Streamfeed: AI-Mediated Feedback Reflection for Live Streamers

Keri Mallari(✉), Sharayu Kute, Bryan McCarthy, and Mark Zachry

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

Abstract. Community feedback is critical for content creators' growth, yet live streamers often struggle to obtain meaningful qualitative input from their audiences. To address this, we developed Streamfeed, a chatbot-based system designed to elicit constructive viewer feedback, paired with a streamer-facing dashboard that organizes and summarizes audience responses. We evaluated Streamfeed using a mixed-methods approach with 11 streamers and 166 viewers who engaged with the tool both during and outside live broadcasts. Our findings demonstrate that Streamfeed provided streamers with novel insights and reinforced existing ideas, supporting reflective practices and community engagement, while also revealing design tensions: balancing the benefits of anonymous candid feedback with streamers' needs for richer contextual information, and integrating feedback from newcomers without undermining established community norms. These insights advance HCI research by demonstrating how conversational agents can actively mediate participatory feedback processes and by illuminating the complex role of feedback as relational infrastructure within online communities.

Keywords: Live Streaming · Social Media · AI-Mediated

1 Introduction

Live streaming has emerged as a dominant form of digital content creation, with platforms like Twitch, YouTube Live, and TikTok Live collectively engaging billions of viewers. Unlike traditional media, live streaming is inherently participatory: creators and audiences co-construct content in real time through chat interactions, donations, and shared rituals. However, this immediacy creates unique challenges for streamers who must simultaneously entertain, moderate, manage technical production, and reflect on their performance, all while broadcasting live. Central to streamers' growth is feedback from their audiences, yet the fast-paced nature of streaming makes it difficult to obtain constructive qualitative input beyond fragmented chat reactions or platform analytics that illuminate what is happening but rarely why or how to improve.

To address this feedback gap, we developed Streamfeed, an AI-mediated system consisting of a conversational chatbot that guides viewers through structured feedback prompts across content production, community management,

A. Coman and S. Vasilache (Eds.): HCII 2026, LNCS 16716, pp. 49–66, 2026.
https://doi.org/10.1007/978-3-032-31418-5_4

and marketing strategy, paired with a streamer-facing dashboard that aggregates and summarizes responses using large language models. We evaluated Streamfeed through a mixed-methods study with 11 Twitch streamers and 166 viewers, examining not only whether the tool provided useful feedback but how it integrated into streamers' practices and revealed tensions in community dynamics. Additionally, we compared feedback from established community members versus newcomers, eight recruited individuals unfamiliar with the participating streamers, to understand how feedback source affects perceived value.

Our study investigates the design and impact of Streamfeed, a feedback system tailored to live streamers. We explore the following research questions:

RQ1: How do streamers perceive the value and relevance of structured feedback from Streamfeed?

RQ2: In what ways does Streamfeed integrate into and shape streamers' ongoing content practices, decision-making, and community-building efforts?

RQ3: What tensions or opportunities arise in engaging feedback from both regular and newcomer viewers?

Our findings demonstrate that Streamfeed provided streamers with novel insights and affirmed existing ideas, supporting reflection and concrete practice changes. However, we also surfaced important design tensions: streamers valued anonymous feedback for its candor but desired richer contextual information, and while newcomer feedback offered fresh perspectives, it sometimes conflicted with established community norms. These insights advance HCI research by demonstrating how conversational agents can actively mediate participatory feedback processes and by illuminating the complex role of feedback as relational infrastructure within online communities.

1.1 Feedback in Interactive and Creative Contexts

Feedback plays a critical role in the iterative process of creative work, and recent research highlights strategies that enhance its usefulness, specificity, and emotional appropriateness in digital environments. Anonymity improves crowdsourced feedback quality, with anonymous contributors offering more specific praise and criticism that is rated as more useful than identified feedback [13]. Additionally, scaffolding, guiding individuals through structured subtasks [22], produces more actionable critiques, as demonstrated by Critiki's use of structured prompts with crowdworkers [10] and CrowdCrit's finding that aggregated crowd input approximates expert-level feedback [18]. Interactive guidance tools with adaptive suggestions and reusable templates help novice reviewers generate feedback that is specific, actionable, and justifiable [20], yet feedback quality also depends on relational dynamics: online fanfiction writers highly valued feedback from individuals with authentic, ongoing relationships that fostered trust, contextual understanding, and emotional support [6]. These studies underscore that effective feedback in creative digital contexts requires integrating structural support (anonymity and scaffolding), technological guidance (interactive tools), and social connection (relational continuity), while recognizing that platform infrastructures and cultures shape what kinds of feedback are encouraged, making it essential to study feedback mechanisms alongside the affordances and social dynamics of their mediating platforms.

1.2 Audience Engagement and Viewer-Streamer Dynamics

Audience engagement in digital environments spans from intentional community-building in institutional media to emergent, platform-mediated practices in creator-driven spaces, with creators navigating different forms of participation shaped by technological and social affordances. In public service media like Flemish television, producers strategically cultivate engagement through immersive and interactive practices supported by curator, facilitator, and observer roles [27], contrasting with more decentralized creator-centered platforms where algorithmic controls imposed by YouTube and Twitch create tension between platform metrics and creator autonomy [4]. Twitch exemplifies a participatory model where live streams as "third places," informal social hubs fostering co-creative culture through real-time interaction, chat dynamics, and communal rituals [11], while on algorithmically driven platforms like TikTok, influencers deploy charisma, content quality, and relational strategies to cultivate engagement through emotional performance and identity bonding [28]. Together, these cases reveal that audience engagement is strategically curated, algorithmically constrained, and socially co-constructed depending on platform context and the roles assumed by creators and audiences.

1.3 Labor with Online Content Creation

Live streaming exemplifies the increasingly complex nature of digital content creation labor, where creators operate simultaneously as entertainers, community managers, technical producers, and entrepreneurs. Recent scholarship illuminates how this work is characterized by precarity, emotional demands, and constant self-optimization within algorithmically governed platforms, with streamers engaging in "playbour," the blurring of play and work, while managing parasocial relationships and experiencing "algorithmic anxiety" as they attempt to reverse-engineer platform recommendation systems [4]. Feedback plays a critical yet underexamined role in this labor ecosystem: constructive audience input can validate creative decisions and provide emotional support, yet the lack of structured feedback mechanisms means streamers rely on fragmented signals that provide incomplete pictures of audience sentiment, intensifying the already-demanding work of content creation by forcing consequential decisions with limited information.

1.4 Designing Feedback Tools for Content Creators

As creative expression increasingly takes place within digital and algorithmically mediated spaces, feedback has evolved into both a community-building mechanism and a computational challenge, with contemporary research reflecting this duality: platforms shape norms through interface cues and participatory culture, while advanced AI systems simulate, interpret, and augment feedback in real time. Feedback in creative communities often functions as social scaffolding, platforms like Mosaic foster reflective peer feedback on works-in-progress rather than finished products [16], while systemic approaches like Feedback Shaping predict creator sensitivity to rebalance platform incentives supporting creator well-being [26], and features like YouTube's creator heart guide comment dynamics by signaling approval and curating tone [7]. Recent advances in LLMs have enabled interactive, anticipatory feedback systems: Proxona transforms comment data into structured audience personas for simulated conversations [8], while SimTube generates pre-release audience comments using multimodal video analysis and persona modeling [14]. However, these systems primarily operate in static or anticipatory contexts, leaving a gap in dynamic, high-pressure environments like livestreaming where creators must adapt rapidly without structured feedback mechanisms. To address this, we built Streamfeed to embed constructive, structured feedback into the livestreaming workflow through conversational guidance, feedback structuring, and synthesis tools that support streamers in continuously engaging with audience perspectives.

2 Streamfeed Development

2.1 Design Rationale

The design of Streamfeed is grounded in recurring themes identified across related work: the importance of structured feedback, the relational and affective dimensions of audience interaction, and the influence of platform-specific dynamics on engagement. We chose a chatbot interface because it supports guided, conversational feedback, a format shown to reduce cognitive load and improve clarity among novice reviewers [10], and embedded real-time follow-up logic to clarify vague comments and redirect inappropriate inputs. To ensure relevance to the streaming context, we adopted a "strengths and improvements" framework [3], providing streamers with both affirmation and actionable advice organized into three categories, content production, community management, and marketing strategy [19], that reflect the core dimensions of streamer success and support reflection and resilience in this inherently fast-paced, high-pressure, and socially mediated work.

2.2 Conversational Agent Design

To address the needs surfaced in prior work and our formative study, we developed a conversational agent deployed during livestreams via a shared link, guiding viewers through constructive feedback using prompt engineering strategies centered on clarity, constraint, and adaptivity. The chatbot follows a structured conversational prompt across three targeted domains, **marketing strategies**, **content production**, and **community management**, asking one question at a time and assessing input quality in real time using heuristics such as minimum character length, presence of vague or overly negative language, and relevance; unclear or off-topic responses (e.g., “okay,” “boring”) trigger follow-up requests for clarification as seen in Fig. 1. The system employs a strengths-and-improvements framework [3] and aims to collect **at least one strength and one area for improvement per category** before concluding. The dashboard (Fig. 2) presents feedback using an inverted pyramid approach [21]: a top-level “tl;dr” section summarizes viewer count, key motivations, and actionable suggestions, while detailed breakdowns of each category below include synthesized summaries and two representative viewer quotes, allowing streamers to quickly grasp insights and access deeper context as needed.

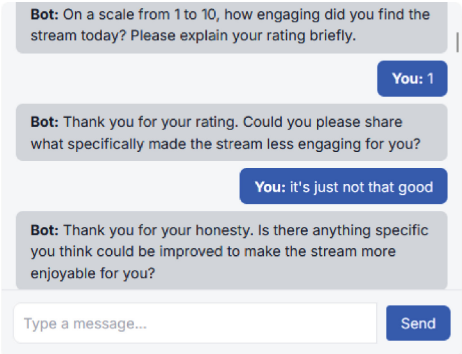

Fig. 1. Chatbot interface example. The bot is asking the user for feedback about the stream, and is asking for further clarifications when the user is not responding clearly.

2.3 System Development

The system consisted of two main components: a chatbot interface and a dashboard for presenting the data. Streamers shared a unique chatbot link (/chat/<username>) with their audience, who interacted with the chatbot in real time. These interactions were captured and stored in a PostgreSQL database. The dashboard, accessible via a tokenized URL (/chat/?token=<unique_token>), displayed a summary of viewer interactions. It was initially populated with data upon the first visit after engagement and was subsequently updated manually on a weekly basis by the lead researcher. The front end of the system was developed using React, while the backend leveraged Node.js and Express for routing and API handling. The system diagram is presented in Fig. 3.

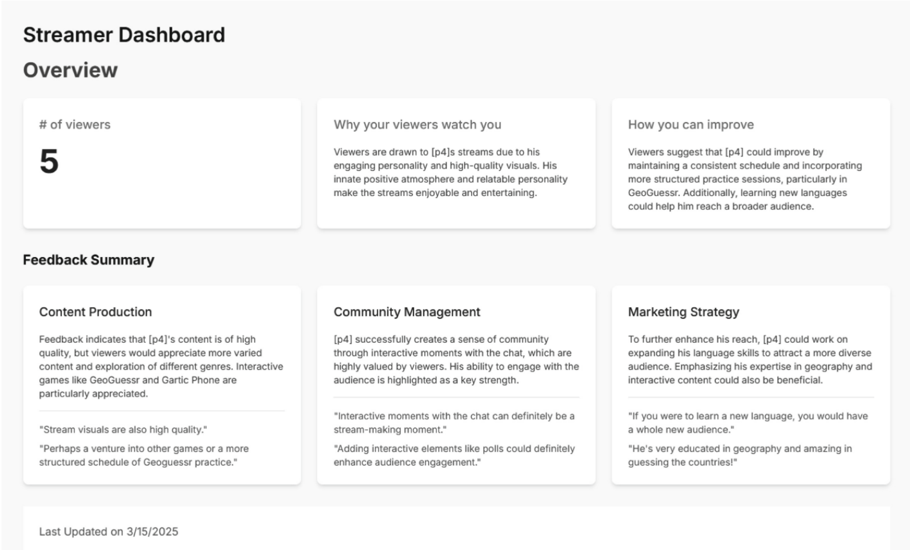

Fig. 2. Streamer Dashboard for participant [p4], displaying a summary of viewer engagement and qualitative feedback. The dashboard highlights a viewer count of 5, reasons viewers enjoy the stream (e.g., personality and visuals), suggestions for improvement (e.g., consistent schedule, language learning), and categorized feedback on content production, community management, and marketing strategy. Each section includes quotes and insights drawn from viewer comments.

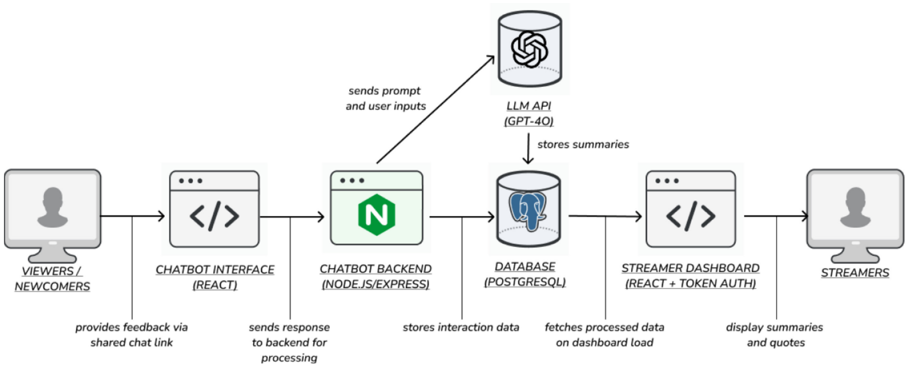

Fig. 3. System architecture of the Streamfeed platform. Viewers provide feedback through a React-based chatbot interface during livestreams. Inputs are processed by a Node.js/Express backend, which interacts with a PostgreSQL database and a GPT-4 API for both adaptive prompting and summarization. Summarized feedback is then displayed to streamers via a secure dashboard interface.

3 Methods

This study aimed to evaluate the usability, challenges, and benefits of a chatbot based feedback system designed for streamers. The primary objective was to determine whether the system could provide streamers with new and relevant constructive feedback to improve their content and stream planning. To achieve this, we recruited streamers to promote usage of the chatbot with their viewers during, and after their live streams. These interactions generated feedback that was filtered, aggregated, and summarized in a dashboard for streamers to review. The study consisted of three phases: onboarding, active use of the system during live streams, and a follow-up evaluation.

3.1 Participants

This usability study consisted of 11 streamers (Table 1) recruited through social media platforms and personal networks. Eligibility criteria required participants to have prior experience in live streaming, and must have streamed the week prior the onboarding call. The streamers represented diverse content focuses, including gaming, art, and just chatting, with audience sizes ranging from 5 to 100 viewers. Most participants were based in North America. Our participants also had a range of streaming experience, three of whom were Partnered on Twitch, which means they get a bigger portion of the revenue from ads, subscriptions and bits in contrast to an Affiliate streamer [1].

Table 1. Participant Demographics

Participant ID (PID)	Location	Followers	Streamer Status	Average Viewer (past 30 days)	Hours Streamed (past 30 days)
p1	Sweden	3,145	Affiliate	18	86
p2	USA	49,676	Partner	82	133
p3	USA	106	Affiliate	3	21
p4	Canada	304	Affiliate	4	21
p5	USA	1,448	Affiliate	18	249
p6	Mexico	19,046	Partner	69	173
p7	USA	144	Affiliate	1	4
p8	Australia	14,123	Partner	88	127
p9	USA	7,395	Affiliate	55	120
p10	USA	456	Affiliate	1	49
p11	Philippines	166	Affiliate	9	86

In addition to feedback from existing community members, we conducted a focused case study involving a group of eight newcomers (Table 2) recruited from a university research group. These individuals were familiar with live streaming, either as viewers or former content creators, but had no prior relationship with the participating streamers. Their role was to provide fresh, outsider perspectives

Table 2. Newcomer Demographics

Newcomer ID	Age Range	Streamer Status	Stream Engagement Frequency	Preferred Stream Genre
n1	25–34	Yes, Affiliate	Daily	Games
n2	25–34	No	Monthly	Events
n3	18–24	Yes, Affiliate	Daily	Games
n4	25–34	Yes, Partner	Daily	Games
n5	18–24	No	Rarely	Games
n6	25–34	No	Rarely	Games
n7	18–24	No	Rarely	Games
n8	25–34	No	Daily	Games
n9	25–34	No	Daily	Games

by engaging with streams and submitting feedback through the same chatbot interface used by regular viewers.

This study received institutional IRB approval (STUDY00013563) and all participants provided informed consent prior to participation. Streamers received a $50 gift card as compensation for their time and contributions, while newcomers participated voluntarily as part of their involvement in the research group.

3.2 Study Phases

The study consisted of three phases described below.

Onboarding Phase. Each streamer began with a 20–30 min onboarding call. Researchers explained the study purpose, obtained consent, and asked streamers to complete a baseline survey assessing current feedback experiences. Researchers then demonstrated the tool via screen sharing, showing both the chatbot interface and an example dashboard. Streamers were asked to promote the chatbot to their community and aim for at least five viewer interactions before the evaluation call.

System Engagement. After onboarding, streamers incorporated the chatbot into their streams by sharing its link with viewers during or outside broadcasts. The chatbot prompted viewers with questions about content, community, and marketing. In addition to community members, the newcomer research group provided feedback through the same interface, allowing comparison between established viewers and newcomers. Dashboards were populated after initial viewer engagement and updated weekly by researchers to reflect new feedback submissions.

Evaluation Phase. Once five viewers had engaged with the chatbot, streamers participated in a 30–40 min semi-structured interview. The interview explored system usability, dashboard review with screen sharing, implementation of feedback suggestions, and differences between community and newcomer feedback. Streamers also completed a post-survey repeating the baseline quality items using their Streamfeed dashboard as reference.

3.3 Newcomer Feedback Protocol

Assessing live streamers from newcomers' perspectives serves as an exploratory approach to understanding how feedback from individuals with varying familiarity yields distinct insights [11], informed by literature showing that feedback-givers' roles, identities, and relational positioning significantly shape content and reception [15]. We focused on newcomers because streaming growth depends heavily on first-time impressions, streamers' ability to attract and retain new viewers hinges on how quickly content communicates value [11], offering actionable insights into growth barriers that long-time community members might overlook. To ensure structured evaluations, newcomers followed research team-developed heuristics, watching two to three recent clips, skimming full-length recordings, observing live streams when possible, reviewing About and Schedule pages for brand clarity, and examining past broadcasts and community elements (chat rules, follower counts, subscriber benefits) to gauge consistency and accessibility, providing feedback that complemented organic community input by highlighting areas invisible to established viewers.

3.4 Quantitative Measures and Analysis

Survey Instrument Development. Survey items were custom-designed based on Steelman et al.'s Feedback Environment Scale [24], adapted to the streaming context. The pre-survey assessed feedback frequency, sources, and baseline quality perceptions across nine dimensions (e.g., "provides clear steps for improving streams," "helps improve community interaction"). The post-survey repeated these nine quality items, with streamers rating them based on their Streamfeed dashboard rather than typical feedback sources. All items used 5-point Likert scales (1=Strongly Disagree to 5=Strongly Agree). The survey was pilot-tested to assess clarity and completion time (3–5 minutes).

Quantitative Analysis Approach. Given our sample size (n = 11 streamers) and the exploratory nature of this study, we employed descriptive statistics as the primary quantitative method to characterize feedback patterns and perceived quality changes. We calculated means and standard deviations for pre- and post-survey responses across the nine quality dimensions, examining directional trends rather than conducting formal hypothesis testing. This approach is consistent with mixed-methods research where qualitative findings provide primary insights and quantitative data offers supporting context [9]. All 11 streamers completed both surveys with no missing data. For viewer chatbot data, we required greater than 10 total messages (viewer + chatbot combined) for inclusion, ensuring meaningful engagement beyond superficial interactions. This threshold was determined through pilot testing to filter out sessions with only single-word responses. Chatbot engagement metrics tracked unique participants, message counts, drop-off rates, and completion rates to characterize system usage patterns.

3.5 Qualitative Coding and Thematic Analysis

Interviews were recorded and then transcribed using Otter.ai. The first author cleaned typographical errors and mistranscriptions in all transcripts before storing them for analysis. Three members of the research group were involved in the qualitative coding process described below.

The team employed a grounded theory approach to qualitative coding [5]. The process began with a collaborative coding session on one interview transcript, which facilitated the emergence of initial themes and informed the development of a preliminary codebook. Each coder was then assigned three additional interviews to independently code. At subsequent meetings, the team discussed discrepancies and updated the codebook to reflect emerging themes and refinements. This iterative process continued until the final codebook was established and all interview transcripts were fully coded. To assess consistency across coders, we calculated inter-rater reliability using Fleiss's Kappa, achieving a score of 0.855, indicating near-perfect agreement.

The final codebook organized eight codes into three themes. The first theme encompassed Ease of Use and Platform and Technical Considerations, capturing participants' experiences with tool functionality and constraints. The second included Feedback Characteristics, Newcomer Perspectives, and Newcomer vs. Community Feedback, addressing perceptions of feedback quality and source differences. The third comprised Feature Desires, Tool Integration, and Viewer-Centered Improvements, reflecting how streamers adopted the tool and envisioned enhancements.

These categories informed the thematic organization of the qualitative findings presented in Sect. 5.

4 Quantitative Findings: Viewer and Tool Engagement

4.1 Survey Findings

The survey results reveal diverse patterns in feedback frequency and sources for live streamers. Regarding frequency, a majority of respondents (7 out of 11) reported that they "rarely receive feedback when I stream," while a smaller group (4 out of 11) stated they "receive feedback most of the time when I stream."

Feedback sources varied, with viewers being the most common, cited by 10 of 11 participants. Analytics tools, such as built-in Twitch analytics, were the second most common source, mentioned by 7 respondents.

The content of feedback ranged from specific suggestions to general improvements. Several streamers reported viewers suggesting games to play, with one mentioning, “I have received suggestions for games to play” (p1). Another streamer received advice to expand their content “and expand to Youtube at least for short form videos” (p2). Stream improvements were also common, with one participant noting a “viewer recently asked if [they] could add certain rewards for viewers/chat experience” (p4).

Comparing Feedback Quality: Pre vs. Post-Streamfeed. To understand how Streamfeed-generated feedback compared to streamers’ typical feedback sources, we examined changes in perceived quality across nine dimensions. As shown in Fig. 3, average scores improved across all assessed categories when rating Streamfeed feedback compared to baseline feedback sources. The largest improvements were observed in “provides clear steps for improving streams” (baseline M=3.00, SD=1.18; Streamfeed M=4.45, SD=0.52; difference=+1.45) and “helps me improve with my community” (baseline M=3.55, SD=1.04; Streamfeed M=4.64, SD=0.50; difference=+1.09). More modest improvements appeared in dimensions like “is actionable” (difference=+0.55) and “well-timed and relevant to my needs” (difference=+0.64).

While our small sample size (n=11) limits generalizability, these descriptive patterns suggest that streamers perceived Streamfeed-generated feedback as notably more structured, clear, and actionable compared to their typical feedback sources. These quantitative trends align with qualitative interview data (Sect. 5), where streamers explicitly valued the tool’s categorization, specificity, and constructive framing.

4.2 Viewer and Chatbot Engagement Metrics

The study included 11 streamers and 161 viewers. Viewer engagement with the chatbot (Table 3 varied across streams, ranging from 5 to 26 active participants per stream. To measure sustained engagement, we filtered for meaningful interactions by excluding conversations with fewer than 10 total messages (combined from both viewers and the chatbot). Analysis revealed 25 viewer drop-offs across all 11 streamer sessions, with a total of 89 message drop-offs observed throughout the study.

Table 3. Chatbot Engagement Statistics

PID	Viewer Engagement Metrics			Message Activity Metrics		
	Original Count	Filtered Count	Drop-Offs	Original Count	Filtered Count	Drop-Offs
p1	7	5	2	80	72	8
p2	22	21	1	390	385	15
p3	5	5	0	84	84	0
p4	5	5	0	56	56	0
p5	20	16	4	242	228	14
p6	22	15	7	230	216	14
p7	6	5	1	60	56	4
p8	23	23	0	330	330	0
p9	26	19	7	425	401	24
p10	19	14	2	210	200	10
p11	6	6	0	110	110	0

5 Streamer Reflections on Streamfeed

5.1 Perceived Value of Streamfeed

Streamfeed was widely perceived as highly accessible and user-friendly by both streamers and viewers. During the onboarding process, the researcher guided streamers through the initial setup, which included screen sharing, entering a preferred display name, and accessing their personal dashboard and chatbot link for viewer engagement. While streamers did not complete the setup independently, they found navigating the dashboard and sharing relevant links to be straightforward and uncomplicated. Participants consistently emphasized the simplicity of the process. As one streamer explained, "I think it was very simple because people didn't have to do anything, like, there's no sign up required, so you can directly, just put what you think... So there's no sign up, you can be there anonymously and just tell what you think" (p9). A viewer echoed this sentiment, noting, "Like, it's very easy to kind of interact with. It's like, I feel like it's in a format where people intuitively know how to use it" (p8). Overall, the lack of required registration and the intuitive design were identified as key factors contributing to Streamfeed's ease of use for both streamers and viewers.

5.2 Integration Into Streamer Practices

Streamers used various methods to share the Streamfeed link and encourage participation, integrating it into Discord servers,"It was mainly on my stream. I did put in my Discord as well" (p9), and pinning it in chat: "I also pinned it in my chat. So it was always on top of the chat that, like, if you want to provide feedback, here it is" (p1). Some created custom commands or automated reminders: "I kind of made a command in my stream title, and was actually like having some time to encourage people to do it" (p6), while another used

NightBot to periodically remind viewers about the feedback opportunity (p5). Streamers also encouraged feedback verbally, "And then throughout the stream, I would just ask people if they had time, and they felt like they're up to it, if they would mind leaving some feedback. Good and bad" (p2), or prompted individuals directly: "I just, like, urged individual viewers, like, hey, fill out the entire... fill out the stream feed link thing" (p3).Some placed the link in panels below their streams ensuring persistent accessibility through these combined strategies of Discord integration, chat features, verbal encouragement, and panel placement. This range of adoption strategies suggests the tool accommodated different streaming styles and community sizes without requiring major workflow disruption.

5.3 Novelty and Constructiveness of Feedback

The tool enabled streamers to receive novel insights they had never considered and affirmations of ideas they had contemplated but not acted on, with one noting: "I liked the feedback. I think there was some actually really insightful feedback in there. It was just kind of highlighting what I'd already kind of self-assessed. But I do think a streamer that isn't as self-aware, it can be very valuable" (p8). The feedback was actionable, several streamers implemented suggestions between onboarding and evaluation, such as one who started raiding other streams after receiving that feedback: "that is easily fixable. And I just started doing that and implementing that" (p5) though not all feedback was adopted due to required labor, with one resisting Discord and schedule requests: "I could, technically, because they asked me to make a Discord server and to have a stream schedule. I'm like, No, I'm not doing that" (p3). Streamers appreciated the structured interface with categorization, concise summaries, and direct quotes that lent credibility and specificity. The anonymous nature enabled honesty: "I think people were honest there. So I think I like that. Because usually people are a little bit like, too damn scared if they want to say something" (p9), with another adding: "Some people aren't going to just outright say what they would like to see, so I think this provides some anonymity for them to be able to give feedback without feeling like they're going to be targeted" (p5).

5.4 Reactions to Newcomer Feedback

Streamers' responses to newcomer feedback varied, balancing appreciation for fresh perspectives with skepticism about contextual understanding. Many found newcomer feedback brought critical observations existing community members might overlook: "It's interesting to see their aspect of things, since they haven't been around. . . they're not necessarily familiar with how the stream goes. So it is interesting to see from a fresh person's aspect of things" (p5). However, this freshness came with limitations, streamers raised concerns about engagement depth and validity: "I would be interested to know how much engagement in chat they did" (p5), with some deprioritizing newcomer input: "It's informative, it's not useless, but it's not going to influence what my next stream is" (p8), noting "The community is a little bit more specific... because they don't know me, obviously they don't have context, they're just making an impression" (p8). Language and cultural barriers further limited usefulness in some cases, as one Spanish-speaking streamer noted that non-Spanish-speaking newcomers faced disconnects: "It was kind of awkward... that would be a big barrier for them to overcome the content and just actually enjoy it" (p6). In such cases, newcomer feedback risked misinterpreting intentional creative choices as errors.

Despite these reservations, streamers recognized that both feedback sources served distinct purposes: "I give value to both groups kind of the same way, because you want newcomers, of course, but you also want to nurture the community you have. So those two are very, very valuable for me" (p4). The challenge was not determining which source was "better," but rather understanding when each type of feedback was most relevant. Newcomer feedback proved valuable for growth-oriented questions (e.g., "How can I attract new viewers?"), while community feedback better addressed retention and deepening engagement (e.g., "How can I better serve my existing audience?").

6 Discussion

Current feedback systems for live streamers prioritize analytics and numeric metrics or offer raw comment streams with minimal structure. Our findings suggest feedback systems should be reconceptualized as sociotechnical tools that scaffold streamers' reflection, community engagement, and long-term development.

6.1 Design Tensions: Structure, Context, and Anonymity

Streamfeed's structured categorization into content production, community management, and marketing strategy proved valuable, streamers could quickly identify relevant insights and the "strengths and improvements" framework provided both validation and direction. The conversational chatbot format successfully elicited more substantive feedback than typical chat interactions, aligning with prior work showing that scaffolding produces more actionable feedback [22].

However, participants desired greater control over feedback organization and presentation. Streamers sought more granular customization options and content-specific subcategories beyond the three broad categories. This reflects findings that Twitch streamers frequently seek or build unique tools when platforms fall short [19], and echoes Ackerman's argument that rigid technical systems fail to accommodate users' evolving practices [2]. Since streamers simultaneously act as entertainers, community leaders, and strategists, giving them control over feedback mechanisms increases ownership and trust.

A critical tension emerged around anonymity. Anonymous feedback encouraged candor, as one streamer noted, "I think people were honest there. So I think I like that. Because usually people are a little bit like, too damn scared if they want to say something" (p9). Anonymity proved valuable given power dynamics where viewers may avoid honest criticism. However, anonymity also removed crucial interpretive context. Streamers wanted temporal and situational details: "I would be interested to know how much engagement in chat they did" (p5). This suggests feedback systems should offer graduated transparency, allowing viewers to choose identification levels while automatically capturing contextual metadata (timestamp, activity, stream segment) that helps streamers interpret feedback without compromising viewer comfort.

6.2 Community Dynamics: Newcomers, Norms and AI Mediation

Our comparison of newcomer versus community member feedback exposed a fundamental challenge in online communities: integrating new voices while preserving established culture [17]. Streamers valued fresh perspectives from newcomers, "It's interesting to see their aspect of things, since they haven't been around" (p5), particularly for identifying accessibility barriers or onboarding friction. However, this freshness came with limitations. Streamers questioned whether newcomers engaged deeply enough to provide informed feedback, with some explicitly deprioritizing newcomer input: "It's informative, it's not useless, but it's not going to influence what my next stream is" (p8).

Cultural and language barriers further complicated matters. One Spanish-speaking streamer noted that non-Spanish-speaking newcomers faced fundamental disconnects: "It was kind of awkward... that would be a big barrier for them to overcome the content and just actually enjoy it" (p6). In such cases, newcomer feedback risked misinterpreting intentional creative choices as errors. Well-intentioned suggestions (like changing format or game) sometimes disregarded practices that veteran viewers valued, echoing Kraut and Resnick's observation that newcomers "in their ignorance, may act in ways that offend other group members or undercut the smooth functioning of the group" [17].

Despite these reservations, streamers recognized both feedback sources served distinct purposes: "I give value to both groups kind of the same way, because you want newcomers, of course, but you also want to nurture the community you have" (p4). This suggests feedback systems should explicitly distinguish source and allow differential weighting, newcomer feedback proves valuable for growth-oriented questions while community feedback better addresses retention. Beyond source dynamics, streamers' comfort with the chatbot revealed expanded roles for AI in streaming communities. Rather than viewing the feedback system merely as a technical utility, participants treated the bot as a legitimate community participant facilitating dialogue, extending prior work positioning bots as social actors [23]. This suggests AI agents can mediate constructive feedback processes when designed for social compatibility rather than just efficiency.

6.3 Reconceptualizing Feedback as Relational Infrasturcutre

This work extends prior conceptualizations of online feedback by reframing it as relational infrastructure, socially negotiated processes tightly coupled to identity work and community maintenance, rather than static information transfer oriented around performance metrics. By embedding feedback solicitation into live streams and using conversational AI to structure input, Streamfeed enabled a participatory model that shifted control from platform-centric metrics to community-grounded dialogue. Feedback became material for reflection and relationship-building: streamers used it to validate creative choices, understand audience perspectives, and make decisions honoring both personal vision and community desires.

Our findings highlight substantial interpretive labor streamers perform to make feedback meaningful. Comments lacking contextual detail were difficult to act upon; streamers needed to triangulate multiple sources, assess credibility based on familiarity, and weigh suggestions against their creative vision and resource constraints. This interpretive labor is invisible in most discussions of creator analytics but represents a significant component of content creation work. Feedback systems must account for this by providing context-rich data and supporting sensemaking, not just aggregating comments.

The distinction between newcomer and community feedback revealed that feedback is filtered through social position, relational history, and mutual understanding. Long-time community members provided feedback grounded in shared context and trust developed over time; newcomers offered valuable outsider perspectives but lacked this relational foundation. We argue that feedback in creator-audience contexts should be understood as relational infrastructure, the sociotechnical scaffolding supporting creators' navigation of identity, labor, and growth in public digital spaces. Future feedback systems in online communities should reflect this complexity through nuance, context-awareness, and support for the interpretive practices that make feedback meaningful.

7 Future Work and Limitations

Our work could be expanded to other creative domains such as art live streams, educational content, and more. Additionally, scaling the system to accommodate larger communities would enable broader insights and a more robust evaluation. Improving integration with host platforms could also improve engagement, as asking users to go to an external site may hinder engagement. Finally, a longitudinal study could provide additional insights on how the tool affects creator-audience dynamics over time, including changes in engagement and content development.

We acknowledge the limitations of this work such as its limited scope, focusing on a single platform, Twitch. Other live streaming platforms like Youtube or TikTok have distinct interaction models that would require additional analyses. From a technical standpoint, the current implementation faces challenges common to language models, such as potential hallucinations or rigid adherence to prompt templates [12]. These issues may affect quality and relevance of the summaries and future iterations should address such limitations through fine-tuning the model or using human-in-the-loop mechanisms [25].

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://www.twitch.tv/p/en/partners/
2. Ackerman, M.S.: The intellectual challenge of cscw: the gap between social requirements and technical feasibility. Hum.-Comput. Interact. **15**(2–3), 179–203 (2000)
3. Aguinis, H., Gottfredson, R.K., Joo, H.: Delivering effective performance feedback: the strengths-based approach. Bus. Horiz. **55**(2), 105–111 (2012)
4. Bishop, S.: Algorithmic experts: selling algorithmic lore on youtube. Soc. Media+ Soc. **6**(1), 2056305119897323 (2020)
5. Charmaz, K.: Grounded theory. Qualit. Psychol. Pract. Guide Res. Methods **3**, 53–84 (2015)
6. Cheng, R., Frens, J.: Feedback exchange and online affinity: a case study of online fanfiction writers. Proc. ACM Hum.-Comput. Interact. **6**(CSCW2), 1–29 (2022)
7. Choi, F., Lambert, C., Koshy, V., Pratipati, S., Do, T., Chandrasekharan, E.: Creator hearts: Investigating the impact of positive signals from youtube creators in shaping comment section behavior. arXiv preprint arXiv:2404.03612 (2024)
8. Choi, Y., Kang, E.J., Choi, S., Lee, M.K., Kim, J.: Proxona: supporting creators' sensemaking and ideation with llm-powered audience personas. arXiv preprint arXiv:2408.10937 (2025)
9. Creswell, J.W., Clark, V.L.P.: Designing and Conducting Mixed Methods Research. Sage Publications (2017)
10. Greenberg, M.D., Easterday, M.W., Gerber, E.M.: Critiki: a scaffolded approach to gathering design feedback from paid crowdworkers. In: Proceedings of the 2015 ACM SIGCHI Conference on Creativity and Cognition, pp. 235–244 (2015)
11. Hamilton, W.A., Garretson, O., Kerne, A.: Streaming on twitch: fostering participatory communities of play within live mixed media. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, pp. 1315–1324 (2014)
12. Huang, L., et al.: A survey on hallucination in large language models: principles, taxonomy, challenges, and open questions. ACM Trans. Inf. Syst. **43**(2), 1–55 (2025)
13. Hui, J., Glenn, A., Jue, R., Gerber, E., Dow, S.: Using anonymity and communal efforts to improve quality of crowdsourced feedback. In: Proceedings of the AAAI Conference on Human Computation and Crowdsourcing, vol. 3, pp. 72–82 (2015)
14. Hung, Y.K., et al.: Simtube: generating simulated video comments through multimodal ai and user personas. In: Proceedings of the 30th International Conference on Intelligent User Interfaces (IUI) (2025)
15. Keiler, L.S., Diotti, R., Hudon, K., Ransom, J.C.: The role of feedback in teacher mentoring: how coaches, peers, and students affect teacher change. Mentoring Tutoring: Partnership Learn. **28**(2), 126–155 (2020)
16. Kim, J., Agrawala, M., Bernstein, M.S.: Mosaic: designing online creative communities for sharing works-in-progress. In: Proceedings of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing (CSCW), pp. 1740–1751 (2017). https://doi.org/10.1145/2998181.2998195
17. Kraut, R.E., Resnick, P.: Building Successful Online Communities: Evidence-Based Social Design. Mit Press (2012)

18. Luther, K., et al.: Crowdcrit: crowdsourcing and aggregating visual design critique. In: Proceedings of the Companion Publication of the 17th ACM Conference on Computer Supported Cooperative Work & Social Computing, pp. 21–24 (2014)
19. 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)
20. Ngoon, T.J.: Overcoming satisficing: scaffolds for amplifying creativity. In: Proceedings of the 2019 Conference on Creativity and Cognition, pp. 670–674 (2019)
21. Pö ttker, H.: News and its communicative quality: the inverted pyramid—when and why did it appear? Journalism Stud. **4**(4), 501–511 (2003)
22. Reiser, B.J.: Scaffolding complex learning: the mechanisms of structuring and problematizing student work. In: Scaffolding, pp. 273–304. Psychology Press (2018)
23. Seering, J., Flores, J.P., Savage, S., Hammer, J.: The social roles of bots: evaluating impact of bots on discussions in online communities. Proc. ACM Hum.-Comput. Interact. **2**(CSCW), 1–29 (2018)
24. Steelman, L.A., Levy, P.E., Snell, A.F.: The feedback environment scale: construct definition, measurement, and validation. Educ. Psychol. Measur. **64**(1), 165–184 (2004)
25. Tonmoy, S., et al.: A comprehensive survey of hallucination mitigation techniques in large language models. arXiv preprint arXiv:2401.01313 6 (2024)
26. Tu, Y., Lo, C., Yuan, Y., Chatterjee, S.: Feedback shaping: a modeling approach to nurture content creation. arXiv preprint arXiv:2106.11312 (2021)
27. te Walvaart, M., Dhoest, A., Van den Bulck, H.: Production perspectives on audience engagement: community building for current affairs television. Media Ind. J. **6**(1), 43–66 (2019). https://doi.org/10.3998/mij.15031809.0006.103
28. Wang, Y., Feng, D.W.: Identity performance and self-branding in social commerce: a multimodal content analysis of chinese wanghong women's video-sharing practice on tiktok. Discourse, Context Media **50**, 100652 (2022)