction effect was found ( $F_{1,23} = 0.268, p = .610$ ). Participants reported a heightened sense of competence when (1) presented with a familiar audience appearance ( $M = 5.58, SE = 0.16$ ) in contrast to an unfamiliar appearance ( $M = 4.92, SE = 0.24$ ), (2) exposed to a familiar audience voice ( $M = 5.46, SE = 0.18$ ) compared to unfamiliar voices ( $M = 5.04, SE = 0.20$ ).

Our analysis found significant main effects of Audience Appearance ( $F_{1,23} = 22.255, p = .000, \eta_p^2 = 0.492$ ) and Audience Voice ( $F_{1,23} = 21.203, p = .000, \eta_p^2 = 0.480$ ) on Autonomy. No significant interaction effect of these two factors was found ( $F_{1,23} = 1.184, p = .288$ ). Participants expressed a stronger sense of autonomy (1) when facing a familiar audience appearance ( $M = 5.43, SE = 0.15$ ) than an unfamiliar appearance ( $M = 4.69, SE = 0.20$ ) and (2) when exposed to a familiar audience voice ( $M = 5.43, SE = 0.15$ ) compared to unfamiliar voice ( $M = 4.69, SE = 0.19$ ).

ANOVA tests indicated significant main effects of Audience Appearance ( $F_{1,23} = 50.013, p = .000, \eta_p^2 = 0.685$ ) and Audience Voice ( $F_{1,23} = 45.091, p = .000, \eta_p^2 = 0.662$ ) on Relatedness. No significant interaction effect was observed for Audience Appearance  $\times$  Audience Voice ( $F_{1,23} = 0.003, p = .957$ ). Participants reported a stronger sense of relatedness (1) when encountering familiar audience appearance ( $M = 4.88, SE = 0.18$ ) compared to unfamiliar appearance ( $M = 3.63, SE = 0.19$ ) and (2) when exposed to familiar audience voice ( $M = 4.83, SE = 0.19$ ) compared to unfamiliar voice ( $M = 3.68, SE = 0.18$ ).

For Presence/Immersion, significant main effects were found for Audience Appearance ( $F_{1,23} = 16.152, p = .001, \eta_p^2 = 0.413$ ) and Audience Voice ( $F_{1,23} = 16.565, p = .000, \eta_p^2 = 0.419$ ), but not for Audience Appearance  $\times$  Audience Voice ( $F_{1,23} = 3.871, p = .061$ ). Participants experienced a higher level of immersion (1) with a familiar audience appearance ( $M = 5.06, SE = 0.16$ ) in contrast to an unfamiliar appearance ( $M = 4.64, SE = 0.17$ ) and (2) when presented with a familiar audience voice ( $M = 5.00, SE = 0.16$ ) compared to unfamiliar voice ( $M = 4.70, SE = 0.17$ ). Notably, no main nor interaction effects were found for Intuitive Controls.

Furthermore, Pearson's correlation revealed significant associations between participants' self-consciousness and their perceived Competence when faced with a familiar audience appearance ( $r = .431, p = .036$ ) and familiar audience voice ( $r = .424, p = .039$ ). Additionally, we found a significant correlation between participants' self-consciousness and their perceived Relatedness when exposed to

familiar audience appearance ( $r = .494, p = .014$ ) and familiar audience voice ( $r = .460, p = .024$ ). No other significant correlations were observed.

### 5.2.2 Emotion

The ratings for PAM in each condition can be observed in Figure 4. Regarding Energized-Pleasant, Our analysis showed a significant main effect of Audience Voice ( $F_{1,23} = 19.959, p = .000, \eta_p^2 = 0.465$ ), but not for Audience Appearance ( $F_{1,23} = 3.119, p = .091$ ) and interaction effects ( $F_{1,23} = 0.371, p = .548$ ). Participants reported a heightened sense of energized-pleasantness when exposed to a familiar audience voice ( $M = 5.28, SE = 0.22$ ) as opposed to an unfamiliar audience voice ( $M = 4.49, SE = 0.25$ ).

In terms of Energized-Unpleasant, ANOVA tests revealed a significant main effect of Audience Appearance ( $F_{1,23} = 9.543, p = .005, \eta_p^2 = 0.293$ ), but not for Audience Voice ( $F_{1,23} = 0.466, p = .502$ ) and interaction effects ( $F_{1,23} = 3.363, p = .080$ ). Participants expressed a stronger sense of energized-unpleasantness when confronted with an unfamiliar audience voice ( $M = 2.75, SE = 0.21$ ) than a familiar audience voice ( $M = 2.85, SE = 0.27$ ).

For Calm-Unpleasant, ANOVA tests yielded a significant main effect of Audience Appearance ( $F_{1,23} = 9.425, p = .005, \eta_p^2 = 0.291$ ), but not for Audience Voice ( $F_{1,23} = 3.955, p = .059$ ) and Audience Appearance  $\times$  Audience voice ( $F_{1,23} = 0.070, p = .793$ ). Participants reported a greater degree of calm-unpleasantness when exposed to an unfamiliar audience voice ( $M = 3.05, SE = 0.26$ ) compared to a familiar audience voice ( $M = 2.77, SE = 0.28$ ).

No significant effects were observed for the Calm-Pleasant category. Moreover, our analysis did not reveal any significant correlations between the four emotional categories and participants' self-consciousness.

### 5.2.3 Simulator Sickness

None of our participants had an SSQ higher than 20 among four conditions—UAUV ( $M = 3.27, SE = 3.87$ ), UAFV ( $M = 2.81, SE = 4.01$ ), FAUV ( $M = 2.96, SE = 4.41$ ), and FAFV ( $M = 2.96, SE = 4.41$ ). Our analysis also did not show significant results on Nausea, Oculomotor, Disorientation, as well as total SSQ scores.

## 5.3 Exertion

Values for each condition for exertion measures can be found in Figure 5. ANOVA tests revealed significant main effects of Audience Appearance ( $F_{1,23} = 5.374, p = .030, \eta_p^2 = 0.189$ ) and Audience Voice ( $F_{1,23} = 5.979, p = .015, \eta_p^2 = 0.231$ ) on perceived exertion (Borg RPE; after data were transformed via ART). No significant interaction effects were found ( $F_{1,23} = 0.702, p = .411$ ). Post-hoc pairwise comparisons indicated that participants experienced a higher level of exertion when (1) exposed to unfamiliar audience appearance ( $M = 6.02, SE = 0.59$ ) compared to familiar

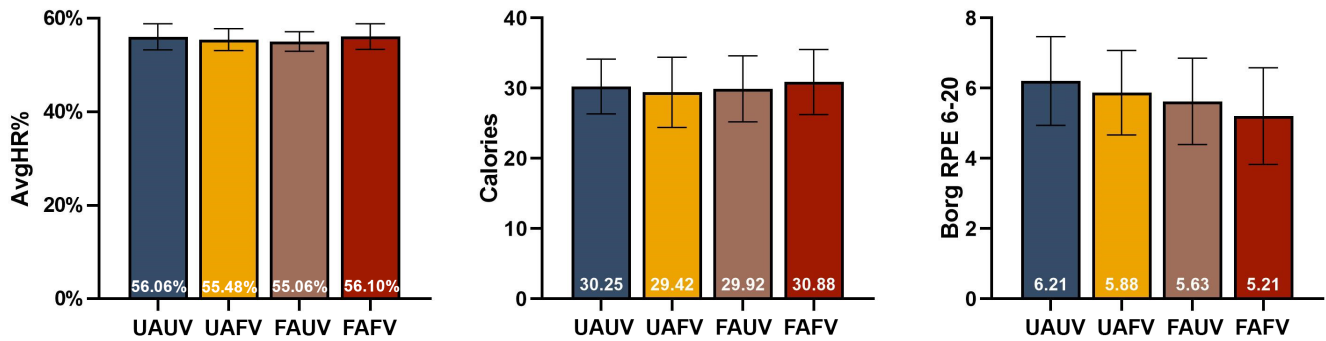


Figure 5: AvgHR%, Calories burned, and Borg ratings for each condition. Error bars indicate 95% confidence intervals.

audience appearance ( $M = 5.34, SE = 0.59$ ) (2) confronted with unfamiliar audience voice ( $M = 5.98, SE = 0.60$ ) than familiar audience voice ( $M = 5.44, SE = 0.57$ ). We did not observe any significant main or interaction effects on avgHR% and calories burned.

Pearson’s correlation revealed a significant correlation between participants’ self-consciousness and their avgHR% when exposed to unfamiliar audience appearance ( $r = .428, p = .037$ ) and unfamiliar audience voice ( $r = .459, p = .024$ ).

#### 5.4 User Rankings and Feedback

Out of the 24 participants, 21 ranked Familiar Appearance with Familiar Voice (FAFV) as the best condition while the remaining 3 participants chose each of the other three conditions respectively. For the second place, 14 participants ranked Unfamiliar Appearance with Familiar Voice (UAFV) as their second choice, while 7 participants ranked Familiar Appearance with Unfamiliar Voice (FAUV) as their second choice. When asked about which factor (appearance or voice) had a more significant effect, 17 participants indicated that they were more easily influenced by audience voice, particularly when they were exposed to familiar audience voice.

When queried about additional details pertaining to their gaming experiences, 17 participants (P1-5, P11-P17, P19, P21-24) expressed a proclivity towards familiar elements, perceiving games featuring these elements as more “*enjoyable/interesting*”. 7 participants (P3-P5, P8, P10, P16, P22) conveyed that the presence of familiar elements instilled a sense of “*comfort/companionship*” during gameplay. Moreover, 4 participants (P9, P15, P19-P20) reported feeling “*relaxed*” when confronted with familiar elements. Notably, a subset of 14 participants (P4-P6, P8-P10, P12-P13, P15, P18-P19, P21, P23-P24) opined that familiar elements fostered “*focus*” and “*enhanced performance*,” albeit 4 participants (P1, P3, P11, P22) perceived such elements as “*distracting*” and impeding their gaming experience.

There was a mixed attitude towards positive and negative feedback, 7 participants (P1-P2, P5, P7, P9, P14, P17) perceived positive feedback as having a positive impact, whereas negative feedback was regarded as having a negative impact. In contrast, 6 participants (P3-P4, P13, P15-P16, P23) recognized both types of feedback being positive and identified negative feedback as a stimulus for “*competitiveness*”. 6 participants (P6, P8, P10, P12, P18, P21) affirmed that both positive and negative feedback from familiar audience served as “*motivation*,” whereas feedback from unfamiliar sources elicited no impact.

Regarding the potential of familiar audiences in future VR exergames to promote long-term engagement, 19 participants expressed a positive outlook. They conveyed a desire for NPC audiences that could be “*customized*” and demonstrate heightened “*intelligence*.” However, the remaining 5 participants (P7, P13-P14, P17, P22) who responded negatively held the view that the sense

of “*comfort*” and “*novelty*” elicited by familiar audiences was only ephemeral.

## 6 DISCUSSION

### 6.1 Impact of Appearance Familiarity on Player Performance and Experience

To address RQ1, our research findings indicate that while appearance familiarity does not significantly impact player performance, it does demonstrate effects in enhancing player experience, promoting positive emotions, and reducing perceived exertion levels. This aligns with prior research findings [23], which indicate that individuals tend to feel more relaxed and pleasant and have enhanced experiences when facing familiar audiences. These positive effects can be attributed to the familiarity and social connection established between the individuals and the audience, creating a more comfortable and engaging atmosphere [6].

However, while previous studies [1, 25] have shown the influence of appearance familiarity with the audience on performance outcomes in VR speech training tools, our findings did not reveal a significant effect in the context of exergames. One possible explanation for this discrepancy is the nature of the visually dominant gameplay, which requires players to maintain a visual focus on the tasks at hand. Several participants (P2, P4, P7, P16) mentioned that at the start of the game, they would purposefully gaze at the NPC’s faces, but as the game progressed, they would focus on the gameplay itself with occasional glances towards them. It is worth noting that individuals have limited attentional capacity [18], and they may not have enough available cognitive resources to allocate to the appearance aspects of NPC audiences.

### 6.2 Impact of Voice Familiarity on Player Performance and Experience

Regarding RQ2, we found that voice familiarity played a crucial role in enhancing both gameplay performance and subjective experiences. Participants demonstrated notable improvements in their total score, fruit cut rate, and maximum combo count when exposed to familiar voices of NPC audiences. Similar to the appearance aspect, familiar voices also have a positive impact on player experience, emotions, and perceived level of exertion.

One possible reason for the influence of voice familiarity on both performance and experience can be attributed to the characteristics of audio stimuli. Audios are typically provided for shorter durations than visuals, enabling immediate and continuous auditory cues that offer a constant stream of feedback throughout the interaction [41]. Moreover, audios are emitted omnidirectionally and constantly, as our ears cannot be closed, ensuring effective feedback delivery regardless of user orientation or visual focus [35]. Voice feedback does not occupy valuable visual space, allowing users to fully utilize their

visual attention for other important aspects of the interaction [5, 28]. While previous research [1, 25, 31] on audience familiarity in virtual environments has rarely involved voice familiarity, our findings indicate that the presence of familiar audience voice can have a meaningful impact on both the performance and overall engagement of users in VR exergames.

### 6.3 Relationship between Self-Consciousness and the Impact of Audience Familiarity

To answer **RQ3**, while there was a limited association between users' self-consciousness and audience familiarity in terms of performance, there was a significant correlation in terms of experience and average heart rates. Our results indicate that the presence of NPCs does not appear to impose pressure on participants with higher levels of self-consciousness, nor does it significantly impact their performance. However, the inclusion of familiar NPCs enhances their subjective experience and contributes to a more relaxed feeling. Interestingly, in the presence of unfamiliar audience stimuli, participants with higher self-consciousness levels may experience increased physiological responses, as reflected by elevated heart rates. On the other hand, they felt a greater sense of relatedness and competence when interacting with familiar audiences. The presence of familiar audiences seemed to provide them with a sense of comfort and social support.

The results showed that participants with higher levels of self-consciousness demonstrated a greater sensitivity to the perception of audience familiarity, both in terms of appearance and voices. This finding aligns with previous research [37, 39] that suggests individuals with higher self-consciousness are more attentive to social cues and the perception of others' judgments. Previous research [16, 20, 50] has shown that the presence and feedback of an audience can enhance player performance and experience with limited exploration of players' self-consciousness. Our findings indicate that for individuals with self-conscious tendencies, the presence of familiar audiences may have a positive influence, resulting in a more relaxed and enjoyable gaming experience.

### 6.4 Design Guidelines

#### 6.4.1 Incorporating Audience Familiarity into Exergames

One of the key findings of our study is the positive impact of incorporating audience familiarity into gameplay. By integrating familiar NPC audiences, players not only improved performance but also enjoyed a heightened sense of game experience and positive emotions. Notably, our findings shed light on the influence of personality traits, particularly self-consciousness, on the effects of audience interactions. Understanding the role of personalities in audience interactions enables developers to tailor the game experience to individual differences.

To facilitate familiarity, we recommend the inclusion of customizable features (1. appearance—face, height, and body shape and 2. voice) for NPC audiences in VR exergames. By allowing players to import and incorporate their own network of acquaintances into the game, such as their friends and family members, NPC avatars can resemble familiar individuals from their personal lives. This recommendation acknowledges the significance of individual differences and empowers players to enhance their enjoyment and immersion in VR exergames.

#### 6.4.2 Prioritizing Voice Familiarity

We recommend prioritizing integrating familiar voices in the design of NPC audiences in VR exergames. Voice familiarity played a pivotal role in influencing player performance and experience in VR exergames, while appearance familiarity only impacts experience. In addition, voice-only familiarity has a higher rating than appearance-only familiarity based on the ranking data and participants' interview feedback. Most importantly, the key advantage

of emphasizing voice familiarity is the relative ease of capturing and incorporating familiar voices compared to facial features. Voice recordings can be obtained with greater convenience and accuracy, facilitating seamless integration into the game. This accessibility empowers games to easily leverage the power of familiar voice cues, enabling games with a more personalized and engaged gameplay experience for players.

### 6.5 Limitations and Future Work

Our study has provided valuable insights into the impact of audience familiarity on player performance and experience in VR exergames. However, there are opportunities for further exploration and improvement that can build upon our findings. One potential direction for future research is to expand the range of variables considered beyond facial appearance and voice. In particular, incorporating other elements, such as facial expressions and clothing, can offer a more comprehensive understanding of how different aspects of audience familiarity influence player experiences.

Another area of improvement lies in the feedback provided by the NPC audience. While our study included a limited set of voice prompts and animations, there is room to enhance the immersive experience by incorporating a broader range of vocal expressions and animated responses. By allowing for customized vocal expressions and animations, future research can enable more dynamic and personalized interactions between players and the NPC audience.

The study participants mainly came from a young population. Exploring a wider range of age groups and demographic characteristics in the future will enhance the overall comprehensiveness of our findings. Furthermore, our study was conducted as a short-term experiment, limiting our understanding of the long-term effects of audience familiarity. Future research could address this limitation by conducting longitudinal studies that examine the sustained effects of audience familiarity over an extended period, such as a 6-week experiment in [42]. This would provide valuable insights into how familiarity with the virtual audience develops and evolves over time and how it impacts player engagement and performance in the long run. Additionally, while our results did not reveal significant issues with appearance and auditory mismatches—as evidenced by participants' ranking preferences—further investigation into factors like player acceptance and familiarity perception could be beneficial.

## 7 CONCLUSION

Our research illuminates the importance of audience familiarity in the context of VR exergames. We have demonstrated that both appearance and voice familiarity contributes to the player experience and perceived exertion, with voice familiarity exerting a more prominent influence on player performance. The positive impact of familiar NPC audiences on player experiences, particularly in terms of creating a relaxed and enjoyable atmosphere, suggests the potential for enhancing engagement and satisfaction in virtual reality environments. Furthermore, we discovered that individuals with higher self-consciousness are more sensitive to the influence of audience familiarity. This underscores the importance of considering individual differences and personal traits when designing virtual environments to ensure optimal player engagement and comfort. By emphasizing the integration of appearance and voice familiarity in NPC audience design, our study contributes to the advancement of immersive VR exergames that cater to individual preferences and promote positive player outcomes. Our findings offer valuable insights to developers, designers, and researchers in creating engaging VR exergames that harness the power of audience familiarity.

## ACKNOWLEDGMENTS

The authors want to express gratitude to the participants who took part in the study, as well as to the reviewers, for their valuable insights and constructive suggestions that helped improve the paper.



## REFERENCES

- [1] A. Barrett, A. Pack, D. Monteiro, and H.-N. Liang. Exploring the influence of audience familiarity on speaker anxiety and performance in virtual reality and real-life presentation contexts. *Behaviour & Information Technology*, pp. 1–13, 2023.
- [2] R. F. Baumeister. Choking under pressure: self-consciousness and paradoxical effects of incentives on skillful performance. *Journal of personality and social psychology*, 46(3):610, 1984.
- [3] C. Belletier, A. Normand, and P. Huguet. Social-facilitation-and-impairment effects: From motivation to cognition and the social brain. *Current Directions in Psychological Science*, 28(3):260–265, 2019.
- [4] G. A. Borg. Psychophysical bases of perceived exertion. *Medicine & science in sports & exercise*, 1982.
- [5] I. d. V. Bosman, O. uruk, K. Jørgensen, and J. Hamari. The effect of audio on the experience in virtual reality: a scoping review. *Behaviour & Information Technology*, pp. 1–35, 2023.
- [6] E. Burkitt, D. Watling, and H. Message. Expressivity in children’s drawings of themselves for adult audiences with varied authority and familiarity. *British Journal of Developmental Psychology*, 37(3):354–368, 2019.
- [7] C. S. Carver and M. F. Scheier. Self-focusing effects of dispositional self-consciousness, mirror presence, and audience presence. *Journal of personality and social psychology*, 36(3):324, 1978.
- [8] P. Caserman, A. Garcia-Agundez, A. Gámez Zerbán, and S. Göbel. Cybersickness in current-generation virtual reality head-mounted displays: systematic review and outlook. *Virtual Reality*, 25(4):1153–1170, 2021.
- [9] P. M. Desmet, M. H. Vastenburger, and N. Romero. Mood measurement with pick-a-mood: review of current methods and design of a pictorial self-report scale. *Journal of Design Research*, 14(3):241–279, 2016.
- [10] H. D. Ellis, D. M. Jones, and N. Mosdell. Intra-and inter-modal repetition priming of familiar faces and voices. *British Journal of Psychology*, 88(1):143–156, 1997.
- [11] L. K. Epting, K. N. Riggs, J. D. Knowles, and J. J. Hanky. Cheers vs. jeers: Effects of audience feedback on individual athletic performance. *North American Journal of Psychology*, 13(2), 2011.
- [12] A. Fenigstein, M. F. Scheier, and A. H. Buss. Public and private self-consciousness: Assessment and theory. *Journal of consulting and clinical psychology*, 43(4):522, 1975.
- [13] T. Garcia-Marques, D. M. Mackie, H. M. Claypool, and L. Garcia-Marques. Is it familiar or positive? mutual facilitation of response latencies. *Social Cognition*, 28(2):205–218, 2010.
- [14] T. Garcia-Marques, M. Prada, and D. M. Mackie. Familiarity increases subjective positive affect even in non-affective and non-evaluative contexts. *Motivation and Emotion*, 40:638–645, 2016.
- [15] P. S. Goodman and D. P. Leyden. Familiarity and group productivity. *Journal of applied psychology*, 76(4):578, 1991.
- [16] J. C. Haller, Y. H. Jang, J. Haller, L. Shaw, and B. C. Wünsche. Hiit the road: Using virtual spectator feedback in hiit-based exergaming. In *Proceedings of the Australasian Computer Science Week Multiconference*, pp. 1–9, 2019.
- [17] E. Harmon-Jones and J. J. Allen. The role of affect in the mere exposure effect: Evidence from psychophysiological and individual differences approaches. *Personality and Social Psychology Bulletin*, 27(7):889–898, 2001.
- [18] B. Hubert-Wallander, C. S. Green, and D. Bavelier. Stretching the limits of visual attention: the case of action video games. *Wiley interdisciplinary reviews: cognitive science*, 2(2):222–230, 2011.
- [19] T. Kaneko, H. Tarumi, K. Kataoka, Y. Kubochi, D. Yamashita, T. Nakai, and R. Yamaguchi. Supporting the sense of unity between remote audiences in vr-based remote live music support system ksa2. In *2018 IEEE International Conference on Artificial Intelligence and Virtual Reality (AIVR)*, pp. 124–127. IEEE, 2018.
- [20] D. L. Kappen, P. Mirza-Babaei, J. Johannsmeier, D. Buckstein, J. Robb, and L. E. Nacke. Engaged by boos and cheers: the effect of co-located game audiences on social player experience. In *Proceedings of the first ACM SIGCHI annual symposium on Computer-human interaction in play*, pp. 151–160, 2014.
- [21] R. S. Kennedy, N. E. Lane, K. S. Berbaum, and M. G. Lilienthal. Simulator sickness questionnaire: An enhanced method for quantifying simulator sickness. *The international journal of aviation psychology*, 3(3):203–220, 1993.
- [22] M. R. Leary, J. B. Nezlek, D. Downs, J. Radford-Davenport, J. Martin, and A. McMullen. Self-presentation in everyday interactions: effects of target familiarity and gender composition. *Journal of personality and social psychology*, 67(4):664, 1994.
- [23] P. D. MacIntyre and K. A. Thivierge. The effects of audience pleasantness, audience familiarity, and speaking contexts on public speaking anxiety and willingness to speak. *Communication Quarterly*, 43(4):456–466, 1995.
- [24] C. Mesagno, J. T. Harvey, and C. M. Janelle. Choking under pressure: The role of fear of negative evaluation. *Psychology of Sport and Exercise*, 13(1):60–68, 2012.
- [25] D. Monteiro, A. Wang, L. Wang, H. Li, A. Barrett, A. Pack, and H.-N. Liang. Effects of audience familiarity on anxiety in a virtual reality public speaking training tool. *Universal Access in the Information Society*, pp. 1–12, 2023.
- [26] B. M. Nes, I. Janszky, U. Wisløff, A. Støylen, and T. Karlsen. Age-predicted maximal heart rate in healthy subjects: The hunt fitness study. *Scandinavian journal of medicine & science in sports*, 23(6):697–704, 2013.
- [27] R. J. Nesbit and D. Watling. The role of audience familiarity and activity outcome in children’s understanding of disclaimers. *British Journal of Developmental Psychology*, 37(2):230–246, 2019.
- [28] H. Palani and N. A. Giudice. Evaluation of non-visual panning operations using touch-screen devices. In *Proceedings of the 16th international ACM SIGACCESS conference on Computers & accessibility*, pp. 293–294, 2014.
- [29] D.-P. Pertaub, M. Slater, and C. Barker. An experiment on public speaking anxiety in response to three different types of virtual audience. *Presence*, 11(1):68–78, 2002.
- [30] P. Raimbaud, A. Jovane, K. Zibrek, C. Pacchierotti, M. Christie, L. Hoyet, J. Pettre, and A.-H. Olivier. The stare-in-the-crowd effect in virtual reality. In *2022 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*, pp. 281–290. IEEE, 2022.
- [31] F. Raja. Anxiety level in students of public speaking: Causes and remedies. *Journal of education and educational development*, 4(1), 2017.
- [32] H. T. Reis, M. R. Maniaci, P. A. Capriarello, P. W. Eastwick, and E. J. Finkel. Familiarity does indeed promote attraction in live interaction. *Journal of personality and social psychology*, 101(3):557, 2011.
- [33] R. M. Ryan, C. S. Rigby, and A. Przybylski. The motivational pull of video games: A self-determination theory approach. *Motivation and emotion*, 30:344–360, 2006.
- [34] S. R. Schweinberger, H. Kawahara, A. P. Simpson, V. G. Skuk, and R. Zäske. Speaker perception. *Wiley Interdisciplinary Reviews: Cognitive Science*, 5(1):15–25, 2014.
- [35] S. Serafin, M. Geronazzo, C. Erkut, N. C. Nilsson, and R. Nordahl. Sonic interactions in virtual reality: State of the art, current challenges, and future directions. *IEEE computer graphics and applications*, 38(2):31–43, 2018.
- [36] P. M. Strojny, N. Duźmańska-Misiarczyk, N. Lipp, and A. Strojny. Moderators of social facilitation effect in virtual reality: Co-presence and realism of virtual agents. *Frontiers in Psychology*, 11, 2020. doi: 10.3389/fpsyg.2020.01252
- [37] D. M. Tice, J. Buder, and R. F. Baumeister. Development of self-consciousness: At what age does audience pressure disrupt performance? *Adolescence*, 20(78):301, 1985.
- [38] D. M. Tice, J. L. Butler, M. B. Muraven, and A. M. Stillwell. When modesty prevails: Differential favorability of self-presentation to friends and strangers. *Journal of personality and social psychology*, 69(6):1120, 1995.
- [39] J. Wang, D. Marchant, T. Morris, and P. Gibbs. Self-consciousness and trait anxiety as predictors of choking in sport. *Journal of science and medicine in sport*, 7(2):174–185, 2004.
- [40] G. Weary and R. M. Arkin. Attributional self-presentation. *New directions in attribution research*, 3:223–246, 1981.
- [41] A. Westerberg and H. Schoenau-Fog. Categorizing video game audio: An exploration of auditory-psychological effects. In *Proceedings of the 19th International Academic Mindtrek Conference*, pp. 47–54, 2015.

- [42] W. Xu, H.-N. Liang, N. Baghaei, X. Ma, K. Yu, X. Meng, and S. Wen. Effects of an immersive virtual reality exergame on university students' anxiety, depression, and perceived stress: Pilot feasibility and usability study. *JMIR Serious Games*, 9(4):e29330, Nov 2021. doi: 10.2196/29330
- [43] W. Xu, H.-N. Liang, Q. He, X. Li, K. Yu, Y. Chen, et al. Results and guidelines from a repeated-measures design experiment comparing standing and seated full-body gesture-based immersive virtual reality exergames: Within-subjects evaluation. *JMIR serious games*, 8(3):e17972, 2020.
- [44] W. Xu, H.-N. Liang, X. Ma, and X. Li. Virusboxing: A hiit-based vr boxing game. In *Extended Abstracts of the 2020 Annual Symposium on Computer-Human Interaction in Play, CHI PLAY '20*, p. 98–102, 2020. doi: 10.1145/3383668.3419958
- [45] W. Xu, H.-N. Liang, K. Yu, and N. Baghaei. Effect of gameplay uncertainty, display type, and age on virtual reality exergames. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*, pp. 1–14, 2021.
- [46] W. Xu, H.-N. Liang, K. Yu, S. Wen, N. Baghaei, and H. Tu. Acceptance of virtual reality exergames among chinese older adults. *International Journal of Human-Computer Interaction*, pp. 1–15, 2022.
- [47] W. Xu, K. Yu, X. Meng, D. Monteiro, D. Kao, and H.-N. Liang. Exploring the effect of the group size and feedback of non-player character spectators in virtual reality exergames. *Frontiers in Psychology*, 14, 2023.
- [48] H. Yakura and M. Goto. Enhancing participation experience in vr live concerts by improving motions of virtual audience avatars. In *2020 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*, pp. 555–565. IEEE, 2020.
- [49] A. W. Young, S. Frühholz, and S. R. Schweinberger. Face and voice perception: Understanding commonalities and differences. *Trends in cognitive sciences*, 24(5):398–410, 2020.
- [50] K. Yu, S. Wen, W. Xu, M. Caon, N. Baghaei, and H.-N. Liang. Cheer for me: effect of non-player character audience feedback on older adult users of virtual reality exergames. *Virtual Reality*, pp. 1–17, 2023.
- [51] M. Zuo and Y. Shen. How features and affordances of a metaverse portal engage users? evidence from exergames. *Internet Research*, 2023.