Computers in Human Behavior Reports 22 (2026) 101107

Contents lists available at ScienceDirect

Computers in Human Behavior Reports

journal homepage: www.sciencedirect.com/journal/computers-in-human-behavior-reports

# “Did you see it?” visual anchoring and attentional effectiveness in the context of live-streaming using eye-tracking

Sebastiano Accardi [a,b,*], Margherita Zito [a,b], Vincenzo Russo [a,b]

[a] *Department of Business Law Economics and Consumer Behaviour Carlo A. Ricciardi IULM University, Italy*
[b] *Behavior and Brain Lab IULM Neuromarketing Research Center IULM University, Milan, Italy*

A B S T R A C T

Cognitive and marketing research shows that human faces act as strong attentional magnets for consumers. Yet, communicative effectiveness also depends on spatial configuration, with proximity between faces and advertising elements emerging as a critical factor. In this regard, Social Live Streaming Services (SLSS) platforms offer an ecologically valid context for examining these dynamics, due to their capacity to generate prolonged user exposure, the growing interest of companies investing in streamer sponsorships, and the streamers' freedom to position on-screen elements as they see fit. The present study investigated the effects of spatial brand placement in video game live-streaming content, comparing brand positioning near the streamer's face with peripheral placement, across both esports and single-player gaming formats. A mixed experimental design was integrated with 52 participants, and eye-tracking methodology was used to objectively quantify the attention distribution, along with surveys to measure its rational implications. Results indicate that advertising banners placed close to the streamer's face capture significantly greater visual attention than those located in peripheral screen areas, but it does not automatically lead to a persuasive path. Spatial contiguity with the face thus enhances attention to sponsored content, marking an early perceptual stage that may precede higher-level advertising outcomes and inform optimal overlay design. This distinction offers practical insights for companies and content creators aiming to optimize overlay design and maximize viewers' exposure to branded elements.

## 1. Introduction

In today's media landscape, the exponential growth of available content parallels a decline in audience attention spans. This dynamic is effectively captured by the Limited Capacity Model of Motivated Mediated Message Processing (LC4MP), which explains that individuals have finite cognitive resources (Lang, 2006). As a consequence, not all promotional messages encountered by the audience are perceived, encoded, or retained. It has become essential to identify stimuli that reduce attention dispersion and facilitate visual perception. For this reason, the principle of contiguity (Mayer & Anderson, 1992) provides the theoretical background for identifying a strong attentional stimulus capable of competing with increasingly dense video stimuli. In digital contexts, this is particularly evident in Social Live Streaming Services (SLSS), where viewers are exposed to multiple stimuli, such as content, real-time chat interactions, and overlay advertisements (Mancini et al., 2022), all competing for limited attentional resources. Identifying key elements that can optimize brand placement within media content gains importance. The presence of the creator strongly influences consumers' evaluation of the endorsed brand (Deska et al., 2022). Most notably, the visibility of the human face plays a central role, not only in capturing visual attention independent of content type (Shi et al., 2024), but also in effectively directing that attention toward the advertised product or brand (Adil et al., 2018). Therefore, by exploiting the principle of spatial contiguity, the face represents a strategic and stable element within media content as a key visual resource (Harrison, 2003). In a fast-paced environment, this simple, human element could be the game-changer for improving visual attention to sponsored elements. Despite extensive evidence on face salience, it remains unclear whether spatial proximity to a face systematically guides visual attention in complex media environments, such as live-streaming.

This study is framed as a contribution to the investigation of attentional mechanisms in mediated environments, emphasizing early-stage perceptual processes rather than directly assessing advertising effectiveness. The theoretical and structural design of this work aims to investigate the role of the human face as a catalyst for attention to sponsored brands in a modern context with a growing user base, such as SLSS platforms, using mixed-effects models to account for individual-level variability in visual attention.

* Corresponding author. Department of Business Law Economics and Consumer Behaviour Carlo A. Ricciardi IULM University, Italy.
*E-mail address:* sebastiano.accardi@studenti.iulm.it (S. Accardi).

https://doi.org/10.1016/j.chbr.2026.101107
Received 30 September 2025; Received in revised form 18 April 2026; Accepted 13 May 2026
Available online 14 May 2026

### *1.1. The face as a stable visual anchor*

The nature of human cognition limitation, as postulated by the LC4MP model, identifying the most effective cue to guide visual attention toward specific elements is a functional strategy. Human faces represent evolutionarily prioritized stimuli that automatically attract attention, engaging scarce cognitive resources even within the visually saturated ecosystem of social media (Palermo & Rhodes, 2007). The scientific literature has extensively demonstrated how the human face is not merely a narrative or decorative feature but rather a crucial perceptual anchor that directs attention in meaningful ways (Finzi et al., 2021; Kanwisher et al., 1997; Ro et al., 2001; Visconti di Oleggio Castello and Gobbini, 2015). This remarkable efficiency in detecting faces within a visual field could make the face a fundamental cue for facilitating the operation of the contiguity principle. Specifically, this principle posits that information is processed more rapidly and effectively when spatially close rather than distant (Mayer & Fiorella, 2014), thereby minimizing the need to split attentional resources across disparate areas of content (Schroeder & Cenkci, 2018). This principle differs from the Gestalt principle of proximity (Brunswik & Kamiya, 1953), which describes the perceptual tendency to group nearby elements as belonging to a single visual unit. In contrast, the contiguity principle considers spatial proximity as a condition that facilitates the joint processing of distinct elements.

Following this framework, the face may function as a bottom-up attentional trigger that biases attention toward nearby elements, thereby facilitating subsequent voluntary orienting compared to more spatially distant elements. This mechanism is further supported by evidence that faces are processed in specialized neural regions such as the fusiform face area (FFA) (Kanwisher et al., 1997), which are functionally distinct from object-processing areas such as the lateral occipital complex (LOC) (Grill-Spector, 2003), reducing direct competition for cognitive resources. However, it remains unclear whether the face, by leveraging the principle of spatial contiguity, may effectively function as an attentional cue that guides viewers' gaze toward adjacent elements. This mechanism should be particularly relevant when a face, an inherently salient and socially meaningful stimulus, could amplify attention toward nearby elements, including those strategically integrated for communicative or commercial purposes.

### *1.2. Eye tracking in consumer neuroscience*

The face's attentional role and the integration of branded elements in live-streaming make the consumer neuroscience methodology particularly suited to exploring the phenomenon of attentional effectiveness. Neuroscientific methodologies provide a decisive contribution by employing tools that enable a more thorough investigation of consumers' implicit reactions (Russo et al., 2023), allowing researchers to monitor implicit processes in real time (Mileti et al., 2016). In this way, they overcome the inherent constraints of self-report measures, which rely on post-hoc rationalization through questionnaires administered after exposure to target stimuli (Missaglia et al., 2017). Moreover, the findings using eye-tracking methodology indicate that banner ads featuring faces elicit a longer average fixation duration than those without (Casado-Aranda et al., 2023). Among the neuroscience tools, the eye tracker is the one most capable and accurate for recording cognitive and attentional responses (Santos et al., 2015), especially when linked to logos or brands (Russo et al., 2023). Within consumer neuroscience, eye-tracking provides an appropriate empirical approach for capturing fine-grained attentional dynamics and for exploring how the spatial contiguity principle is reflected in eye movements.

The eye-tracker provides essential output measures for quantifying users' visual attention, including fixations (Zhou & Xue, 2021), revisits (Ou et al., 2025), and dwell time expressed in milliseconds (Russo et al., 2023), which indicate both the activation and engagement within specific Areas of Interest (AOIs). Moreover, transition matrices between AOIs enable an objective analysis of gaze shifts, revealing when and how viewers shift their focus from one defined region to another (Burch & Timmermans, 2020). Eye movement analysis serves as a fundamental methodology for investigating implicit attention processes. In contrast to questionnaires, which assess post-stimulus effects, eye trackers enable real-time examination of attentional processes as reflected in eye movements (Wedel & Pieters, 2008). This approach yields essential information for decoding attentive gaze, particularly when participants are exposed to visually complex stimuli (Mancini et al., 2022; Seo et al., 2018). The eye tracker is widely used to examine the spatial contiguity principle, particularly in studies on learning outcomes (Johnson and Mayer, 2012; Makransky et al., 2019; Zhang, 2021). However, research in consumer neuroscience remains limited, despite the fact that this principle, originating in the learning domain, could be highly relevant within this framework. Through the analysis of eye movements, consumer neuroscience provides valuable insights into the mechanisms underlying selective attention (Resnick & Albert, 2014). Nevertheless, prior attention to an element does not automatically imply its subsequent encoding or conscious perception. From a consumer neuroscience perspective, attentional capture represents only the first stage of message processing and must be distinguished from later phases of memory consolidation and explicit evaluative judgment (Missaglia et al., 2017). This distinction is crucial when the stimuli involve fictitious or unfamiliar brands (Folkvord et al., 2020).

Taken together, these findings underscore the central role that faces may have in persuasive communication strategies through three steps. First, the face acting as an attentional mediator, may elicit an automatic orienting response triggered by the number of fixations on human facial features. Second, the focal-anchor effect occurs around the face, which consequently could facilitate greater visual attention toward proximity elements, measuring by revisits and the time spent in the face or around it. Finally, as a result of these preceding stages, a reduction of banner blindness could emerge, driven by the increased visual attention elicited by the face's attentional cue, and measured by transitions of the elements' proximity. In this chain of cognitive mechanisms governing visual attention, the salience of the scene or event itself can play a crucial role in facilitating visual attention toward specific elements (Mancini et al., 2022), particularly those close to the face. The face may operate as a bottom-up attentional trigger, acting as a visual stable anchor detected through fixations. This cue may then facilitate a top-down orienting process toward spatially proximal elements, increasing transitions toward the nearby objects. When this attentional shift elicits interest, it may sustain engagement through longer dwell times or renewed attention reflected in revisits, thereby counteracting habituation effects and mitigating banner blindness.

Although previous research has consistently shown that faces are powerful attentional anchors, less is known about how this bias operates when branded elements appear in close spatial proximity to the face. SLSS platforms such as Twitch offer an ideal environment for investigating this mechanism, where the streamer's constant presence on-screen could allow both brands and creators to strategically leverage this attention dynamic to improve the communicative impact and attentional effectiveness of on-screen sponsored elements.

### *1.3. Live-stream overlays: constraints & opportunities*

SLSS platforms have attracted many users by allowing them to watch a variety of content and share it with the rest of the community in real-time. Twitch. tv is among the most popular social live-streaming platforms, extending beyond gaming content (Sjöblom et al., 2017) and attracting growing attention from psychology and business research (De Wit et al., 2020; Johnson & Woodcock, 2019a; Pollack et al., 2021). Professional streamers utilize advertising through sponsorships with medium to large corporations, often featuring brand overlays during live-streaming (Johnson & Woodcock, 2019b). Editing the overlay of a live-streaming allows content creators to incorporate ads easily

(Hamilton et al., 2014). Although faces strongly attract attention, even in static ads (Sajjacholapunt & Ball, 2014), it remains unclear whether positioning ads near the facecam enhances perception across different content types such as esports or single-player games. To develop evidence-based guidelines for effective overlay design, it is crucial to understand how users visually engage with streamed content and where attention is most often directed. The live-streaming content largely depends on the game or context, giving the creator limited control over the flow of attention (Mancini et al., 2022; Seo et al., 2018). To date, the scientific literature has primarily focused on the motivations driving users to watch live-streams (Bian & Yang, 2025; Sjöblom & Hamari, 2017; Törhönen et al., 2020), largely overlooking how these users attend to visual content, particularly about branded elements. Recent eye-tracking studies on esports games show that banners attract attention during salient events and influence brand attitude and implicit memory (Min et al., 2025; Seo et al., 2018). Mancini et al. (2022) found that the facecam area receives more attention than peripheral banners, suggesting that facial proximity may enhance sponsorship visibility. However, since their study tested only distant placements and esports games, it remains unclear whether spatial closeness to the face further amplifies attention or brand perception, or whether other game genres exhibit different patterns. An open question crucial for optimizing Overlay Live-Streaming Advertisements (OLSA) and determining the optimal position of overlay elements to increase attentional effectiveness. Moreover, most research focuses almost exclusively on esports, overlooking single-player games and limiting generalizability to the broader streaming audience.

### 1.4. Research gap and hypotheses development

Although numerous studies have confirmed the role of the human face as a powerful attractor of visual attention (Finzi et al., 2021; Kanwisher et al., 1997; Ro et al., 2001; Visconti di Oleggio Castello and Gobbini, 2015), limited research has explored how this attentional bias can be strategically leveraged for the spatial placement of sponsored elements within complex media environments (Mancini et al., 2022; Min et al., 2025; Seo et al., 2018). Moreover, research has focused mainly on esports, leaving single-player games underexplored despite their relevance in the wider live-streaming ecosystem. Our study aims to address these gaps by applying the principle of spatial contiguity within a consumer neuroscience framework and examining whether the content creator's face functions as an attentional facilitator guiding viewers' visual attention toward nearby elements in dynamic live-streaming contexts. This is investigated through two distinct research cores. The primary core examines the principle of contiguity by investigating how the face, as a salient attentional cue, can implicitly facilitate gaze transitions and guide attentional allocation toward nearby sponsored elements, as measured through eye-tracking indices (H1–H4). Subsequently, the second core investigates whether the contiguity principle translates into explicit brand outcomes, including memory, interest and purchase intention, assessed via self-report measures (H5).

More specifically, H1 tests how content type modulates the spatial distribution of visual attention (dwell time, fixations, revisits) and assesses the face as a visually stable anchor. H2 tests whether gaze transitions from facecam to banner confirm the face's role as an attentional cue toward nearby elements. H3 quantifies visual attention facilitated by spatial proximity between the banner and the face using the same indices. H4 examines how scene salience modulates this proximity effect, enhancing attention during emotionally charged or goal-related moments. Finally, H5 investigates the subsequent stage of brand impact (interest, purchase intention, and recall) through self-report measures (see Table 1). The conceptual model is presented in Fig. 1.

**Table 1**
Summary of research questions and hypotheses, associated method, stimuli and metrics.

| Research Question (RQ) | Hypotheses | Method | Associated stimuli | Metrics |
|---|---|---|---|---|
| **RQ1**: Do viewers' visual exploration during gaming content vary depending on the type of in-game camera used (platformer vs. esports) in live-streaming? | **H1a**: In Platformer content, characterized by a centrally fixed camera, viewers' attention is expected to converge predominantly within the central regions of the screen. | Eye-Tracker | Gridded AOI | Dwell time (ms) |
| | **H1b**: In Esports content adopting an isometric perspective, viewers' attention is expected to be more widely distributed across peripheral areas. | Eye-Tracker | Gridded AOI | Dwell time (ms) |
| | **H1c**: In Esports, the viewers' attention is expected to be more focused on the facecam then platformer content. | Eye-Tracker | Facecam | Fixations (count)<br>Revisits (count)<br>Dwell time (ms) |
| **RQ2**: Does placing a sponsored banner in proximity to the facecam facilitate transitions from the facecam toward the banner due to the presence of the face? | **H2**: Advertising banners positioned in proximity to the facecam receive a significantly higher number of visual transitions compared to those placed in areas more distant from the streamer's face. | Eye-Tracker | Facecam to Banner | Transitions (count) |
| **RQ3**: Does the proximity of a banner to the facecam significantly increase visual attention toward it? | **H3**: Advertising banners positioned in proximity to the facecam elicit a significantly higher number of fixations, revisits, and total dwell time compared to those placed at a greater distance from the facecam. | Eye-Tracker | Banner | Fixations (count)<br>Revisits (count)<br>Dwell time (ms) |
| **RQ4**: During salient scenes (e.g., when the streamer scores a point), does the proximity of the banner to the facecam further enhance visual attention compared to non-salient scenes? | **H4**: Advertising banners positioned in proximity to the facecam elicit a significantly higher number of fixations, revisits, and dwell time during scenes characterized by a salient scene. | Eye-Tracker | Banner | Fixations (count)<br>Revisits (count)<br>Dwell time (ms) |
| **RQ5**: Does placing a | **H5a**: Brands associated with | Survey | Brand | Brand Recognition |

(continued on next page)

**Table 1** (*continued*)

| Research Question (RQ) | Hypotheses | Method | Associated stimuli | Metrics |
|---|---|---|---|---|
| banner near the facecam in live-streamed content improve brand recall and interest? | banners positioned in proximity to the facecam achieve significantly higher levels of brand recognition compared to those placed in more distant locations. | | | |
| | **H5b:** Brands associated with banners positioned in proximity to the facecam elicit significantly greater brand interest compared to those placed farther from the facecam. | Survey | Brand | Brand Interest |
| | **H5c:** Brands associated with banners positioned in proximity to the facecam generate significantly higher levels of purchase intention compared to those placed in more distant areas of the screen. | Survey | Brand | Advertised Brand Purchase Intention |

## 2. Methods

### 2.1. Sample

The sample comprised 52 participants (27 men, 25 women; M_age = 29.50, SD = 8.34), aged 18–49, reflecting Twitch's active age range (Statista, n.d.). Most were gamers (65.4%) and regular gaming content viewers (59.6%), with 61.5% reporting energy drink consumption. Sample characteristics are in Table 2.

Participants were randomly assigned to experimental conditions. A 2-way ANOVA confirmed, using age as the dependent variable, the distribution across experimental conditions ($F(1, 48) = 1.51$, $p = .224$, $\eta^2 = 0.03$), gender ($F(1, 48) = 0.01$, $p = .909$, $\eta^2 = 0.00$), and experimental conditions × gender interaction ($F(1, 48) = 0.55$, $p = .462$, $\eta^2 = 0.01$). A priori power analysis (GPower 3.1; Faul et al., 2009) indicated 34 participants were required for detecting medium interaction effects ($f = 0.25$; $\alpha = .05$; power = 0.80; Cohen, 2013) in repeated-measures ANOVA with a within–between design. The sample was increased to 54 to account for potential data loss, aligning with previous eye-tracking studies on live-streaming (Baek et al., 2025; Chen et al., 2022, 2023; Duhaime et al., 2020; Mancini et al., 2022). Two participants were later excluded due to data quality issues, resulting in a final sample of 52. The study received ethics committee approval (n. 0033687, 24/04/2025) and was conducted per the Declaration of Helsinki and GDPR. Participants provided informed consent.

### 2.2. Materials

#### 2.2.1. Content

Two games representing distinct content types were selected: Crash Bandicoot: N. Sane Trilogy (Platformer; Swacha, 2025 and EA FC 2025 (Esports; Jenny et al., 2024). Selection criteria: (1) authentic streamer engagement plausibility, and (2) limited emotional variability. The Platformer's third-person view and Esports' isometric football simulation ensured ecological validity. Foreign league teams avoided national bias. Game characteristics defined video durations for authenticity. Preliminary trials showed Platformer completion times of 2:50–3:35, and 3:14 clip was selected reflecting realistic completion and pacing. Esports' minimum half-match duration (4:00) determined stimulus length, ensuring comparable ecological validity across content types.

#### 2.2.2. Streamer interpretation

To ensure ecological validity while maintaining experimental control, an ex-streamer actor performed gameplay segments, delivering consistent nonverbal behavior throughout. The actor signed a written informed consent and image release form. Video versions were analyzed with FaceReader (iMotions) to classify emotional valence (positive, neutral, negative). In a pre-test, ninety-eight participants rated streamer tone (9-point scale), confirming significant differences among conditions for both games (Platformer: $F(2,95) = 36.60$, $p < .001$; Esports: $F(2,95) = 18.00$, $p < .001$). The neutral version was selected to minimize emotional and nonverbal confounds and isolate brand placement effects

**Table 2**
Demographic characteristics of the final sample (N = 52), including gender, age groups, gaming habits, and energy drink consumption.

| Demographic characteristics | | | |
|---|---|---|---|
| Variables | Category | N | Percentage |
| Gender | Male | 27 | 51.9% |
| | Female | 25 | 48.1% |
| Age | 18 - 24 | 14 | 26.9% |
| | 25 - 34 | 26 | 50.0% |
| | 35 - 44 | 8 | 15.4% |
| | 45 - 49 | 4 | 7.7% |
| Play Videogames | Yes | 36 | 65.4% |
| | Not | 18 | 34.6% |
| Watch Gaming content | Yes | 32 | 59.6% |
| | Not | 22 | 40.4% |
| Energy drink | Yes | 33 | 61.5% |
| | Not | 21 | 38.5% |

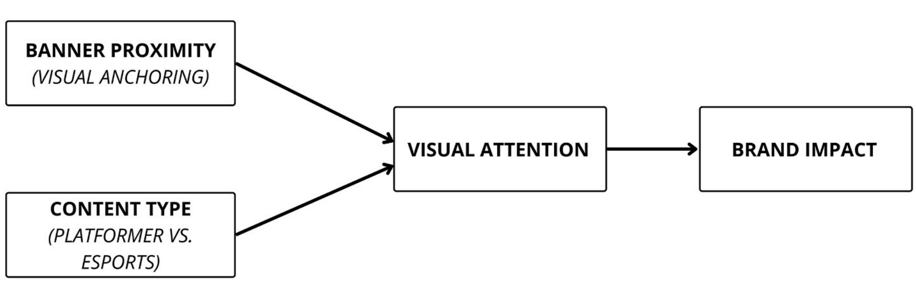


**Fig. 1.** Conceptual model of the study. Banner proximity (visual anchoring) and content type (Platformer vs. Esports) influence visual attention, which in turn affects Brand Impact.

(Zeelenberg et al., 2006). Details are in Appendix A.

#### 2.2.3. Brand

Energy drinks were chosen for brand sponsorship due to live-streaming prevalence (Edwards et al., 2022). Five fictitious brands were created and pre-tested with 45 participants (Perneger et al., 2015) for authenticity and category congruence. Energy Boost and Apex Fuel were consistently recognized as energy drinks ($p < .001$) and selected. Both logos were analyzed for luminance, contrast, and colorfulness to confirm perceptual equivalence. Fictitious yet credible brands allowed controlled sponsorship manipulation, ensuring internal validity without compromising ecological realism aligned with Twitch practices (Holden & Vanhuele, 1999; Hsu & Chiu, 2025). Details are in Appendix B.

#### 2.2.4. Experimental stimuli

The rigorous development and validation process, including pre-testing of individual elements used in developing and editing the final stimuli, enabled us to maintain the study's internal validity. Final video stimuli were edited within a custom platform that replicated Twitch. tv to increase ecological validity and simulate the platform's content. Viewer count was set to 38 and chat was disabled via cinema mode (Shi et al., 2024). Facecam was stably positioned in the upper-right for all the stimuli. Brands' ads were incorporated per experimental conditions, in proximity (upper right) or at a distance (lower left) from the facecam. This design aligns with independent streamer practices and brand integration platforms, such as inStreamly (2025a). Banner duration was 25 s, following Min et al. (2025), who used 15-s exposures in multi-brand esports contexts. Increased duration accounts for single-brand use. Banners appeared five times in Platformer (3:14 min) and six times in Esports (4:00 min), covering ~56% of total duration with 25-s exposures (10 s for final insertion). Banners appeared every 20 s, starting 5 s after onset, with constant size (X = 29%, Y = 29%) to avoid perceptual bias. In Esports, GOAL (50 s) and NO-GOAL (75 s) segments enabled testing banner effects across scenes. Time and banner exposure are in Table 3. All these procedures were refined to strike a balance between experimental control and ecological validity, allowing for the systematic manipulation of highly dynamic and variable social media content.

### 2.3. Experimental design and conditions

A 2×2 mixed factorial design was employed: Group (Proximity vs Distant) was manipulated between subjects, whereas Content (Platformer vs. Esports) was manipulated within subjects (Fig. 2). Participants were randomly assigned to one of two experimental Group conditions:

- PROXIMITY (P): the banner ad was placed directly below the facecam;
- DISTANT (D): the banner was located in the lower-left part of the screen.

Participant assignment to between-subject group conditions and the sequence of brand presentation (Energy Boost followed by Apex Fuel, or vice versa) were randomized and counterbalanced at the participant level using Microsoft Excel (RANDOM function), resulting in balanced exposure across the two possible brand sequences. Within-subject content order (Platformer and Esports videos) was randomized automatically by iMotions software for each participant, thereby preventing systematic order effects. These procedures ensured balanced exposure and appropriate counterbalancing across experimental conditions.

**Table 3**
Summary of stimulus duration and banner exposure across content types, showing total clip length, banner visibility time, and relative exposure percentage for each video.

| Stimulus | Clip Duration (s) | Banner Exposure (s) | Total Banner (%) | Goal Scene (s/%) | No-Goal Scene (s/%) |
|---|---|---|---|---|---|
| Platformer | 194 | 110 | 56.70% | — | — |
| Esports | 240 | 135 | 56.25% | 50/20.8% | 75/31.3% |

### 2.4. Eye-tracker data

iMotions software v. 10.0 was used to deliver the stimuli, surveys and synchronize the eye-tracker data. Raw gaze data were recorded at 150 Hz using a Tobii Pro Spectrum bar (Tobii LLC), processed using iMotions' R notebooks default parameters (iMotions, 2022) and analyzed across Areas of Interest (AOI).

#### 2.4.1. Signal preprocessing

The I-VT (Identification by Velocity Threshold) algorithm (threshold = 30°/s; window = 20 ms) was used to identify fixations and saccades as in previous studies (Lal et al., 2024; Minen et al., 2023; Sqalli et al., 2023). Fixations shorter than 60 ms were excluded, and adjacent fixations within 0.5° and 75 ms were merged to prevent fragmentation. Gaps shorter than <75 ms were linearly interpolated to determine blink/track loss and no additional smoothing was applied. A 9-point calibration ensured an accuracy of <0.5°. After extracting the data, a data quality check was conducted, and two participants with <85% (Wang and Cho, 2022) were excluded from the sample, resulting in a 3.7% data loss. For the remaining valid participants (52), the gaze data had an average of 97.3% valid AOI samples.

#### 2.4.2. Areas of Interest (AOI)

Dynamic AOIs were defined for the banner (active only during visibility) and the facecam (active throughout the stimuli), allowing quantitative extraction of visibility and saliency metrics (Russo et al., 2021). Appendix C presents the timelines for each stimulus of the analyzed AOIs. The dimensions and temporal transitions of the dynamic AOIs corresponded to the on-screen appearance of the elements on the banner screen (as reported in the section on experimental stimuli, reaching a maximum of approximately 320 × 70 pixels) and the facecam (present throughout the duration of the content, covering a fixed area of 360 × 205 pixels). No spatial overlap occurred between dynamic AOIs, as the facecam and banner regions were positioned in distinct, non-contiguous areas of the screen. Consequently, no hierarchical ordering rule was required for AOI overlap resolution (Fig. 3). Each stimulus was additionally divided into nine gridded AOIs (3×3: Left–Center–Right × 1–3 from bottom to top) to map spatial gaze distribution across all the content (Fig. 4). From an experimental perspective, the use of dynamic AOIs constitutes a methodological strength of the study. This approach not only preserves ecological validity, by reproducing the transient and variable nature of overlays in real live-streaming contexts, but also enhances data quality and interpretability.

#### 2.4.3. Metrics

In order to account for unequal exposure across stimuli (Platformer = 110 s; Esports = 135 s; Goal = 50 s; No-Goal = 75 s), all gaze metrics were analyzed using generalized linear mixed models (GLMMs) and linear mixed models (LMMs), including a log-transformed offset proportional to banner visibility time. This normalization enabled the comparability of attention rates across content and scene durations. The following metrics were analyzed:

- Fixations: the total number of fixations within an AOI after the first gaze entry, indicating the initial attentional capture and frequency of visual sampling;


**Fig. 2.** Experimental Conditions. A) Proximity B) Distant; 1) Platformer content; 2) Esports content. The actor's face and Twitch profiles are blurred for privacy.


**Fig. 3.** Dynamic Areas of Interest (AOIs) for facecam and banner in: A) Proximity; B) Distant; 1) Platformer content; 2) Esports content.

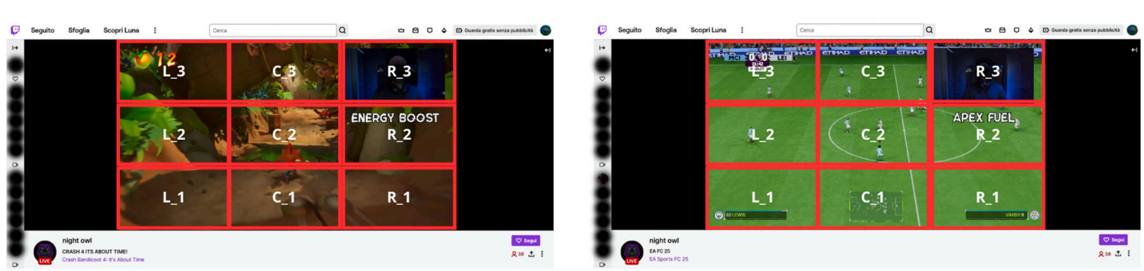


**Fig. 4.** 3 × 3 gridded AOI layout applied to Platformer and Esports content to analyze spatial distribution of visual attention.

- Revisits: the number of gazes returns to the same AOI following the initial visual contact, reflecting sustained interest and re-engagement after attentional shifts;
- Dwell Time (ms): the absolute amount of time, in milliseconds, that participants spent within the AOI, measuring the cumulative depth of visual processing and encoding;

- Transitions: defined as direct saccadic shifts from a fixation within the facecam AOI to the subsequent fixation within the banner AOI, automatically extracted via iMotions, suggesting the efficacy of attentional cue in the stimuli.

The adoption of these eye-tracking metrics enables the objective quantification of visual attention in live-streaming environments with multiple concurrent on-screen elements. To address hypotheses 1-4, these gaze measures examine how attention is distributed across the content and assess whether the face, as a cue, facilitates viewing of target elements.

### 2.5. Self-report

Participants completed four validated self-report scales directly within the iMotions software platform (Appendix D). The Brand Recognition (BR) scale, adapted from Min et al. (2025), was employed to assess whether the position of the banner advertisement influenced participants' ability to recall the fictitious brand among a set of distractors correctly. A response was coded as correct when the participant selected the brand displayed in the previously viewed stimulus. The Advertising Intrusiveness (AI) scale (Li et al., 2013), through 7-point likert scale (1 = strongly disagree to 7 = strongly agree) across seven items, was used to evaluate the perceived intrusiveness of the banner advertisement during content viewing. The Advertised Brand Purchase Intention (ABPI) scale, as developed by Anubha and Jain (2024), through 7-point likert scale across three items, was administered to assess the participants' intention to purchase the featured brand within a gaming context. The Brand Interest (BI) scale (Machleit et al., 1990), through 7-point likert scale across four items, was used to measure the participant's curiosity and openness to learning more about a previously unknown brand. This scale is particularly suited to fictitious brands, as participants cannot express pre-existing attitudes toward them. Consequently, using a conventional brand attitude measure would not be a methodologically appropriate approach. Each scale was administered twice, in relation to the gaming content. All measurement instruments were translated into the target language following the back-translation procedure (Brislin, 1970), ensuring both linguistic and conceptual equivalence across languages.

The use of these dimensions enables the quantification of the sponsored brand impact on screen. Within the experimental design, these measurements allow us to answer hypothesis 5 by examining how much proximity to the face can increase brand perception.

### 2.6. Protocol

The participant was welcomed and guided to the laboratory, where they received research information and signed an informed consent form. Afterwards, they were seated in front of a 23.8″ PC monitor (FlexScan EV2451 by Eizo KK, Hakusan, Japan) at an appropriate distance from the Eye Tracker device. Next, the calibration phase for the eye tracker tool was established. Participants were then instructed: 'A friend invited you to watch some live-streaming of a streamer. Relax and enjoy the show.' The Platformer and the Esports videos followed in randomized order. Once the experimental phase was completed, the participants filled out a questionnaire for self-report measurements. The entire procedure took about 20 min per participant. Fig. 5 shows a diagram of the protocol divided into experimental conditions (see Fig. 6).

### 2.7. Statistical analysis

Given the study's repeated-measures design, GLMM and LMM were used to account for within-subject dependencies. Compared to a repeated-measures ANOVA, mixed models were chosen to incorporate random intercepts to model baseline individual differences in visual attention, thereby avoiding inflated Type I errors that occur when observations are not independent. Furthermore, the distributional flexibility of mixed models, through appropriate distribution-link function pairs, allows for handling the different nature of dependent variables (count and continuous measures). Statistical analysis was performed in RStudio (2025.05.0, Build 496) using R packages lme4 (Bates et al., 2015) for GLMMs and LMMs, and glmmTMB for zero-inflated models. Distribution–link function pairs were selected based on outcome nature: Poisson GLMMs with log link were fitted using glmer() for count variables (fixations, revisits, transitions); zero-inflated Poisson (ZIP) models were fitted using glmmTMB() for banner transitions, given the high proportion of zeros; Gaussian LMMs were fitted using lmer() for continuous questionnaire variables; and binomial GLMMs with logit link were fitted using glmer() for brand recognition. For dwell time, given its right-skewed distribution values were log-transformed, with Gaussian LMMs fitted using lmer().

All models included content type (Platformer, Esports) and ad position (Proximity, Distant) as fixed factors, with their interaction term. Random-effects structure was initially tested using maximal models (1 + Content | Participant), but due to non-convergence and singularity issues across multiple analyses, simplified random intercept structures (1 | Participant) were adopted, following the principle of parsimony (Bates et al., 2015). Model assumptions and residual diagnostics were systematically assessed using DHARMa. All models converged without singularity warnings. Eye-tracking metrics were analyzed, incorporating an offset term for stimulus exposure time, with exposure duration converted from seconds to minutes and entered as offset (LOG(EXPOSURE_TIME)). This approach ensures results are interpreted as rates per minute of exposure, enabling direct comparability across videos of different lengths. The internal reliability of the averaged questionnaire items was assessed using Cronbach's alpha (Nunnally, 1978) and

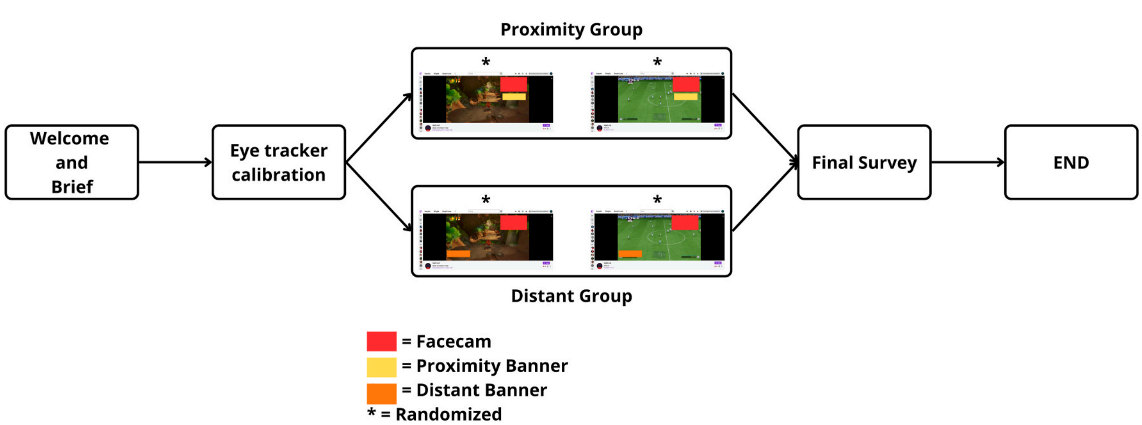


**Fig. 5.** Protocol sequence of the study.

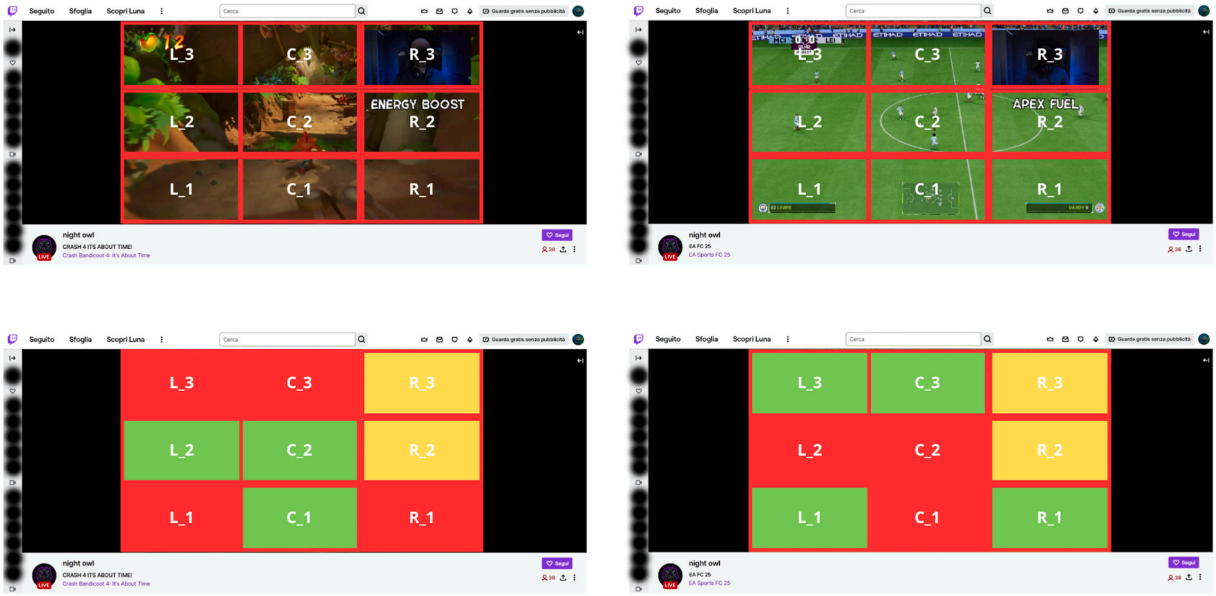


**Fig. 6.** Gridded AOI results. A) Platformer; B) Esports. Red = non-significant, Green = significant, Yellow = non-significant for both contents. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

McDonald's Omega (McDonald, 1999) for each construct (Appendix D). Shapiro–Wilk tests were used to assess normality assumptions for both eye-tracker and self-report measures. The results of the normality tests and descriptive statistics tables are included in the supplementary materials. Statistical significance threshold was set at $\alpha = .05$, and Holm's sequential correction (Holm, 1979) was applied to control family-wise error rate within each dependent variable family.

## 3. Results

### 3.1. Gridded AOI

Visual attention patterns across the entire content were analyzed using nine LMMs, one for each grid area. The DHARMa diagnostic tests indicated issues of uniformity for C_2 and the presence of outliers for both C_2 and R_2. A refitted model using Gamma distribution with a log link did not improve the diagnostics. These anomalies likely reflect the high visual salience of these specific areas, which attracted a disproportionate amount of gaze time, leading to deviations from the expected distribution. The model used to predict the time spent on the gaming content was specified as follows:

$$\textbf{\textit{GRIDDED AREA}} \sim \textbf{\textit{GROUP}} * \textbf{\textit{CONTENT}} + \textbf{\textit{OFFSET}}(\textbf{\textit{LOG}}(\textbf{\textit{EXPOSURE\_TIME}})) + (\mathbf{1} \mid \textbf{\textit{PARTICIPANT}})$$

In the center row, PLATFORMER elicited longer dwell rates on C_1 ($\exp(\beta) = 0.55$, $p_{holm} < 0.001$, 95% CI [0.46, 0.64]) and C_2 ($\exp(\beta) = 0.84$, $p_{holm} < 0.001$, 95% CI [0.78, 0.90]), while ESPORTS showed higher dwell rate on C_3 ($\exp(\beta) = 2.36$, $p_{holm} < 0.001$, 95% CI [1.94, 2.87]). Right areas of the content showed ESPORTS increased attention on R_1 ($\exp(\beta) = 2.11$, $p_{holm} = 0.009$, 95% CI [1.31, 3.39]), but not on R_2 ($\exp(\beta) = 1.19$, $p_{holm} = 0.057$, 95% CI [1.00, 1.41]) and R_3 ($\exp(\beta) = 1.41$, $p_{holm} = 0.057$, 95% CI [1.04, 1.92]) after Holm correction. Left areas of the content showed ESPORTS increased attention on L_1 ($\exp(\beta) = 2.26$, $p_{holm} = 0.003$, 95% CI [1.44, 3.55]), L_3 ($\exp(\beta) = 3.17$, $p_{holm} = < 0.001$, 95% CI [2.28, 4.40])) and PLATFORMER longer dwell rates in L_2 ($\exp(\beta) = 0.78$, $p_{holm} = 0.009$, 95% CI [0.66, 0.91]). No Group or Group × Content interaction emerged. Table 4 reports the effects of content on eye-tracking measures within the gridded AOI.

### 3.2. Facecam

Visual attention across the entire content was analyzed using GLMMs, with Poisson distributions applied to count measures and Gaussian LMMs to continuous measures. No violations were detected in any of the DHARMa diagnostic tests. The models followed the structure:

$$\textbf{\textit{FIXATION COUNT}} \ / \ \textbf{\textit{REVISITS}} \ / \ \textbf{\textit{DWELL TIME}} \sim \textbf{\textit{GROUP}} * \textbf{\textit{CONTENT}} + \textbf{\textit{OFFSET}}(\textbf{\textit{LOG}}(\textbf{\textit{OFFSET\_MIN}})) + (\mathbf{1} \mid \textbf{\textit{RESPONDENT}})$$

On the Fixation Count, the fixed effect of Content was statistically significant (IRR = 1.31, SE = 0.06, $p_{holm} < 0.001$, 95% CI [1.16, 1.48]). The Group effect (IRR = 1.12, SE = 0.17, $p_{holm} = 1.000$, 95% CI [0.81, 1.56]) and the Group × Content interaction (IRR = 0.96, SE = 0.08, $p_{holm} = 1.000$, 95% CI [0.81, 1.13]) were not significant. The marginal $R^2 = 0.04$ confirmed that the Content fixed effect, while the conditional $R^2 = 0.70$. For the number of Revisits, the fixed effect of Content was statistically significant (IRR = 1.24, SE = 0.08, $p_{holm} = 0.041$, 95% CI [1.07, 1.45]). Neither the Group effect (IRR = 1.13, SE = 0.18, $p_{holm} = 1$

.000, 95% CI [0.79, 1.61]) nor the interaction (IRR = 0.96, SE = 0.10, $p_{holm}$ = *1* .000, 95% CI [0.78, 1.17]) reached significance. The marginal $R^2$ = 0.02, while the conditional $R^2$ = 0.62. For Dwell Time (ms), the fixed effect of Content was not statistically significant after Holm correction (exp(β) = 1.55, SE = 0.16, $p_{holm}$ = 0.061, 95% CI [1.12, 2.14]), neither for Group effect (exp(β) = 1.18, SE = 0.22, $p_{holm}$ = *1* .000, 95% CI [0.76, 1.83]) nor the interaction term (exp(β) = 1.02, SE = 0.23, $p_{holm}$ = *1* .000, 95% CI [0.65, 1.60]) showed statistical significance. The marginal $R^2$ = 0.08, while the conditional $R^2$ = 0.53. Table 5 reports the content effects on the eye-tracker measures on the facecam.

### 3.3. Transitions

Multiple count distributions were compared, including Poisson, Zero-Inflated Poisson (ZIP), Zero-Inflated Negative Binomial (ZINB), and Hurdle models. The standard Poisson model provided the best fit (AIC = 201.90; ΔAIC <2 vs ZIP), confirming that the simpler specification adequately captured the data structure. The model used to predict the number of transitions toward the banner was specified as follows:

$$\textbf{\textit{TRANSITION}} \sim \textbf{\textit{GROUP}} * \textbf{\textit{CONTENT}} + \textbf{\textit{OFFSET}}(\textbf{\textit{LOG}}(\textbf{\textit{OFFSET\_MIN}})) + (\mathbf{1} \mid \textbf{\textit{RESPONDENT}})$$

The Group effect showed that banners placed in Proximity to the facecam generated substantially more transitions than those placed Distant (*IRR* = 29.22, *SE* = 1.03, *z* = 3.27, *p* = .001, 95% *CI* [3.87, 220.75]). Neither the Content effect (*IRR* = 0.82, *SE* = 1.41, *z* = −0.145, *p* = .885, 95% *CI* [0.05, 13.03]) nor the Group × Content interaction (*IRR* = 1.56, *SE* = 1.43, *z* = 0.312, *p* = .755, 95% *CI* [0.09, 25.88]) reached statistical significance. The marginal $R^2$ = 0.68, while the conditional $R^2$ = 0.75. Fig. 7 shows the distribution of gaze transitions from the facecam to the banner across groups, while Table 6 reports the main fixed effects of the transitions. Fig. 8 illustrates an example sequence of fixations depicting the transition between the facecam and the banner.

### 3.4. Banner

Visual attention across the entire content was analyzed using GLMMs, with Poisson distributions applied to count measures and Gaussian LMMs to continuous measures. DHARMa diagnostic tests revealed distributional issues in the Poisson model for the Revisits variable. Therefore, a Zero-Inflated Poisson (ZIP) model was adopted, which successfully converged and showed no violations in the DHARMa diagnostics. The models followed the structure:

$$\textbf{\textit{FIXATION COUNT}} \ / \ \textbf{\textit{REVISITS}} \ / \ \textbf{\textit{DWELL TIME}} \sim \textbf{\textit{GROUP}} * \textbf{\textit{CONTENT}} + \textbf{\textit{OFFSET}}(\textbf{\textit{LOG}}(\textbf{\textit{OFFSET\_MIN}})) + (\mathbf{1} \mid \textbf{\textit{RESPONDENT}})$$

In terms of Fixation Count, the fixed effects of Group (IRR = 1.64, SE = 0.16, $p_{holm}$ = 0.010, 95% CI [1.19, 2.25]) and Content (IRR = 1.65, SE = 0.10, $p_{holm}$ < 0.001, 95% CI [1.35, 2.00]) were statistically significant. The Group × Content interaction did not reach significance (IRR = 0.81, SE = 0.13, $p_{holm}$ = 0.135, 95% CI [0.63, 1.04]). The marginal $R^2$ = 0.17 while the conditional $R^2$ = 0.62. Turning to Revisits, the Group effect (IRR = 2.22, SE = 0.19, $p_{holm}$ < 0.001, 95% CI [1.53, 3.23]) and Content fixed effect (IRR = 1.84, SE = 0.17, $p_{holm}$ = 0.001, 95% CI [1.33, 2.55]) were significant. The Group × Content interaction did not reach significance (IRR = 0.69, SE = 0.20, $p_{holm}$ = 0.135, 95% CI [0.47, 1.01]). The marginal $R^2$ = 0.10 while the conditional $R^2$ = 0.16. For Dwell Time (ms), the fixed effect of Group (exp(β) = 3.20, SE = 0.22, $p_{holm}$ < 0.001, 95% CI [2.10, 4.88]) and Content (exp(β) = 2.64, SE = 0.17, $p_{holm}$ < 0.001, 95% CI [1.88, 3.71]) were significant. The Group × Content interaction did not reach significance (exp(β) = 0.61, SE = 0.24, $p_{holm}$ = 0.135, 95% CI [0.38, 0.98]). The marginal $R^2$ = 0.38 while the conditional $R^2$ = 0.60. Table 7 presents all the main effects of the Banner during the live-streaming.

**Table 4**
Effect Size exp(β) for DWELL_TIME (Log-Normal LMM) for ESPORTS vs PLATFORMER.

| Content effects on dwell time across nine gridded areas | | | | | |
|---|---|---|---|---|---|
| AOI | Content Effect | exp(β) | 95% CI | | p(Holm) |
| | | | Lower | Upper | |
| **C_1** | PLATFORMER > ESPORTS | 0,55 | 0,46 | 0,64 | <0.001*** |
| **C_2** | PLATFORMER > ESPORTS | 0,84 | 0,78 | 0,90 | <0.001*** |
| **C_3** | ESPORTS > PLATFORMER | 2,36 | 1,94 | 2,87 | <0.001*** |
| **R_1** | ESPORTS > PLATFORMER | 2,11 | 1,31 | 3,39 | 0.009** |
| **R_2** | ESPORTS > PLATFORMER | 1,19 | 1,00 | 1,41 | 0.057 |
| **R_3** | ESPORTS > PLATFORMER | 1,41 | 1,04 | 1,92 | 0.057 |
| **L_1** | ESPORTS > PLATFORMER | 2,26 | 1,44 | 3,55 | 0.003** |
| **L_2** | PLATFORMER > ESPORTS | 0,78 | 0,66 | 0,91 | 0.009** |
| **L_3** | ESPORTS > PLATFORMER | 3,17 | 2,28 | 4,40 | <0.001*** |

**Notes:**Effect Size exp(β) for ESPORTS vs PLATFORMER.
p_holm values are Holm-corrected for multiplicity. ***p < .001, **p < .01, *p < .05.

### 3.5. Salient scene

Visual attention across the entire content was analyzed using GLMMs, with Poisson distributions applied to count measures and Gaussian LMMs to continuous measures. No violations were detected in any of the DHARMa diagnostic tests. The models followed the structure:

$$\textbf{\textit{FIXATION COUNT}} \ / \ \textbf{\textit{REVISITS}} \ / \ \textbf{\textit{DWELL TIME}} \sim \textbf{\textit{GROUP}} * \textbf{\textit{SALIENT SCENE}} + \textbf{\textit{OFFSET}}(\textbf{\textit{LOG}}(\textbf{\textit{OFFSET\_MIN}})) + (\mathbf{1} \mid \textbf{\textit{RESPONDENT}})$$

For Fixations, neither the effect of Group (IRR = 1.27, $p_{holm}$ = 0.705, 95% CI [0.92, 1.75]), Salient scene (IRR = 1.20, $p_{holm}$ = 0.705, 95% CI [0.94, 1.54]), nor their interaction (Group × Salient scene: IRR = 1.08, $p_{holm}$ = *1* .000, 95% CI [0.79, 1.49]) reached statistical significance. The marginal $R^2$ = 0.08, while the conditional $R^2$ = 0.56. For Revisits, the effect of Salient scene (IRR = 1.72, $p_{holm}$ = 0.272, 95% CI [1.03, 2.89])

**Table 5**
Fixed effects of Content (Esports vs. Platformer) on visual attention to the facecam. Effect sizes are expressed as incidence rate ratios (IRR) for Revisits and Fixations, and as exponentiated coefficients (exp β) for Dwell Time (ms).

| Facecam Content Effects | | | | | |
|---|---|---|---|---|---|
| Variable | Parameter | Effect Size | 95% CI | | p(Holm) |
| | | | Lower | Upper | |
| **FIXATIONS** | ESPORTS > PLATFORMER | 1,31 | 1,16 | 1,48 | <0.001** |
| **REVISITS** | ESPORTS > PLATFORMER | 1,24 | 1,07 | 1,45 | 0.041* |
| **DWELL TIME** | ESPORTS = PLATFORMER | 1,55 | 1,12 | 2,14 | 0.061 |

**Notes:**Effect Size: IRR for FIXATIONS and REVISITS (Poisson models); exp(β) for DWELL_TIME (Log-Normal LMM).
p_holm values are Holm-corrected for multiplicity. ***p < .001, **p < .01, *p < .05.

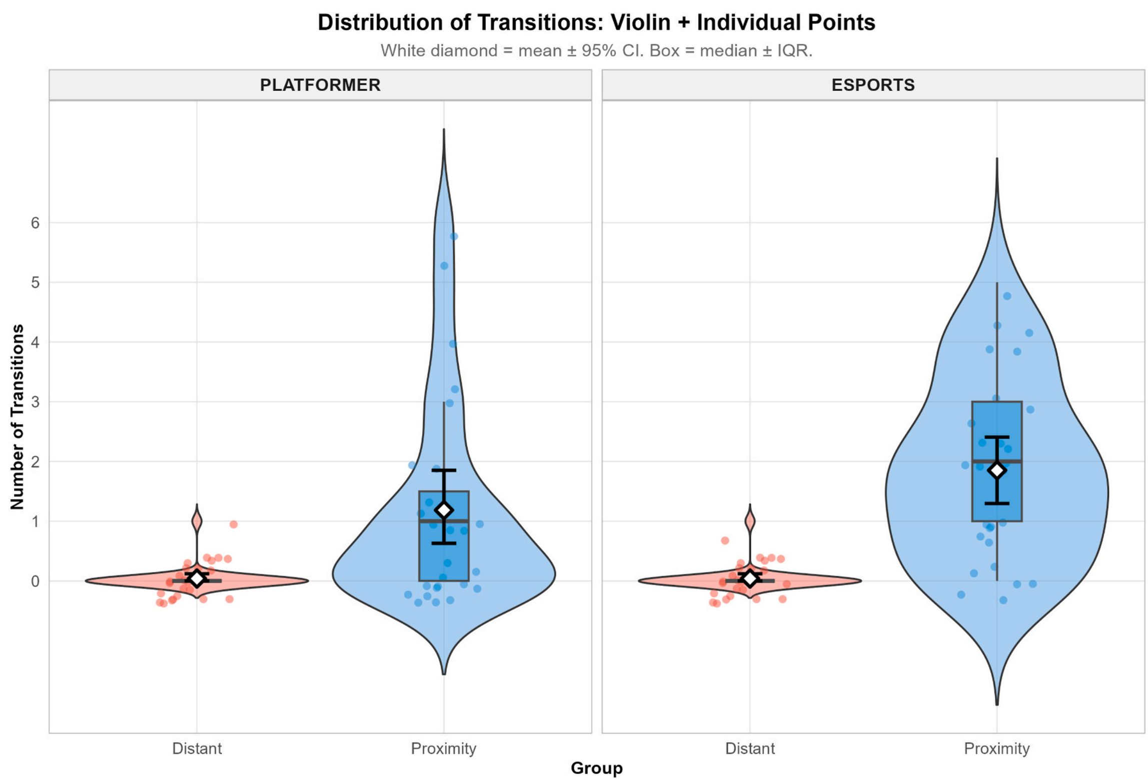


**Fig. 7.** Distribution of gaze transitions from facecam to banner by Group (Distant vs Proximity) and content type (Platformer vs Esports).

was statistically significant before correction but did not survive the Holm adjustment, indicating only a trend toward higher revisits during goal-oriented scenes. Neither the Group effect (IRR = 1.56, $p_{holm}$ = 0.645, 95% CI [0.91, 2.67]) nor the interaction (IRR = 1.39, $p_{holm}$ = 0.920, 95% CI [0.74, 2.63]) reached significance. The marginal $R^2$ = 0.23, while the conditional $R^2$ = 0.42. For Dwell Time, both the Group (exp(β) = 1.90, $p_{holm}$ = 0.071, 95% CI [1.20, 3.02]) and the Salient scene (exp(β) = 1.84, $p_{holm}$ = 0.090, 95% CI [1.16, 2.93]) effects were significant at the uncorrected level but did not remain so after Holm correction. The Group × Salient scene interaction was not significant (exp(β) = 0.81, $p_{holm}$ = *1* .000, 95% CI [0.43, 1.52]). The marginal $R^2$ = 0.17, while the conditional $R^2$ = 0.23.

### 3.6. *Self-report*

To analyze the brand impact after viewing the stimuli, questionnaires were analyzed using Gaussian LMMs for continuous measures (AI, ABPI, and BI) and Binomial GLMM with logit link for binomial measures (BR). Models for ABPI and BI exhibited singular fits (ICC = 0), indicating negligible between-subject variability. All models converged successfully, and DHARMa residual diagnostics indicated deviations from uniformity and dispersion in some of the self-report models (AI, ABPI, and BI; ps ≤ 0.005), while Brand Recognition showed a satisfactory fit (all ps > 0.20). Despite these warnings, no evidence of zero-inflation or outliers was detected. To assess the robustness of the findings, the analyses were

**Table 6**
Effect Size IRR for Transitions (Poisson) from facecam to banner.

| Transitions effects | | | | | | |
|---|---|---|---|---|---|---|
| Parameter | IRR | 95% CI | | SE | z-value | p-value |
| | | Lower | Upper | | | |
| **GROUP EFFECT (Proximity > Distant)** | 29,22 | 3,87 | 220,75 | 1,03 | 3,27 | 0.001** |
| **CONTENT EFFECT (ESPORTS > PLATFORMER)** | 0,82 | 0,05 | 13,03 | 1,41 | −0,15 | 0.885 |
| **GROUP × CONTENT** | 1,56 | 0,09 | 25,88 | 1,43 | 0,31 | 0.755 |

**Fig. 8.** Prototypical fixation-by-fixation sequence of the transition from the facecam to the banner in the Platformer stimulus.

replicated using GLMM with log link distribution and non-parametric tests. The model used to predict the attention on the banner during the salient scene followed the structure:

$$AI \ / \ ABPI \ / \ BI \ / \ BR \sim GROUP * CONTENT + (1 \mid RESPONDENT)$$

For the Advertising Intrusiveness, the fixed effects of Group ($\beta$ = 0.16, *SE* = 0.28, $p_{holm}$ = *1* .000, *95% CI* [−0.40, 0.71]*)*, Content (β = 0.09, SE = 0.14, $p_{holm}$ = *.1* .000, 95% CI [−0.19, 0.37]) and Group x Content ($\beta$ = *-0.09, SE = 0.20,* $p_{holm}$ = *1* .000, *95% CI [−0.48, 0.29]*) were not statistically significant. The marginal $R^2$ = 0.004, while the conditional $R^2$ = 0.761. For the Advertised Brand Purchase Intention, the fixed effects of Group ($\beta$ = - 0.05 *SE* = 0.24, $p_{holm}$ = *1* .000, *95% CI* [−0.53, 0.42]), Content ($\beta$ = - 0.16, *SE* = 0.25, $p_{holm}$ = *1* .000, *95% CI* [-0.64, 0.32]), and Group × Content interaction ($\beta$ = 0.16, *SE* = 0.34, $p_{holm}$ = *1* .000, *95% CI* [-0.51, 0.83]) were not statistically significant. The marginal $R^2$ = 0.004, while the conditional $R^2$ = 0.004. For Brand Interest, the fixed effects of Group ($\beta$ = 0.14, *SE* = 0.29, $p_{holm}$ = *1* .000, *95% CI* [-0.42, 0.71]), Content ($\beta$ = - 0.39, *SE* = 0.29, $p_{holm}$ = *1* .000, *95% CI* [-0.97, 0.19]), and Group × Content interaction ($\beta$ = 0.13, *SE* = 0.41, $p_{holm}$ = *1* .000, *95% CI* [-0.67, 0.93]) were not statistically significant. The marginal $R^2$ = 0.034, while the conditional $R^2$ = 0.034. The same pattern of results emerged from GLMM and non-parametric tests, indicating that deviations from normality or distributional assumptions did not drive the absence of significance. For Brand Recognition, the fixed effects of Group (*OR* = 54.60, *SE* = 3.25, $p_{holm}$ = *1* .000, *95% CI [*0.09, 32187.87*]*), Content (*OR* = 27.99, *SE* = 1.98, $p_{holm}$ = *1* .000, *95% CI [*0.57, 1365.24*]*), and Group × Content interaction (*OR* = 0.08, *SE* = 2.01, $p_{holm}$ = *1* .000, *95% CI [*0.00, 4.36*]*) were not statistically significant. The marginal $R^2$ = 0.13, while the conditional $R^2$ = 0.12.

## 4. Discussion

The findings provide two key insights related to visual attention, derived from the eye-tracking data, and one key insight from explicit self-report measures concerning memory and brand interest. First, the facecam consistently attracts visual attention, regardless of the type of

**Table 7**
Fixed effects on visual attention to the Banner. Effect sizes are expressed as incidence rate ratios (IRR) for Revisits and Fixations, and as exponentiated coefficients (exp β) for Dwell Time (ms).

| Banner Fixed effects | | | | | |
|---|---|---|---|---|---|
| Variable | Parameter | Effect Size | 95% CI | | p(Holm) |
| | | | Lower | Upper | |
| **FIXATIONS** | GROUP EFFECT (Proximity > Distant) | 1,64 | 1,19 | 2,25 | 0.010* |
| **FIXATIONS** | CONTENT EFFECT (ESPORTS > PLATFORMER) | 1,65 | 1,35 | 2,00 | <0.001*** |
| **FIXATIONS** | GROUP × CONTENT | 0,81 | 0,63 | 1,04 | 0.135 |
| **REVISITS** | GROUP EFFECT (Proximity > Distant) | 2,22 | 1,53 | 3,23 | <0.001*** |
| **REVISITS** | CONTENT EFFECT (ESPORTS > PLATFORMER) | 1,84 | 1,33 | 2,55 | 0.001** |
| **REVISITS** | GROUP × CONTENT | 0,69 | 0,47 | 1,01 | 0.135 |
| **DWELL** | GROUP EFFECT (Proximity > Distant) | 3,20 | 2,10 | 4,88 | <0.001*** |
| **DWELL** | CONTENT EFFECT (ESPORTS > PLATFORMER) | 2,64 | 1,88 | 3,71 | <0.001*** |
| **DWELL** | GROUP × CONTENT | 0,61 | 0,38 | 0,98 | 0.135 |
| **Notes:** | | | | | |

Effect Size: IRR for FIXATIONS (Poisson) and REVISITS (ZIP); exp(β) for DWELL_TIME (Log-Normal LMM).
p_holm values are Holm-corrected for multiplicity. ***p < .001, **p < .01, *p < .05.

video game being streamed. Second, placing sponsorship close to the facecam significantly increases visual attention toward the brand. Third, such proximity does not automatically translate into higher brand recognition, indicating a clear distinction between visual exposure and brand recall. These results collectively indicate an increase in attentional effectiveness toward elements positioned near the streamer's face, as the human face serves as a strong attentional cue. Yet, this effect does not necessarily translate into advertising effectiveness. Table 8 summarizes the responses to all research hypotheses addressed in the study.

The type of gaming content distinctly shaped viewers' visual exploration. In Platformer videos, the fixed third-person perspective concentrated attention on central screen areas. In contrast, the isometric view of Esports promoted a broader and more information-driven distribution of gaze toward peripheral regions. The facecam area and the region immediately below it showed no significant Content effects, indicating that these areas received comparable visual attention across both content types. The non-significant Content effect for facecam dwell time aligns with findings from the gridded area analysis, confirming that this region maintained stable attention regardless of content. From a theoretical standpoint, these findings align with the visual saliency model (Itti & Koch, 2001), particularly in the context of dynamic visual content where motion serves as the primary driver of attentional allocation (Mital et al., 2011). When applied to the SLSS context with gaming stimuli, this framework may explain how isometric esports content, characterized by continuously shifting action, creates a distributed pattern of visual attention across the screen. In contrast, third-person platformers maintain centralized attention due to gameplay concentrated in central screen regions where the avatar is located. The streamer's face region consistently served as a stable focal point of attention, with greater visual engagement in the Esports condition, as indicated by higher fixation and revisit rates. No significant Group or interaction effects emerged, indicating that banner position alone did not influence gaze allocation toward the facecam. This further reinforces the role of the facecam as a stable attentional anchor within the streaming interface. The methodological use of a neutral, unfamiliar streamer eliminated potential confounds related to familiarity or parasocial interaction (Wulf et al., 2021), isolating the face's role as an

**Table 8**
The table reports all research hypotheses, their corresponding dependent measures, statistical outcomes, and whether each was supported, partially supported, or not supported according to model results.

| Hypothesis Summary | | | | | |
|---|---|---|---|---|---|
| Hypothesis | Description | Dependent measure(s) | Model | Statistical outcome | Supported/ Rejected |
| H1a | In Platformer content, attention converges in central areas | Dwell Time | LMM | p < .01 | Supported |
| H1b | In Esports content, attention is distributed across the periphery | Dwell Time | LMM | p < .01 | Supported |
| H1c | More attention on Facecam in Esports vs Platformer | Fixations, Revisits | LMM and GLMM | p < .05 | Supported |
| H2 | Proximity to facecam increases transitions Facecam to Banner | Transitions | GLMM | p = .001 | Supported |
| H3 | Proximity to facecam increases attention on the banner | Fixations, Revisits, Dwell Time | LMM and GLMM | p < .01 | Supported |
| H4 | Proximity × Salient scenes (Goal) increase banner attention | Fixations, Revisits, Dwell Time | LMM and GLMM | n.s. | Not supported |
| H5a | Proximity improves Brand Recognition | Brand Recognition | GLMM | n.s. | Not supported |
| H5b | Proximity increases Brand Interest | Brand Interest | LMM | n.s. | Not supported |
| H5c | Proximity increases Purchase Intention | Advertised Brand Purchase Intention | LMM | n.s. | Not supported |

attentional magnet. While most prior studies have focused on players' gaze behavior during active gameplay (Sundstedt, 2012; Wang et al., 2024; Yang et al., 2025), the present results clarify how audiences allocate visual attention during streamed gaming content. Also, prior research suggests that the facecam satisfies viewers' social and authenticity needs (Ghosh & Tripathi, 2025), confirming this area as a perceptual and a fixed attentional reference that may structure the gaze allocation across the scene. Within SLSS, streamers design personalized overlays that become part of their visual signature, among which the facecam stands out as the most stable and distinctive element. It operates both as a constant feature of the interface and as a perceptual anchor for viewers. In this framework, content type shapes overall attentional dynamics, whereas proximity to the facecam amplifies the visibility of adjacent branded elements. This anchoring effect provides the necessary first step foundation for spatial contiguity, explaining why elements placed near the facecam capture disproportionate visual attention. These findings address H1, demonstrating that attentional patterns vary according to content type while maintaining a stable focus on the facecam area.

The findings from gaze transitions between the facecam and the banner revealed a systematic effect of proximity. Participants in the P condition performed significantly more transitions toward the banner than those in the D condition. The use of fictitious brands eliminates confounds related to familiarity or prior interest (Ou et al., 2025), while luminance and contrast analyses confirmed that perceptual differences between the stimuli did not drive the effect. From an attentional standpoint, the facecam not only serves as a stable visual anchor but also draws attention to adjacent branded elements (Sajjacholapunt & Ball, 2014). This effect aligns with evidence that faces firmly guide visual attention (Kanwisher et al., 1997; Langton et al., 2008; Palermo and Rhodes, 2007) and that spatial contiguity enhances the processing of nearby stimuli (Mayer & Fiorella, 2014; Schroeder & Cenkci, 2018). Overall, these findings fully address H2, demonstrating that banners placed near the streamer's face significantly increase visual transitions, confirming the facilitating role of facial proximity in guiding viewers' attention and reflecting the second step of the proposed mechanism that orients the shift from the face toward adjacent elements.

The P condition not only made a significantly higher number of transitions toward the banner, but also displayed more visual attention on it in terms of Fixations, Revisits and Dwell Time. This effect was consistent across groups and contents, with Esports eliciting overall higher fixation counts. This can be partly explained by the denser informational layout of Esports gameplay, which also displays key in-game information. Yet, participants in the Proximity condition consistently maintained higher attention levels, confirming that facial proximity remains the dominant driver of gaze allocation (Adil et al., 2018; Mañas-Viniegra et al., 2020; Plassmann et al., 2012). This suggests that the attentional pull of the facecam combines with the heightened intensity and competitiveness of Esports (Qian et al., 2019), further extending visual attention toward the overlayed message. Although this alone cannot be considered evidence of full advertising effectiveness, it represents an early perceptual phase that may facilitate subsequent stages of cognitive and evaluative processing. Thus, the findings confirm H3 that placing the banner near the content creator's face leads to a significant increase in visual attention, helping to ensure a perceptual stage. These results confirm the third step of the proposed attentional mechanism. After the initial bottom-up anchoring elicited by the facecam and the subsequent orienting shift toward nearby elements, attention naturally consolidates on the adjacent banner, reflecting enhanced attentional effectiveness through spatial contiguity.

Focusing on the key moment of the Salience scene, the findings do not support the hypotheses. Contrary to Mancini et al. (2022), the results do not show increased fixations on banners during salient scenes, and placing it near the facecam does not significantly enhance visual attention to the advertising element. The lack of consistent fixation and revisit effects could indicate that viewers' attention was primarily directed toward the salient events themselves rather than the banner (Seo et al., 2018). When considering Dwell Time, this measure was higher, although not statistically significant after Holm correction, in the Proximity condition and during Goal scenes. This result is not sufficient to conclude that the proximity of the facecam, combined with a salient event, enhances visual attention. Further research incorporating the emotional valence of the streamer and their real-time reactions could provide deeper insights into how the salience of an event interacts with the use of the facecam as an attentional cue. These findings do not support H4, highlighting the need for further research on this pattern. The contradictions in the literature, with different insights emerging from streaming contexts (Mancini et al., 2022; Seo et al., 2018), highlight the complexity of attentional dynamics during salient events and underscore the need for a more systematic investigation of the interplay between spatial proximity, scene salience, and content type in live-streaming environments.

Unlike eye-tracking findings, self-report measures revealed no significant effects for banner position or content type. Advertising Intrusiveness results indicate banners (near or distant) were not perceived as disturbing, suggesting both positions serve as neutral advertising spaces without disrupting viewing (Li et al., 2013). Brand Interest, Purchase Intention, and Brand Recognition did not differ between conditions, a pattern that is theoretically consistent with the use of fictitious brands. In the absence of prior brand memory or emotional associations, exposure alone is insufficient to predict purchase propensity (Folkvord et al., 2020). Though facial proximity enhances visual attention, it does not translate into deeper memory encoding or brand interest. This divergence reflects the implicit-declarative distinction in consumer neuroscience (Missaglia et al., 2017). Faces trigger implicit neurophysiological responses that automatically capture attention (Harris et al., 2019; Russo et al., 2023), but do not necessarily imply explicit brand recall. These findings do not support H5. While facecam proximity mitigates banner blindness at the attentional level, it does not foster the deeper cognitive processing required for brand recall and interest in unknown brands.

In summary, the findings demonstrate that leveraging the contiguity principle the face functions as an attentional cue within live streaming directing gaze and attention toward adjacent elements. However, the lack of significant results in the questionnaire data shows that attentional effectiveness alone does not guarantee brand impact. In live-streaming contexts, characterized by rapid visual dynamics, high information density, and fleeting exposure, capturing attention through spatial proximity reflects only the first perceptual stage of persuasion. Once the attentional cue that effectively directs viewers' gaze has been identified, this does not automatically imply brand memory or recall. Without emotional engagement or meaningful brand associations, attention alone does not progress into memory consolidation or attitudinal change. Thus, although facial proximity effectively enhances visibility, it does not ensure lasting advertising effectiveness, highlighting the need for integrated strategies that combine attentional cues with affective and cognitive drivers of brand processing.

## 5. Implications

The facecam serves not only as an attentional attractor but also as an attentional catalyst, guiding viewers' gaze toward nearby overlay elements. The IRR effect of transition confirms that spatial proximity enhances visual attention, underscoring the facecam's role as a cognitive anchor. Given streamers' flexibility in overlay design (Hamilton et al., 2014), positioning overlay elements, such as OLSA, near the facecam may offer a clear attentional advantage, which may represent a prerequisite for downstream brand recognition processes (Phua et al., 2023). The stronger effect sizes for Fixations, Revisits, and Dwell Time in Esports compared to Platformer content highlight how the emotional intensity of competitive gameplay and viewers' curiosity toward the streamer's real-time reactions further amplify the facecam's attentional value. From an applied perspective, these findings may support dynamic overlays, such as event-triggered or voice-activated banners (inStreamly, 2025b), which synchronize ad visibility with salient in-game moments. Integrating brand cues with content-driven triggers may contribute to the evolution of static placements toward more interactive formats, aligning sponsorship with emotional peaks and enhancing the perceived value of high-attention areas near the streamer's face. In conclusion, these findings tentatively suggest that streamers' awareness of their audience's visual attention patterns may offer potential advantages in brand partnership negotiations, as they could demonstrate how overlay positioning might influence gaze direction toward sponsored elements. Correspondingly, brands may view such evidence-based placement strategies as a factor when considering investments in streamer collaborations, though the extent to which attentional capture translates into downstream marketing outcomes remains to be established.

## 6. Limitations and future research

Several limitations of this study warrant consideration and suggest new directions for consumer neuroscience research. The findings offer clear evidence regarding the distribution of visual attention during live-streaming with overlay elements; however, these results pertain only to early perceptual processes and do not address memory encoding or persuasion. Subsequent research may build on these findings to further examine the persuasive mechanisms in live-streaming environments. A minor asymmetry in clip duration and banner insertions between Platformer and Esports videos reflects the ecological authenticity of gameplay but was statistically controlled through an exposure-time offset; nonetheless, future studies should replicate these findings with balanced stimuli, larger samples, and real brands to strengthen statistical power and robustness, particularly regarding thin transitions that, while ecologically valid, require confirmation under more controlled conditions. The use of fictitious brands enabled isolation of attentional mechanisms from prior familiarity, though it limits external validity. Future research should incorporate authentic brands and recruit a large sample of participants with varying degrees of brand familiarity to examine how prior knowledge affects attentional allocation, memory, and emotional processing. Similarly, the streamer's neutral and emotionally restrained performance minimized parasocial effects but reduced ecological realism. Subsequent research should explore how different expressive styles, engagement levels, or authentic influencers with established audiences influence attentional and affective responses, as well as brand–content–streamer congruence. Moreover, the use of Twitch's “cinema mode,” which excluded chat interactions, ensured internal validity but reduced ecological realism. Future studies should test overlay advertising within interactive environments that include social and parasocial cues to evaluate both attentional and persuasive outcomes. The study's focus on two gaming genres allowed differentiation of attentional patterns but represents only a limited subset of gaming and live streaming contexts. To investigate attentional patterns further, additional research should consider additional genres with different types of salience events and different stream layout conventions integrating neurophysiological measures such as EEG or GSR to elucidate how attentional capture translates into emotional and mnemonic processing.

## 7. Conclusions

The study demonstrates, through a rigorously controlled experimental design conducted within a realistic live-streaming simulation, that the strategic placement of overlay elements, such as OLSA sponsored content, near the content creator's face significantly influences viewers' gaze behavior. In live-streaming contexts, the streamer's face operates as a stable visual anchor that directs attention toward adjacent elements through the principle of spatial contiguity, thereby enhancing attentional efficiency. This effect represents an early perceptual stage of message processing, facilitating visual exposure without necessarily implying persuasion or attitude change, which depend on higher-level cognitive and affective mechanisms. Theoretically, the findings deepen our understanding of the face's role as an attentional catalyst capable of guiding visual attention toward spatially proximal elements and extend this mechanism to the underexplored domain of entertainment live-streaming platforms. Although not intended to provide prescriptive solutions to banner blindness, the study identifies naturally occurring features within SLSS, particularly the facecam, that can reduce perceptual neglect of overlay advertising. The study therefore serves as a basis for future research into high-level outcomes, starting from the perceptual attentional basis of the face in media content. Practically, the results provide actionable insights for brands and content creators. Recognizing which overlay zones most effectively capture attention enables more informed negotiations of promotional space and data-driven optimization of sponsorship layouts. Companies can strategically decide where to position branded elements and with whom to collaborate, based on the type of content and viewers' attentional patterns. Collectively, these insights contribute to a more evidence-based approach to designing and integrating branded content within live-streamed media.

## CRediT authorship contribution statement

**Sebastiano Accardi:** Writing – review & editing, Writing – original draft, Visualization, Validation, Supervision, Software, Resources, Project administration, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Margherita Zito:** Writing – review & editing, Visualization, Supervision. **Vincenzo Russo:** Writing – review & editing, Visualization, Supervision.

## Declaration of generative AI and AI-assisted technologies in the writing process

During the preparation of this work, the authors used Claude in order to improve the readability and language of the manuscript. After using this service, we reviewed and edited the content as needed and take full responsibility for the content of the publication.


## Funding statement

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.


## Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

## Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.chbr.2026.101107.

## Data availability

Data will be made available on request.

## References


Adil, S., Lacoste-Badie, S., & Droulers, O. (2018). Face presence and gaze direction in print advertisements: How they influence consumer responses—an eye-tracking study. *Journal of Advertising Research, 58*(4), 443–455. https://doi.org/10.2501/JAR-2018-004

Anubha, & Jain, A. (2024). In-game advertising and brand purchase intentions: An SOR perspective. Global knowledge. *Memory and Communication, 73*(1–2), 24–44. https://doi.org/10.1108/GKMC-02-2022-0050

Baek, K. D., Hwang, E., & Lee, J. (2025). The influence of the valence and evaluation type of social feedback on game streamers' emotion, attention, and performance. *International Journal of Human-Computer Interaction*, 1–17. https://doi.org/10.1080/10447318.2025.2465871

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software, 67*(1), 1–48. https://doi.org/10.18637/jss.v067.i01

Bian, X., & Yang, A. (2025). From spectatorship to loyalty: Unraveling the influence of game streaming watch and gaming-related social connectivity on MOBA gamers. *Computers in Human Behavior, 162*, Article 108433. https://doi.org/10.1016/j.chb.2024.108433

Brislin, R. W. (1970). Back-translation for cross-cultural research. *Journal of Cross-Cultural Psychology, 1*(3), 185–216. https://doi.org/10.1177/135910457000100301

Brunswik, E., & Kamiya, J. (1953). Ecological cue-validity of “proximity” and of other gestalt factors. *American Journal of Psychology, 66*(1), 20–32. https://doi.org/10.2307/1417965

Burch, M., & Timmermans, N. (2020). Sankeye: A visualization technique for AOI transitions. In *ACM symposium on eye tracking research and applications* (pp. 1–5). https://doi.org/10.1145/3379156.3391833

Casado-Aranda, L. A., Sánchez-Fernández, J., & Ibáñez-Zapata, J.Á. (2023). Evaluating communication effectiveness through eye tracking: Benefits, state of the art, and unresolved questions. *International Journal of Business Communication, 60*(1), 24–61. https://doi.org/10.1177/2329488419893746

Chen, Y., Tong, X., Yang, S., & Zhou, S. (2023). Effects of intrinsic and extrinsic cues on customer behavior in live streaming: Evidence from an eye-tracking experiment. *Industrial Management & Data Systems, 123*(9), 2397–2422. https://doi.org/10.1108/IMDS-10-2022-0606

Chen, H.-C., Wang, C.-C., Hung, J. C., & Hsueh, C.-Y. (2022). Employing eye tracking to study visual attention to live streaming: A case study of Facebook live. *Sustainability, 14*(12), 7494. https://doi.org/10.3390/su14127494

Cohen, J. (2013). *Statistical power analysis for the behavioral sciences*. Routledge. https://doi.org/10.4324/9780203771587

De Wit, J., Van der Kraan, A., & Theeuwes, J. (2020). Live streams on Twitch help viewers cope with difficult periods in life. *Frontiers in Psychology, 11*, Article 586975. https://doi.org/10.3389/fpsyg.2020.586975

Deska, J. C., Hingston, S. T., DelVecchio, D., Stenstrom, E. P., Walker, R. J., & Hugenberg, K. (2022). The face of the brand: Spokesperson facial width-to-height ratio predicts brand personality judgments. *Psychology and Marketing, 39*(8), 1487–1503. https://doi.org/10.1002/mar.21666

Duhaime, V., Prom Tep, S., & Arcand, M. (2020). The effectiveness of implicit brand placement in online videogame streaming services: An eye tracking study. *Journal of Marketing Development & Competitiveness, 14*(3). https://doi.org/10.33423/jmdc.v14i3.3066

Edwards, C. G., Pollack, C. C., Pritschet, S. J., Haushalter, K., Long, J. W., & Masterson, T. D. (2022). Prevalence and comparisons of alcohol, candy, energy drink, snack, soda, and restaurant brand and product marketing on Twitch, Facebook gaming and YouTube gaming. *Public Health Nutrition, 25*(1), 1–12. https://doi.org/10.1017/s1368980021004420

Faul, F., Erdfelder, E., Buchner, A., & Lang, A. G. (2009). Statistical power analyses using G* power 3.1: Tests for correlation and regression analyses. *Behavior Research Methods, 41*(4), 1149–1160. https://doi.org/10.3758/BRM.41.4.1149

Finzi, D., Gomez, J., Nordt, M., Rezai, A. A., Poltoratski, S., & Grill-Spector, K. (2021). Differential spatial computations in ventral and lateral face-selective regions are scaffolded by structural connections. *Nature Communications, 12*, 2278. https://doi.org/10.1038/s41467-021-22524-2

Folkvord, F., Roes, E., & Bevelander, K. (2020). Promoting healthy foods in the new digital era on Instagram: an experimental study on the effect of a popular real versus fictitious fit influencer on brand attitude and purchase intentions. *BMC Public Health, 20*(1), 1677. https://doi.org/10.1186/s12889-020-09779-y.

Ghosh, A., & Tripathi, S. D. (2025). Understanding viewer fascination with reaction videos: A discourse analysis of YouTube comments. *Media Watch (Taipei), 16*(2), 230–249. https://doi.org/10.1177/09760911241313236

Grill-Spector, K. (2003). The neural basis of object perception. *Current Opinion in Neurobiology, 13*(2), 159–166. https://doi.org/10.1016/S0959-4388(03)00040-0

Hamilton, W. A., Garretson, O., & Kerne, A. (2014). Streaming on Twitch: Fostering participatory communities of play within live mixed media. In *Proceedings of the SIGCHI conference on human factors in computing systems* (pp. 1315–1324). https://doi.org/10.1145/2556288.2557048

Harris, J. M., Ciorciari, J., & Gountas, J. (2019). Consumer neuroscience and digital/social media health/social cause advertisement effectiveness. *Behavioral Sciences, 9*(4), 42. https://doi.org/10.3390/bs9040042

Harrison, C. (2003). Visual social semiotics: Understanding how still images make meaning. *Technical Communication, 50*(1), 46–60.

Holden, S. J., & Vanhuele, M. (1999). Know the name, forget the exposure: Brand familiarity versus memory of exposure context. *Psychology and Marketing, 16*(6), 479–496. https://doi.org/10.1002/(SICI)1520-6793(199909)16:6%3C479::AID-MAR3%3E3.0.CO;2-Y

Holm, S. (1979). A simple sequentially rejective multiple test procedure. *Scandinavian Journal of Statistics*, 65–70. https://www.jstor.org/stable/4615733.

Hsu, Y. C., & Chiu, T. P. (2025). Uncertainty makes beautiful? The cross-cultural study of consumers' aesthetic preference on Instagram posts of multinational brands. In P. L. P. Rau (Ed.), *Lecture notes in computer science: 15784. Cross-cultural design*. Cham: Springer. https://doi.org/10.1007/978-3-031-93739-2_3. HCII 2025.

iMotions. (2022). Fixation and saccade classification with the I-VT filter. *iMotions*. July 23, 2025 https://help.imotions.com/v2/docs/en/fixation-and-saccade-classification-with-the-i-vt-filter.

inStreamly. (2025a). Snickers is in the game: Artwork displays react to in-game events. ***inStreamly***. https://instreamly.com/posts/snickers-is-in-the-game-artwork-displays-react-to-in-game-events/.

inStreamly. (2025b). InStreamly monte snack AI VRM campaign: Display reaction case study. ***inStreamly***. https://instreamly.com/posts/instreamly-monte-snack-ai-vrm-campaign-case-study/.

Itti, L., & Koch, C. (2001). Computational modelling of visual attention. *Nature Reviews Neuroscience, 2*(3), 194–203. https://doi.org/10.1038/35058500

Jenny, S. E., Besombes, N., Brock, T., & Cote, A. C. (Eds.). (2024). *Routledge handbook of esports*. Routledge. https://doi.org/10.4324/9781003410591.

Johnson, C. I., & Mayer, R. E. (2012). An eye movement analysis of the spatial contiguity effect in multimedia learning. *Journal of Experimental Psychology: Applied, 18*(2), 178. https://psycnet.apa.org/doi/10.1037/a0026923.

Johnson, M. R., & Woodcock, J. (2019a). The impacts of live streaming and Twitch. tv on the video game industry. *Media, Culture & Society, 41*(5), 670–688. https://doi.org/10.1177/0163443718818363

Johnson, M. R., & Woodcock, J. (2019b). "And today's top donator is": How live streamers on Twitch. tv monetize and gamify their broadcasts. *Social Media+ Society, 5*(4), Article 2056305119881694.

Kanwisher, N., McDermott, J., & Chun, M. M. (1997). The fusiform face area: A module in human extrastriate cortex specialized for face perception. *Journal of Neuroscience, 17*(11), 4302–4311. https://doi.org/10.1523/JNEUROSCI.17-11-04302.1997

Lal, S., Eysink, T. H. S., Gijlers, H. A., Veldkamp, B. P., Steinrücke, J., & Verwey, W. B. (2024). Explicit metrics for implicit emotions: Investigating physiological and gaze indices of learner emotions. *Frontiers in Psychology, 15*, Article 1440425. https://doi.org/10.3389/fpsyg.2024.1440425

Lang, A. (2006). Using the limited capacity model of motivated mediated message processing to design effective cancer communication messages. *Journal of Communication, 56*(suppl_1), S57–S80. https://doi.org/10.1111/j.1460-2466.2006.00283.x

Langton, S. R., Law, A. S., Burton, A. M., & Schweinberger, S. R. (2008). Attention capture by faces. *Cognition, 107*(1), 330–342. https://doi.org/10.1016/j.cognition.2007.07.012

Li, H., Edwards, S. M., & Lee, J. H. (2013). Measuring the intrusiveness of advertisements: Scale development and validation. *Journal of Advertising, 31*(2), 37–47. https://doi.org/10.1080/00913367.2002.10673665

Machleit, K. A., Madden, T. J., & Allen, C. T. (1990). Measuring and modeling brand interest as an alternative aad effect with familiar brands. *Advances in Consumer Research, 17*(1).

Makransky, G., Terkildsen, T. S., & Mayer, R. E. (2019). Role of subjective and objective measures of cognitive processing during learning in explaining the spatial contiguity effect. *Learning and Instruction, 61*, 23–34. https://doi.org/10.1016/j.learninstruc.2018.12.001

Mañas-Viniegra, L., Núñez-Gómez, P., & Tur-Viñes, V. (2020). Neuromarketing as a strategic tool for predicting how instagramers have an influence on the personal identity of adolescents and young people in Spain. *Heliyon, 6*(3), Article e03578. https://doi.org/10.1016/j.heliyon.2020.e03578

Mancini, M., Cherubino, P., Cartocci, G., Martinez, A., Di Flumeri, G., Petruzzellis, L., Cimini, M., Aricò, P., Trettel, A., & Babiloni, F. (2022). Esports and visual attention: Evaluating in-game advertising through eye-tracking during the game viewing experience. *Brain Sciences, 12*(10), 1345. https://doi.org/10.3390/brainsci12101345

Mayer, R. E., & Anderson, R. B. (1992). The instructive animation: Helping students build connections between words and pictures in multimedia learning. *Journal of Educational Psychology, 84*(4), 444–452. https://doi.org/10.1037/0022-0663.84.4.444

Mayer, R. E., & Fiorella, L. (2014). Principles for reducing extraneous processing in multimedia learning: Coherence, signaling, redundancy, spatial contiguity, and temporal contiguity principles. *Cambridge University Press*, 279–315. https://doi.org/10.1017/CBO9781139547369.015

McDonald, R. P. (1999). Test theory: A unified treatment. *Lawrence Erlbaum, Mahwah, NJ*.

Mileti, A., Guido, G., & Prete, M. I. (2016). Nanomarketing: A new frontier for neuromarketing. *Psychology and Marketing, 33*(8), 664–674. https://doi.org/10.1002/mar.20907

Min, B., Kim, D., & Ko, Y. J. (2025). Effect of in-game situations and advertisement animation in eSports on visual attention, memory, brand attitude and behavioral intentions. *Internet Research*. https://doi.org/10.1108/INTR-03-2024-0491

Minen, M., Gregoret, L., Seernani, D., & Wilson, J. (2023). Systematic evaluation of driver's behavior: A multimodal biometric study. In C. Stephanidis, M. Antona, S. Ntoa, & G. Salvendy (Eds.), *Communications in computer and information science: Vol. 1836. HCI international 2023 posters. HCII 2023*. Cham: Springer. https://doi.org/10.1007/978-3-031-36004-6_9.

Missaglia, A. L., Oppo, A., Mauri, M., Ghiringhelli, B., Ciceri, A., & Russo, V. (2017). The impact of emotions on recall: An empirical study on social ads. *Journal of Consumer Behaviour, 16*(5), 424–433. https://doi.org/10.1002/cb.1642

Mital, P. K., Smith, T. J., Hill, R. L., & Henderson, J. M. (2011). Clustering of gaze during dynamic scene viewing is predicted by motion. *Cognitive Computation, 3*(1), 5–24. https://doi.org/10.1007/s12559-010-9074-z

Nunnally, J. C. (1978). *Psychometric theory*. New York, NY: McGraw-Hill.

Ou, Y., Lyu, X., & Liu, X. (2025). A study of the impact of brand logo familiarity on visual search: Based on an eye-tracking experiment. *Current Psychology*. https://doi.org/10.1007/s12144-025-07839-3

Palermo, R., & Rhodes, G. (2007). Are you always on my mind? A review of how face perception and attention interact. *Neuropsychologia, 45*(1), 75–92. https://doi.org/10.1016/j.neuropsychologia.2006.04.025

Perneger, T. V., Courvoisier, D. S., Hudelson, P. M., et al. (2015). Sample size for pre-tests of questionnaires. *Quality of Life Research, 24*, 147–151. https://doi.org/10.1007/s11136-014-0752-2

Phua, P., Page, B., Trinh, G., Hartnett, N., & Kennedy, R. (2023). Does childhood exposure to a brand improve brand name recognition? Comparing age-of-acquisition effects with ongoing brand exposure and experience. *Journal of Advertising Research, 63*(4), 370–383. https://doi.org/10.2501/JAR-2023-025

Plassmann, H., Ramsøy, T. Z., & Milosavljevic, M. (2012). Branding the brain: A critical review and outlook. *Journal of Consumer Psychology, 22*(1), 18–36. https://doi.org/10.1016/j.jcps.2011.11.010

Pollack, C. C., Gilbert-Diamond, D., Emond, J. A., Eschholz, A., Evans, R. K., Boyland, E. J., & Masterson, T. D. (2021). Twitch user perceptions, attitudes and behaviours in relation to food and beverage marketing on Twitch compared with YouTube. *Journal of nutritional science, 10*, Article e32. https://doi.org/10.1017/jns.2021.22

Qian, T. Y., Zhang, J. J., Wang, J. J., & Hulland, J. (2019). Beyond the game: Dimensions of esports online spectator demand. *Communication & Sport, 8*(6), 825–851. https://doi.org/10.1177/2167479519839436

Resnick, M., & Albert, W. (2014). The impact of advertising location and user task on the emergence of banner Ad blindness: An eye-tracking study. *International Journal of Human-Computer Interaction, 30*(3), 206–219. https://doi.org/10.1080/10447318.2013.847762

Ro, T., Russell, C., & Lavie, N. (2001). Changing faces: A detection advantage in the flicker paradigm. *Psychological Science, 12*(1), 94–99. https://doi.org/10.1111/1467-9280.00317

Russo, V., Bilucaglia, M., Casiraghi, C., Chiarelli, S., Columbano, M., Fici, A., Rivetti, F., Rossi, C., Valesi, R., & Zito, M. (2023). Neuroselling: Applying neuroscience to selling for a new business perspective. An analysis on teleshopping advertising. *Frontiers in Psychology, 14*, Article 1238879. https://doi.org/10.3389/fpsyg.2023.1238879

Russo, V., Milani Marin, L. E., Fici, A., Bilucaglia, M., Circi, R., Rivetti, F., & Zito, M. (2021). Strategic communication and neuromarketing in the fisheries sector: Generating ideas from the territory. *Frontiers in Communication, 6*, Article 659484. https://doi.org/10.3389/fcomm.2021.659484

Sajjacholapunt, P., & Ball, L. J. (2014). The influence of banner advertisements on attention and memory: Human faces with averted gaze can enhance advertising effectiveness. *Frontiers in Psychology, 5*, 166. https://doi.org/10.3389/fpsyg.2014.00166

Santos, R. D. O. J., de Oliveira, J. H. C., Rocha, J. B., & Giraldi, J. D. M. E. (2015). Eye tracking in neuromarketing: A research agenda for marketing studies. *International Journal of Psychological Studies, 7*(1), 32. https://doi.org/10.5539/ijps.v7n1p32

Schroeder, N. L., & Cenkci, A. T. (2018). Spatial contiguity and spatial split-attention effects in multimedia learning environments: A meta-analysis. *Educational Psychology Review, 30*, 679–701. https://doi.org/10.1007/s10648-018-9435-9

Seo, Y. N., Kim, M., Lee, D., & Jung, Y. (2018). Attention to eSports advertisement: Effects of ad animation and in-game dynamics on viewers' visual attention. *Behaviour & Information Technology, 37*(12), 1194–1202. https://doi.org/10.1080/0144929x.2018.1488993

Shi, R., Wang, M., Qiao, T., & Shang, J. (2024). The effects of live streamer's facial attractiveness and product type on consumer purchase intention: An exploratory study with eye tracking technology. *Behavioral Sciences, 14*(5), 375. https://doi.org/10.3390/bs14050375

Sjöblom, M., & Hamari, J. (2017). Why do people watch others play video games? An empirical study on the motivations of Twitch users. *Computers in Human Behavior, 75*, 985–996. https://doi.org/10.1016/j.chb.2016.10.019

Sjöblom, M., Törhönen, M., Hamari, J., & Macey, J. (2017). Content structure is king: An empirical study on gratifications, game genres and content type on Twitch. *Computers in Human Behavior, 73*, 161–171. https://doi.org/10.1016/j.chb.2017.03.036

Sqalli, M. T., Al-Thani, D., Elshazly, M. B., Al-Hijji, M., Alahmadi, A., & Houssaini, Y. S. (2023). An eye-tracking based machine learning model towards the prediction of visual expertise for electrocardiogram interpretation. In R. Su, Y. Zhang, H. Liu, & A. F Frangi (Eds.), *Lecture notes in electrical engineering: 810. Medical imaging and computer-aided diagnosis. MICAD 2022*. Singapore: Springer. https://doi.org/10.1007/978-981-16-6775-6_25.

Statista. (n.d.). Distribution of Twitch users worldwide as of July 2025, by age group. Stat.Retrieved July 19, 2025, from https://www.statista.com/statistics/634057/twitch-user-age-worldwide/.

Sundstedt, V. (2012). Eye tracking in video games. In *Gazing at games. Synthesis lectures on computer graphics and animation*. Cham: Springer. https://doi.org/10.1007/978-3-031-79552-7_4.

Swacha, J. (2025). The relative popularity of video game genres in the scientific literature: A bibliographic survey. *Multimodal Technologies and Interaction, 9*(6), 59. https://doi.org/10.3390/mti9060059

Törhönen, M., Sjoblom, M., Vahlo, J., & Hamari, J. (2020). View, play and Pay?–the relationship between consumption of gaming video content and video game playing and buying. In *Proceedings of the 53rd Hawaii international conference on system sciences (Hawaii)*.

Visconti di Oleggio Castello, M., & Gobbini, M. I. (2015). Familiar face detection in 180ms. *PLoS One, 10*(8), Article e0136548. https://doi.org/10.1371/journal.pone.0136548

Wang, Z. Y., & Cho, J. Y. (2022). Older adults' response to color visibility in indoor residential environment using eye-tracking technology. *Sensors, 22*(22), 8766. https://doi.org/10.3390/s22228766

Wang, H., Yang, J., Hu, M., Tang, J., & Yu, W. (2024). A comparative analysis for eye movement characteristics between professional and non-professional players in FIFA eSports game. *Displays, 81*, Article 102599. https://doi.org/10.1016/j.displa.2023.102599

Wedel, M., & Pieters, R. (2008). A review of eye-tracking research in marketing. *Review of Marketing Research, Naresh K. Malhotra*. https://doi.org/10.1108/S1548-6435(2008)0000004009

Wulf, T., Schneider, F. M., & Queck, J. (2021). Exploring viewers' experiences of parasocial interactions with videogame streamers on Twitch. *Cyberpsychology, Behavior, and Social Networking, 24*(10), 648–653. https://doi.org/10.1089/cyber.2020.0546

Yang, L., Zhang, W., Li, P., Tang, H., Chen, S., & Jin, X. (2025). The aiming advantages in experienced first-person shooter gamers: Evidence from eye movement patterns. *Computers in Human Behavior, 165*, Article 108573. https://doi.org/10.1016/j.chb.2025.108573

Zeelenberg, R., Wagenmakers, E. J., & Rotteveel, M. (2006). The impact of emotion on perception: Bias or enhanced processing? *Psychological Science, 17*(4), 287–291. https://doi.org/10.1111/j.1467-9280.2006.01700.x. PMID: 16623684.

Zhang, X. (2021). Using eye tracking to analyze the effects of spatial contiguity in MOOC video subtitles. In D. Harris, & W. C. Li (Eds.), *Lecture notes in computer science: 12767. Engineering psychology and cognitive ergonomics. HCII 2021*. Cham: Springer. https://doi.org/10.1007/978-3-030-77932-0_27.

Zhou, L., & Xue, F. (2021). Show products or show people: An eye-tracking study of visual branding strategy on Instagram. *The Journal of Research in Indian Medicine, 15*(4), 729–749. https://doi.org/10.1108/JRIM-11-2019-0175