AI vs. Human Live streaming Host: A Comparative Study on Fatigue Resistance and Audience Retention in Live-streaming E-commerce

Wang Chao

University of Pisa, Pisa, Italy,
c.wang@studenti.unipi.it

Wu Yue

Obuda University, Doctoral School of Security Studies, Bánki Donát Faculty of Mechanical and Safety Engineering, Budapest, Hungary,
wu.yue@bgk.uni-obuda.hu

Abstract: With the advancement of 5G network technology and artificial intelligence, live streaming has become one of the increasingly important channels in the e-commerce field, including not only live streaming by human streamers, but also live streaming by AI streamers.This study compares the dynamics of AI-powered and human streamers in live-streaming e-commerce, with an emphasis on audience retention patterns, fatigue resilience, and algorithmic interactions. The study presents a multi-dimensional approach to assessing streamers' endurance and viewer engagement sustainability. The study examines at how the operational stability of AI streamers and the adaptive capabilities of human streamers differently affect long-tail audience retention and algorithmic visibility through observational data capturing real-time viewership metrics and platform-level interventions. The framework presents the idea of "algorithmic acclimation" to quantify platform-driven traffic compensation mechanisms triggered by consistent performance metrics. Preliminary results reveal that AI and human streamers have different resilience characteristics, which has consequences for content strategy optimization. By developing a fatigue-inclusive evaluation model, this study advances the understanding of human-AI coexistence in digital retail ecosystems while also providing organizations with strategic insights for streamers deployment.

1 Introduction

The rapid integration of artificial intelligence (AI) into live-streaming e-commerce has introduced transformative opportunities. However, many enterprises remain hesitant to adopt AI streamers due to limited empirical insights into their comparative efficacy and operational dynamics. A significant knowledge gap is exposed by current industry practices: e-commerce companies lack frameworks to assess whether AI-driven broadcasts can equal or exceed the engagement capabilities of human hosts, particularly in maintaining audience retention and adapting to real-time interactions, even though AI streamers promise cost efficiency and scalability. This uncertainty extends to the strategic allocation of resources, as companies struggle to balance the cost-effectiveness and viewer engagement of dividing streaming hours between AI and human streamers. Current research mostly concentrates on discrete performance indicators (such peak viewership), ignoring comprehensive evaluations of audience attrition trends, fatigue resilience, and platform-driven traffic compensation mechanisms—elements essential to sustained operational success. There are still significant gaps in spite of these developments. First, current research does not systematically assess the ability of AI and human streamers to adjust to changing audience demands, such adjusting material in real time amid unforeseen encounters (e.g., resolving technical issues or emotional appeals). Second, there is still a lack of research on the long-term trade-offs between algorithmic consistency and human spontaneity, especially in situations that need for continuous audience involvement after peak hours. Third, despite the widespread use of platform-driven traffic compensation methods (such as recommendation tags), little is known about how they affect AI streamers differently than human streamers, particularly how algorithmic biases may unintentionally prioritize AI's stability above human innovation. This work fills these gaps by offering practical insights into the feasibility of AI streamers as long-term substitutes or enhancements to human streamers. In order to promote a healthy and profitable live-streaming environment, the findings seek to address urgent industry issues. The purpose of the study is to determine how customer behavior and the decision-making process in a live streaming setting are impacted by the live host selection.

2 Literature Review

AI virtual streamers have been incorporated into live-streaming e-commerce through a series of interrelated research phases. Initial research concentrated on foundational technologies that improved product display in virtual streaming environments, such as multimodal content production platforms like AliMe Avatar [1]. Later developments brought specific e-commerce animation frameworks, which allowed virtual anchors to interact with millions of users by responding in real time [2]. The foundation for researching the behavioral effects of virtual hosts was created by these technological advancements..

After establishing this technical foundation, research turned to examining social processes. Research found human streamers greatly increase purchase intention through superior perceived intimacy and responsiveness, according to a comprehensive investigation of streamer types [3]. The research is conducted by using the social cognitive theory. These effects were more noticeable for customers who were not very interested in trying new things [3].Through quasi-social ties, research has shown that streamers' social capital increases purchasing intentions by reducing information asymmetry based on a questionnaire survey of live streaming viewers [4]. Subsequent research showed threshold effects in the way that social qualities of virtual streamers increase experience value, especially when mediated by communicative and environmental elements [5]. This signaled an evolution from technical validation to the psychological processes that influence consumer choices.

Later, the factors that led customers to switch from human to AI virtual streamers were closely examined. Adoption willingness is influenced by personality factors and shopping motives, which are mediated by perceived innovation obstacles, according to asymmetric modeling [6]. Notably, research revealed possible "uncanny valley" consequences during hyper-realistic interactions, emphasizing the necessity of striking a balance between user comfort and human-likeness [6]. Concurrently, expectancy violation theory highlighted the intricacy of human-AI engagement by highlighting how repurchase behavior is dynamically reshaped by deviations from user expectations in AI-streamers interactions [7].

Hybrid collaboration methods are given priority in recent studies. Research shows that systems that combine the emotional adaptability of human streamers with the operational constancy of AI virtual streamers perform better than single-mode methods in maintaining engagement [8]. Credibility measures play a key role in refining such systems; research shows that AI virtual streamers' perceived relatability and knowledge outperform traditional indicators like viewer counts in terms of sales prediction [9]. Standardized assessment frameworks now incorporate technical performance (such as latency), emotional resonance, and conversion rates to operationalize these insights and provide useful standards for industry adoption [10].

3 Methodology

This study used publicly available data from the Douyin platform for AI broadcasters and manually recorded observational data for human streamers to examine audience retention patterns of three AI streamers (A, B, and C) and three human streamers (A, B, and C), as shown in Table 1 and Table 2 . While AI streamers were tracked for longer periods of time (t = 1–10 or t = 1–11), human broadcasters were observed for six hours (t = 0–5). Preprocessing the data involved keeping the raw hourly viewer counts to maintain temporal dynamics and visually inspecting audience sequences to detect abnormalities (such as abrupt spikes or dips). Linear regression models were used For fixed-rate audience attrition (y= a+ bt , where y= audience count, t= time, b= slope). However, exponential decay models exponential decay models for AI streamers with smooth percentage-based decay (y=y0ekt) , with parameters estimated using natural logarithm transformation. R2 was used to assess model fitness, with linear models being given priority unless exponential fits demonstrated noticeably higher R2 (e.g., AI Streamer C: R 2 = 0.68 vs. FCP= 0.54 for linear).

Streamer	Optimal Model	Attenuation Rate	R^2	FCP	Platform Rule
Human Streamer A	Linear	5.3%/hour	0.73	0.5	Initial audience > 150
Human Streamer B	Linear	7.3%/hour	0.65	0.33	Reduced exposure if <100
Human Streamer C	Segmented Linear	10%/hour (first 4h)	0.85	0.5	Initial audience > 200
AI Streamer A	Linear	7.7%/hour	0.7	0.7	Initial audience > 300
AI Streamer B	Segmented Linear	7.1%/hour(3-8h)	0.89	0.625	No support for low initial
AI Streamer C	Exponential	7.9%/hour	0.68	0.54	Intense promotion if >1000

Table 1
Model Comparison
Source： from publicly available live-streaming records on Douyin (China’s TikTok) and manual observational recordings

Fatigue resilience (FCP) was computed as the ratio of the overall streaming duration to the first hour when viewer counts fell below the average audience for the streamer. FCP=First Time Below Average Audience/ Total Streaming Duration, where Average Audience is calculated per streamer. Threshold Crossing is the earliest hour when viewer count fell below the average. Based on abrupt audience increases (≥10% in a single hour) and the lack of outside promotions, traffic compensation events were found. This was confirmed by platform metadata (e.g., "Recommended" tags for human streamers; unexplained spikes for AI streamers).

In order to determine significance, Welch's t-test compared the FCP values between the human and AI groups after descriptive data (mean audience, attrition rates, and FCP) were calculated for each streamer.

Traffic compensation events were detected based on two criteria: (1) abrupt audience surge (≥10% rise within one hour) and (2) lack of external promotions (e.g., no tags for external links, host announcements, or sponsored advertisements). "Confirmed Compensation" was the classification given to surges for human streams that included platform-generated labels (such as "Recommended" or "Top 100" tags), whereas "Suspected Compensation" was given to surges for AI streamers that lacked explanatory metadata (such as no discernible promotional triggers). The distinction between algorithm-driven platform interventions and natural audience variations was guaranteed by this dual-tagging method.

Streamer Type	**Streamer**	**Time Points (Audience Counts)**	**Avg Audience**	**Slope**
Human	A	T0:171, T1:167, T2:143, T3:161, T4:144, T5:118	150.5	-9.03
Human	B	T0:121, T1:100, T2:109, T3:99, T4:85, T5:79	98.5	-8.8
Human	C	T0:235, T1:230, T2:154, T3:143, T4:139, T5:161	178.2	-23.5
AI	A	T1:450, T2:387, T3:459, T4:553, T5:386, T6:440, T7:300, T8:330, T9:348, T10:134	372.6	-34.7
AI	B	T1:59, T2:71, T3:33, T4:31, T5:27, T6:12, T7:13, T8:12	29.6	-7.5
AI	C	T1:1357, T2:1417, T3:1402, T4:1238, T5:1065, T6:908, T7:917, T8:841, T9:790, T10:778, T11:814	1074.5	-62.3

Table 2.
Hourly Audience Dynamics and Attenuation Rates, by Streamer Type
Source: from publicly available live-streaming records on Douyin (China’s TikTok) and manual observational recordings

Linear models is applied to quantify fixed-rate audience attrition using the formula

(1) :

$$y=a+bt \tag{1}$$

where y= audience count, t= time, b = slope (attrition rate).

Exponential Regression is used as formula

(2) shown for AI streamers with smooth percentage-based decay:

$$y=y_0e^{kt} \tag{2}$$

Linearized via natural logarithm transformation for parameter estimation.

Evaluated using R2 Linear models were prioritized unless exponential models showed significantly higher R2 (e.g., AI Streamer C: R2=0.68 vs. 0.54 for linear).

4 Results

4.1 Integrated Analysis of Human Streamers

As illustrated in Table 3, Human streamers predominantly follow linear or segmented linear attenuation models, driven by interactive fluctuations and content adjustments. Streamer A (linear model, R^2=0.73) exhibits an audience decline of 9 viewers per hour (5.3% attenuation rate), with moderate fatigue resilience (FCP = 0.5) reflected in the audience rebound at t=3. Streamer B (linear model, R^2=0.65) demonstrates a faster decay rate (7.3%/hour), but low FCP (0.33) and platform penalties for audiences below 100 necessitate compressed streaming durations to control costs. Streamer C (segmented linear model, R^2=0.85) shows a steep initial decay of 10%/hour for t=0−4, followed by anomalous recovery at t=5 (FCP = 0.5), highlighting the importance of interactive strategies. Overall, human streamers rely on FCP to buffer decay but must avoid platform traffic penalties (e.g., audiences < 100) by optimizing retention through timed interventions.

Category	Human Streamers	AI Streamers
Traffic Compensation	**Confirmed Events**:	**Suspected Events**:
	Human A: Hour 3 (+12.6%, "Recommended" or "Top100" tag)	AI A: Hour 4 (+43.3%, no explanatory metadata)
	Human B: Hour 4 (+15.0%, "Recommended" or "Top100"tag)	AI C: Hour 2 (+4.4%, below threshold)
	Human C: Hour 5 (+15.8%, "Recommended" or "Top100" tag)	
FCP Resilience	Mean FCP: 0.28	Mean FCP: 0.55
	Highest: 0.33 (Human A/C)	Highest: 0.70 (AI A)
	Lowest: 0.17 (Human B)	Lowest: 0.45 (AI C)
Compensation Impact	Human A: 1 event → +12.6% audience surge (no FCP improvement)	AI A: 1 suspected event → +43.3% surge (high FCP maintained via consistency)

Table 3

Integrated Analysis Table: Traffic Compensation, FCP Resilience, and Strategic Recommendations
Source： from publicly available live-streaming records on Douyin (China’s TikTok) and manual observational recordings

4.2 Integrated Analysis of AI Streamers

AI streamers exhibit distinct attenuation patterns: Streamer A (linear model, R2=0.70) loses 35 viewers/hour (7.7%), with high FCP (0.7) indicating algorithmic adaptability (e.g., audience spikes at t=3). Streamer B follows a segmented linear model (R^2=0.89) with a moderated decay rate of 7.1%/hour and improved fatigue resilience (FCP=0.625). While it requires initial traffic support (e.g., paid promotions for audiences below 100), its segmented decay pattern indicates algorithmic adaptability to audience retention phases. Streamer C (exponential model, R2=0.68) decays smoothly at 7.9%/hour, leveraging initial high visibility (1,357 viewers) for platform promotion but requiring truncated low-efficiency periods post t=3. AI streamers generally depend on early traffic windows (e.g., first 3 hours) for high-density conversions. Streamer B’s segmented linear decay (7.1%/hour) demonstrates that algorithmic adaptability can stabilize retention in later phases, reducing reliance on external promotions. When comparing AI and human streamers, the analysis shows a sharp difference in fatigue resilience (FCP), with AI streamers showing a considerably greater mean FCP (0.55 vs. 0.28, $p < 0.01$). The algorithmic advantage of AI in maintaining audience retention through consistent content delivery is highlighted by this, while human broadcasters, especially Human B (FCP=0.17), need focused interventions to slow down rapid attrition. Traffic compensation events failed to improve FCP even if they momentarily increased audiences (e.g., +12.6% for Human A at Hour 3). This suggests that platform algorithms give short-term visibility precedence over long-term retention. However, because automated material is inherently stable, AI streams were able to maintain high FCP without heavily relying on compensation.

In order to overcome these results, a mixed operational approach is suggested: AI broadcasters are most stable when used in low-competition times (such as late-night slots), while human streamers should enhance interaction components during crucial attrition phases (such as Hour 1 for Human B). In order to provide transparent traffic distribution and maintain parity between human and AI hosts, platforms must adjust their compensation algorithms. Future research should evaluate the cross-platform validity of these dynamics and investigate real-time FCP changes using adaptive AI algorithms.

5 Discussion

5.1 Concluding remarks

The findings offer practical insights for e-commerce cost management. The outcomes of this study support the findings of Xu et al., who found that trust functions as a mediator and that parasocial relationships—such as viewers' emotional identification with streamers—significantly increase purchase intention [4]. With an FCP of 0.33 and a Attenuation Rate of 5.3%/hour, Human Streamer A's comparatively steady engagement rhythm (such as hourly Q&A sessions) created parasocial ties and supported the idea that trust is important for maintaining audience retention. However, by introducing the Fatigue Resilience Critical Point (FCP), this work departs from previous studies like Hu et al., which concentrated on the technical implementation of AI streamers (e.g., virtual avatar generation) [2], and Gao et al., which examined the effects of streamers on purchase intent [3]. The algorithmic consistency of AI streamers makes up for greater early attrition rates, according to this metric, which measures sustained retention capabilities (e.g., AI Streamer A’s FCP=0.7). Based on above research, human streamers should prioritize FCP-driven strategies (e.g., timed promotions) to extend effective streaming durations, while AI streamers benefit from early traffic exploitation and algorithmic tuning to mitigate rapid decay. Streamer B’s segmented linear decay

(7.1%/hour) suggests that adaptive algorithms can stabilize retention beyond initial phases, offering cost-efficient long-term streaming strategies. Platform operators could refine recommendation thresholds (e.g., adjusting visibility penalties) to balance fairness and efficiency. For researchers, this study proposes a hybrid modeling framework integrating linear, segmented, and exponential decay, adaptable to diverse streaming scenarios. Future work should expand datasets to 24-hour cycles and incorporate real-time external variables (e.g., ad spend, viewer demographics) for robust validation.

5.2 Limitations and Future Research Directions

This study has several limitations. First, the small sample size (e.g., 6-hour observations for human streamers) restricts generalizability. Second, external factors (e.g., holidays, platform algorithm updates) were not controlled, potentially confounding decay trends. Third, FCP metrics rely on manual recovery event counting, introducing subjectivity. Finally, exponential models for AI streamers were calculated using linearized log-transformations, which may underestimate nonlinear dynamics compared to direct nonlinear regression

To improve forecast accuracy, future research could include external elements like ad expenditure and viewer demographics using machine learning approaches (e.g., random forests) and extend temporal analysis to 24-hour cycles to evaluate the effects of circadian rhythm on audience attrition. Furthermore, partnerships with streaming services to obtain real-time recommendation logs may enhance algorithmic transparency and allow for accurate traffic compensation system validation. Lastly, to generalize results and guarantee robustness across different user behaviors and platform algorithms, cross-platform validation on several ecosystems (such as Twitch and TikTok) is essential.

Conclusion

This study reveals distinct audience attenuation patterns between human and AI streamers. Human streamers predominantly follow linear or segmented linear models due to interactive fluctuations, with attenuation rates ranging from 5.3% to 10% per hour. Fatigue resilience (FCP) values (0.33–0.5) highlight their capacity to recover audiences through engagement, though platform penalties necessitate strategic adjustments. In contrast, AI streamers exhibit hybrid trends: AI Streamer C aligns with an exponential decay model (7.9%/hour, R2=0.68). AI Streamer B aligns with a segmented linear model (R^2=0.89), showing moderated decay (7.1%/hour) and improved FCP (0.625), indicating algorithmic adaptability in sustaining audience retention. High initial audiences (e.g., 1,357 for AI Streamer C) trigger platform traffic boosts, but rapid decay post-critical timepoints demands optimized conversion timing.

References

[1] F.-L. Li et al., “AliMe Avatar: Multi-modal Content Production and Presentation for Live-streaming E-commerce,” in Proceedings of the 44th International ACM SIGIR Conference on Research and Development in Information Retrieval, Virtual Event Canada: ACM, Jul. 2021, pp. 2635–2636. doi: 10.1145/3404835.3464922.

[2] L. Hu et al., “A virtual character generation and animation system for E-commerce live streaming,” in Proceedings of the 29th ACM International Conference on Multimedia, Virtual Event China: ACM, Oct. 2021, pp. 1202–1211. doi: 10.1145/3474085.3481547.

[3] J. Gao, X. Zhao, M. Zhai, D. Zhang, and G. Li, “AI or Human? The Effect of Streamer Types on Consumer Purchase Intention in Live Streaming,” Int. J. Human–Computer Interact., vol. 41, no. 1, pp. 305–317, Jan. 2025, doi: 10.1080/10447318.2023.2299900.

[4] P. Xu, B. Cui, and B. Lyu, “Influence of Streamer’s Social Capital on Purchase Intention in Live Streaming E-Commerce,” Front. Psychol., vol. 12, Jan. 2022, doi: 10.3389/fpsyg.2021.748172.

[5] R. Wu, J. Liu, S. Chen, and X. Tong, “The effect of E-commerce virtual live streamer socialness on consumers’ experiential value: an empirical study based on Chinese E-commerce live streaming studios,” J. Res. Interact. Mark., vol. 17, no. 5, pp. 714–733, Oct. 2023, doi: 10.1108/JRIM-09-2022-0265.

[6] Z. Shao, “Understanding the switching intention to virtual streamers in live streaming commerce: innovation resistances, shopping motivations and personalities,” J. Res. Interact. Mark., vol. 19, no. 3, pp. 333–357, 2025.

[7] Y. Chen and X. Li, “Expectancy Violations and Discontinuance Behavior in Live-Streaming Commerce: Exploring Human Interactions with Virtual Streamers,” Behav. Sci., vol. 14, no. 10, p. 920, Oct. 2024, doi: 10.3390/bs14100920.

[8] Y. Zhang, X. Wang, and X. Zhao, “Supervising or assisting? The influence of virtual anchor driven by AI–human collaboration on customer engagement in live streaming e-commerce,” Electron. Commer. Res., Nov. 2023, doi: 10.1007/s10660-023-09783-5.

[9] X. Ji, “Influence of Virtual Live Streamers’ Credibility on Online Sales Performance,” Sage Open, vol. 14, no. 3, p. 21582440241271171, Jul. 2024, doi: 10.1177/21582440241271171.

[10] L. Zhang, J. Zhang, D. Wang, and J. Mu, “Development and Validation of an AI Virtual Streamer Scale for Live-Streaming E-Commerce,” Int. J. Human–Computer Interact., pp. 1–14, Oct. 2024, doi: 10.1080/10447318.2024.2411088.