future internet

MDPI

Article

Eye Tracking and Human Influence Factors' Impact on Quality of Experience of Mobile Gaming

Omer Nawaz [1,*], Siamak Khatibi [1,*], Muhammad Nauman Sheikh [2] and Markus Fiedler [1]

[1] Department of Technology and Aesthetics, Blekinge Tekniska Högskola, 371 79 Karlskrona, Sweden; markus.fiedler@bth.se
[2] Dubizzlelabs, Lahore 58000, Pakistan; nauman.sheikh@dubizzlelabs.com
* Correspondence: omer.nawaz@bth.se (O.N.); siamak.khatibi@bth.se (S.K.)

Abstract: Mobile gaming accounts for more than 50% of global online gaming revenue, surpassing console and browser-based gaming. The success of mobile gaming titles depends on optimizing applications for the specific hardware constraints of mobile devices, such as smaller displays and lower computational power, to maximize battery life. Additionally, these applications must dynamically adapt to the variations in network speed inherent in mobile environments. Ultimately, user engagement and satisfaction are critical, necessitating a favorable comparison to browser and console-based gaming experiences. While Quality of Experience (QoE) subjective evaluations through user surveys are the most reliable method for assessing user perception, various factors, termed influence factors (IFs), can affect user ratings of stimulus quality. This study examines human influence factors in mobile gaming, specifically analyzing the impact of user delight towards displayed content and the effect of gaze tracking. Using Pupil Core eye-tracking hardware, we captured user interactions with mobile devices and measured visual attention. Video stimuli from eight popular games were selected, with resolutions of 720p and 1080p and frame rates of 30 and 60 fps. Our results indicate a statistically significant impact of user delight on the MOS for most video stimuli across all games. Additionally, a trend favoring higher frame rates over screen resolution emerged in user ratings. These findings underscore the significance of optimizing mobile gaming experiences by incorporating models that estimate human influence factors to enhance user satisfaction and engagement.

Keywords: gaze tracking; QoE; mobile gaming; eye tracking in mobile gaming; human IFs and multimedia

Citation: Nawaz, O.; Khatibi, S.; Sheikh, M.N.; Fiedler, M. Eye Tracking and Human Influence Factors' Impact on Quality of Experience of Mobile Gaming. *Future Internet* **2024**, *16*, 420. https://doi.org/10.3390/fi16110420

Academic Editor: Gianluigi Ferrari

Received: 30 September 2024
Revised: 8 November 2024
Accepted: 11 November 2024
Published: 13 November 2024

1. Introduction

As of 2023, multimedia content, particularly video, accounts for over 80% of internet usage, with more than 5 billion subscribers accessing the internet via mobile connectivity, according to mobile data statistics and forecasts [1,2]. Meanwhile, there is a trend towards increased consumption of multimedia content, particularly video streaming and online gaming on mobile devices. The global mobile gaming software market exceeded 190 billion US dollars in 2023, with mobile devices accounting for over 50% of this revenue, and this trend is projected to triple within the next four years [3–7]. The main reasons are the availability of high-speed internet access at reduced cost on mobile networks, portability, and hardcore gaming titles becoming available on mobile devices. The popularity and success of mobile gaming titles are contingent upon the optimization of gaming applications for the specific hardware constraints of mobile devices, such as smaller displays and lower computational power relative to gaming consoles. Additionally, these applications must adapt dynamically to variations in network speed, a consequence of the inherent mobility paradigm. Ultimately, the most critical factor is user engagement and satisfaction, which must be favorably compared to that of browser and console-based gaming experiences.

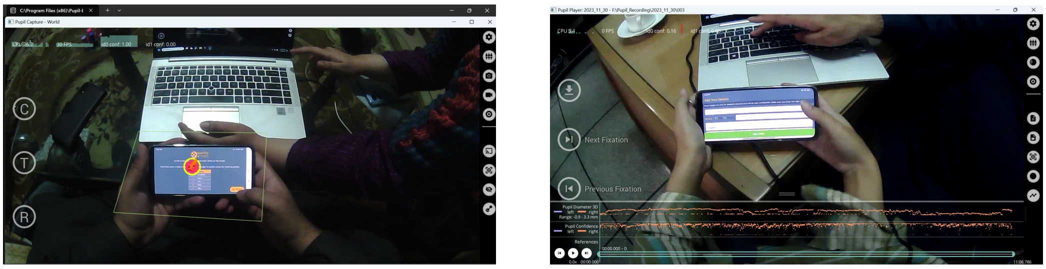
To evaluate the user opinion towards an application, the most reliable method for assessing user perception involves subjective evaluations through user surveys on stimulus quality. However, many factors can influence user ratings regarding the quality of the stimuli, ranging from human to context, etc. The term influence factor (IF) is defined as '*Any characteristic of a user, system, service, application, or context whose actual state or setting may have influence on the Quality of Experience for the user*' [8,9]. The Mean Opinion Score (MOS), as defined by the International Telecommunication Union Standardization Sector (ITU-T) [10], was originally developed to gauge user opinions on speech quality in telecommunications. However, it has become a de facto standard due to its extensive applicability across various multimedia applications. The effectiveness of the MOS for evaluating video stimuli has been questioned in numerous studies, particularly regarding whether the distinction between "excellent" and "good" is equivalent to that between "fair" and "poor" [11,12]. To address these concerns, alternative metrics such as %Good or Better (GoB) and %Poor or Worse (PoW) have been proposed to provide a more transparent and meaningful interpretation of user feedback [13,14].

The objective of this article is to examine the human influence factors in mobile gaming. Specifically, we analyze the impact of user delight (liking) towards displayed content and the effect of gaze tracking on mobile gaming. Previous studies on multimedia streaming stimuli have explored various influence factors such as delight, mood, gender, and frequency of watching online videos, etc., establishing the role of delight in the user ratings [15–18]. Our observations revealed that user delight towards the displayed content affects the MOS for certain stimuli, though the results were not statistically significant across the entire dataset [19,20]. A notable finding was that the nature of content significantly influences the MOS, particularly when the same videos with minor impairments are repeatedly shown due to packet loss or jitter degradation. At high jitter levels and packet loss exceeding 1%, MOS values predominantly fell into the poor or worst range, rendering meaningful interpretation difficult.

While numerous studies have investigated the role of eye tracking and its effect on user perception in online gaming [21–25], there remains a research gap concerning mobile gaming and user perception of offered services. This study utilizes Pupil Core eye-tracking hardware developed by Pupil Labs Berlin, Germany [26,27] to capture user interactions with mobile devices and simultaneously measure their visual attention. We employed the Pupil Core fixation detector module [28] to track user gaze within the calibrated area encompassing the mobile screen during subjective assessments, as depicted in Figures 1 and 2a.

Figure 1. Subjective assessment with Pupil Core.

We selected video stimuli from eight popular games, with resolutions of 720p and 1080p, the highest achievable on mobile gaming devices, even with 5G connections. It is noteworthy that most common gaming consoles and major gaming titles are also limited to 1080p resolution [29]. The chosen stimuli featured frame rates of 30 and 60 frames per second (fps). Our results indicate a statistically significant impact of user delight for most video stimuli across all games. Additionally, an interesting trend emerged favoring higher frame rates over screen resolution in terms of user ratings. Finally, we observed some influence of user gaze based on gaze confidence values.



(**a**) (**b**)

Figure 2. Pupil Core software v3.5.1. (**a**) Pupil Capture for session recording with fixation detector in the calibrated area; (**b**) Pupil Player screen for replaying the individual session and exporting data.

The structure of this paper is as follows: Section 2 provides an overview of the background and a brief exploration of relevant technologies. Section 3 details the experimental setup, including all parameters and methods applied to extract the results. Section 4 presents the assessment outcomes with necessary explanations. Finally, Section 5 outlines the conclusions drawn from our study.

2. Background and Methodological Overview

This section provides the related work and a brief overview of the Quality of Experience (QoE) influence factors and corresponding metrics.

Regarding the role of gaze and visual attention on user experience and consequent ratings, Capozzi et al. [30] have examined the impact of gaze and visual attention on user experience and subsequent ratings, identifying perceptions, interpretations, and evaluations as key elements that shape social interactions. They argue that non-verbal social communication, particularly related to perception, is typically mediated by gaze or gestures, indicating that attention is integral to social behavior and can significantly influence user perception. Similarly, Chica et al. [31] analyze the interplay between visual attention and conscious perception, concluding that, unlike internal factors, exogenous orienting may have substantial implications for conceptual perception. Dalmaso also demonstrates that saccadic eye movements (rapid shifts of fixation from one point to another) can be guided by facial cues, thus influencing attention in social scenarios [32]. Mulckhuyse et al. [33] further assert that attention serves as a fundamental mechanism for selecting objects and locations within an environment. They explore how distracting stimuli in complex visual settings can capture attention, a factor particularly relevant to fast-paced gaming contexts, where player focus is frequently redirected by dynamic on-screen events.

In the pursuit of quantifying user gaze, blinks, facial expressions, and visual attention, diverse methodologies have been employed. One approach involves capturing user sessions via conventional cameras and subsequently extracting gaze data using tools such as the open-face toolkit and related software [34,35]. An alternative method, also used in virtual reality gaming headsets, gauges attention by presenting pop-up stimuli during gaming sessions, potentially accompanied by auditory cues [36,37]. However, both methodologies have inherent limitations, as the introduction of pop-ups disrupts the user's focus on the content, thereby affecting their attentional allocation [38,39].

Mesfin et al. [40] applied the multiple sensorial media (Mulsemedia), which integrates various human senses, arguing that it can enhance QoE in digital environments. Their study presented a range of video stimuli with attributes such as color, shape, and brightness, supplemented by cross-modal sounds. These stimuli were displayed on a monitor while an external EyeTribe camera tracked user gaze, and heart rate monitoring was also conducted. Although the study is not related to our research and the experimental setup limits user mobility due to the external eye tracker, which affects user experience in mobile gaming, it offers valuable insights for future research regarding Mulsemedia using mobile gaming stimuli.

Wibirama et al. [41] conducted a study based on the different screen sizes of mobile devices with negligible variation in the resolution and its impact on gaming. It was observed that the participants were more immersed during gaming on a mobile screen with a larger screen size. The participants were evaluated on their involvement, enjoyment, attention, and challenge while they were playing on different screen sizes.

Jiang et al. [42] conducted an experiment in which participants played four different games while their eye movements and gaze patterns were observed and documented using heat maps to indicate the user's focus throughout the experiment. However, the games used in the experiment were lesser-known and not mainstream, without considering factors such as resolution or personal bias. In contrast, our research primarily focuses on human interaction factors (IFs). Therefore, we specifically chose popular gaming titles to explore the impact of external factors, such as user delight (liking and disliking), resolution, and frame rate, on the overall user experience.

Gunawardena et al. [25] analyzed 36 publications published after 2010 related to gaze tracking in mobile phones and tablets and proposed an edge computing-based eye-tracking solution. The study encompasses the publications related to using commercial devices specifically designed for eye tracking using external glasses like Tobii, Pupil Core, etc., and screen-based eye-tracking solutions that depend on the mobile device's front-facing camera. The authors mentioned the high cost associated with external hardware-based eyeglasses and the complex calibration process but also highlighted the limitations with head pose and movements with mobile device cameras. The authors did not conduct any benchmarking between hardware-based solutions and mobile device cameras prior to drawing conclusions about both methods.

In our study, the focus was on gaze tracking during mobile gaming, and putting limitations on user movement compromises the user's comfort and experience related to mobile gaming. This resulted in the selection of dedicated external eye-tracking glasses.

2.1. Pupil Labs Core

This paper employs the Pupil Labs Core device for gaze estimation along with its accompanying software. The device utilized in this study features two cameras: a world-view camera capable of recording 720p videos at 60 Hz, focusing on the user's front view, and an eye camera focusing on the user's gaze at 200 Hz. The world-view camera can be equipped with either a wide-angle or narrow-angle lens; we opted for the narrow-angle lens due to the limited calibration area available on the mobile viewing device. While the device can operate with multiple eye-tracking cameras targeting both eyes, we used a monochrome setup for the eye camera to ensure that the video viewing experience was not compromised, as illustrated in Figure 1. Detailed technical specifications of the Core device are available in this paper and on the manufacturer's website [26,27].

The Core software package includes three components: Pupil Capture, Pupil Service, and Pupil Player [43], with version 3.5.1 being used in this study. Pupil Capture is utilized to calibrate both cameras to enhance the confidence of gaze estimation and to establish the required field of view for the world camera. It also enables various plugins and records the sessions. Pupil Player is employed for post-processing, visualizing data, exports, and replaying recordings, as depicted in Figure 2a.

Pupil and Gaze Positions

It is crucial to understand that the confidence level in pupil detection does not indicate the position where the subject looked at the calibrated area. Instead, it reflects the accuracy of the algorithm in identifying the pupil from eye camera images, thereby indicating the reliability of baseline data (pupil position). A confidence value of 0 means the pupil was not detected at all, rendering the data unreliable, while a value of 1 indicates complete certainty in the pupil's position. For meaningful analysis, values above 0.6 are generally considered to be reliable [44].

2.2. Quality of Experience and User Delight

The *Quality of Experience* (QoE) is defined by ITU-T as '*The degree of delight or annoyance of the user of an application or service*' [45], with reference to the full definition that continues with '*It results from the fulfillment of his or her expectations with respect to the utility and/or enjoyment of the application or service in the light of the user's personality and current state*' [8]. QoE encompasses a range of complex, interrelated dimensions, including psychological factors, networks, applications, etc. Numerous influencing factors (IFs) affect QoE, which are generally classified into three categories: human, system, and contextual factors [8]. Although user satisfaction with an application is integral to the definition of QoE, there is disagreement among researchers regarding the influence of human behavioral tendencies and how to classify emotions triggered by various stimuli. Robert et al. [46] have explored the emotional impact of visual stimuli and the evaluation of this experience.

Methods for quantifying Quality of Experience (QoE) are typically categorized into objective and subjective measures. Objective metrics primarily consist of tools that assess the quality of stimuli by comparing them to the original (full-reference), reduced-reference, or no-reference benchmarks, depending on the availability of the stimuli for comparison. Conversely, while subjective assessment is more resource- and time-intensive, it remains the most reliable method, as it involves direct input from users regarding their perceptions. The MOS is commonly used with various scales to quantify user perception. The most widely adopted method for measuring the MOS is the five-point Absolute Category Rating (ACR) scale, as recommended by the ITU-T [47]. Originally developed for scaling voice quality in telecommunications, the MOS has been studied extensively for its limitations in evaluating multimedia content [11,12,48]. Tobias et al. [13,14] have shown that QoE distribution in a system might not correlate with MOS distribution. They proposed that on a five-point ACR scale where 1 = Bad, 2 = Poor, 3 = Fair, 4 = Good, and 5 = Excellent, the mapping of Good or Better and Poor or Worse ratios provides a more accurate measure for QoE mappings.

QoE Subjective Metrics

The MOS provides the average ratings of participants in a subjective assessment, and in the case of assessing the delight towards the shown content of two different groups on a binary scale, the MOS can be formulated as follows:

$$\mathrm{MOS} = \frac{1}{N_A + N_B}\left(\sum_{i=1}^{N_A} R_{A,i} + \sum_{i=1}^{N_B} R_{B,i}\right) \tag{1}$$

where:

$$\begin{aligned} N_A &: \text{Total number of ratings in Group A.} \\ N_B &: \text{Total number of ratings in Group B.} \\ R_{A,i} &: \text{Rating given by the } i\text{-th user in Group A.} \\ R_{B,i} &: \text{Rating given by the } i\text{-th user in Group B.} \end{aligned}$$

This combined equation allows for a comprehensive assessment of the overall MOS across both user groups. In the case of the MOS for each group where $x \in \{A, B\}$, we can calculate it as:

$$\mathrm{MOS}_x = \frac{1}{N_x}\sum_{i=1}^{N_x} R_{x,i} \tag{2}$$

where:

$$N : \text{Total number of ratings}$$
$$R_i : \text{Rating given by the } i\text{-th user}$$

Similarly, GoB or PoW ratios can be easily calculated; the formula for the GoB ratio over an ACR scale can be written as:

$$\text{GoB} = \frac{N_4 + N_5}{N_\text{T}}$$

where:

$$N_4 : \text{Number of ratings with a score of 4 (Good)}$$
$$N_5 : \text{Number of ratings with a score of 5 (Excellent)}$$
$$N_T : \text{Total number of ratings}$$

In this study, we explored how user preference (delight) towards presented content influences their video quality ratings, resulting in the formation of various sub-groups. Beyond the typical statistical analyses using the MOS and confidence intervals in QoE subjective studies, we employed one-way ANOVA [49] to assess the differences and statistical significance when comparing the sub-categories.

3. Experimental Setup

This section explains the rationale behind the selection of video stimuli for this study, the configuration and calibration of the Pupil Core device for gaze data collection for every individual assessment [50], the customization of the mobile application used to obtain user ratings, and the methodology employed to synchronize the continuous data extracted from Pupil Core with Google Firebase [51] to obtain the gaze confidence values for specific videos.

3.1. Video Selection

In recent research, we developed a video stimuli database for online video streaming and conducted subjective assessments on a mobile device using the latest codecs [20]. However, due to the focus of this study being on mobile gaming, we utilized the mobile gaming database provided by the Laboratory for Image and Video Engineering (LIVE) [52], which is based on YouTube gaming video clips [53]. This database targets online user-generated mobile gaming content, and further details can be found in Yu et al. [54]. A within-group experiment was conducted using 8 popular gaming titles, each with 720p and 1080p resolutions, and frame rates of 30 and 60 fps. The technical specifications of the videos are provided in Table 1.

A common issue encountered with subjective assessments is participants losing focus due to watching a loop of video stimuli of similar types and very short duration [55,56]. To mitigate this and ensure sustained user attention towards the video content, we selected a total of 40 video clips. For most games, a quartet of 30 and 60 fps frame rates at both resolutions was unavailable; for instance, *Code Vein* video clips were not available in 1080p resolution at 30 fps, as shown in Table 1. Despite limiting the selection of video stimuli to maintain an appropriate assessment duration ensuring that subjects did not lose focus due to the reasons stated above or fatigue from wearing a hardware-based eye tracker, we included a diverse range of game genres to gather comprehensive user feedback. Most of the selected games represent an amalgamation of multiple genres, including *Animal Crossing* (social simulation), *Counter-Strike 2* (tactical, first-person shooter), *Call of Duty* (military first-person shooter), *Code Vein* (action, role-playing), *Fortnite* (survival, battle royale), *Minecraft* (sandbox, survival), *PUBG* (battle royale), and *Rocket League* (vehicular soccer), as illustrated in Figure 3.

Table 1. Reference videos' specifications.

Name	Length, seconds	fps	Resolution
Animal Crossing	9 9	60 30	1920 × 1080, 1280 × 720 1280 × 720
Counter-Strike 2	9, 8 (720p) 9	60 30	1920 × 1080, 1280 × 720 1280 × 720
Call of Duty	9 9	60 30	1920 × 1080, 1280 × 720 1280 × 720
Code Vein	9 9	60 30	1920 × 1080, 1280 × 720 1920 × 1080
Fortnite	9 9	60 30	1920 × 1080, 1280 × 720 1280 × 720
Minecraft	9, 8 (720p) 9	60 30	1920 × 1080, 1280 × 720 1280 × 720
PUBG	9 9	60 30	1920 × 1080, 1280 × 720 1280 × 720
Rocket League	9 9	60 30	1920 × 1080, 1280 × 720 1280 × 720

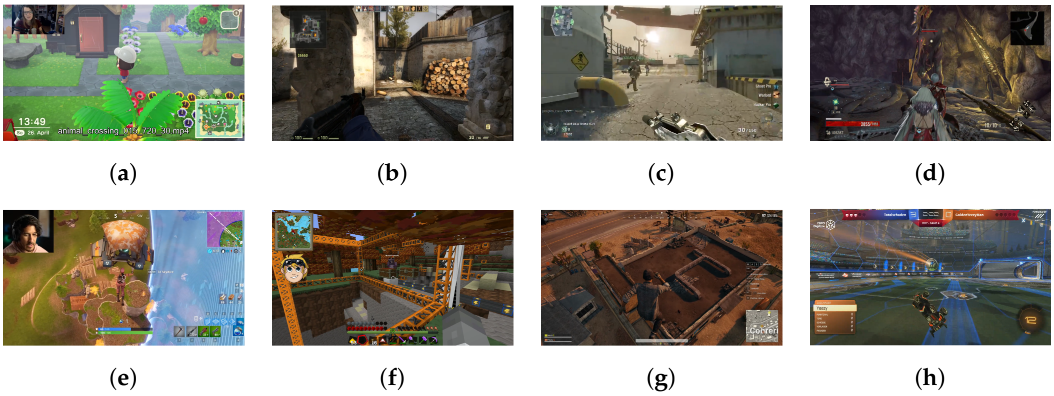


Figure 3. Frames of video stimuli. (**a**) Animal Crossing; (**b**) Counter-Strike 2; (**c**) Call of Duty; (**d**) Code Vein; (**e**) Fortnite; (**f**) Minecraft; (**g**) PUBG; (**h**) Rocket League.

3.2. Gaze Settings

Eye tracking on mobile devices can be achieved through various methods: using the mobile phone's camera with a software toolkit, positioning an external camera at an optimal angle to track gaze movements, or employing a dedicated hardware-based commercial solution mounted on the head with multiple cameras to record the user's fixations based on calibrations and the user's view. Given that this study focuses on evaluating user's experience in mobile gaming, it was important to make sure that the user interaction with a mobile device while playing a normal game should be preserved. Thus, the restrictions on user movement, head position, or even device orientation should not affect gaze tracking. Thus, we utilized the Pupil Core with a world-view camera and a monochrome camera setup to capture gaze fixations.

Following the mapping of the world view and gaze camera within the calibrated region using the gaze estimation function discussed in Section 2.1, the fixations plugin was employed to detect fixations. The plugin was configured with the following settings:

- Maximum Dispersion (degrees): 1.50
- Minimum Duration (milliseconds): 80
- Maximum Duration (milliseconds): 220

The minimum and maximum durations are confined to our requirements to ensure that fixations are neither missed nor duplicated, as users can quickly shift their gaze multiple times per second. The calibration of the mobile screen was performed using the built-in calibration function in Pupil Capture. Since subjects have varying head and face sizes, we ensured during calibration that the device was worn properly to maintain a confidence value close to 1 before commencing the subjective assessment. Pupil Player was utilized to view the world view and eye recordings and to export the data, as shown in Figure 2b. The exported data from each assessment consisted of CSV files containing information such as gaze and pupil positions, fixations, etc.

It is important to note that two independent sets of data were obtained from the experiment: the exported files and recordings from individual sessions via Pupil Player, and the mobile application data, which included user information and quality ratings from the Google Firebase database. The gaze confidence values are continuous. To calculate the mean gaze confidence for a specific time when the user viewed a stimulus and provided ratings, we needed to map the gaze data with timestamps stored in the mobile application database during the rating of each stimulus.

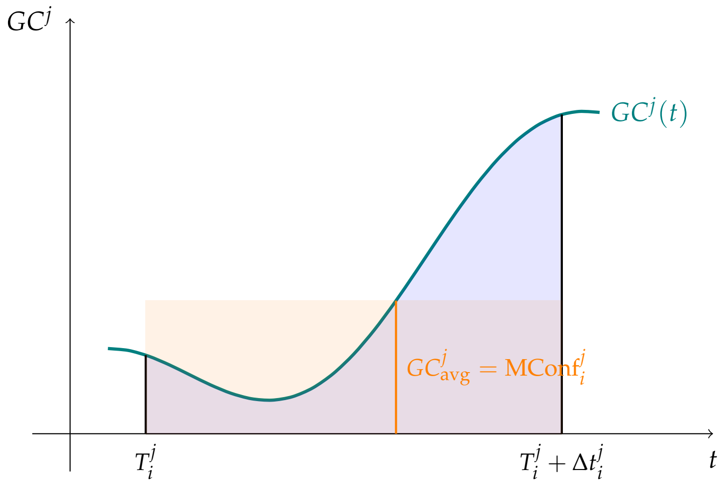
Let us assume that the gaze confidence of the j-th user at any time t is represented by $GC^j(t)$.

The value of confidence ranges between 0 and 1 with 1 being the highest value of confidence concerning the fixation marker. The Pupil Core stores both system and pupil time, which are convertible using the "start_time_synced_s" key. The system time was also stored in mobile application data after every rating. It is important to note that the timestamps obtained from Pupil Core data are not equidistant. To obtain the mean confidence for a specific video, the following formula can be derived:

After cropping the interval for a specific video of the j-th user, let $[T_i^j, T_i^j + \Delta t_i^j]$ denote the time interval of the i-th video of the j-th user, with T_i^j being the start time and Δt_i^j the time span of the i-th video of the j-th user, respectively. The mean confidence MConf_i^j of the i-th video of the j-th user can be determined as follows:

$$\mathrm{MConf}_i^j = \frac{1}{\Delta t_i^j} \int_{T_i^j}^{T_i^j + \Delta t_i^j} GC^j(t)dt \tag{3}$$

The above equation can be represented graphically as follows:



Thus, using the above equation, we calculated the average confidence for gaze estimation for each video clip viewed by a subject during the assessment.

3.3. Subjective Assessment

The subjective assessment was conducted using an Android application specifically designed to collect additional user information alongside video playback and quality ratings in our previous studies [20]. For this experiment, the application was modified to record timestamps for every user input. The application was run on a Redmi K60 Ultra mobile phone, equipped with a Dimensity 9200+ processor and 12 GB of RAM. As the users were wearing a Pupil Core headset and providing ratings on a mobile application as shown in Figure 1, it was not possible to adhere to all the recommendations provided in ITU-T P.910 [47] and ITU-R BT.500 [57]. However, the methodology for subjective assessment regarding the number of users, rating scales, etc., adhered to the ITU-T P.910 [47] recommendations. Before each assessment, users participated in a training session during which they received detailed instructions both verbally and in writing. Additionally, the subjective assessments adhered to the research ethics principles established by the Swedish Research Council [58].

Subsequently, users were shown one video clip from each game and asked to indicate whether they liked the game content on a binary scale ('Yes' or 'No'). They were also requested to rank their delight towards the game content on a 1–9 point Likert scale. After the delight ratings, the users were shown 4 videos from one game in a randomized order to obtain quality ratings. The MOS values were obtained using the five-point ACR scale using the Single-Stimulus method.

A total of 40 stimuli were presented to the subjects, with the option for playback available in the mobile application in cases of indecision. However, due to the time required for the training session and particularly for calibrating the Pupil Core headset to accommodate different head sizes and eye positions to achieve a confidence value close to 1, the total duration for individual sessions was approximately 13–15 min excluding the calibration process. Thirty-nine users participated in the assessment, the majority of whom were either bachelor's or master's students in Computer Science and familiar with quality ratings. The ratings of one subject were discarded due to inconsistency between the binary and Likert scales regarding delight; specifically, the subject selected 'Yes' on the binary scale but gave a rating lower than 4 on the Likert scale. Additionally, two outliers were detected based on MOS scores and subsequently excluded from the final calculations, resulting in a total of 36 users. Among these 36 participants, there were 24 males and 12 females, with a mean age of 26 years and a standard deviation of 11.21.

4. Results and Discussion

Given the selection of the most popular games, the majority of subjects had already played at least one of the eight games. Additionally, the short duration of the assessment ensured that subjects remained attentive throughout. However, the colors in some games, such as *Animal Crossing* and *Minecraft*, were not as vibrant compared to others. We also observed that users were polarized in their ratings for sandbox survival games like *Minecraft*, leading to wide error bars based on confidence intervals and statistically significant results in terms of delight.

The GoB and PoW values for video stimuli are presented in Table 2.

The results for Code Vein at 30fps are with a 1080p resolution, as shown in Table 1 and discussed in Section 3.1. The results reveal a dispersion in user ratings despite identical video quality in terms of resolution and frame rate. This dispersion, evident from the %GoB and %PoW metrics, indicates that factors other than video quality influence user experience across different game titles. One potential reason is that certain video stimuli, such as *Animal Crossing* and *Minecraft*, exhibit more muted colors compared to other games. Additionally, some clips may appear slightly faded depending on the specific in-game context.

Furthermore, the data suggest that the video resolution in close ranges, specifically 1080p and 720p, has a negligible effect on user ratings. In contrast, frame rate emerges as a significant determinant of user satisfaction. Consequently, under conditions of low-speed internet connections, a lower resolution paired with a higher frame rate is likely to yield higher user satisfaction. Typically, game developers program their games to automatically adjust to the user's native resolution or allow the user to select a lower or higher resolution based on device capabilities and internet connection speed. Our study offers insights suggesting that developers might consider prioritizing frame rate over resolution for an optimized user experience.

Table 2. GoB and PoW metrics of video stimuli.

	1080p, 60 fps		720p, 60 fps		720p, 30 fps	
Game Name	**GoB**	**PoW**	**GoB**	**PoW**	**GoB**	**PoW**
Animal Crossing	58.3%	0%	77.8%	2.8%	16.7%	11.1%
CSGO	91.7%	0%	88.9%	0%	77.8%	0%
Call of Duty	63.9%	8.3%	25%	13.9%	19.4%	41.7%
Code Vein	69.4%	0%	80.6%	2.8%	55.6% *	13.9% *
Fortnite	69.4%	5.6%	50%	5.6%	13.9%	69.4%
Minecraft	72.2%	5.6%	38.9%	33.3%	22.2%	27.8%
PUBG	91.7%	2.8%	66.7%	2.8%	33.3%	19.4%
Rocket League	77.8%	0%	66.7%	0%	61.1%	11.1%

Code Vein*: 1080p, 30 fps.

4.1. Impact of Delight of Shown Video Content

Besides video quality, various human and system factors can influence user ratings. To evaluate the human influence factors, we asked the subjects to give their opinion about liking a specific game in terms of graphics, content, gameplay, etc. We collected user feedback regarding their satisfaction with the displayed content using both binary and ordinal scales, as detailed in Section 3. Figure 4 presents the MOS values for four games on a binary scale in terms of user satisfaction, along with 95% confidence intervals (CI).

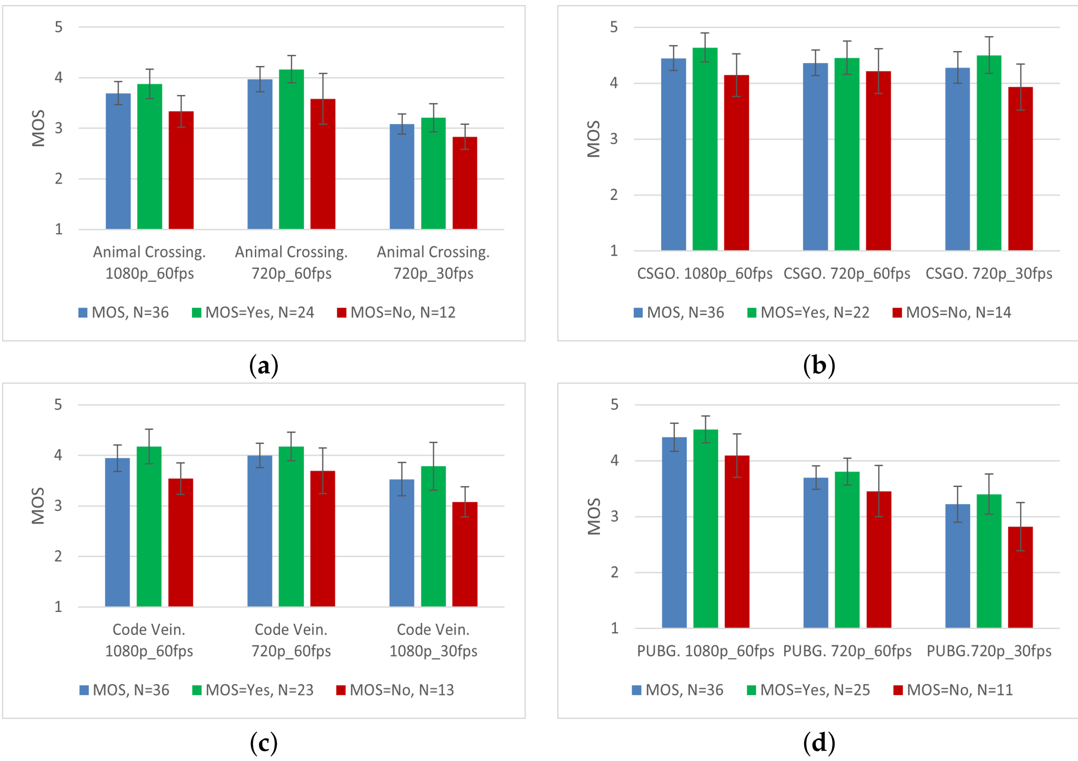


Figure 4. MOS of subjective assessment based on delight with 95% CI. (**a**) MOS_Animal Crossing; (**b**) MOS_Counter-Strike 2; (**c**) MOS_Code Vein; (**d**) MOS_PUBG.

The results demonstrate the effect of user delight (liking) on the presented content across various games. With the exception of the *PUBG* results, most outcomes from the remaining games are either statistically significant or nearly so. It is important to note that the sample size for user delight is inconsistent across the outcomes as shown in Figure 4, making it challenging to draw statistical conclusions. Therefore, after confirming the homogeneity of variance by using Levene's test [59], we also conducted a one-way ANOVA using SPSS across all scenarios to gain more relevant insights into the results. The one-way ANOVA outcomes for all games are provided in Table 3.

Table 3. Results of one-way ANOVA based on user delight.

	1080p, 60 fps		720p, 60 fps		720p, 30 fps	
Game Name	**F**	**Sig.**	**F**	**Sig.**	**F**	**Sig.**
Animal Crossing	6.004	0.020	5.696	0.023	3.290	0.079
CSGO	5.533	0.025	1.062	0.310	4.650	0.038
Call of Duty	4.208	0.048	5.822	0.021	1.985	0.168
Code Vein	6.896	0.013	4.075	0.051	4.877 *	0.034 *
Fortnite	4.333	0.045	2.188	0.148	4.294	0.046
Minecraft	6.149	0.018	11.565	0.002	10.536	0.003
PUBG	3.348	0.076	2.435	0.128	2.967	0.094
Rocket League	4.310	0.046	2.005	0.166	2.336	0.136

Code Vein*: 1080p, 30 fps.

The results highlight the impact of user delight on MOS ratings, underscoring the importance of selecting appropriate stimuli to assess the influence of human factors. Furthermore, they reveal that video stimuli designed with varying quality levels, as typically found in traditional QoE databases based on quality-of-service (QoS) metrics, are inadequate for investigating human influence factors. This is because quality degradation can significantly alter users' perceptions of the stimuli, overshadowing the intended focus on factors such as delight or mood. In our previous research [19,20], we could not achieve statistically significant results because the databases contained multiple stimuli with degradations due to QoS metrics like packet loss, jitter, and other factors. Thus, illustrating the impact of user delight was challenging, as the inconsistent content and user frustration with highly impaired stimuli led to insignificant outcomes.

4.2. Impact of User Gaze on Mobile Gaming Experience

Figure 5 shows a histogram of gaze confidence values derived from eye-tracking data for all users across the full set of stimuli, based on the MOS. According to the guidelines in the 'pupil_gaze_positions_info' file, which is generated during the export of recordings using Pupil Player, the confidence value reflects the accuracy of the pupil detector's measurements. A value of 0 indicates no confidence, meaning the position data should be ignored as no fixations are detected, while a value of 1 indicates perfect confidence. It is noted that useful data typically have a confidence value greater than approximately 0.6 as discussed in Section 2.1. The figure presents the MOS values across all games, stratified by confidence values and the number of subjects. It is important to note that this analysis only utilizes fixation data from the calibrated area. The results indicate that a greater number of users with higher gaze confidence were more likely to assign a higher MOS to videos with elevated resolution and frame rates. However, the ratings remain dispersed, as the distribution of subjects is relatively balanced between high and low opinion score categories across different resolutions. Additionally, it is important to note that the average gaze confidence level accounts for both the time spent watching the video and the time taken to provide a rating, as users had the option to replay certain stimuli if they felt uncertain.

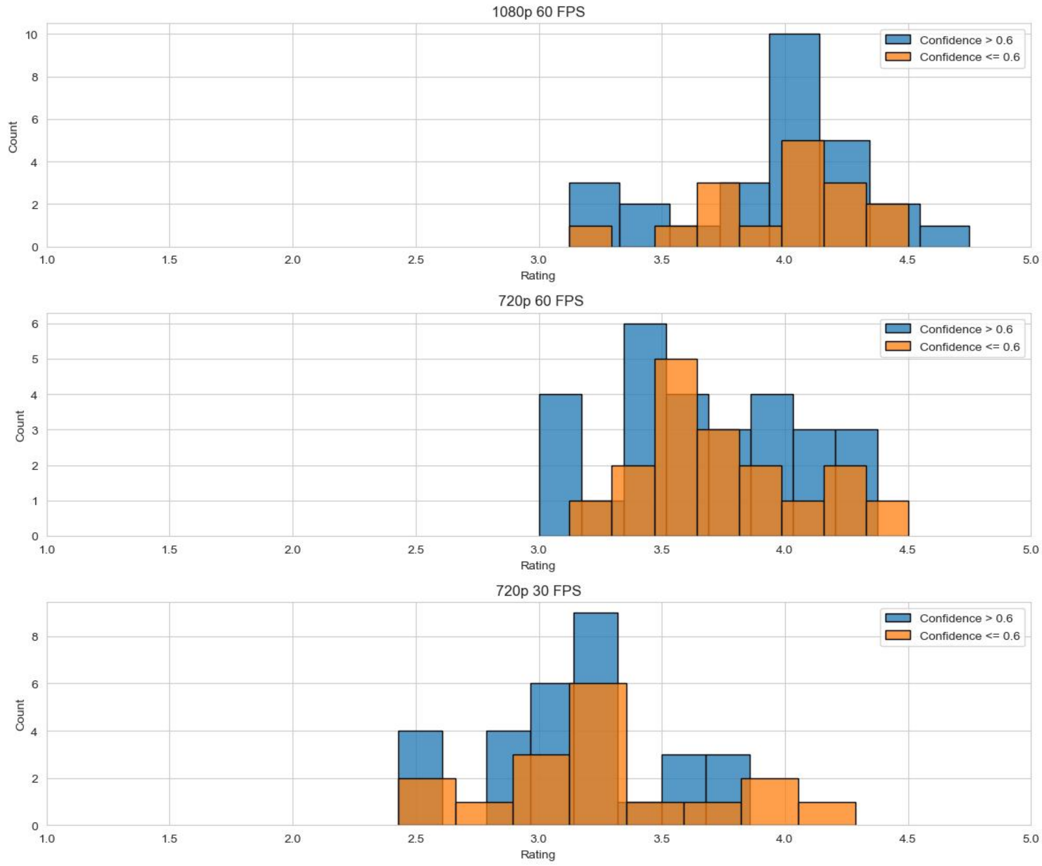


Figure 5. Histogram of user gaze.

Figure 6 presents the relative frequency [60] of user ratings for all games, categorized by the %GoB and %PoW rating scales, along with *Neutral* ratings on the ACR scale.

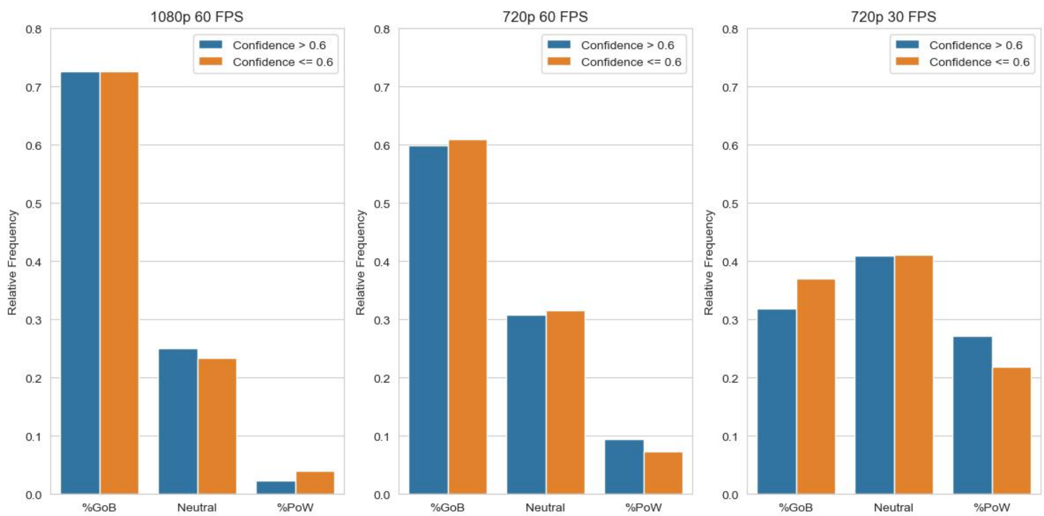


Figure 6. Relative frequency of gaze based on %GoB and %PoW ratings.

A noticeable trend emerged in which users with lower gaze confidence tended to rate the highest quality video (1080p at 60 frames per second) as either poor or bad. Conversely, for the lowest-quality video, users with higher gaze confidence provided feedback that was more closely aligned with the actual quality of the stimuli. Although this difference is evident but not statistically significant, it aligns with our expectations given the experimental design. The small area of interest on the mobile screen, the selection of gaming-related stimuli, and the brief assessment duration effectively focused participants' visual attention within the calibrated area, as observed during the assessments.

5. Conclusions

This study highlights the critical influence of human factors on mobile gaming experiences. As multimedia content, particularly video, dominates internet usage and the mobile gaming market continues to expand, optimizing gaming applications for mobile hardware constraints and adapting to network speed variations remain essential. Our research emphasizes that user delight significantly impacts the user ratings for most video stimuli, with higher frame rates preferred over screen resolution. These findings suggest that developers should prioritize higher frame rates to enhance user experience. Additionally, while gaze tracking shows the effect of visual attention, further investigation using advanced tracking modules along with multisensory methods is recommended.

Author Contributions: Conceptualization, O.N. and S.K.; methodology, O.N., S.K. and M.F.; software, O.N. and M.N.S.; subjective assessment, O.N. and M.N.S.; results compilation, O.N., S.K. and M.N.S.; validation, M.F. and S.K.; writing—original draft, O.N.; writing—review and editing, S.K. and M.F.; supervision, S.K. and M.F. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: Data are contained within the article. The Pupil Core exported files, SPSS data file, and Firebase database can be obtained by contacting the authors and under the fair use policy. Any publication using the data should acknowledge the DITE, Blekinge Tekniska Hogskola, Karlskona, Sweden, and the current article needs to be cited in the paper. The mobile application can also be made available for subjective assessment, incorporating additional collaboration factors, by contacting the authors.

Conflicts of Interest: Author Muhammad Nauman Sheikh was employed by the company Dubizzle Labs. The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Abbreviations

The following abbreviations are used in this manuscript:

ACR	Absolute Category Rating
ANOVA	Analysis of Variance
GoB	Good or Better
IFs	Influence factors
ITU-T	International Telecommunication Union Telecommunication Standardization Sector
MOS	Mean Opinion Score
Mulsemedia	Multiple Sensorial Media
PoW	Poor or Worse
QoE	Quality of Experience
QoS	Quality of Service

References

1. Cisco Annual Cisco Annual Internet Report (2018–2023)-White Paper. Available online: https://www.cisco.com/c/en/us/solutions/collateral/executive-perspectives/annual-internet-report/white-paper-c11-741490.html (accessed on 21 February 2024).
2. Ericsson Mobility Report November 2022. Available online: https://www.ericsson.com/en/reports-and-papers/mobility-report/reports/november-2022 (accessed on 21 February 2024).
3. Precedence Research: Mobile Gaming Market Size. Available online: https://www.precedenceresearch.com/mobile-gaming-market (accessed on 16 June 2024).
4. Console vs. Mobile Gaming–Understanding Trends Across 18 Markets. Available online: https://business.yougov.com/content/48108-console-vs-mobile-gaming-understanding-trends-across-18-markets (accessed on 14 April 2024).
5. New Free Report: Explore the Global Games Market in 2023. Available online: https://newzoo.com/resources/blog/explore-the-global-games-market-in-2023 (accessed on 14 April 2024).

6. Infographic: Mobile and Console Games Dominate Video Game Market. Available online: https://www.statista.com/chart/30560/estimated-global-video-game-revenues-by-segment (accessed on 12 April 2024).
7. Mobile Games-Worldwide | Statista Market Forecast. Available online: https://www.statista.com/outlook/dmo/digital-media/video-games/mobile-games/worldwide (accessed on 14 April 2024).
8. Qualinet. Qualinet White Paper on Definitions of Quality of Experience. Technical Report. Library Catalog. 2013. Available online: www.qualinet.eu (accessed on 20 August 2024).
9. Reiter, U.; Brunnström, K.; De Moor, K.; Larabi, M.C.; Pereira, M.; Pinheiro, A.; You, J.; Zgank, A. Factors Influencing Quality of Experience. In *Quality of Experience: Advanced Concepts, Applications and Methods*; Möller, S., Raake, A., Eds.; T-Labs Series in Telecommunication Services; Springer International Publishing: Cham, Switzerland, 2014; pp. 55–72.
10. ITU-T Rec. P.800.1. Mean Opinion Score (MOS) Terminology. Available online: https://www.itu.int/rec/T-REC-P.800.1-201607-I/en (accessed on 27 May 2024).
11. Streijl, R.C.; Winkler, S.; Hands, D.S. Mean Opinion Score (MOS) Revisited: Methods and Applications, Limitations and Alternatives. *Multimed. Syst.* **2016**, *22*, 213–227. [CrossRef]
12. Pinson, M.H.; Janowski, L.; Pepion, R.; Huynh-Thu, Q.; Schmidmer, C.; Corriveau, P.; Younkin, A.; Le Callet, P.; Barkowsky, M.; Ingram, W. The Influence of Subjects and Environment on Audiovisual Subjective Tests: An International Study. *IEEE J. Sel. Top. Signal Process.* **2012**, *6*, 640–651. [CrossRef]
13. Hoßfeld, T.; Heegaard, P.E.; Varela, M.; Möller, S. QoE beyond the MOS: An in-depth look at QoE via better metrics and their relation to MOS. *Qual. User Exp.* **2016**, *1*, 2. [CrossRef]
14. Hoßfeld, T.; Heegaard, P.E.; Skorin-Kapov, L.; Varela, M. Deriving QoE in systems: From fundamental relationships to a QoE-based Service-level Quality Index. *Qual. User Exp.* **2020**, *5*, 7. [CrossRef]
15. Moor, K.D.; Fiedler, M.; Reichl, P.; Varela, M. Quality of Experience: From Assessment to Application (Dagstuhl Seminar 15022). *Dagstuhl Rep.* **2015**, *5*, 57–95.
16. Zhu, Y.; Heynderickx, I.; Redi, J.A. Understanding the role of social context and user factors in video Quality of Experience. *Comput. Hum. Behav.* **2015**, *49*, 412–426. [CrossRef]
17. Zhu, Y.; Chandra, G.S.; Lin, W.; Ghinea, G.; Redi, J.A. Measuring Individual Video QoE. In *ACM Transactions on Multimedia Computing, Communications, and Applications (TOMM)*; ACM PUB27: New York, NY, USA, 2018.
18. Minhas, T.N.; Nawaz, O.; Fiedler, M.; Khatibi, S. The Effects of Additional Factors on Subjective Quality Assessments. In Proceedings of the 2019 2nd International Conference on Advancements in Computational Sciences (ICACS), Lahore, Pakistan, 18–20 February 2019; pp. 1–5. [CrossRef]
19. Nawaz, O.; Fiedler, M.; Khatibi, S. Impact of Human and Content Factors on Quality of Experience of Online Video Streaming. In Proceedings of the 17th International Joint Conference on e-Business and Telecommunications, Online and Lieusaint-Paris, France, 8–10 July 2020; pp. 59–66. [CrossRef]
20. Nawaz, O.; Fiedler, M.; Khatibi, S. QoE-Based Performance Comparison of AVC, HEVC, and VP9 on Mobile Devices with Additional Influencing Factors. *Electronics* **2024**, *13*, 329. [CrossRef]
21. Sundstedt, V.; Bernhard, M.; Stavrakis, E.; Reinhard, E.; Wimmer, M. Visual Attention and Gaze Behavior in Games: An Object-Based Approach. In *Game Analytics: Maximizing the Value of Player Data*; Seif El-Nasr, M., Drachen, A., Canossa, A., Eds.; Springer: London, UK, 2013; pp. 543–583.
22. Sundstedt, V.; Navarro, D.; Mautner, J. Possibilities and challenges with eye tracking in video games and virtual reality applications. In Proceedings of the SIGGRAPH ASIA 2016 Courses, New York, NY, USA, 5 December 2016; SA '16, pp. 1–150.
23. Zalewski, D. Eye Tracking and Video Games Research. Available online: https://tryevidence.com/blog/eye-tracking-and-video-games-research/ (accessed on 2 May 2024).
24. Cheng, S.; Ping, Q.; Wang, J.; Chen, Y. EasyGaze: Hybrid eye tracking approach for handheld mobile devices. *Virtual Real. Intell. Hardw.* **2022**, *4*, 173–188. [CrossRef]
25. Gunawardena, N.; Ginige, J.A.; Javadi, B. Eye-tracking Technologies in Mobile Devices Using Edge Computing: A Systematic Review. *ACM Comput. Surv.* **2022**, *55*, 158:1–158:33. [CrossRef]
26. Kassner, M.; Patera, W.; Bulling, A. Pupil: An open source platform for pervasive eye tracking and mobile gaze-based interaction. In Proceedings of the 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing: Adjunct Publication, UbiComp '14 Adjunct, New York, NY, USA, 13–17 September 2014; pp. 1151–1160. [CrossRef]
27. Pupil Core-Open Source Eye Tracking Platform. Available online: https://pupil-labs.com/products/core (accessed on 28 May 2024).
28. Pupil Core-Fixations. Available online: https://docs.pupil-labs.com/core/terminology/#fixations (accessed on 14 July 2024).
29. Sony Play Station 4 Tech Specs. Available online: https://www.playstation.com/en-us/ps4/tech-specs/ (accessed on 17 July 2024).
30. Capozzi, F.; Ristic, J. How attention gates social interactions. *Ann. N. Y. Acad. Sci.* **2018**, *1426*, 179–198. [CrossRef] [PubMed]
31. Chica, A.B.; Lasaponara, S.; Chanes, L.; Valero-Cabré, A.; Doricchi, F.; Lupiáñez, J.; Bartolomeo, P. Spatial attention and conscious perception: The role of endogenous and exogenous orienting. *Atten. Percept. Psychophys.* **2011**, *73*, 1065–1081. [CrossRef] [PubMed]

32. Dalmaso, M. Exploring the Social Environment with the Eyes: A Review of the Impact of Facial Stimuli on Saccadic Trajectories. *Int. J. Environ. Res. Public Health* **2022**, *19*, 16615. [CrossRef] [PubMed]
33. Mulckhuyse, M.; Theeuwes, J. Unconscious attentional orienting to exogenous cues: A review of the literature. *Acta Psychol.* **2010**, *134*, 299–309. [CrossRef] [PubMed]
34. Baltrušaitis, T.; Robinson, P.; Morency, L.P. OpenFace: An open source facial behavior analysis toolkit. In Proceedings of the 2016 IEEE Winter Conference on Applications of Computer Vision (WACV), Lake Placid, NY, USA, 7–10 March 2016; pp. 1–10. [CrossRef]
35. Kreiensieck, E.; Ai, Y.; Zhang, L. A Comprehensive Evaluation of OpenFace 2.0 Gaze Tracking. In Proceedings of the Human Computer Interaction, Copenhagen, Denmark, 23-28 July 2023; Springer: Cham, Switzerland, 2023; pp. 532–549. [CrossRef]
36. Jeon, M.; Lee, J.H. The Ecological AUI (Auditory User Interface) Design and Evaluation of User Acceptance for Various Tasks on Smartphones. In Proceedings of the Human-Computer Interaction. Interaction Modalities and Techniques, Las Vegas, NV, USA, 21–26 July 2013; Kurosu, M., Ed.; Springer: Berlin/Heidelberg, Germany, 2013; pp. 49–58. [CrossRef]
37. Kim, H.; Lee, I.K. Studying the Effects of Congruence of Auditory and Visual Stimuli on Virtual Reality Experiences. *IEEE Trans. Vis. Comput. Graph.* **2022**, *28*, 2080–2090. [CrossRef] [PubMed]
38. Bahr, G.S.; Ford, R.A. How and why pop-ups don't work: Pop-up prompted eye movements, user affect and decision making. *Comput. Hum. Behav.* **2011**, *27*, 776–783. [CrossRef]
39. Kwok, C.K. Understanding User Engagement Level During Tasks via Facial Responses, Eye Gaze and Mouse Movements. Ph.D. Thesis, Hong Kong Polytechnic University, Hong Kong, 2018.
40. Mesfin, G.; Hussain, N.; Covaci, A.; Ghinea, G. Using Eye Tracking and Heart-Rate Activity to Examine Crossmodal Correspondences QoE in Mulsemedia. *ACM Trans. Multimed. Comput. Commun. Appl.* **2019**, *15*, 34:1–34:22. [CrossRef]
41. Wibirama, S.; Nugroho, H.A. Towards understanding addiction factors of mobile devices: An eye tracking study on effect of screen size. In Proceedings of the 2017 39th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), Jeju, Republic of Korea, 11–15 July 2017; pp. 2454–2457. [CrossRef]
42. Jiang, J.Y.; Guo, F.; Chen, J.H.; Tian, X.H.; Lv, W. Applying Eye-Tracking Technology to Measure Interactive Experience Toward the Navigation Interface of Mobile Games Considering Different Visual Attention Mechanisms. *Appl. Sci.* **2019**, *9*, 3242. [CrossRef]
43. Pupil Core-Pupil Player. Available online: https://docs.pupil-labs.com/core/software/pupil-player/ (accessed on 14 July 2024).
44. Pupil Core-Confidence. Available online: https://docs.pupil-labs.com/core/software/pupil-player/#raw-data-exporter (accessed on 14 July 2024).
45. P.10/G.100, I. Vocabulary for Performance, Quality of Service and Quality of Experience. Recommendation (11/2017), ITU-T. 2017. Available online: https://www.itu.int/rec/T-REC-P.10 (accessed on 20 August 2024).
46. Schleicher, R.; Antons, J.N. Evoking Emotions and Evaluating Emotional Impact. In *Quality of Experience: Advanced Concepts, Applications and Methods*; Möller, S., Raake, A., Eds.; T-Labs Series in Telecommunication Services; Springer International Publishing: Cham, Switzerland, 2014; pp. 121–132.
47. ITU-T Rec. P.910. Subjective Video Quality Assessment Methods for Multimedia Applications. Available online: https://www.itu.int/rec/T-REC-P.910-202310-I/en (accessed on 27 May 2024).
48. Le Maguer, S.; King, S.; Harte, N. The limits of the Mean Opinion Score for speech synthesis evaluation. *Comput. Speech Lang.* **2024**, *84*, 101577. [CrossRef]
49. One Way Analysis of Variance (ANOVA)-StatsDirect. Available online: https://www.statsdirect.com/help/analysis_of_variance/one_way.htm (accessed on 14 August 2024).
50. Pupil Core-Calibration. Available online: https://docs.pupil-labs.com/core/software/pupil-capture/#calibration (accessed on 14 June 2024).
51. Google Firebase. Available online: https://firebase.google.com/ (accessed on 2 November 2023).
52. Laboratory for Image and Video Engineering-The University of Texas at Austin. Available online: https://live.ece.utexas.edu/research/LIVE-YT-Gaming/index.html (accessed on 20 June 2024).
53. YouTube Gaming. Available online: https://www.youtube.com/gaming (accessed on 20 June 2024).
54. Yu, X.; Ying, Z.; Birkbeck, N.; Wang, Y.; Adsumilli, B.; Bovik, A.C. Subjective and Objective Analysis of Streamed Gaming Videos. *IEEE Trans. Games* **2023**, *16*, 445–458. [CrossRef]
55. Fröhlich, P.; Egger, S.; Schatz, R.; Mühlegger, M.; Masuch, K.; Gardlo, B. QoE in 10 seconds: Are short video clip lengths sufficient for Quality of Experience assessment? In Proceedings of the 2012 Fourth International Workshop on Quality of Multimedia Experience, Melbourne, VIC, Australia, 5–7 July 2012; pp. 242–247. [CrossRef]
56. Nawaz, O.; Minhas, T.N.; Fiedler, M. QoE based comparison of H.264/AVC and WebM/VP8 in an error-prone wireless network. In Proceedings of the Integrated Network and Service Management (IM), 2017 IFIP/IEEE Symposium, Lisbon, Portugal, 8–12 May 2017; pp. 1005–1010.
57. ITU-R Rec. BT.500. Methodologies for the Subjective Assessment of the Quality of Television Images. Available online: https://www.itu.int/rec/R-REC-BT.500-15-202305-I/en (accessed on 27 May 2024).
58. *Good Research Practice*; Swedish Research Council: Stockholm, Sweden, 2017; ISBN 978-91-7307-354-7

59. NIST. Levene Test for Equality of Variances. In *Engineering Statistics Handbook*; National Institute of Standards and Technology, U.S. Department of Commerce: Gaithersburg, MD, USA, 2012.
60. Jambu, M. Chapter 3-1-D Statistical Data Analysis. In *Exploratory and Multivariate Data Analysis*; Jambu, M., Ed.; Academic Press: Boston, MA, USA, 1991; pp. 27–62. [CrossRef]