PostChat: Supporting Reflective Practice and Well-Being Among Live Streamers Through AI-Mediated Post-Broadcast Analytics

Keri Mallari(✉), Michelle H. Nguyen, and Mark Zachry

Department of Human Centered Design and Engineering, University of Washington, Seattle, WA 98105, USA
{kmallari,mnguyenh,zachry}@uw.edu

Abstract. Live streaming platforms require streamers to interpret constant chat flows while managing complex community dynamics. This paper introduces PostChat, an AI-mediated reflection system with dual-granularity analytics: message-level tools for detecting ambiguous negativity and supporting retroactive moderation, and persona-based tools for understanding audience segments. Through usability testing with 11 Twitch streamers, we find that message-level analytics effectively reduce emotional labor and inform content planning, while persona-based analytics receive mixed reactions regarding actionability. Our findings reveal that effective post-stream tools must balance strategic insights with affective support, validating streamers' relational work while providing actionable guidance. We contribute design implications for creator support systems that address both community understanding and emotional well-being.

Keywords: Live Streaming · Social Media · AI-Mediated · Human Computer Interaction

1 Introduction

Live streaming has become a rapidly growing medium for content creation, characterized by real-time, interactive engagement between streamers and their audiences on platforms like Twitch and YouTube Live. However, these platforms introduce distinct sociotechnical challenges: streamers must continuously interpret fast-moving viewer chat messages, some supportive, others critical, many ambiguous [23,34], while maintaining on-camera performance, leaving little cognitive capacity for deep reflection on audience dynamics or message intent during broadcasts [24]. The effects of these interactions linger beyond the live moment, shaping streamers' emotional well-being, understanding of their community, and decisions about future content [10,27,35].

Prior work highlights the considerable emotional labor involved in streaming, particularly around managing community dynamics, processing viewer feedback,

A. Coman and S. Vasilache (Eds.): HCII 2026, LNCS 16716, pp. 67–84, 2026.
https://doi.org/10.1007/978-3-032-31418-5_5

and sustaining parasocial relationships at scale [10,27,35]. Existing platform tools, primarily focused on real-time moderation, flag or remove overtly harmful content but often fail to account for context-dependent language (sarcasm, banter, in-group slang) common in streaming cultures. This leaves streamers to manually review hours of chat logs post-stream or forgo review entirely due to time constraints and emotional taxation. Moreover, standard analytics emphasize surface-level engagement metrics like view counts or chat message frequency while offering little insight into who participated, what they valued, or how different audience segments responded [8,22]. These limitations make it difficult for streamers to understand their communities at a meaningful depth, plan content strategically, or process emotionally charged interactions in healthy ways.

This paper presents PostChat, a system designed to support streamers' post-stream reflection through two complementary approaches operating at different granularities: (1) message-level analysis that helps streamers identify and contextualize negative or ambiguous comments they might otherwise miss or misinterpret, and (2) persona-based audience analytics that cluster viewers into composite profiles to surface patterns in community engagement and viewer motivations. By exploring both conceptual and functional prototypes with the same participants, we investigate how reflection tools operating at different scales, from individual message interpretation to audience-wide pattern recognition, affect streamers' emotional labor, content planning practices, and understanding of their communities.

Through qualitative usability testing with active streamers, we examine three interconnected research questions:

- **RQ1**: What types of post-stream insights do streamers find most valuable for content planning and audience relationship building?
- **RQ2**: How do automated analysis tools affect streamers' emotional labor related to chat review and negativity management?
- **RQ3**: How does granularity of reflection (message-level vs. audience-level analytics) shape streamers' sense-making practices?

Our findings reveal that effective post-stream reflection systems must address both strategic and affective needs: streamers valued tools that not only provided actionable insights for content decisions but also validated their relational work and helped them feel more connected to their communities. Automated analysis was appreciated not for eliminating engagement with difficult content but for making comprehensive review feasible within the time and emotional constraints streamers face. Crucially, message-level and audience-level tools served complementary rather than competing functions, the former addressing interpretive ambiguity and negativity detection, the latter enabling pattern recognition and strategic planning that individual message review cannot surface.

We contribute a layered understanding of how post-stream reflection can be supported through sociotechnical design, with implications for developing tools that mitigate emotional burdens while enriching community understanding in live content creation. We begin with literature on live streaming, emotional labor, and feedback exchange, describe the design and development of PostChat's dual systems, present our usability testing methodology and thematic analyses, and conclude with findings and discussion of key insights for future streaming support tools.

1.1 Audience Analytics and Viewer Understanding

Recent scholarship highlights the importance of audience analytics in shaping creators' strategies and self-perceptions. Studies show that streamers use real-time metrics and chat interactions to gauge viewer preferences, adapt content, and foster community engagement [22,31]. However, analytics often provide only a partial picture, leading to potential misinterpretations of audience intent and satisfaction.

Content creators actively use analytics to adapt their content strategies in response to audience behaviors and platform algorithms. Creators often experience tension between algorithmic demands and genuine audience engagement, regularity versus audience autonomy, and analytics versus decision-making autonomy [12]. Prior work has shown that these are some strategies that enable creative control over their content while also optimizing for platform and algorithmic distribution.

One popular approach has been sentiment analysis for content adaptation. Sentiment analysis has allowed creators to gauge audience reactions to their content and inform content optimization strategies, which can be used for tailoring future content to audience preference [26]. This broadly applies to content creators in general, such as YouTube videos and social media posts. For example, platform analytics tools like YouTube Studio provide metrics on view counts, watch time, and demographic data to support content decisions; similarly, Twitch offers these same metrics and a suggestions page to help streamers with scheduling and recommended categories for future planning. However, these analytics have shown significant limitations for live streamers, leading many creators to turn to external analytics tools. [19].

For live streamers, engagement metrics are often referred to as the currency of success for those hoping to grow and resonate with their community [5]. The technical features of social media platforms also facilitate the formation of dynamic and networked communities, not just within streamer-viewer relationships but also between streamers and between viewers. Many content creators often try to understand how the recommendation algorithms work so they can "work the system" to potentially increase their visibility and profit [7,12,16,20,21].

While analytics offer valuable insights, they also introduce new forms of labor and emotional investment for creators, as explored in the following section.

1.2 Labor in Live Streaming

Live streaming blurs the boundaries between work and play, demanding both emotional and physical labor from creators. Streamers face constant pressure to perform, manage their communities, and maintain an authentic presence, which can often lead to burnout or emotional exhaustion [27,30].

Live streaming requires creators to engage in intensive emotional labor beyond just appearing on camera. Streamers often have to be constantly "on," friendly, or witty, while also staying in character for hours at a time [25,35]. This labor manifests through various performances required of creators, such as being friendly to viewers, soliciting donations and subscriptions, and more [2,35]. Emotional labor is critical as streamers need to express the right emotions to foster viewers' active attitudes and behaviors in communication and transactions [37]. This requires streamers to display organizationally desired emotions while also actively suppressing others that might negatively affect current, live, and ongoing content [37].

There are also gendered dimensions to emotional labor. For example, female streamers are often expected and labeled to be "nicer" and to intersect with societal expectations of women as attentive and caring [27]. However, this becomes intertwined with sexualized expectations, where viewers expect female streamers to entertain with sexual innuendos and fantasies to maintain streamer-viewer relationships and encourage gift giving, often hiding their personal relationships behind the camera [27].

Prior research has identified and highlighted hidden labor as a key component of streaming work, including the day-to-day maintenance for potential earnings, responding to collaborations and emails, checking statistics, planning future content, community management, and more to ensure that each stream is well run. These tasks vary among different types of viewers and streamers. For those who are successful, streamers spend many hours adhering to a regular streaming schedule, properly configuring technical aspects, and promoting streams across social media [22,35].

Labor dynamics are further complicated by the ambiguous and interpretive nature of online interactions, which can heighten uncertainty and stress for both creators and viewers.

1.3 Limitations of Existing Moderation and Feedback Tools

Although platforms provide various moderation and feedback tools, research indicates that these systems are often inadequate for addressing the nuanced and evolving challenges of live streaming. Automated moderation can miss context, while manual moderation places additional strain on creators and their communities [3,18,29]. Feedback mechanisms may fail to capture the full spectrum of viewer sentiment, limiting creators' ability to respond adaptively.

Recent advances in large language models (LLMs) and multimodal AI have enabled interactive, real-time feedback systems at scale. Proxona, for example, leverages static comments to build structured audience personas, allowing creators to simulate conversations with diverse viewer segments and use feedback for ideation and content refinement [6]. SimTube extends this by generating predictive audience responses before a video's release, combining video data with persona modeling to offer creators context-rich, demographically varied feedback scenarios [13]. These tools frame feedback as a creative asset rather than a reactive metric.

While such systems highlight new possibilities for creator support, they also underscore a critical gap in our understanding of how creators navigate feedback, labor, and moderation in practice. While these systems offer creators simulated perspectives and audience insight, they primarily operate in a static or anticipatory context.

2 PostChat Design and Development

In this section, we describe the details of the iterative design and development of two complementary designs of PostChat, aimed at supporting after-stream reflection for live streamers: a chat message-level analysis and an audience-level persona analysis dashboard. Our approach is grounded in user-centered design principles, developing solutions iteratively and in cooperation with their target users [11].

2.1 Chat Message-Level Reflection and Reframing

Existing moderation tools primarily address surface-level disruptions in real-time, but rarely facilitate the deeper interpretive work that streamers undertake after the broadcast. Our design investigates whether streamers value structured message-level reflection, how ambiguous negativity affects them, and what features are most beneficial.

Drawing on the framework from "Navigating Ambiguous Negativity: A Case Study of Twitch.tv Live Chats," we operationalized three tiers of harmful chat behavior: surface-level disruptions (e.g., spam, slurs), relational/dialogic behaviors (e.g., trolling, banter), and meta-discourse/ambiguous negativity (e.g., subtle social tensions or controversial topics) [23]. Representative chat messages were generated across these layers, informed by this typology.

In the diagram below (Fig. 1), we present a screenshot of the prototype for chat negativity with the following components: On the left side, we start with the metadata of the stream including category, stream date, duration, and number of messages to allow users to easily contextualize which stream is being represented. Below, we show the frequency of chat messages in a bar chart segmented in 5-minute intervals, highlighting segments with colors relating to themes that showed up in the analysis. Finally, there is a scrollable chat transcript that provides a line-by-line chat message history, with respective messages highlighted with the theme color. On the top right, a pop-up tool tip provides information about this page. Below, are the list of themes as collapsible/expandable sections to get a description of the highlighted negativity themes, some findings related to the chat, and associated recommendations. This is repeated for the other themes found in this chat history.

For this part of the prototype, we simply designed a static HTML page utilizing sample data. Given that different streamers may experience varying degrees of "ambiguous negativity" from little to none, we presented them with this synthetic page to get their feedback across these three themes.

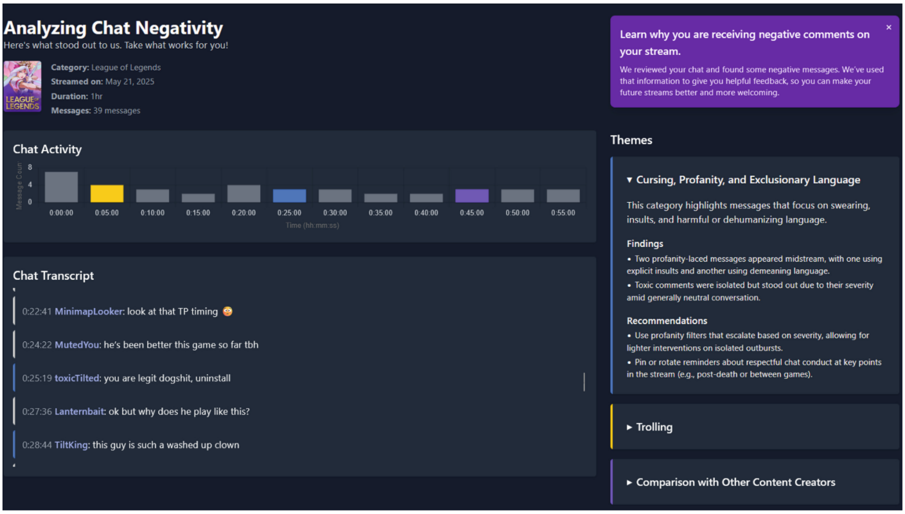

Fig. 1. Screenshot of Prototype Analyzing Chat Negativity

2.2 Audience-Level Reflection with Persona-Based Analytics

While Twitch offers a real-time chat for a continuous stream of audience feedback, this data is often noisy, overwhelming, and difficult to interpret. These platforms also provide basic quantitative metrics (i.e., viewer count, chat messages), but they offer little support for understanding composition, communication styles, and motivations of the audience. Prior research has emphasized the emotional and attentional demands of sustaining live engagement, yet tools for post-hoc reflection on audience behavior and community dynamics remain limited.

Building on systems like Proxona [6] and SimTube [13], that transform raw feedback into persona-based insights, we also explore how persona-based analytics can support deeper reflection for streamers. By abstracting chat into composite audience profiles, this approach surfaces patterns in tone, recurring themes, and styles of interaction.

A Chain-of-Thought (CoT) prompt guides the model through interpretable steps: identifying patterns, assigning labels, selecting representative quotes, and describing behavioral themes [33]. We adopted a grounded typology-based coding framework, following categories from Schuck [28], to classify each chat message into one of three recurring viewer personas:

– System Alterers (SA): Messages attempting to influence or direct the streamer (e.g., advice, critique).
– Financial Sponsors (FS): Messages referencing financial support mechanisms (e.g., subscriptions, donations).
– Social Players (SP): Messages engaging in playful, interactive, or socially bonding behaviors (e.g., emote spam, memes).

All chat messages were included via forced classification to ensure comprehensive representation.

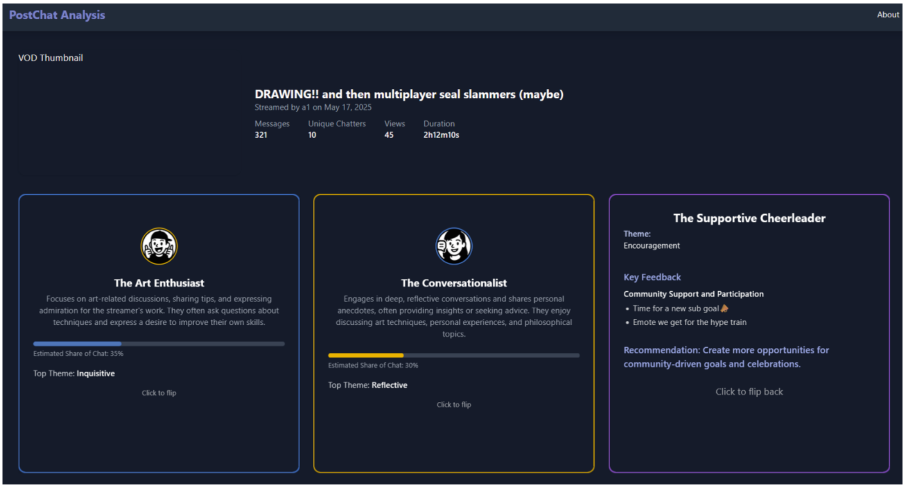

Fig. 2. Viewer Persona Prototype Tailored to Streamers Using Twitch VODs

To support audience-level reflection, we developed a prompt-based inference pipeline leveraging GPT-4o via the OpenAI API [14]. Chat logs were segmented into batches of 200 messages, with each batch analyzed independently to maintain granularity and scalability. For each batch, we employed structured prompting to guide the model in inferring approximately three distinct viewer personas. Each persona profile (Fig. 2) included a concise descriptive label, a summary description, an estimated proportion of messages, and dominant communicative theme (such as encouragement or critique). When users click the card, the profile also includes representative quotes pulled from chat accompanied by actionable feedback for streamers.

To ensure interpretability and reduce redundancy, similar persona profiles were aggregated and merged using Levenshtein-based fuzzy matching on both name similarity and behavioral traits [17]. This process allowed us to surface the most salient and distinct personas for each stream, supporting nuanced audience insights. Importantly, all model prompts included strict constraints to minimize hallucination and speculation, ensuring that all generated insights were directly traceable to the underlying chat data [36].

From a technical perspective, the system was configured with a temperature setting of 0.6 and a maximum token limit of 3000 per prompt. Outputs from the language model were validated against a predefined JSON schema to ensure structural consistency and facilitate downstream processing. Finally, the resulting personas were visualized in an interactive dashboard, with each persona represented as a card displaying sentiment bars on the front and qualitative summaries on the back. This design enabled streamers to quickly assess the emotional landscape and communicative dynamics of their audience, supporting reflective practice and informed community management.

Both prototypes were developed iteratively, beginning with low-fidelity wireframes created via Figma and advancing to interactive prototypes using Python and Flask. This process enabled rapid exploration of the design space, early user feedback, and refinement of core interaction flows [32].

3 Methods

This study employed remote usability testing sessions to investigate how streamers reflect on and manage their post-stream activities, audience dynamics, and interactions with chat data. We used personalized audience persona pages and a message-level prototype to evaluate participants' perceptions of these tools and gather feedback on their potential utility (Table 1).

Table 1. Participant Demographics. Note: Avg. Viewers and Hours Streamed are for the past 30 days

PID	Location	Followers	Status	Avg. Viewers	Hours Streamed
p1	USA	109	Affiliate	2	11
p2	USA	1580	Affiliate	8	48
p3	Philippines	230	Affiliate	8	82
p4	USA	3569	Affiliate	3	25
p5	USA	5205	Affiliate	12	28
p6	India	7576	Affiliate	56	122
p7	USA	1685	Affiliate	17	217
p8	India	128	Affiliate	5	25
p9	Australia	3133	Affiliate	16	244
p10	Indonesia	7633	Partner	46	51
p11	USA	143	Affiliate	7	19

3.1 Participants

We had recruited 11 participants, recruited through a combination of personal networks, public Discord servers, and relevant organizations. All recruitment activities were conducted with appropriate permissions, including moderator approval when necessary. Eligibility criteria required participants to have at least 2 publicly available recent VODs (video on demand) for generating their personalized audience personas, and have streamed recently. The streamers represented a range of content types including gaming, art, and just chatting, with followers ranging from 109 to 7633. Participants were recruited globally with six from USA, four from Asia, and one from Australia. We also required our participants to be at least Twitch Affiliates, the minimum level at which creators can earn revenue on the platform, as a signal that they were serious about streaming [1]. Participants received a $20 gift card upon completion of the study. This study was approved by the Institutional Review Board (IRB) at the authors' university.

3.2 Interviews

The usability testing sessions included a semi-structured interview. Prior to the study sessions, participants were asked to share links to two of their previous and recent Twitch VODs (videos on demand), enabling the research team to generate personalized audience personas ahead of each interview. Participants were first asked about their typical post-stream workflows and the ways in which they currently analyze their stream content. Next, we introduced the personalized audience-level personas generated from their own Twitch data and invited participants to reflect on the personas' effectiveness in helping them understand their audiences. In the final stage of the session, we presented a prototype that visualized individual chat messages using synthetic data, focusing on potential applications for identifying negativity and supporting content analysis. Participants were asked about the prototype's utility, design, and potential extensions.

3.3 Data Analysis

All interviews were recorded and transcribed using Otter.ai. The first author manually reviewed each transcript to correct typographical errors and address any transcription inaccuracies. Transcripts were then shared with the research team for qualitative analysis and coding via Google Sheets and Miro.

We adopted a grounded theory approach to analyze the interview data [4]. The coding process began with a collaborative session on two transcripts, during which the team identified initial themes and drafted a preliminary codebook. In subsequent meetings, team members reconciled coding discrepancies and iteratively refined the codebook to incorporate emerging concepts. This process continued until all transcripts were fully coded. To assess coding consistency, we calculated inter-rater reliability using Cohen's Kappa, achieving a score of 0.83.

These thematic categories structure the presentation of our findings in the following section.

4 Findings

4.1 Post-Stream Insights for Content Planning and Audience Relationships

Participants identified persona-based analytics and contextual recommendations as the most valuable insights for shaping future content direction and deepening audience understanding. P1 explained their motivation: "I'm mostly curious about the new viewers, what their preferences are, and how to keep them engaged without changing my entire personality." P1 articulated the desire for authentic engagement:"I'd rather have someone come to my stream and want to stay here whether they like my conversations or the content I create."

This value stems from addressing a fundamental gap in existing analytics: while platforms provide surface-level metrics like view counts, streamers need to understand why viewers engage and who their audience segments are. Beyond strategic planning, participants valued insights that validated their relational work. P7 summarized: "The recommendations are nice, and it's really good to see that kind of thing."

The actionable nature of recommendations, particularly when linked to specific chat quotes, made them especially valuable. P4 reflected: "It shows exactly what they're saying, like, 'Oh, you didn't press R.' Okay, fine. I'll press R next time." P6 appreciated theme-specific recommendations to create a dedicated segment of sharing tips and strategy to enhance [the] player segment, particularly in games like *Genshin Impact*. However, P5 suggested the tool would be more helpful if it was providing more data, such as emotional or gameplay context when negative feedback occurred.

4.2 Automated Analysis and Emotional Labor

PostChat's automated analysis capabilities were framed by participants as tools for mitigating the emotional burden of comprehensive chat review, particularly around identifying and processing negative content. P1 reflected: "If I want to do any reflections, I do it, like, the next day I try to, like, see what's going on and how I can improve." P8 captured post-stream exhaustion: "I just die in bed [after stream] because I'm so tired. "

One major application was retroactive moderation, giving streamers time to make thoughtful decisions. P1 described evaluating user behavior: "This person is a regular and they're only saying this to make fun of me not in bad faith," but added, "If it's a new person and they're suddenly saying that stuff, of course, I'm like, 'Get out of here."' P1 also recalled: "Why are you telling me about all your dates? Okay, I'm gonna ban you," illustrating the emotional burden streamers face when establishing boundaries to preserve their emotional wellbeing.

For larger channels, tools that filter messages and identify hate speech were essential. P7 noted, "this would be more useful for someone who has a huge viewership so they can have the analysis even post stream." P10 reflected on using analytics to flag homophobic messages: "I think maybe I can use it to report it to Twitch." P11 emphasized that LGBTQ+ streamers may face disproportionate hate, making robust detection tools especially critical. Across these contexts, the ability to systematically identify harmful messages serves both the operational needs of high-volume channels and the safety needs of marginalized streamers facing targeted harassment.

Participants appreciated how the platform offered contextual cues that existing bots might miss. P4 commented, "it shows a lot more sections and what specifically is being said," highlighting the tool's depth. P3 said, "I didn't know what to do with the dude who's commenting on the streamer's appearance," noting that recommendations could provide actionable steps. P3 described their role as a moderator when they "try to catch the message before [streamer] read it so it doesn't ruin their mood while they're streaming."

However, several participants expressed concerns about how negative feedback was presented. P6 felt the platform initially was a page of everything that they hate about you. P1 requested "I would love to see this for the positive comments." P11 reflected that banning can feel like an extreme move, even when justified. These concerns reveal a tension in moderation analytics: while identifying harmful content is necessary for streamer safety, an exclusive focus on negativity risks creating an emotionally taxing experience that may discourage engagement with protective tools.

4.3 Granularity of Reflection and Sense-Making

The distinction between message-level and audience-level analysis tools served different but complementary sense-making functions. Message-level features addressed immediate interpretive challenges. P5 noted language barriers "English is like a second or third language for a lot of them." P3 reflected "I don't remember all my conversations."

A key use was identifying peak moments for clips and highlights. P4 liked to "see where [they] make them talk to see what hypes someone." P2 emphasized that they are "always trying to get more clips [from my stream]. It is like a major thing that [they] do." P1 noted timestamps acted as markers "even if you cannot get the context of the video itself you'll be like, 'Oh, that was this topic."' P1 also expressed a way to recall community members: "Oh yeah, that conversation did happen. Oh yeah, this person did say that." Collectively, these uses highlight how PostChat helps streamers overcome the cognitive challenge of remembering specific moments from lengthy streams, surfacing content that would otherwise be lost to the ephemeral nature of live broadcasting.

In contrast, audience-level persona analytics supported higher-order pattern recognition. P1 envisioned using persona histories: "If this user has, like, a history of being like an art enthusiast or a supportive person, then I won't ban them or something like that." P1 explained interest in tracking shifts: "When did this persona suddenly become different kinds of personas? Which sections of the video were mostly this?"

P4 appreciated quantitative summaries: "That means I did something good during that moment when learning subscriber engagement at specific times." P10 highlighted surfacing missed moments: "A very heartfelt message because sometimes, honestly, I miss chat or totally forget about them." However, participants could recall specific instances but not their entirety, indicating that audience-level analytics fill a memory gap that message-level review alone cannot address. The complementarity suggests post-stream reflection tools should be designed as integrated systems supporting both interpretive depth and synthetic breadth.

5 Discussion

Our findings reveal that effective post-stream reflection systems must operate simultaneously at multiple levels of granularity, each addressing distinct but interdependent challenges in streamers' sense-making practices. While prior work has documented the emotional labor of live streaming [27,35] and the limitations of platform analytics [8,22], our study demonstrates that the granularity at which reflection tools operate fundamentally shapes both their utility and their effect on streamers' emotional and cognitive work.

5.1 Complementary Granularities Challenge Single-Level Analytics

The complementarity between message-level and audience-level tools challenges prevailing assumptions in platform design about providing uniform analytics solutions. Current streaming platforms predominantly offer either real-time moderation tools focused on individual message filtering [29] or aggregate metrics emphasizing surface-level engagement statistics [22]. Our findings contest this assumption: participants valued message-level analytics specifically for addressing interpretive ambiguity and retroactive moderation, while audience-level personas served entirely different functions related to pattern recognition and strategic planning. This division maps onto a fundamental tension in live streaming work: the cognitive impossibility of comprehensive real-time sense-making during broadcasts versus the need for strategic understanding that emerges only through synthesis. Streamers retain fragments of interactions but lack holistic understanding, they remember individual topics but not complete context, struggle to recall all conversations, and find that individual messages leave vivid impressions without systematic structure. Conversely, aggregate metrics tell streamers what happened without explaining why viewers engaged or who comprises their community. PostChat's dual-granularity approach addresses both gaps: message-level tools recover the interpretive work impossible during live performance, helping streamers identify peak moments and understand engagement drivers, while persona-based analytics provide the synthetic overview that isolated message review cannot generate, revealing audience segment shifts and content resonance patterns. Rather than pursuing increasingly granular data capture or more sophisticated aggregation independently, our findings suggest reflection tools should be designed as integrated ecosystems where different granularities inform one another.

5.2 Redistributing Rather Than Eliminating Emotional Labor

PostChat's automated analysis capabilities were consistently framed by participants as tools for redistributing emotional labor rather than simply reducing it, a crucial distinction with implications for how we design creator support systems. Prior research has documented the substantial emotional work required to maintain on-camera performance while processing viewer feedback [25,27], leading to assumptions that automation should minimize creators' engagement with challenging content. However, our participants valued PostChat not for avoiding negative content but for making engagement with it more manageable and comprehensive. This redistribution manifests in three specific ways. First, automation enables comprehensive rather than selective review: where manual chat analysis might focus only on obviously problematic messages, automated flagging surfaces ambiguous negativity and behavioral patterns that streamers might otherwise overlook, such as repeated low-level harassment across multiple messages. Second, structured presentation reduces the cognitive load of sense-making itself, offering deeper analysis than traditional keyword filters by showing complete sections and contextual information. This is particularly valuable given post-stream exhaustion, where streamers must balance reflection with self-care and often delay analysis until the next day when they have emotional capacity to engage thoughtfully. Third, temporal separation allows for emotional distance and context-aware decision-making. Streamers described making more nuanced evaluations when reviewing chat post-stream, distinguishing between regular viewers engaging in playful banter versus new users making inappropriate comments, and considering user histories to inform fair moderation decisions. These findings complicate narratives of automation as straightforward labor-saving technology in creative work [9]: PostChat transforms work from scattered real-time processing to structured reflection, from emotionally reactive to analytically supported interpretation, potentially increasing the time streamers spend engaging with chat data while decreasing its emotional taxation.

5.3 The Affective Dimensions of Analytical Tools

Perhaps most surprising was the extent to which participants valued PostChat for affective rather than purely analytical functions. Streamers described the prototype's findings as validating their efforts, framing audience insights not merely as strategic intelligence but as recognition of their relational work. This affective valuation reveals how professional creative tools operate simultaneously in technical and emotional registers, particularly in fields characterized by parasocial relationships and platform precarity [15,35]. The importance of validation connects to broader literature on parasocial dynamics in streaming communities. Streamers invest substantial emotional labor in cultivating relationships with viewers [27], yet this investment occurs under conditions of uncertainty. They desire viewers who stay for authentic conversations and content rather than performative personas, yet struggle to understand new viewer preferences without compromising their identity. Platform analytics that focus exclusively on growth metrics fail to validate this relational work because they measure outcomes rather than relationships. In contrast, PostChat's persona-based analytics operationalize what streamers already intuitively know about their communities, surfacing emotional moments they missed during live broadcasts and confirming their interpretive competence while making it actionable for content planning. However, the affective impact can be negative if feedback is poorly framed. Participants expressed concern when the platform felt like a catalog of criticisms, requesting more emphasis on positive comments and validation. This tension highlights a critical design challenge: tools must balance transparency about community dynamics with emotional support for creators. The affective value of audience analytics illuminates how knowing your audience constitutes both strategic and emotional work in content creation, enabling tailored content decisions while providing the connection and motivation necessary for sustained creative labor under precarious conditions.

5.4 Design Implications

Based on our findings, we propose four key design implications for post-stream analytics tools:

Integrate Multi-granularity Analytics as Complementary Systems. Tools should seamlessly connect message-level and audience-level views, allowing streamers to drill down from persona patterns to specific chat moments and vice versa. This addresses complementary needs: message-level for interpretive depth and retroactive moderation, audience-level for strategic planning and pattern recognition. Rather than offering these as separate features, design them as an integrated ecosystem where insights at one level inform exploration at another.

Balance Negative and Positive Feedback Through Configurable Framing. Analytics interfaces must thoughtfully frame critique alongside affirmation. Unbalanced negative feedback risks discouragement, with participants describing overwhelming experiences when confronted primarily with criticism. Design configurable modes that allow streamers to prioritize positive insights, emotional highlights, or critical feedback based on their current emotional capacity. Tools should recognize that effective analytics serve both strategic and affective functions.

Provide Contextual, Actionable Recommendations Grounded in Evidence. Recommendations should be explicitly linked to specific chat evidence and stream context. Participants valued concrete guidance showing exact viewer comments and how to respond, rather than generic suggestions. Move beyond abstract advice to concrete guidance tied to actual viewer behavior and community patterns, particularly for newer streamers seeking support in handling challenging situations like inappropriate comments.

Support Temporal Moderation with Behavioral Context. Enable retroactive moderation with full user history and behavioral patterns. Streamers described making more nuanced evaluations when they could distinguish between regular viewers engaging in playful banter versus new users making inappropriate comments, and when they could detect patterns of repeated low-level harassment rather than isolated incidents. Tools should surface patterns that inform graduated responses rather than forcing immediate binary decisions, reducing the emotional burden of real-time judgment calls where banning can feel like an extreme action.

These implications highlight that post-stream analytics must address both the strategic work of content planning and community management, and the affective work of maintaining emotional well-being and connection with audiences. Tools that attend to only one dimension risk either providing emotionally unsustainable insights or offering strategically hollow validation.

6 Future Work and Limitations

Several limitations should be acknowledged. First, our participant sample (N = 11 Twitch Affiliates, 109-7,633 followers) may not generalize to larger professional streamers, hobbyist creators, or non-English-speaking communities. Second, findings are based on prototype reactions rather than longitudinal deployment data, reflecting initial impressions rather than sustained workflow integration. Third, we evaluated message-level and persona-level analytics separately rather than as an integrated system. Finally, ethical considerations surrounding automated persona modeling, algorithmic bias in negativity detection, and privacy implications of chat data analysis warrant deeper investigation.

Future research should pursue longitudinal deployment studies to reveal how sustained use shapes content strategies, community health, and streamer well-being over time. Comparative studies across platform scales would illuminate how analytics needs differ by channel size. Cross-platform analytics could provide holistic understanding of multi-platform communities. Expanding scope to include moderators and community members could foster collaborative community management. Finally, ethical frameworks for creator analytics, addressing consent, algorithmic accountability, and data governance, demand interdisciplinary attention.

7 Conclusion

This study demonstrates the value of dual-granularity analytics for supporting post-stream reflection among live content creators. Through qualitative usability testing with 11 active Twitch streamers, we found that message-level and audience-level analytics serve complementary functions: message-level tools enable retroactive moderation and interpretive depth, while audience-level personas support pattern recognition and strategic planning.

Our findings reveal that effective analytics tools must address both strategic and affective dimensions of creator work. Streamers valued PostChat not only for actionable insights but also for validation of their relational labor. Crucially, automated analysis was appreciated not for eliminating engagement with difficult content but for redistributing emotional labor, transforming scattered real-time processing into structured post-stream reflection with contextual support.

We contribute empirical insights into how different analytical granularities support distinct aspects of streamer sense-making, design implications for AI-mediated creator support tools, and a working prototype demonstrating integrated analytics. By designing tools that respect both the strategic and emotional dimensions of creative labor, we can better support streamers in building sustainable, healthy relationships with their audiences while managing the inherent challenges of real-time interactive content creation.

Acknowledgments. This research was supported by a doctoral research grant from the Ramey Research Fund and the University of Washington's Department of Human Centered Design & Engineering.

References

1. https://help.twitch.tv/s/article/joining-the-affiliate-program?language=en_US
2. Cabeza-Ramírez, L.J., Fuentes-García, F.J., Muñoz-Fernandez, G.A.: Exploring the emerging domain of research on video game live streaming in web of science: State of the art, changes and trends. Int. J. Environ. Res. Public Health **18**(6), 2917 (2021)
3. Cai, J., Wohn, D.Y.: Categorizing live streaming moderation tools: An analysis of twitch. International Journal of Interactive Communication Systems and Technologies (IJICST) **9**(2), 36–50 (2019)
4. Charmaz, K.: Grounded theory in global perspective: Reviews by international researchers. Qual. Inq. **20**(9), 1074–1084 (2014)
5. Chen, Y., Hua, H., Wu, B.: The impact of audience numbers on tipping behavior in user-generated live streaming. In: International Conference on Management Science and Engineering Management. pp. 312–323. Springer (2024)
6. Choi, Y., Kang, E.J., Choi, S., Lee, M.K., Kim, J.: Proxona: Supporting creators' sensemaking and ideation with llm-powered audience personas. In: Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems. pp. 1–32 (2025)
7. DeVito, M.A.: How transfeminine tiktok creators navigate the algorithmic trap of visibility via folk theorization. Proceedings of the ACM on Human-Computer Interaction **6**(CSCW2), 1–31 (2022)
8. Flores-Saviaga, C., Hammer, J., Flores, J.P., Seering, J., Reeves, S., Savage, S.: Audience and streamer participation at scale on twitch. In: Proceedings of the 30th ACM Conference on Hypertext and Social Media. pp. 277–278 (2019)
9. Gray, M.L., Suri, S.: Ghost work: How to stop Silicon Valley from building a new global underclass. Harper Business (2019)
10. Guajardo, A.M.: "it sucks for me, and it sucks for them": The emotional labor of women twitch streamers. In: Proceedings of DiGRA 2022 conference: Bringing worlds together (2022)
11. Gulliksen, J., Göransson, B., Boivie, I., Blomkvist, S., Persson, J., Cajander, Å.: Key principles for user-centred systems design. Behaviour and Information Technology **22**(6), 397–409 (2003)
12. Hödl, T., Myrach, T.: Content creators between platform control and user autonomy: the role of algorithms and revenue sharing. Business & Information Systems Engineering **65**(5), 497–519 (2023)
13. Hung, Y.K., Huang, Y.C., Su, T.Y., Lin, Y.T., Cheng, L.P., Wang, B., Sun, S.H.: Simtube: Simulating audience feedback on videos using generative ai and user personas. In: Proceedings of the 30th International Conference on Intelligent User Interfaces. pp. 1256–1271 (2025)
14. Hurst, A., Lerer, A., Goucher, A.P., Perelman, A., Ramesh, A., Clark, A., Ostrow, A., Welihinda, A., Hayes, A., Radford, A., et al.: Gpt-4o system card. arXiv preprint arXiv:2410.21276 (2024)

15. Johnson, M.R., Woodcock, J.: "and today's top donator is": How live streamers on twitch. tv monetize and gamify their broadcasts. Social Media+ Society **5**(4), 2056305119881694 (2019)
16. Karizat, N., Delmonaco, D., Eslami, M., Andalibi, N.: Algorithmic folk theories and identity: How tiktok users co-produce knowledge of identity and engage in algorithmic resistance. Proceedings of the ACM on human-computer interaction **5**(CSCW2), 1–44 (2021)
17. Kaufman, A.R., Klevs, A.: Adaptive fuzzy string matching: How to merge datasets with only one (messy) identifying field. Polit. Anal. **30**(4), 590–596 (2022)
18. London, T.M., Crundwell, J., Eastley, M.B., Santiago, N., Jenkins, J.: Finding effective moderation practices on twitch. In: Digital ethics, pp. 51–68. Routledge (2019)
19. Luo, M., Hsu, T.W., Park, J.S., Hancock, J.T.: Emotional amplification during live-streaming: Evidence from comments during and after news events. Proceedings of the ACM on human-computer interaction **4**(CSCW1), 1–19 (2020)
20. Ma, R., Gui, X., Kou, Y.: Multi-platform content creation: the configuration of creator ecology through platform prioritization, content synchronization, and audience management. In: Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems. pp. 1–19 (2023)
21. Ma, R., Kou, Y.: "i am not a youtuber who can make whatever video i want. i have to keep appeasing algorithms": Bureaucracy of creator moderation on youtube. In: Companion Publication of the 2022 Conference on Computer Supported Cooperative Work and Social Computing. pp. 8–13 (2022)
22. Mallari, K., Williams, S., Hsieh, G.: Understanding analytics needs of video game streamers. In: Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems. pp. 1–12 (2021)
23. Mihailova, T.: Navigating ambiguous negativity: A case study of twitch. tv live chats. New Media & Society **24**(8), 1830–1851 (2022)
24. Mihailova, T.M., Hall, J.A.: Effects of negativity type and active involvement on the likelihood of responding to negativity in stream live chats. Available at SSRN 4553848
25. Mok, T., Au Yueng, C.M., Tang, A., Oehlberg, L.: Talk like somebody is watching: Understanding and supporting novice live streamers. In: Proceedings of the 2020 ACM International Conference on Interactive Media Experiences. pp. 154–159 (2020)
26. Radaideh, A., Dweiri, F.: Sentiment analysis predictions in digital media content using nlp techniques. International Journal of Advanced Computer Science & Applications **14**(11) (2023)
27. Ruberg, B., Cullen, A.L.: Feeling for an audience: The gendered emotional labor of video game live streaming. Digital Culture & Society **5**(2), 85–102 (2019)
28. Schuck, P., Altmeyer, M., Krüger, A., Lessel, P.: Viewer types in game live streams: questionnaire development and validation. User Model. User-Adap. Inter. **32**(3), 417–467 (2022)
29. Seering, J., Kairam, S.R.: Who moderates on twitch and what do they do? quantifying practices in community moderation on twitch. Proceedings of the ACM on Human-Computer Interaction **7**(GROUP), 1–18 (2023)
30. Shi, Y., Ma, C., Zhu, Y.: The impact of emotional labor on user stickiness in the context of livestreaming service–evidence from china. Front. Psychol. **12**, 698510 (2021)
31. Sjöblom, M., Törhönen, M., Hamari, J., Macey, J.: The ingredients of twitch streaming: Affordances of game streams. Comput. Hum. Behav. **92**, 20–28 (2019)

32. Snyder, C.: Paper prototyping: The fast and easy way to design and refine user interfaces. Morgan Kaufmann (2003)
33. Wei, J., Wang, X., Schuurmans, D., Bosma, M., Xia, F., Chi, E., Le, Q.V., Zhou, D., et al.: Chain-of-thought prompting elicits reasoning in large language models. Adv. Neural. Inf. Process. Syst. **35**, 24824–24837 (2022)
34. Wohn, D.Y., Freeman, G.: Audience management practices of live streamers on twitch. In: Proceedings of the 2020 ACM International Conference on Interactive Media Experiences. pp. 106–116 (2020)
35. Woodcock, J., Johnson, M.R.: The affective labor and performance of live streaming on twitch. tv. Television & New Media **20**(8), 813–823 (2019)
36. Yeo, W.J., Satapathy, R., Goh, R.S.M., Cambria, E.: How interpretable are reasoning explanations from prompting large language models? arXiv preprint arXiv:2402.11863 (2024)
37. Zhao, H., Fan, Y., Zheng, L.J., Liang, X., Chiu, W., Jiang, X., Liu, W.: The effect of social support on emotional labor through professional identity: evidence from the content industry. Journal of Global Information Management (JGIM) **30**(1), 1–19 (2022)