

# Characterizing Collective Attention in YouTube Live Chats of Brazilian Football Rivalries

Lucas Emanuel Cota Carneiro

lucas.ecc@aluno.ufop.edu.br

Department of Computing and Systems, UFOP  
Brazil

Josemar Alves Caetano

josemar@ufop.edu.br

Department of Computing and Systems, UFOP  
Brazil

## ABSTRACT

Online live chats have become an important space for collective participation during live sports broadcasts, enabling large audiences to react synchronously to unfolding events. Despite growing interest in collective attention in online systems, relatively little is known about how attention evolves during live sports discussions and how different attention levels are reflected in audience language. In this paper, we investigate collective attention dynamics in YouTube live chats associated with three major Brazilian football rivalries: Atlético Mineiro versus Cruzeiro (*Super Clássico Mineiro*), Corinthians versus Palmeiras (*Derby Paulista*), and Flamengo versus Fluminense (*Fla-Flu*). We model each match as a temporal sequence of relative attention levels derived from minute-level message activity and analyze both their temporal dynamics and linguistic characteristics. First, we characterize collective attention through attention-level distributions, transition probabilities, the duration of peak-attention episodes, and their relationship with salient match events. Second, we investigate the linguistic and emotional patterns associated with different attention levels. Our results reveal that all rivalries share a common attention structure in which most match minutes are concentrated in low- and moderate-attention levels, while peaks are comparatively rare. The rivalries differ mainly in how strongly goals drive the chat toward peak attention, whereas apparent differences in how long peak-attention episodes last are largely explained by volume. Across all rivalries, high-attention levels share a distinctive linguistic profile marked by informal language, emotional expression, and markers of anxiety, anger, and profanity. These findings suggest that rivalries diverge mainly in how peak collective attention is triggered rather than in their baseline distributions, even as the linguistic signature of intense attention remains largely shared. Beyond contributing to the understanding of user behavior during football matches, this work offers a methodology for analyzing collective attention in other contexts, including political debates, breaking news coverage, and other live-streamed events.

## KEYWORDS

Collective Attention, Live Chat, YouTube, Football, Stream

## 1 INTRODUCTION

Online live chats have become an integral component of contemporary sports consumption, enabling thousands of viewers to collectively react to unfolding events during live broadcasts [31, 45]. Platforms such as YouTube allow audiences to comment, celebrate, criticize, and interact in real time while matches are being played [31, 39]. The scale of contemporary sports livestreaming highlights the growing importance of these environments. For instance, Brazil's first match in the 2026 FIFA World Cup attracted more than 12 million simultaneous viewers on a popular Brazilian YouTube broadcast [29]. Events of this magnitude generate large volumes of user interactions, making live chats a valuable source for studying how groups collectively allocate attention to unfolding events in real time.

Collective attention refers to the phenomenon in which multiple individuals direct their cognitive resources toward the same stimulus, piece of information, or event [2, 10, 30, 49]. This concept has been studied across disciplines such as psychology, sociology, and computer science, where it is explored through diverse theoretical approaches that emphasize shared focus and group dynamics [3, 8, 44]. In online environments, collective attention is commonly observed through behavioral traces such as user activity, information consumption, and communication patterns, which reveal how groups collectively react to salient events and dynamically allocate attention over time [19, 22, 43, 50].

While collective attention has been studied across a variety of online platforms, comparatively less research has addressed highly synchronous environments such as YouTube live chats. These environments are characterized by synchronous, fast-paced exchanges that often resemble offline group conversations [6, 8]. Participants frequently engage in overlapping dialogues, shift attention rapidly between topics, and contribute to the collaborative construction of ongoing conversations [7, 33]. Because message production constitutes the primary observable form of interaction in live chats, temporal fluctuations in message activity provide an observable indicator of how collective attention emerges, intensifies, and declines throughout an event [8, 43]. Among the many contexts in which live chats are used, sports broadcasts provide particularly rich opportunities for observing collective attention dynamics.

Football constitutes a particularly interesting setting for studying collective attention in online environments. Matches are characterized by uncertainty, emotional intensity, and rapidly changing competitive contexts [12]. Goals, missed opportunities, referee decisions, substitutions, and controversial plays frequently trigger large-scale audience reactions, producing sudden fluctuations in communication activity [12, 26]. Among football matches, regional rivalries, commonly referred to as *classics*, represent some of the

most emotionally charged events in Brazilian sports [41]. Rivalries such as Atlético Mineiro versus Cruzeiro, Corinthians versus Palmeiras, and Flamengo versus Fluminense are historically associated with strong social identities, intense competition, and elevated fan engagement, making them ideal scenarios for investigating collective behavior in synchronous online discussions [34].

Previous research has examined sports-related discussions through sentiment analysis, fan engagement, toxicity, and interaction dynamics on social media platforms [9, 18, 20, 28]. More broadly, studies of collective attention have shown that online audiences often concentrate their activity around salient events, generating bursts of communication and engagement [14, 22, 43]. However, these two research streams remain largely disconnected. While sports-related studies typically focus on engagement and content characteristics, collective attention studies rarely investigate how attention evolves and persists in highly synchronous sports discussions or how different attention levels are reflected in audience language [6, 8]. Consequently, our understanding of both the temporal dynamics of collective attention and its linguistic manifestations during large-scale sports broadcasts remains limited.

In this work, we address these questions through a large-scale study of YouTube live chats collected during three major Brazilian football rivalries: Atlético Mineiro versus Cruzeiro (*Super Clássico Mineiro*), Corinthians versus Palmeiras (*Derby Paulista*), and Flamengo versus Fluminense (*Fla-Flu*), broadcast between 2022 and 2026. We model each match as a temporal sequence of relative attention levels derived from minute-level message activity. Unlike previous studies that focus primarily on isolated events or aggregate engagement measures [18, 50], our approach represents collective attention as a sequence of attention levels, enabling the characterization of both temporal attention dynamics and the linguistic and emotional properties associated with different levels of audience engagement. This representation allows us to investigate not only how collective attention evolves, but also how peak-attention periods emerge, persist, and relate to salient match events.

Moreover, because it depends only on minute-level message activity, our attention-level representation generalizes beyond football: it can support comparisons of collective attention across events with different audience sizes, platforms, and languages, such as e-sports tournaments, political debates, breaking news coverage, and other live-streamed events.

Our study is guided by the following research questions:

**RQ1:** How does collective attention evolve and persist during online discussions of football rivalries?

**RQ2:** What linguistic and emotional characteristics are associated with different collective attention levels?

To answer these questions, we first characterize the temporal dynamics of collective attention by analyzing attention-level distributions, transition probabilities, how long peak-attention episodes last, and the relationship between goals and attention peaks. We then investigate how language changes across attention levels through the analysis of linguistic and emotional categories [11].

Our results reveal that football rivalries share a common attention structure in which most of the match is spent in low- and moderate-attention levels, while peaks are comparatively rare. The

rivalries differ mainly in how strongly goals drive the chat toward peak attention, whereas apparent differences in how long peak-attention episodes last are largely explained by chat volume. Furthermore, we show that high-attention levels exhibit a distinctive linguistic profile marked by informal language, emotional expression, and markers of anxiety, anger, and profanity.

The remainder of this paper is organized as follows. Section 2 reviews related work on collective attention, synchronous live chats, and football fandoms. Section 3 describes our data collection and the attention-level representation. Section 4 presents the results for both research questions, and Section 5 discusses their implications. Section 6 concludes and outlines limitations and directions for future work. Finally, Section 7 discusses the ethical aspects related to data collection and use in this work.

## 2 RELATED WORK

Our study lies at the intersection of three research lines: collective attention in online environments, synchronous live chats and live-streaming audiences, and computational analyses of football fandoms. We review these areas and then position our contribution relative to prior work.

*Collective Attention in Online Environments.* Collective attention describes how many individuals simultaneously direct their cognitive resources toward the same stimulus or event [10, 49]. In online environments, it has been commonly studied through behavioral signals such as clicks, downloads, hashtags, shares, and content-consumption patterns [19, 35, 38]. These signals have been used to characterize how attention emerges, propagates, and decays in response to news, viral content, and social phenomena [10, 30, 40]. A recurring finding is that sudden increases in activity often correspond to bursts of collective focus triggered by salient events [10, 40], while models of novelty, decay, and competition help explain how users allocate attention across competing stimuli [1, 30, 38].

Much of this literature, however, focuses on asynchronous platforms with explicit engagement mechanisms, treating collective attention primarily as a matter of activity volume. Closer to synchronous settings, Caetano et al. [8] propose a three-pronged characterization of collective attention in WhatsApp, Telegram, and YouTube live chats, mapping it onto three cognitive dimensions: alertness, processing capacity, and selectivity, rather than reducing it to a single volume-based signal. In a recent journal extension of that work, Caetano et al. [5] build on this mapping to introduce a computational model that classifies each conversation period as *focused*, when participants concentrate intensively on a single discussion, or *unfocused*, when attention is distributed across multiple parallel discussions. Together, these two studies establish that collective attention in synchronous chats is a multidimensional, platform-comparable phenomenon, but characterize it structurally, without examining how audience language shifts as attention intensifies or how attention relates to external triggers such as goals.

*Synchronous Live Chats and Streaming Audiences.* Live-stream chats differ from forums and microblogs because messages are short, noisy, rapidly produced, and often visible only briefly. These characteristics make live chats highly interactive but also difficult to

analyze using models designed for asynchronous platforms. Moon et al. [36], for example, show that models trained on asynchronous platforms perform poorly on live-stream chats and that incorporating synchronous chat context improves performance. Related studies also highlight the fast-paced and conversational nature of online chats [27, 33, 46], supporting the use of message activity as a meaningful proxy for collective attention when explicit engagement signals are unavailable [6–8].

A growing body of work has examined engagement and collective behavior in live-streaming environments. Kurtz et al. [26] analyze YouTube live-chat activity during a Brazilian sports broadcast and show that comment peaks align with salient broadcast moments. Jiang et al. [25] study bullet chats in esports livestreaming and relate chat frequency to teams, matches, and toxicity. Ferreira et al. [18] adapt engagement metrics to live streaming across YouTube, Twitch, and Facebook, showing that interaction counts fluctuate over time and should be complemented by the temporal dynamics and semantics of comments. Other studies focus on harmful collective behavior, such as toxicity cascades [28], coordinated attacks [42], and hate raids in real-time chats [9, 15, 21, 48]. Across this literature, live chats are examined mainly through participation counts, engagement peaks, toxicity, or moderation-related phenomena, with comparatively little attention to how transitions between attention regimes unfold or how language varies as attention intensifies.

*Computational Analysis of Football Fandoms.* Football is a natural setting for studying event-driven collective behavior, and prior work has analyzed how fans react to match outcomes, team performance, and salient events. Wang and Lu [47] examine Premier League subreddits and show that the gap between pre-match expectations and actual results shapes sentiment, topics, and voting behavior. Hill [23] analyzes tens of millions of Reddit posts and finds that match outcomes affect sentiment and posting volume, with negative sentiment associated with problematic language and spillover effects. Mazhar et al. [34] track r/PremierLeague over a decade and report increasing negativity and toxicity as the community grows. Related work also shows that football fandoms can be mobilized for political influence [37], reinforcing the importance of fan communities as collective identities.

Closer to our work, Citraro et al. [12, 13] study the temporal dynamics of emotional reactions among Italian soccer fans on Instagram, showing that emotions exhibit different temporal patterns, with joy tending to be more regular and anger or sadness appearing in bursts. This line of work treats fan reactions as temporal signals rather than static aggregates, but analyzes asynchronous comments aggregated over an entire season. Taken together, studies of football fandoms have mostly examined asynchronous communities such as Reddit and Instagram over match-level or season-level timescales, leaving minute-by-minute, within-match discourse comparatively unexplored.

*Positioning of Our Work.* Prior research shows that collective attention online is event-driven and bursty, that live-stream chats provide rich traces of collective behavior, and that football fandoms react strongly to match events and outcomes. Nevertheless, these strands remain only partially connected. Studies of collective attention in synchronous chats characterize it as a multidimensional,

structural phenomenon, but without linking it to language or to independently verified external triggers; studies of live chats often focus on engagement peaks, toxicity, or participation; and studies of football fandoms frequently analyze asynchronous communities such as Reddit or Instagram rather than synchronous match-time interaction.

We bridge these strands by modeling each football match as a sequence of relative attention levels derived from minute-level message activity in YouTube live chats. Rather than broadening the set of structural dimensions used to characterize attention as in [5, 8], we hold the activity dimension fixed as a simple, comparable measure across matches, and use it to analyze the distribution of attention levels, transitions between levels, the persistence of high-attention episodes, and their relationship with goals. We further connect these temporal dynamics to the linguistic and emotional profile of messages using LIWC categories, showing that peak attention carries a distinctive discourse signature rather than merely a volume one. To our knowledge, this is the first study to jointly characterize how collective attention evolves and how audience language shifts across attention levels in synchronous live chats, and to do so comparatively across major Brazilian football rivalries.

### 3 METHODOLOGY

This section describes the methodology adopted in this work, including the data collection process, the analyses performed, and the computational tools used throughout the study to process and evaluate user interactions in YouTube live chats.

#### 3.1 Data Collection

We selected three of the most prominent Brazilian football rivalries, located in the country's three most populous states [4]: São Paulo, Minas Gerais, and Rio de Janeiro. For each rivalry, we used the YouTube channel of a major regional football broadcaster with a large audience and publicly available archived live chats. Regional broadcasters were preferred because they mostly attract supporters of the teams involved, reducing the influence of the neutral audiences common in nationwide broadcasts: Rádio Itatiaia<sup>1</sup> (Minas Gerais, ≈3.3M subscribers) for Atlético Mineiro vs. Cruzeiro, Rádio Bandeirantes<sup>2</sup> (São Paulo, ≈2M) for Corinthians vs. Palmeiras, and Rádio Tupi FM<sup>3</sup> (Rio de Janeiro, ≈1M) for Flamengo vs. Fluminense (subscriber counts as of June 2026).

For each rivalry, we collected the archived live chats of every match between the two clubs across all official competitions in which they met from 2022 to 2026: the Campeonato Brasileiro, the Copa do Brasil, the state championships (Mineiro, Paulista, and Carioca), and continental tournaments such as the Copa Libertadores and Copa Sul-Americana.

Live chat messages were collected using a custom Python crawler built on top of the pytchat library [24], which retrieves archived YouTube live chats through the platform's public API. For each broadcast, the crawler collected all available chat messages together with their timestamps and author information.

<sup>1</sup><https://www.itatiaia.com.br/>

<sup>2</sup><https://www.band.com.br/radio-bandeirantes>

<sup>3</sup><https://www.tupi.fm/category/esportes/>

To relate audience reactions to in-game events, we additionally collected, for each match, the official minute-by-minute timeline of salient events (in particular, goals) from the real-time match coverage provided by a Brazilian sports outlet that records match events at the minute level [16]. Because the live chat and the video stream unfold together as the broadcast is transmitted, the minute-level alignment between chat activity and match events is consistent within each broadcast.

Table 1 summarizes the resulting dataset for each rivalry in terms of the number of collected messages and unique users. The final dataset comprises 59 matches broadcast on YouTube.

| Rivalry                       | Messages       | Unique users  |
|-------------------------------|----------------|---------------|
| <b>Super Clássico Mineiro</b> |                |               |
| Atlético Mineiro vs. Cruzeiro | 341,947        | 56,907        |
| <b>Derby Paulista</b>         |                |               |
| Corinthians vs. Palmeiras     | 78,199         | 11,900        |
| <b>Fla-Flu</b>                |                |               |
| Flamengo vs. Fluminense       | 125,264        | 20,247        |
| <b>Total</b>                  | <b>545,410</b> | <b>89,054</b> |

Table 1: Live chat statistics per rivalry.

### 3.2 Attention Level Definition

For each match  $m$ , we aggregate all chat messages into one-minute bins. Football match events such as goals, red cards, substitutions, and penalties are conventionally reported and recorded at the minute level [16, 17], making this resolution a natural unit for relating audience reactions to in-game events while preserving enough granularity to capture short-term fluctuations in collective attention.

Let  $x_m(t)$  denote the number of messages posted during minute  $t$  of match  $m$ . This produces a temporal sequence:

$$X_m = \{x_m(1), x_m(2), \dots, x_m(T_m)\},$$

where  $T_m$  is the total number of observed minutes for match  $m$ .

Because matches differ in their overall audience size and chat volume, directly comparing raw message counts may conflate attention dynamics with differences in broadcast reach. To avoid this issue, we define attention levels relative to the maximum activity observed within each match. Let

$$x_m^{\max} = \max_{1 \leq t \leq T_m} x_m(t)$$

be the maximum number of messages observed in any minute of match  $m$ . We then compute the relative activity of each minute as:

$$a_m(t) = \frac{x_m(t)}{x_m^{\max}}.$$

Thus,  $a_m(t) \in [0, 1]$ , where values close to 0 indicate minutes with low activity relative to that match, and values close to 1 indicate minutes near the highest activity level observed in that match.

Based on  $a_m(t)$ , we assign each minute to one of four relative attention levels, denoted  $L1$  to  $L4$ . Each level corresponds to a fixed-width band of relative activity rather than to a sample quartile of the activity distribution:

$$Level_m(t) = \begin{cases} L1, & 0 \leq a_m(t) < 0.25, \\ L2, & 0.25 \leq a_m(t) < 0.50, \\ L3, & 0.50 \leq a_m(t) < 0.75, \\ L4, & 0.75 \leq a_m(t) \leq 1.00. \end{cases}$$

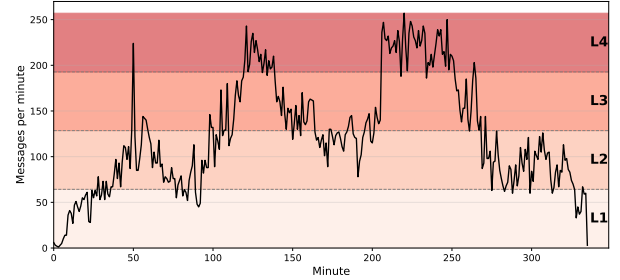


Figure 1: Example of the attention-level representation for a single match (2026 Campeonato Mineiro final). Minute 0 marks the start of the broadcast (instead of match kickoff).

We use four levels as a balance between temporal granularity – enough to capture gradual escalation toward peak attention rather than abrupt low-to-high jumps – and the number of observations available per level and rivalry, particularly for the lower-volume *Derby Paulista* (Table 1); finer bands would leave the highest levels with too few minutes for stable estimates. These levels represent increasing degrees of collective attention relative to each match’s internal activity scale, from low ( $L1$ ) to peak ( $L4$ ) activity.

We adopt the per-match maximum as a simple, interpretable reference for the most intense moment of each broadcast. We note, however, that it makes the whole scale depend on a single reference minute: an anomalous spike compresses all other minutes toward lower levels, possibly inflating the share of time attributed to  $L1$ . We return to this limitation in Section 6.

After assigning attention levels to all minutes, each match is represented as a symbolic sequence of attention levels:

$$\mathcal{L}_m = \{Level_m(1), Level_m(2), \dots, Level_m(T_m)\}.$$

Figure 1 illustrates this representation for the match with the largest number of messages in our dataset: most minutes fall within  $L1$  and  $L2$ , while  $L4$  is reached only during short, sporadic peaks. We treat the entire transmission as the match, spanning the pre-match buildup, both halves, half-time, and post-match, so the match minute extends beyond the 90 regulation minutes.

We use these sequences to analyze collective attention in three ways. First, we compute the proportion of match time spent in each attention level, which indicates whether a rivalry tends to remain predominantly in low-, moderate-, high-, or peak-attention regimes ( $L1$ ,  $L2$ ,  $L3$ , and  $L4$ , respectively). Second, we estimate transition probabilities between levels in consecutive minutes, allowing us to examine how stable or volatile collective attention is during each rivalry. Third, we focus on  $L4$  episodes, defined as maximal contiguous intervals of one or more minutes during which the match remains in the highest relative attention level. These episodes capture periods of sustained collective attention; we report their

durations explicitly, so that single-minute peaks are distinguished from longer, sustained episodes.

Finally, we relate goal events to the attention-level sequence to examine how strongly goals drive collective attention. For each goal, occurring at an official match minute  $t_g$  obtained from the minute-by-minute coverage [16], we consider a forward window of  $w$  minutes,  $[t_g, t_g + w]$ , and identify the highest attention level reached within it. We assign the goal to that level, so that each goal is summarized by the peak collective attention that follows it, ranging from  $L_1$  (the goal is not followed by any noticeable rise in activity) to  $L_4$  (the goal is followed by a peak-attention episode). Aggregating these assignments per rivalry yields, for each rivalry, a distribution of goals over the four attention levels.

We set  $w = 3$  minutes, which corresponds to the typical interval during which the reaction to a goal unfolds in the broadcast, encompassing the goal celebration, the replay, and the immediate audience discussion. This choice balances sensitivity, capturing the reaction that a goal elicits, against specificity, avoiding the attribution to a goal of attention peaks that arise from other salient events. To ensure that our conclusions do not depend on this particular choice, we verified that the results remain stable for  $w \in \{2, 3, 4\}$  minutes.

## 4 RESULTS

In this section, we investigate how collective attention unfolds during football live chats and how its characteristics vary across rivalries. We organize our analyses according to the research questions introduced in Section 1. All reported statistical tests use a significance level of 0.05.

### 4.1 Evolution and Persistence of Collective Attention (RQ1)

To answer RQ1, we model each match as a sequence of relative attention levels derived from minute-level message activity. This representation allows us to investigate both the global structure of collective attention and the dynamics of its most intense periods. Specifically, we analyze (i) the proportion of match time spent in each attention level, (ii) the transition probabilities between consecutive levels, (iii) the duration of peak-attention episodes ( $L_4$ ), and (iv) the relationship between goals and the attention levels they lead to. Together, these analyses provide complementary perspectives on how collective attention emerges, persists, and reacts to salient match events across rivalries.

Figure 2 presents the distribution and transition dynamics of relative attention levels across rivalries. Overall, all rivalries exhibit a common attention structure. Most match minutes are concentrated in the two lowest attention levels ( $L_1$  and  $L_2$ ), whereas  $L_3$  is substantially less frequent and  $L_4$  represents only a small fraction of the total match duration. This pattern is consistently observed across the *Super Clássico Mineiro* (Atlético Mineiro vs. Cruzeiro), the *Derby Paulista* (Corinthians vs. Palmeiras), and the *Fla-Flu* (Flamengo vs. Fluminense).

We compute transition matrices by treating the sequence of per-minute attention levels within each match as a discrete-time Markov chain, with transition probabilities estimated empirically as the proportion of minute-to-minute changes from each level to

every other level, aggregated across all matches within a rivalry. The transition matrices reveal similarly stable dynamics across rivalries: self-transition probabilities are strong, particularly for  $L_1$  (90–94%) and  $L_2$  (67–77%), indicating that attention persists over consecutive minutes. Transitions occur predominantly between adjacent levels, while direct jumps between  $L_1$  and  $L_4$  are virtually absent. We argue that attention shifts progressively, with audiences moving through intermediate levels before reaching peak attention.

A chi-square test indicates that the distribution of attention levels differs significantly across rivalries ( $\chi^2 = 191.68$ ,  $p < 0.001$ ). However, the associated effect size is small (Cramér's  $V = 0.074$ ), suggesting that rivalry type explains only a limited fraction of the observed variability. Therefore, despite statistical differences, the overall structure of collective attention remains largely consistent across rivalries.

Nevertheless, some distinctions emerge: the *Derby Paulista* spends more time in  $L_3$ , while the *Super Clássico Mineiro* allocates slightly more to  $L_1$ . Raw  $L_4$  persistence also differs: 70.0% in the *Super Clássico Mineiro* versus 46.2% (*Derby Paulista*) and 50.0% (*Fla-Flu*). However, as shown below, this raw persistence is sensitive to chat volume, motivating a closer look at  $L_4$  episodes and their relationship with salient events such as goals.

Figure 3 focuses on the peak attention level ( $L_4$ ,  $a_m(t) \geq 0.75$ ). Although  $L_4$  accounts for only a small fraction of match time (Figure 2), substantial differences emerge in episode persistence and association with match events.

Figure 3a shows the distribution of  $L_4$  episode durations across rivalries (based on 82, 78, and 76  $L_4$  episodes in the *Super Clássico Mineiro*, *Derby Paulista*, and *Fla-Flu*, respectively). While all rivalries exhibit short peak-attention bursts, the *Super Clássico Mineiro* presents considerably longer episodes than the other rivalries. On average,  $L_4$  episodes last 3.33 minutes in the *Super Clássico Mineiro*, compared to 1.86 minutes in the *Derby Paulista* and 2.00 minutes in the *Fla-Flu*. Moreover, the *Super Clássico Mineiro* contains extreme episodes lasting up to 18 minutes, against at most 7 minutes in the other two. A Kruskal-Wallis test confirms that the  $L_4$  duration distributions differ significantly across rivalries ( $H = 6.72$ ,  $p = 0.035$ ). At face value, this suggests that the *Super Clássico Mineiro* sustains peak attention for longer once it emerges; we show below, however, that this gap is largely explained by chat volume.

Figure 3b shows the probability that a goal leads the live chat to each attention level within the following three minutes. Each goal is assigned to the highest attention level reached in the post-goal window  $[t_g, t_g + w]$ , with  $w = 3$  (Section 3). The results reveal clear differences across rivalries. In the *Super Clássico Mineiro* ( $n = 26$ ), post-goal reactions are spread across all levels, leading to  $L_1$ ,  $L_2$ ,  $L_3$ , and  $L_4$  in 30.8%, 26.9%, 19.2%, and 23.1% of cases, respectively. Goals therefore do not systematically drive the chat toward peak attention in this rivalry.

In contrast, the *Derby Paulista* ( $n = 33$ ) concentrates strongly in high attention after goals: only 3.0% and 9.1% of goals lead to  $L_1$  and  $L_2$ , while 36.4% lead to  $L_3$  and 51.5% lead to  $L_4$ . The *Fla-Flu* ( $n = 39$ ) presents an intermediate pattern: goals rarely lead only to  $L_1$  (2.6%), and post-goal reactions are distributed across  $L_2$  (30.8%),  $L_3$  (30.8%), and  $L_4$  (35.9%).

A chi-square test of independence confirms that the distribution of post-goal attention levels differs significantly across rivalries

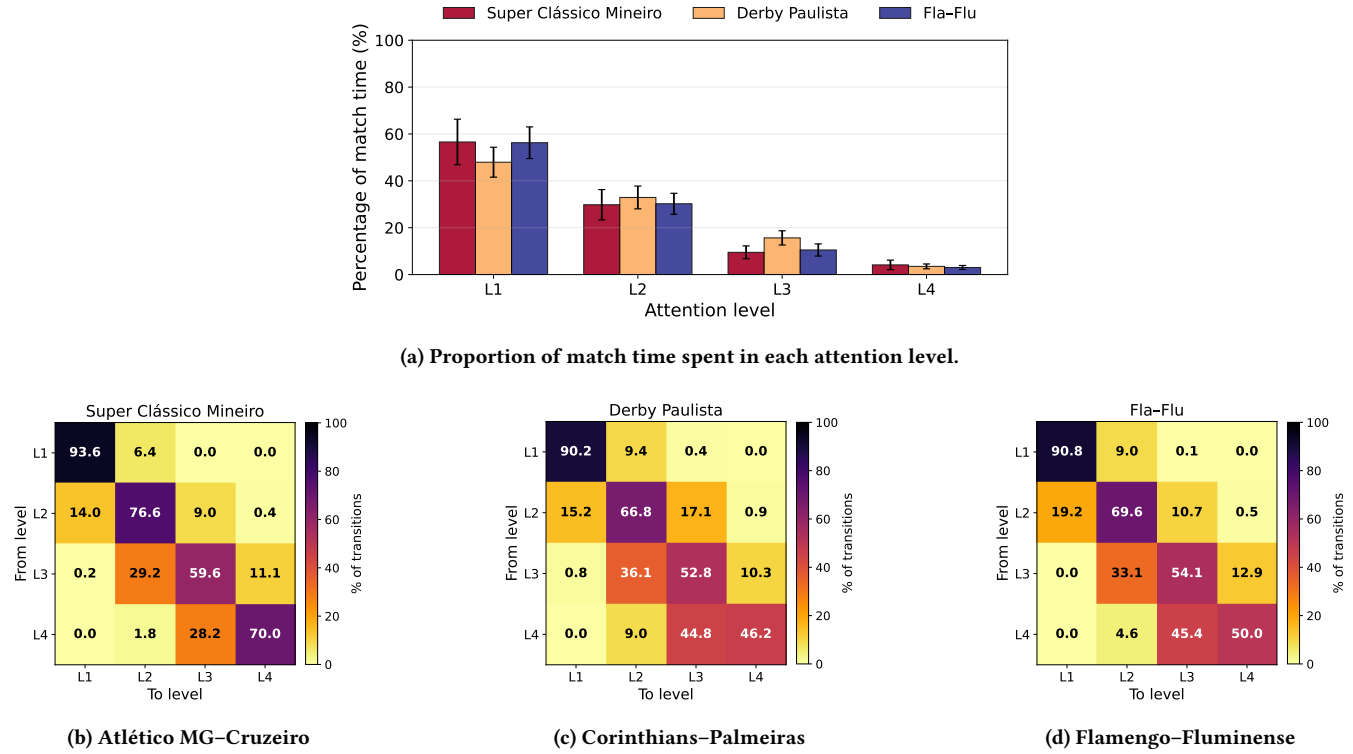


**Figure 2: Collective attention levels during football rivalries. (a) Proportion of match time spent in each relative attention level (L1–L4). (b)–(d) Minute-to-minute transition probabilities between levels ( $L_i \rightarrow L_j$ ) for each rivalry.**

( $\chi^2 = 23.38$ ,  $df = 6$ ,  $p < 0.001$ , Cramér’s  $V = 0.345$ ,  $n = 98$ ), indicating a moderate association between rivalry and the attention level reached after goals. Because a few cells have expected counts below five, we corroborated this result with the Fisher–Freeman–Halton exact test [32], approximated via Monte Carlo simulation ( $2 \times 10^6$  random tables with the same row and column totals as the observed table, drawn under the null hypothesis of independence;  $p$  estimated as the fraction of sampled tables at least as extreme as the observed one), which also indicates a significant association ( $p = 0.0017$ ). Goals thus act as stronger triggers of peak attention in the *Derby Paulista*, and to a lesser extent in the *Fla-Flu*, whereas goal-induced reactions in the *Super Clássico Mineiro* are markedly more heterogeneous.

These results also suggest that goals are only one source of peak attention: since many goals do not lead to  $L_4$ , other factors must also contribute, such as red cards, controversial refereeing, or chat-internal dynamics like banter. Peak attention thus reflects both game-driven and audience-driven processes.

Because the *Super Clássico Mineiro* records substantially more messages per match than the other rivalries (Table 1), denser activity series are statistically smoother and could, in principle, inflate  $L_4$  durations and goal-driven peaks as an artifact of volume rather than fan behavior. To test this, we repeatedly downsampled its messages to the per-minute volume of the *Derby Paulista* and recomputed the  $L_4$  statistics (95% intervals below). The two findings diverge. The weak goal-driven activation is robust: the probability that a goal

leads to  $L_4$  stays low after downsampling (0.21, [0.15, 0.27]), far below the *Derby Paulista* (0.52). The longer  $L_4$  episodes, however, do not survive: mean duration drops from 3.33 to 1.90 minutes ([1.63, 2.21]) and median from 2.0 to  $\approx 1.0$ , converging to the *Derby Paulista* (1.86, 1.0). The duration gap is thus largely a volume effect, whereas the difference in how strongly goals trigger peak attention is volume-independent.

These results refine our answer to RQ1. Football rivalries share a common baseline structure, with most minutes in lower attention levels; the meaningful, volume-independent differences lie in how goals trigger  $L_4$ : weak in the *Super Clássico Mineiro*, strong in the *Derby Paulista*, and intermediate in the *Fla-Flu*, whereas the apparent gap in  $L_4$  duration is largely a volume artifact. Thus, collective attention in football live chats combines a shared temporal structure with rivalry-specific dynamics in how extreme engagement is triggered, a difference that holds even after accounting for chat volume.

## 4.2 Linguistic and Emotional Characteristics Across Attention Levels (RQ2)

To answer RQ2, we investigate whether different attention levels are associated with distinct linguistic and emotional patterns. Specifically, we analyze the prevalence of selected LIWC [11] categories across attention levels ( $L_1$ – $L_4$ ) and rivalries.

LIWC maps words to psychologically and socially meaningful categories, enabling the analysis of dimensions such as cognitive

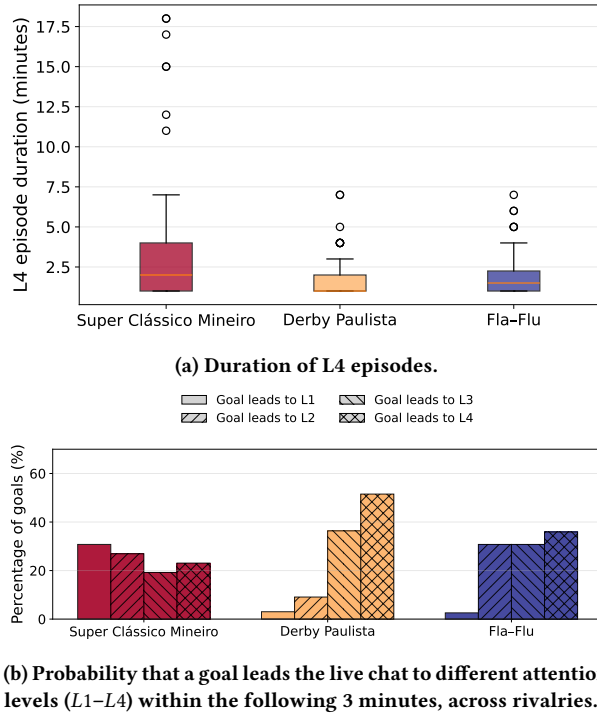


**Figure 3: Characteristics of the attention levels.**

processes and conflict-related language [11]. We use the validated Brazilian Portuguese adaptation of the lexicon, BP-LIWC2015 [11]. For each message, we compute the percentage of its words belonging to each category; we then average these values over all messages assigned to a given attention level.

The selected categories capture dimensions related to informal communication (*netspeak*), emotional expression (*affect*, *negemo*, *anx*, *anger*), offensive language (*swear*), future-oriented discourse (*focusfuture*), and motivational processes (*drives*). Together, these categories provide a broad view of how audience language changes as collective attention intensifies.

Figure 4 reveals a consistent pattern across all rivalries: as collective attention increases from *L1* to *L4*, messages become progressively more informal and emotionally expressive. Informal and affective language rises steadily with attention—*netspeak*, which captures slang and chat conventions, and *affect* both increase from low- to peak-attention levels, indicating that emotionally charged language becomes more prevalent during intense engagement.

The sharpest changes occur in categories tied to negative emotional activation. *Negative emotion* and *anxiety* increase substantially as attention intensifies, while *anger* appears predominantly in *L4*, becoming one of the most salient categories at peak attention. Profanity (*swear*) is also more frequent at higher attention levels, peaking in *L4* for the *Fla-Flu* and the *Super Clássico Mineiro*; in the *Derby Paulista* it is already elevated at intermediate levels and shows wider confidence intervals, making its trend less monotonic. Together, these patterns indicate that peak attention is associated

not merely with higher participation but with heightened emotional arousal. Two further categories remain consistently prominent across all levels: *focusfuture*, reflecting audience anticipation of upcoming plays and outcomes, and *drives*, capturing motivational, competition-oriented language; both show a mild increase toward higher attention.

These findings directly answer RQ2. As collective attention intensifies from *L1* to *L4*, audience discourse becomes not only more active but also markedly more emotional and informal, with greater use of *netspeak*, affective and anxiety-related terms, anger, profanity, and competition-oriented vocabulary. This shift emerges consistently across the three rivalries analyzed, indicating that extreme collective attention carries a common linguistic and emotional signature that transcends individual rivalries.

## 5 DISCUSSION

Our findings indicate that collective attention in football live chats involves both temporal and linguistic dynamics. RQ1 shows how attention evolves and persists throughout the match, while RQ2 shows how audience language changes as attention intensifies. Taken together, these results suggest that high-attention levels should not be interpreted merely as higher message volume, but also as a qualitative shift in audience discourse.

A first implication concerns the temporal organization of collective attention. Football live chats are not uniformly intense: across all rivalries they alternate between long stretches of background activity (*L1-L2*) and short, rare episodes of peak engagement (*L4*), which audiences typically reach gradually, through intermediate levels, rather than abruptly.

The main differences between rivalries emerge in how *L4* is entered rather than in its overall duration: as shown in Section 4.1, the apparent gap in episode length is largely a volume artifact, whereas goal-driven activation remains a robust, volume-independent contrast – weak in the *Super Clássico Mineiro*, strong in the *Derby Paulista*, and intermediate in the *Fla-Flu*. This suggests that rivalry-specific dynamics are less evident in baseline activity levels, and even in raw peak durations, than in how audiences are driven into moments of extreme collective engagement.

The linguistic results complement this temporal interpretation. As attention increases from *L1* to *L4*, messages become more informal, emotional, and expressive, with greater prevalence of *netspeak*, *affect*, *negative emotion*, *anxiety*, *anger*, and *swear*. Thus, *L4* has a distinctive linguistic and emotional signature: when audiences collectively converge around salient match moments, they do not simply post more; they communicate differently. This within-match view complements prior work on emotional dynamics in soccer fandoms: whereas Citraro et al. [12, 13] characterize how emotions such as joy, anger, and sadness unfold across asynchronous Instagram reactions aggregated over a season, our results show that an analogous emotional intensification is already legible minute by minute, within a single match, in the language of synchronous live chats.

These findings have methodological implications for the study of collective attention in synchronous online environments. Activity-based indicators, such as message volume, clicks, hashtags, or engagement bursts, are useful for identifying moments of collective

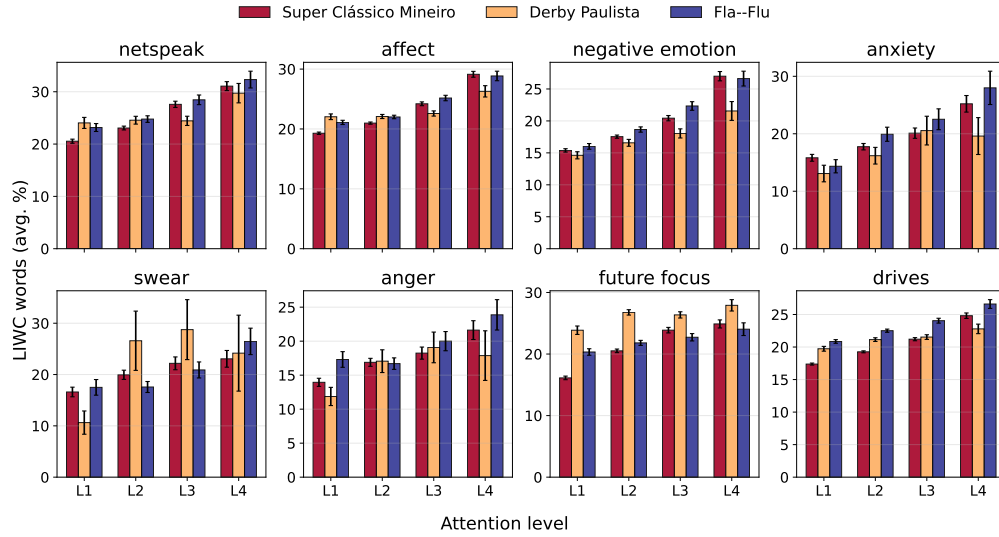


**Figure 4: Average per-message frequency of selected LIWC categories across attention levels (L1–L4) and rivalries. Each panel corresponds to one category; error bars are 95% confidence intervals.**

focus, but they capture only part of the phenomenon. By combining temporal attention levels with linguistic analysis, we can distinguish between periods that are merely active and periods that are emotionally charged.

## 6 CONCLUSION

This paper investigated collective attention in YouTube live chats of three major Brazilian football rivalries: *Super Clássico Mineiro*, *Derby Paulista*, and *Fla--Flu*. Modeling each match as a sequence of relative attention levels derived from minute-level message activity, we combined temporal modeling with LIWC-based linguistic analysis to characterize how collective attention emerges, persists, and manifests in large-scale synchronous sports discussions.

The rivalries share a common attention structure. Most match minutes occur in low- and moderate-attention levels, while peaks are rare, but diverge in how strongly goals drive the chat into peak attention, from largely goal-decoupled in the *Super Clássico Mineiro* to strongly goal-driven in the *Derby Paulista*. This contrast persists after controlling for chat volume, whereas apparent differences in peak-episode duration are largely a volume effect. As attention intensifies, messages also become more informal, emotional, and expressive, so that peak attention reflects not just higher volume but a qualitative shift in audience discourse. By accommodating different audience sizes, the proposed level-based representation can support similar analyses of other event-driven online discussions. Our methodology could support practical applications for broadcasters: if computed as the broadcast unfolds, attention levels could help identify high-attention windows (L3–L4) suitable for advertisements, announcements and as candidates for short clips for social media.

**Limitations.** This study has some limitations. First, as discussed in Section 3.2, defining attention levels relative to each match’s maximum activity makes the scale sensitive to a single reference

minute; a percentile-based reference could offer an alternative to define attention levels. Second, the Brazilian Portuguese LIWC dictionary [11] captures lexical prevalence but not irony, sarcasm, or context-dependent meanings common in live chats. Third, goals were aligned using the official match timeline [16], while audiences react to a stream that may carry transmission delays; our post-goal window ( $w = 3$  minutes) absorbs part of this lag, but finer synchronization remains open. Fourth, our analysis covers the entire broadcast, including pre-match buildup, half-time, and post-match, whose duration varies across channels; since attention levels are relative to each match’s peak activity, long non-gameplay segments can inflate the share of time attributed to L1, so restricting the analysis to the gameplay window is a natural refinement. Finally, our analysis covers 59 matches from three rivalries on a single platform, so the results characterize these communities rather than football fandom in general.

**Future Work.** Future work can extend this study in several directions. The analysis could be broadened to additional Brazilian rivalries and to international ones, testing whether the temporal and linguistic patterns observed here generalize to other countries and fan communities [41]. Collective attention could also be modeled in a more granular, multidimensional way, following prior work that characterizes it through dimensions such as alertness, processing capacity, and selectivity [5, 8], moving beyond message volume to capture richer aspects of audience engagement. Another promising direction is to move beyond goals and model how a broader set of match events (e.g., red cards, substitutions) and audience-driven behaviors (e.g., banter, coordinated reactions) drive transitions between attention levels. Whereas the present work relates goals to the peak attention they elicit, future analyses could examine both upward and downward transitions, disentangling what drives, sustains, and dissipates attention from game-driven stimuli versus chat-internal dynamics.

## 7 ETHICAL CONSIDERATIONS

This work complies with CNS Resolution No. 510/2016, as it consists of the collection and analysis of publicly available discussion data rather than direct contact with human subjects.

The dataset used in this study was collected from YouTube live chats of football broadcasts. Because the data were publicly accessible, individual informed consent was not required for their collection. Nonetheless, to protect personal data in compliance with the Brazilian General Data Protection Law (LGPD, Law No. 13.709/2018), author-identifying information associated with live chat messages (e.g., usernames) was anonymized prior to analysis, and no such information is disclosed in this article.

The use of generative artificial intelligence technologies in this work was limited to spell-checking and grammar review of the text and was not employed in the generation of results or the writing of any section.

## REFERENCES

- [1] Andrea Auconi, Lorenzo Federico, Gianni Riotta, and Guido Caldarelli. 2024. Fluctuations and extreme events in the public attention on Italian legislative elections. *Scientific Reports* 14, 1 (01 Oct 2024), 22804. doi:10.1038/s41598-024-69354-y
- [2] Christian Bauckhage and Kristian Kersting. 2016. Collective Attention on the Web. *Foundations and Trends® in Web Science* 5, 1-2 (2016), 1–136. doi:10.1561/18000000024
- [3] Jörg R. Bergmann and Anssi Peräkylä. 2023. *The social organization of (in-)attention*. Routledge, 299–322. doi:10.4324/9781003094111-15
- [4] Brazilian Institute of Geography and Statistics (IBGE). 2023. Between 2010 and 2022, Brazilian Population Grows 6.5%, Reaches 203.1 Million. <https://agenciadenoticias.ibge.gov.br/en/agencia-news/2184-news-agency/news/37245-between-2010-and-2022-brazilian-population-grows-6-5-reaches-203-1-million>. IBGE News Agency. Accessed: June 24, 2026.
- [5] Josemar Caetano, Humberto Marques-Neto, and Virgílio Almeida. 2026. Modeling and Analyzing Collective Attention in Online Chats. *Trans. Soc. Comput.* (July 2026). Just Accepted. doi:10.1145/3833212
- [6] Josemar Alves Caetano, Jussara Almeida, Marcos Gonçalves, Wagner Meira, Humberto T. Marques-Neto, and Virgílio Almeida. 2022. Analyzing Topic Attention in Online Small Groups. In *Proceedings of the 2021 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining* (Virtual Event, Netherlands) (ASONAM '21). Association for Computing Machinery, New York, NY, USA, 64–68. doi:10.1145/3487351.3488357
- [7] Josemar Alves Caetano, Gabriel Magno, Marcos Gonçalves, Jussara Almeida, Humberto T. Marques-Neto, and Virgílio Almeida. 2019. Characterizing Attention Cascades in WhatsApp Groups. In *Proceedings of the 10th ACM Conference on Web Science* (Boston, Massachusetts, USA) (WebSci '19). Association for Computing Machinery, New York, NY, USA, 27–36. doi:10.1145/3292522.3326018
- [8] Josemar Alves Caetano, Humberto Torres Marques-Neto, and Virgílio Almeida. 2025. Characterizing Collective Attention on Online Chats: A Three-Pronged Approach. In *Web Information Systems Engineering – WISE 2024*, Mahmoud Barhamgi, Hua Wang, and Xin Wang (Eds.). Springer Nature Singapore, Singapore, 56–66.
- [9] Jie Cai, Sagnik Chowdhury, Hongyang Zhou, and Donghee Yvette Wohn. 2023. Hate Raids on Twitch: Understanding Real-Time Human-Bot Coordinated Attacks in Live Streaming Communities. *Proc. ACM Hum.-Comput. Interact.* 7, CSCW2, Article 342 (Oct. 2023), 28 pages. doi:10.1145/3610191
- [10] Cristian Candia, C. Jara-Figueroa, Carlos Rodriguez-Sickert, Albert-László Barabási, and César A. Hidalgo. 2019. The universal decay of collective memory and attention. *Nature Human Behaviour* 3, 1 (01 Jan 2019), 82–91. doi:10.1038/s41562-018-0474-5
- [11] Flavio Carvalho, Rafael Guimarães Rodrigues, Gabriel Santos, Pedro Cruz, Lilian Ferrari, and Gustavo Paiva Guedes. 2019. Evaluating the Brazilian Portuguese version of the 2015 LIWC Lexicon with sentiment analysis in social networks. In *Proceedings of the Brazilian Workshop on Social Network Analysis and Mining* (BraSNAM). SBC, Belém, Brazil, 24–34.
- [12] Salvatore Citraro, Giovanni Mauro, and Emanuele Ferragina. 2024. Burstiness in Emotions: A Case Study on Collective Affective Responses in Italian Soccer Fandoms. In *International Conference on Advances in Social Networks Analysis and Mining*. Springer, 56–69.
- [13] Salvatore Citraro, Giovanni Mauro, and Emanuele Ferragina. 2026. Temporal dynamics of emotions in Italian online soccer fandoms. *EPJ Data Science* (2026).
- [14] Manlio De Domenico and Eduardo G. Altmann. 2020. Unraveling the Origin of Social Bursts in Collective Attention. *Scientific Reports* 10, 1 (13 Mar 2020), 4629. doi:10.1038/s41598-020-61523-z
- [15] Lukas Dreier and Johanna Pirker. 2023. Toxicity in twitch live stream chats: Towards understanding the impact of gender, size of community and game genre. In *2023 IEEE Conference on Games (CoG)*. IEEE, 1–4.
- [16] Globo Esporte. 2026. Agenda de Jogos de Futebol. Retrieved from Globo Esporte (<https://ge.globo.com/agenda/futebol>). Accessed on June 27, 2026.
- [17] Globo Esporte. 2026. Copa do Mundo FIFA 2026: tempo real, lances e resultados. Retrieved from Globo Esporte (<https://ge.globo.com/futebol/copa-do-mundo/>). Accessed on June 27, 2026.
- [18] Paulo Ferreira, Genildo Gomes, Tayana Conte, and Bruno Gadelha. 2023. A live bombou? Avaliando métricas para engajamento em live streaming. In *Anais do XVIII Simpósio Brasileiro de Sistemas Colaborativos* (Rio de Janeiro/RJ). SBC, Porto Alegre, RS, Brasil, 72–85. doi:10.5753/sbsc.2023.229085
- [19] Kiran Garimella, Gianmarco De Francisci Morales, Aristides Gionis, and Michael Mathioudakis. 2017. The Effect of Collective Attention on Controversial Debates on Social Media. In *Proceedings of the 2017 ACM on Web Science Conference* (Troy, New York, USA) (WebSci '17). Association for Computing Machinery, New York, NY, USA, 43–52. doi:10.1145/3091478.3091486
- [20] Kenia Gonçalves, Flávio Soriano, Renato Vimieiro, Humberto Marques-Neto, and Jussara Almeida. 2025. Caracterização das Recomendações de Streams na Twitch para o Público Infantil e Adolescente no Brasil e nos Estados Unidos. In *Proceedings of the 31st Brazilian Symposium on Multimedia and the Web* (Rio de Janeiro/RJ). SBC, Porto Alegre, RS, Brasil, 131–140. doi:10.5753/webmedia.2025.16053
- [21] Catherine Han, Joseph Seering, Deepak Kumar, Jeffrey T Hancock, and Zakir Durumeric. 2023. Hate raids on Twitch: Echoes of the past, new modalities, and implications for platform governance. *Proceedings of the ACM on Human-Computer Interaction* 7, CSCW1 (2023), 1–28.
- [22] Xingsheng He and Yu-Ru Lin. 2017. Measuring and monitoring collective attention during shocking events. *EPJ Data Science* 6, 1 (22 Nov 2017), 30. doi:10.1140/epjds/s13688-017-0126-4
- [23] Mark Hill. 2025. Catching Stray Balls: Football, fandom, and the impact on digital discourse. In *Proceedings of the The 9th Workshop on Online Abuse and Harms* (WOAH). 192–205.
- [24] Taizan Hokuto. 2019. pyTchat: A Python Library for Fetching YouTube Live Chat. <https://github.com/taizan-hokuto/pytchat>. Accessed: 2026.
- [25] Yukun Jiang, Xinyue Shen, Rui Wen, Zeyang Sha, Junjie Chu, Yugeng Liu, Michael Backes, and Yang Zhang. 2024. Games and beyond: analyzing the bullet chats of esports livestreaming. In *Proceedings of the International AAAI Conference on Web and Social Media*, Vol. 18. 761–773.
- [26] Gabriela Birnfeld Kurtz, Stefano de P Carraro, Carlos Roberto G Teixeira, Leonardo D Bandeira, Bernardo L Muller, Roberto Tietzmann, Milene Selbach Silveira, and Isabel H Manssour. 2025. StreamVis: An Analysis Platform for YouTube Live Chat Audience Interaction, Trends and Controversial Topics.. In *ICEIS* (2). 630–640.
- [27] Zijian Lew, Joseph B Walther, Augustine Pang, and Wonsun Shin. 2018. Interactivity in Online Chat: Conversational Contingency and Response Latency in Computer-mediated Communication. *Journal of Computer-Mediated Communication* 23, 4 (06 2018), 201–221. arXiv:https://academic.oup.com/jcmc/article-pdf/23/4/201/25113924/zmy009.pdf doi:10.1093/jcmc/zmy009
- [28] João Lobo, Daniel Barbosa, and Philipe Melo. 2025. Investigando a Dinâmica da Propagação do Ódio em Cascatas de Toxicidade nos Chats Ao Vivo da Twitch. In *Proceedings of the 31st Brazilian Symposium on Multimedia and the Web* (Rio de Janeiro/RJ). SBC, Porto Alegre, RS, Brasil, 321–329. doi:10.5753/webmedia.2025.15110
- [29] Andre Lopes. 2026. Mesmo com delay, CazêTV bate recorde mundial de audiência no YouTube no jogo do Brasil. Retrieved from Exame (<https://exame.com/esporte/mesmo-com-delay-cazetv-bate-recorde-mundial-de-audiencia-no-youtube-no-jogo-do-brasil/>). Accessed on June 20, 2026.
- [30] Philipp Lorenz-Spreen, Bjarke Mørch Mønsted, Philipp Høvel, and Sune Lehmann. 2019. Accelerating dynamics of collective attention. *Nature Communications* 10, 1 (15 Apr 2019), 1759. doi:10.1038/s41467-019-09311-w
- [31] Mufan Luo, Tiffany W. Hsu, Joon Sung Park, and Jeffrey T. Hancock. 2020. Emotional Amplification During Live-Streaming: Evidence from Comments During and After News Events. *Proc. ACM Hum.-Comput. Interact.* 4, CSCW1, Article 47 (may 2020), 19 pages. doi:10.1145/3392853
- [32] S. Lydersen, V. Pradhan, P. Senchaudhuri, and P. Laake. 2007. Choice of test for association in small sample unordered  $r \times c$  tables. *Statistics in Medicine* 26, 23 (Feb. 2007), 4328–4343. doi:10.1002/sim.2839
- [33] Kate Mannell. 2020. Plural and Porous: Reconceptualizing the Boundaries of Mobile Messaging Group Chats. *Journal of Computer-Mediated Communication* 25, 4 (July 2020), 274–290. doi:10.1093/jcmc/zmaa007
- [34] Muhammad Zeeshan Mazhar, Tolga Buz, and Yiran Su. 2025. Are Online Sports Fan Communities Becoming More Offensive? A Quantitative Review of Topics, Trends, and Toxicity of r/PremierLeague. *arXiv preprint arXiv:2510.27003* (2025).

- [35] Delia Mocanu, Luca Rossi, Qian Zhang, Marton Karsai, and Walter Quattrociocchi. 2015. Collective attention in the age of (mis)information. *Computers in Human Behavior* 51 (2015), 1198–1204. Computing for Human Learning, Behaviour and Collaboration in the Social and Mobile Networks Era. doi:10.1016/j.chb.2015.01.024
- [36] Jihyung Moon, Dong-Ho Lee, Hyundong Cho, Woojeong Jin, Chan Park, Minwoo Kim, Jonathan May, Jay Pujara, and Sungjoon Park. 2023. Analyzing norm violations in live-stream chat. In *Proceedings of the 2023 Conference on Empirical Methods in Natural Language Processing*. 852–868.
- [37] Helen Paffard and Diogo Pacheco. 2025. Kicking Politics: How Football Fan Communities Became Arenas for Political Influence. In *International Conference on Complex Networks and Their Applications*. Springer, 182–194.
- [38] Maria J. Palazzi, Albert Solé-Ribalta, Violeta Calleja-Solanas, Sandro Meloni, Carlos A. Plata, Samir Suweis, and Javier Borge-Holthoefer. 2021. An ecological approach to structural flexibility in online communication systems. *Nature Communications* 12, 1 (29 Mar 2021), 1941. doi:10.1038/s41467-021-22184-2
- [39] Karine Pires and Gwendal Simon. 2015. YouTube live and Twitch: a tour of user-generated live streaming systems. In *Proceedings of the 6th ACM Multimedia Systems Conference (MMSys '15)*. ACM. doi:10.1145/2713168.2713195
- [40] Xiaoyan Qiu, Diego F. M. Oliveira, Alireza Sahami Shirazi, Alessandro Flammini, and Filippo Menczer. 2017. Limited individual attention and online virality of low-quality information. *Nature Human Behaviour* 1, 7 (26 Jun 2017), 0132. doi:10.1038/s41562-017-0132
- [41] Redação. 2024. Quais são os principais clássicos brasileiros? Veja lista. Retrieved from Globo Esporte (<https://ge.globo.com/futebol/lista/2024/04/09/c-qua-is-sao-os-principais-classicos-brasileiros-veja-lista.ghtml>). Accessed on June 27, 2026.
- [42] Mohammad Hammas Saeed, Kostantinos Papadamou, Jeremy Blackburn, Emiliano De Cristofaro, and Gianluca Stringhini. 2024. TUBERAIDER: Attributing coordinated hate attacks on YouTube videos to their source communities. In *Proceedings of the International AAAI Conference on Web and Social Media*, Vol. 18. 1354–1366.
- [43] Kazutoshi Sasahara, Yoshito Hirata, Masashi Toyoda, Masaru Kitsuregawa, and Kazuyuki Aihara. 2013. Quantifying Collective Attention from Tweet Stream. *PLOS ONE* 8, 4 (04 2013), 1–10. doi:10.1371/journal.pone.0061823
- [44] Garriy Shteynberg. 2018. A collective perspective: shared attention and the mind. *Current Opinion in Psychology* 23 (Oct. 2018), 93–97. doi:10.1016/j.copsyc.2017.12.007
- [45] Tiago Ventura, Kevin Munger, Katherine McCabe, and Keng-Chi Chang. 2021. Connective Effervescence and Streaming Chat During Political Debates. *Journal of Quantitative Description: Digital Media* 1 (Apr. 2021), 1. doi:10.51685/jqd.2021.001
- [46] Joseph B. Walther. 2017. The Merger of Mass and Interpersonal Communication via New Media: Integrating Metaconstructs. *Human Communication Research* 43, 4 (June 2017), 559–572. doi:10.1111/hcre.12122
- [47] Yucheng Wang and Zhicong Lu. 2023. Making sense of post-match fan behaviors in the online football communities. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*. 1–17.
- [48] Yiluo Wei, Jiahui He, and Gareth Tyson. 2025. Understanding Fanchuan in Livestreaming Platforms: A New Form of Online Antisocial Behavior. *Proceedings of the ACM on Human-Computer Interaction* 9, 7 (2025), 1–31.
- [49] Fang Wu and Bernardo A. Huberman. 2007. Novelty and collective attention. *Proceedings of the National Academy of Sciences* 104, 45 (2007), 17599–17601. arXiv:<https://www.pnas.org/doi/pdf/10.1073/pnas.0704916104> doi:10.1073/pnas.0704916104
- [50] Siqi Wu. 2022. Measuring Collective Attention in Online Content. *SIGWEB NewsL*. 2022, Summer, Article 4 (jul 2022), 4 pages. doi:10.1145/3545196.3545200