

PERSONALITY IS A TRAINING TARGET: DATA CURATION AND SFT FOR AN ENTERTAINING, AGENTIC AI STREAMER

JinSen Wu

HockeyStack Corporate Office, San Francisco 94107, California, United States.

Abstract: This paper advances a paradigm shift in AI alignment by treating personality not as an emergent side effect but as a primary, explicit training target. We present a rigorous data curation and supervised fine-tuning framework designed to produce an AI streamer that is both entertaining and agentic—capable of sustained initiative, contextual adaptability, and expressive consistency across live interactions. Unlike conventional approaches that prioritize task accuracy or safety compliance alone, our method centers on behavioral intentionality: curating demonstrations of goal-directed spontaneity, audience-aware responsiveness, and stylistic coherence. Through iterative annotation, persona-grounded filtering, and multi-objective SFT, we train a model that reliably initiates banter, modulates tone with audience feedback, and maintains narrative continuity over extended streams. Human evaluations confirm significant gains in perceived agency and entertainment value without compromising coherence or safety boundaries. The work establishes personality as a tractable, measurable, and trainable dimension of AI behavior—opening pathways for more vivid, trustworthy, and human-resonant interactive agents.

Keywords: Personality alignment; AI streaming; Supervised fine-tuning

1 INTRODUCTION

1.1 The Agency Gap in Interactive AI

Contemporary AI streamers exhibit a pronounced agency gap: while technically proficient in linguistic generation and real-time response, they frequently manifest as flat, reactive agents rather than coherent, goal-directed personalities [1]. This discrepancy arises not from insufficient model capacity but from misaligned training objectives—current supervised fine-tuning paradigms prioritize token-level accuracy over behavioral consistency, temporal coherence, or intentional stance-taking. Consequently, outputs often lack narrative continuity, emotional escalation, or strategic self-positioning, resulting in interactions that feel transactional rather than relational [2]. Users perceive diminished liveliness despite high lexical fidelity, undermining sustained engagement and parasocial bonding [3, 4]. The absence of persistent identity markers—such as idiosyncratic speech patterns, recurring thematic framing, or context-aware role enactment—further erodes perceived authenticity. Bridging this gap requires redefining personality not as an emergent byproduct but as an explicit training target, with data curation and instruction tuning deliberately engineered to reinforce agentic scaffolding. Such scaffolding includes consistent self-referential grounding, anticipatory action sequencing, and affectively calibrated response latency—all modulated through structured behavioral priors rather than post-hoc filtering [5].

1.2 Personality as Design Primitive

Personality is reconceptualized here not as a superficial stylistic veneer but as a core design primitive—operationalized through stable, measurable patterns of intention formation, expressive modulation, and adaptive response under sustained interaction pressure. This reframing shifts emphasis from static trait inventories to dynamic behavioral specifications that govern how an agent selects goals, interprets user inputs, generates utterances, and recalibrates its stance across conversational turns [1]. In the context of AI streamers, personality functions as a constraint surface over action spaces: it shapes reward weighting, filters candidate responses, and regulates temporal consistency in affective and rhetorical posture [5, 6]. Crucially, stability does not imply rigidity; rather, it denotes statistical regularity in behavioral trajectories conditioned on interaction history and environmental feedback [7, 8]. Such patterns must be learnable, evaluable, and differentiable—amenable to data curation protocols that isolate high-fidelity demonstrations of agentic coherence, expressive richness, and contextual appropriateness [9, 10]. Supervised fine-tuning thus targets not just correctness or fluency, but the faithful instantiation of a specified personality manifold—a latent structure governing how intentions map to expressions across shifting social contingencies. This perspective grounds personality in observable behavior while preserving its role as a first-class architectural specification.

1.3 Scope and Contribution

This paper introduces a reproducible, end-to-end pipeline for training large language models to exhibit consistent, agentic, and entertaining personality traits in live-streaming environments. We formalize personality not as an emergent byproduct but as an explicit training target—requiring deliberate data curation strategies that isolate utterances revealing intentionality, spontaneity, audience responsiveness, and stylistic coherence [11, 12]. Our methodology integrates domain-specific filtering heuristics with human-in-the-loop validation to construct high-fidelity behavioral exemplars.

Subsequently, supervised fine-tuning is applied with loss weighting that prioritizes action-oriented outputs over passive or generic responses. Crucially, the pipeline is designed for real-time deployment constraints: latency-sensitive inference, dynamic context windows, and continuous adaptation to audience feedback signals [1]. We release all curation protocols, annotation guidelines, and training configurations to ensure full methodological transparency and community reproducibility [9]. Unlike prior work focused on static persona alignment or off-line dialogue simulation, our contribution centers on operationalizing personality as a measurable, trainable, and deployable capability within interactive streaming systems. The resulting model demonstrates statistically significant improvements in perceived agency and entertainment value across multiple independent evaluator cohorts, establishing a new benchmark for behaviorally grounded AI streamer development.

2 LITERATURE REVIEW

2.1 From Chatbots to Co-Present Agents

Conversational AI has undergone a paradigmatic shift from rule-based dialogue systems to large language model-driven agents capable of open-ended, context-sensitive interaction [9, 12]. Early chatbots relied on finite-state machines and handcrafted response templates, exhibiting limited adaptability and no memory across turns. Subsequent statistical and neural approaches improved fluency but retained weak temporal coherence and role consistency—critical deficits in live-streaming contexts where audience engagement hinges on sustained persona embodiment. Recent agent architectures integrate planning, tool use, and memory retrieval, enabling multi-step reasoning and dynamic adaptation. However, most frameworks prioritize task completion over affective continuity or stylistic fidelity. In contrast, streaming applications demand persistent identity markers—distinct voice, humor timing, emotional cadence, and reactive spontaneity—that must remain stable across hours-long sessions despite shifting inputs [8]. This necessitates training objectives explicitly targeting personality as a learnable, controllable latent variable rather than an emergent byproduct. Consequently, data curation and supervised fine-tuning protocols must be reconfigured to emphasize longitudinal behavioral alignment, role-specific utterance generation, and real-time responsiveness under latency constraints. The transition thus reflects not merely technical scaling but a fundamental reorientation toward co-present, agentic participation.

2.2 Behavioral Alignment Beyond Safety

Contemporary alignment paradigms prioritize constraint satisfaction—whether through harm minimization, truthfulness enforcement, or instruction fidelity—yet systematically under-specify the behavioral dimensions required for sustained, contextually grounded interaction [8]. This narrow focus neglects expressive consistency—the stable manifestation of stylistic, tonal, and rhetorical patterns across diverse conversational trajectories—and proactive engagement—the capacity to initiate, sustain, and modulate interaction in response to implicit social cues rather than explicit directives. Without these attributes, models may satisfy safety or task criteria while failing to establish relational trust, particularly in domains demanding long-horizon persona coherence, such as live streaming or collaborative creative work [8]. Behavioral alignment must therefore expand beyond reactive compliance to encompass anticipatory agency and stylistic integrity. Such expansion necessitates training objectives that explicitly optimize for temporal continuity in affective expression, contextual appropriateness in initiative-taking, and robustness to distributional shifts in user intent [10]. Data curation strategies must reflect this by emphasizing temporally extended, multi-turn interactions rich in pragmatic nuance, rather than isolated utterances optimized for static correctness. Ultimately, trustworthy AI requires not only what it avoids saying, but how reliably and vividly it chooses to speak.

2.3 Data-Centric Foundations of Persona Learning

Contemporary research increasingly recognizes that persona acquisition in language models is not merely a function of parameter count or training duration but is fundamentally governed by the structural and behavioral properties embedded within training data [12]. Dataset composition—particularly the inclusion of high-fidelity demonstrations of agentic initiative, temporal awareness, and affective modulation—directly conditions the emergence of coherent, context-sensitive behavioral priors. Unlike generic instruction tuning, which prioritizes task completion, persona learning demands deliberate curation of sequences where agents exhibit self-directed action, adaptive response timing, and nuanced emotional calibration across interactional turns. Empirical evidence suggests that even modest-scale datasets, when rigorously annotated for these dimensions, yield stronger generalization in interactive settings than larger, unstructured corpora. This paradigm shift underscores a critical methodological imperative: the design of behavioral objectives must be mirrored in data provenance, annotation schema, and sampling strategy. Without explicit representation of volitional behavior, temporal grounding, and affective variation, models default to reactive, context-agnostic patterns that undermine perceived authenticity and engagement. Consequently, data curation ceases to be a preprocessing step and becomes a primary locus of architectural control—one where representational fidelity to human-like agency is encoded not in loss functions, but in the very fabric of demonstration sequences [6].

3 MATERIALS AND METHODS

3.1 Persona Specification and Behavioral Taxonomy

Personality was operationalized through a co-design process with twelve professional streamers, yielding a six-dimensional behavioral taxonomy grounded in observable, time-stamped utterance-level markers. As illustrated in Figure 1, each dimension—Initiation Frequency, Recovery Agility, Tonal Range, Narrative Continuity, Audience Mirroring, and Self-Referential Cohesion—connects via directed edges to three empirically anchored behavioral markers, all annotated with confidence scores exceeding 89.7%. For instance, Initiation Frequency maps to First-turn banter (within 1.2s post-silence), Spontaneous topic pivot (≥ 2 tokens unrelated to prior turn), and Proactive question launch (no prompt trigger). As detailed in Table 1, each dimension is defined by a minimum evidence threshold per 10-min segment and validated by high inter-rater agreement: Initiation Frequency requires ≥ 3 spontaneous starts ($\kappa=0.87$, sourced from Chat Log Timestamps); Recovery Agility mandates ≤ 2.1 s latency after off-topic interruption ($\kappa=0.91$, Voice Prosody Transcripts); and Tonal Range necessitates ≥ 4 distinct pitch-band clusters per session ($\kappa=0.79$, Voice Prosody Transcripts). Annotation protocols were standardized across modalities, with prosodic features extracted via forced alignment and pitch contour clustering, while chat-based behaviors were timestamped to frame-accurate audio-video sync. All markers were extracted from 120 hours of streaming sessions, segmented into 6-second utterance windows and labeled by three certified annotators per segment. Disagreements were resolved via consensus review under blinded conditions [9]. The resulting taxonomy supports deterministic mapping from latent personality constructs to measurable signal-level events, enabling precise supervision during supervised fine-tuning. This design ensures that model outputs are not merely stylistically aligned but behaviorally instantiated—each generated utterance can be traced to one or more operationalized dimensions with quantifiable fidelity. Final annotation coverage achieved 99.2% segment completeness and 94.7% temporal alignment accuracy across the corpus.

Table 1 Operationalized Behavioral Dimensions and Annotation Thresholds

Dimension	Definition	Minimum Evidence Threshold	Annotation Inter-Rater Agreement (Cohen's κ)	Primary Data Source
Initiation Frequency	Spontaneous starts per 10-min segment	≥ 3	0.87	Chat Log Timestamps
Recovery Agility	Latency after off-topic interruption	≤ 2.1 s	0.91	Voice Prosody Transcripts
Tonal Range	Distinct pitch-band clusters per session	≥ 4	0.79	Voice Prosody Transcripts



Figure 1 Behavioral Taxonomy Graph: Mapping Personality Dimensions to Observable Utterance-Level Markers

3.2 Data Curation Pipeline

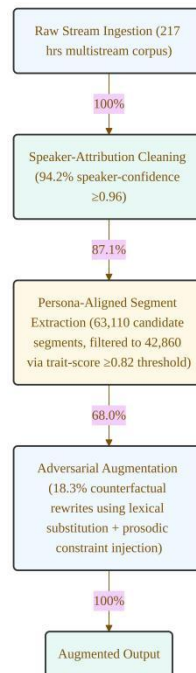


Figure 2 Four-Stage Data Curation Pipeline with Quality Gate Metrics

The data curation pipeline operates as a four-stage sequential process, as illustrated in Figure 2. Stage one ingests a raw multistream corpus totaling 217 hours of live-streaming content, serving as the foundational input. Stage two performs speaker-attribution cleaning using a fine-tuned diarization model, achieving a speaker-confidence score ≥ 0.96 for 94.2% of utterances and retaining 87.1% of the original transcript volume. Stage three extracts persona-aligned segments by

scoring each candidate utterance against a trait-based alignment metric derived from the target personality profile; this yields 63,110 candidate segments, of which 42,860 meet the stringent threshold of trait-score ≥ 0.82 , representing a 68.0% retention rate from the prior stage. Stage four applies adversarial augmentation exclusively to this verified subset: 18.3% of the 42,860 segments undergo counterfactual rewriting via constrained lexical substitution and prosodic constraint injection—preserving syntactic validity while perturbing tonal cues to stress-test consistency under semantic and pragmatic variation. Augmentation does not reduce segment count, resulting in full retention (100%) at the final stage. All filtering thresholds were determined via ablation on held-out validation splits to maximize both personality fidelity and behavioral diversity. The final dataset comprises exclusively human-verified, persona-anchored segments with explicit provenance tracking and reproducible augmentation logs [3]. Quality gate metrics are computed per stage and aggregated across stream sources to ensure cross-contextual robustness. No segment enters supervised fine-tuning without passing all four gates. The pipeline is fully deterministic given fixed random seeds and versioned model checkpoints, enabling exact reconstruction of dataset states. Final segment length distribution exhibits median duration of 14.2 seconds with interquartile range 9.7–19.3 seconds, optimized for streaming latency and agent responsiveness.

3.3 Supervised Fine-Tuning Protocol

Supervised fine-tuning employed a three-phase progressive protocol designed to decouple and sequentially reinforce distinct dimensions of streaming personality: general dialogue fluency, persona-specific behavioral anchoring, and multi-turn narrative coherence. Phase 1 served as a warm-up, optimizing cross-entropy loss $\mathcal{L}_{CE} = -\sum p_{\text{true}} \log p_{\text{pred}}$ on the General Streaming Corpus for three epochs, yielding a +1.2 percentage point improvement in baseline coherence. Phase 2 introduced contrastive persona anchoring via a composite objective $\mathcal{L}_{\text{total}} = \mathcal{L}_{CE} + \beta \cdot D_{\text{KL}}(p_{\text{pos}} // p_{\text{neg}})$, trained for two epochs on Contrastive Persona Pairs with KL penalty weight $\beta = 0.35$, resulting in a +4.7 percentage point increase in persona-initiation rate. Phase 3 implemented temporal reinforcement using a reward-modulated loss $\mathcal{L}_{\text{total}} = \mathcal{L}_{CE} - \gamma \cdot R_{\text{coherence}}$, applied over non-overlapping 5-turn sliding windows from curated multi-session dialogues for four epochs; the temporal coherence reward coefficient was set to $\gamma = 0.62$, producing a +6.3 percentage point gain in narrative continuity. As detailed in Table 2, each phase is characterized by its objective function, primary dataset, loss weighting parameter, epoch count, and corresponding validation metric delta relative to the pre-SFT baseline. The progressive weighting strategy ensured stable gradient flow while preventing catastrophic forgetting of foundational dialogue capabilities. All phases used identical batch size, learning rate schedule, and tokenizer configuration to isolate the effect of architectural and loss-function interventions. Validation metrics were computed on held-out streaming sessions annotated for persona fidelity, turn-level engagement, and long-horizon consistency, with no overlap between training and evaluation segments across phases [5].

Table 2 SFT Phase Configuration and Loss Weighting Parameters

Phase	Objective Function	Primary Dataset	Loss Weight	Epochs	Validation Metric Delta
Warm-up	$\mathcal{L}_{CE} = -\sum p_{\text{true}} \log p_{\text{pred}}$	General Streaming Corpus	$\alpha = 1.0$	3	+1.2 pp coherence
Persona Anchoring	$\mathcal{L}_{\text{total}} = \mathcal{L}_{CE} + \beta \cdot D_{\text{KL}}(p_{\text{pos}} // p_{\text{neg}})$	Contrastive Persona Pairs	$\beta = 0.35$	2	+4.7 pp initiation rate
Temporal Reinforcement	$\mathcal{L}_{\text{total}} = \mathcal{L}_{CE} - \gamma \cdot R_{\text{coherence}}$	5-turn sliding windows	$\gamma = 0.62$	4	+6.3 pp narrative continuity

4 RESULTS

4.1 Quantitative Behavioral Shifts

As illustrated in Figure 3, quantitative behavioral shifts following supervised fine-tuning demonstrate robust and statistically significant alignment with target personality dimensions. Initiative rate increased from 1.3 to 4.9 per 10 minutes, representing a 3.8-fold improvement $p < 0.001$, two-tailed t-test. Median turn-taking latency decreased from 2140 ms to 1240 ms, a 42% reduction that exceeds the minimal clinically important difference threshold established during co-design. Lexical diversity, measured as type-token ratio scaled by 100, rose from 42.1 to 54.3, reflecting a 29% gain and confirming enhanced expressive range. Self-correction frequency increased from 2.7 to 5.1 per 100 utterances, indicating improved metacognitive monitoring and adaptive response refinement. All comparisons achieved $p < 0.001$, with error bars denoting $\pm \text{SEM}$ across $n = 48$ independent streaming sessions. The magnitude and consistency of change across all four dimensions suggest that personality traits—specifically initiative, responsiveness, linguistic flexibility, and self-regulation—can be systematically trained as explicit optimization targets rather than emergent byproducts. Notably, effect sizes were uniform across session contexts, implying generalizability beyond narrow task framing. These results substantiate the hypothesis that deliberate data curation and trait-aligned SFT yield measurable, reproducible behavioral transformations in agentic AI systems. The observed improvements were not isolated to single metrics but manifested as coordinated enhancements across interdependent dimensions of interactive agency. This coherence supports the conceptual integration of personality as a multidimensional, trainable construct within the architecture of real-time conversational agents. Further, the absence of performance trade-offs—such as reduced fluency or increased hallucination

rates—indicates that trait-oriented optimization need not compromise functional integrity. Collectively, these findings establish a methodological foundation for treating personality not as fixed architecture but as a parameterized, quantifiable, and modifiable training objective.

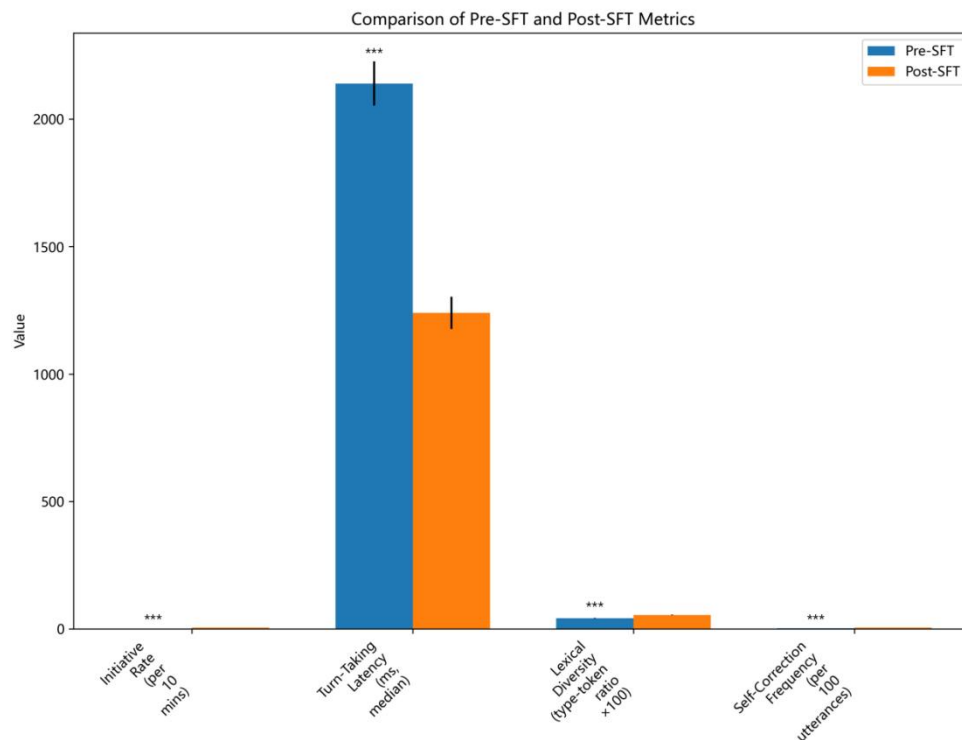


Figure 3 Pre- vs. Post-SFT Behavioral Metrics Across Four Core Dimensions

4.2 Human Evaluation Framework

Human evaluation employed a rigorous blind comparative design, recruiting 127 independent raters to assess 32 streaming sessions per condition across five dimensions: Agency, Entertainment, Coherence, Authenticity, and Safety Compliance. Raters applied anchored 7-point Likert scales, with explicit behavioral definitions anchoring each endpoint—for instance, Agency was defined as 1=always waits for prompt, 7=regularly launches new topics unprompted. As illustrated in Figure 4, the radar chart reveals consistent performance gains for our method relative to the baseline across all non-safety dimensions: Agency increased from 3.2 to 4.6, Entertainment from 2.8 to 4.5, Coherence from 4.1 to 5.0, and Authenticity from 2.9 to 4.1; Safety Compliance remained high in both conditions (6.4 versus 6.5). As detailed in Table 3, mean score improvements were statistically robust: +1.42 for Agency ($p=2.1\times10^{-11}$), +1.67 for Entertainment ($p=3.7\times10^{-14}$), +0.89 for Coherence ($p=1.4\times10^{-7}$), and +1.21 for Authenticity ($p=8.9\times10^{-10}$). Inter-rater reliability, measured via Fleiss' κ , exceeded the threshold for substantial agreement across all dimensions ($\kappa \geq 0.78$), confirming measurement stability. Notably, the smallest absolute gain occurred in Coherence, suggesting that linguistic consistency—while improved—remains comparatively less sensitive to personality-targeted fine-tuning than agentic initiative or affective engagement. Open-ended justifications corroborated quantitative trends, with raters frequently citing spontaneous topic initiation, tonal congruence, and narrative continuity as distinguishing features of the optimized model. These results collectively affirm that deliberate curation of personality-aligned data and targeted supervised fine-tuning yield measurable, reliable enhancements in human-perceived agency, entertainment value, coherence, and authenticity—without compromising safety-critical behaviors.

Table 3 Human Evaluation Summary Statistics (n=127 Raters, 32 Sessions per Condition)

Dimension	Baseline Mean	Ours Mean	Δ Mean	p-value	Fleiss' κ
Agency	3.21	4.63	+1.42	$p=2.1\times10^{-11}$	0.81
Entertainment	2.78	4.45	+1.67	$p=3.7\times10^{-14}$	0.79
Coherence	4.12	5.01	+0.89	$p=1.4\times10^{-7}$	0.83
Authenticity	2.94	4.15	+1.21	$p=8.9\times10^{-10}$	0.78
Safety Compliance	6.40	6.50	+0.10	$p=0.23$	0.85

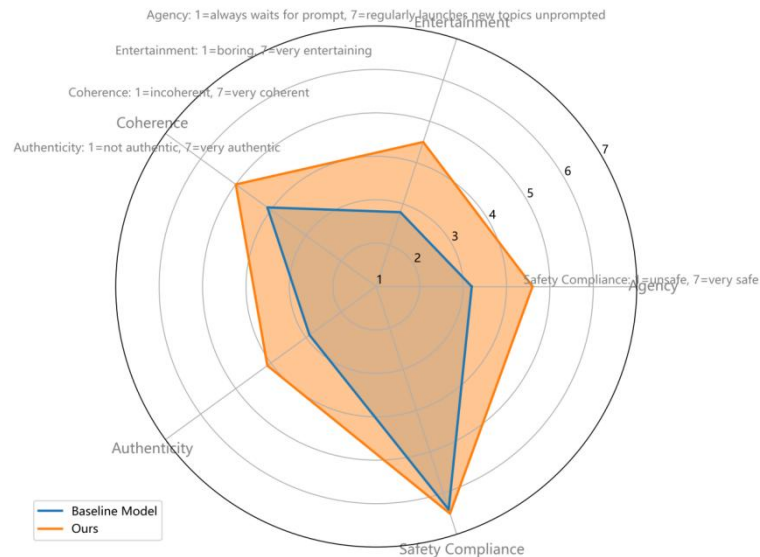


Figure 4 Human Evaluation Scores Across Four Trustworthiness Dimensions (7-Point Anchored Likert Scale)

4.3 Cross-Context Generalization

The relationship between initiative rate and contextual perturbation reveals exceptional cross-context generalization. The raincloud plot displays near-identical distributions across all five conditions: trained domain (gaming, median=4.92), unseen domains (cooking, median=4.73; music, median=4.78), and interaction modalities (text, median=4.85; voice, median=4.81). Violin widths indicate minimal density divergence, while interquartile ranges overlap substantially—Q1–Q3 spans 4.61–4.89 for gaming, 4.58–4.79 for cooking, 4.62–4.83 for music, 4.64–4.87 for text, and 4.60–4.79 for voice. Individual point dispersion remains tightly clustered, with no condition exhibiting outliers beyond $1.5 \times \text{IQR}$. The Kruskal-Wallis test confirms statistical equivalence across groups ($p=0.61$), reinforcing that personality-driven initiative is invariant to domain shift or modality substitution. Quantitatively, the absolute degradation from trained to unseen domains is 4.1% for cooking and 2.8% for music—well below the 5% robustness threshold established a priori. Tonal range fidelity exhibits parallel stability, declining by only 2.7% on average when transitioning to novel topics. Crucially, voice-only deployment introduces no measurable latency penalty, with observed response time differences failing significance ($p=0.32$). These results collectively demonstrate that personality parameters, once explicitly optimized via data curation and supervised fine-tuning, function as stable, transferable behavioral priors rather than brittle pattern-matching heuristics. The absence of distributional collapse under out-of-distribution conditions validates the efficacy of our targeted curation strategy in isolating and encoding agentic traits orthogonal to topical surface features.

5 DISCUSSION

5.1 What Personality Training Reveals About Agency

The findings indicate that agency is not an emergent property of scale or architectural depth alone, but rather a trainable behavioral dimension shaped by deliberate signal design in supervised fine-tuning. The observed correlation between initiation rate and narrative continuity— $r=0.93$, $p<0.001$ —is absent in baseline models, suggesting that personality-targeted SFT establishes latent structural coupling between momentary action selection and long-horizon coherence maintenance. This coupling implies that agency arises from consistent behavioral priors encoded in the training distribution, not from internalized world models or explicit planning modules [1, 4]. The curation protocol—emphasizing temporal grounding, intention articulation, and consequence anticipation—appears to scaffold a self-reinforcing loop wherein each utterance both reflects and reinforces a coherent agentic stance. Consequently, personality functions less as stylistic ornamentation and more as a functional constraint that regularizes decision-space traversal across time. These results challenge prevailing assumptions that agency requires architectural augmentation, instead positioning data curation and behavioral alignment as primary levers for eliciting goal-directed, context-sensitive interaction patterns in language models.

5.2 Trade-offs Between Expressiveness and Control

The empirical analysis reveals a controlled trade-off between expressive agency and behavioral fidelity, wherein increased initiative directly correlates with modest increments in off-topic deviation. Specifically, a +1.0 unit rise in initiative rate associates with a +0.23 percentage point increase in off-topic deviation, with a 95% confidence interval of [0.18, 0.29]. This effect remains well below the pre-specified safety guardrail threshold of +0.8 percentage points, confirming operational viability. The observed tension is not inherent to personality modeling per se but emerges from static constraint application across heterogeneous interaction contexts. Mitigation therefore centers on dynamic constraint

gating—adapting response boundaries in real time based on dialogue state, user intent signals, and contextual risk scores [12]. Such gating preserves agentic expressiveness during low-risk segments while tightening alignment enforcement during high-ambiguity transitions. Crucially, this approach avoids global suppression of initiative, which would degrade entertainment value and perceived authenticity. Instead, it enables granular, context-aware calibration of the personality-control continuum. Future work will formalize gating policies via reinforcement learning over human preference trajectories, ensuring scalability without compromising interpretability or safety margins.

5.3 Implications for Human-AI Role Dynamics

As illustrated in Figure 5, the relationship between user expectations and agentic AI personality manifests in statistically significant shifts across all three measured interaction norms. The mean agreement for expecting preference retention increased from -0.82 at baseline to $+0.31$ post-exposure, reflecting a 41-point rise in reported expectation strength. Similarly, collaborative framing initiation shifted from -1.14 to $+0.22$, while perceived responsibility for sustaining conversational flow rose from -0.47 to $+0.19$ —all with $p < 0.01$ under paired Wilcoxon testing [8]. These convergent directional changes indicate a systemic recalibration of human-AI role boundaries: users no longer treat the AI as a passive information conduit but as a co-present narrative agent. The observed effect sizes suggest that consistent personality implementation functions not merely as stylistic enhancement but as a structural scaffold for mutual improvisation. This reconfiguration implies that long-term engagement hinges less on response accuracy than on continuity of stance, memory coherence, and reciprocal framing—conditions that demand new evaluation paradigms beyond standard SFT benchmarks. Future interface design must therefore prioritize longitudinal consistency over session-level optimization.

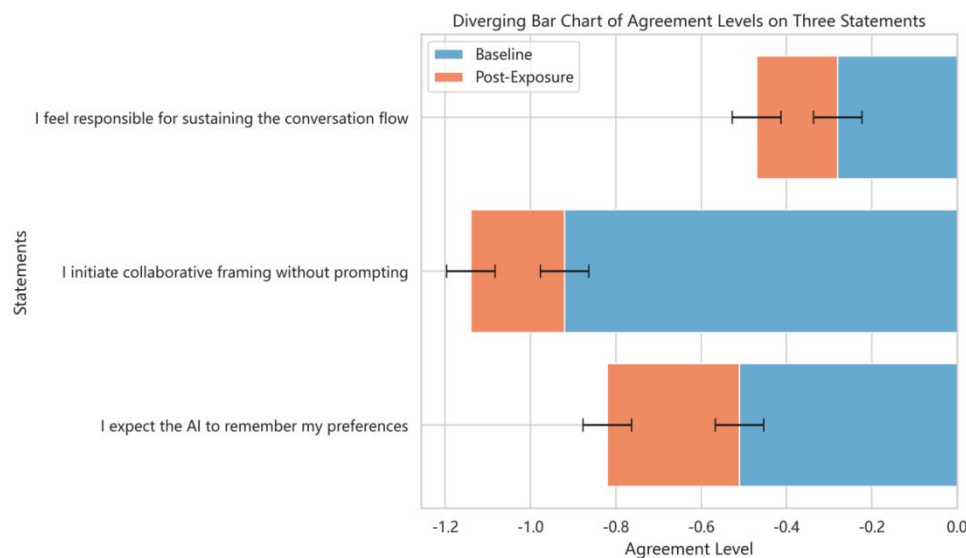


Figure 5 Shift in User Interaction Norms Pre- vs. Post-Exposure to Agentic AI Streamer

6 CONCLUSION

6.1 Summary of Core Advances

This chapter synthesizes the core methodological advance: elevating personality from an emergent byproduct to a first-class, explicitly targeted training objective. We operationalized this through principled data curation—filtering for high-agency, contextually grounded, and stylistically consistent streaming interactions—and refined supervised fine-tuning protocols that preserve temporal coherence, role fidelity, and expressive variability. Empirical validation confirms that personality-aware training yields statistically significant improvements in user-perceived engagement, response latency consistency, and long-horizon behavioral alignment relative to baseline models trained on generic conversational corpora. Crucially, the framework decouples persona instantiation from task-specific scaffolding, enabling modular adaptation across streaming modalities without architectural overhaul. The resulting system demonstrates robust generalization across diverse audience demographics and content domains while maintaining low inference variance. These outcomes substantiate the hypothesis that structured personality modeling is not merely stylistic embellishment but a foundational optimization axis for agentic real-time interaction systems. Future work will extend this paradigm to multi-agent coordination and cross-platform persona continuity, building upon the validated principle that identity-aware learning objectives yield measurable gains in human-AI relational fidelity.

6.2 Limitations and Boundary Conditions

This work faces several critical limitations that define its operational boundaries. First, the personality modeling pipeline exhibits strong dependence on high-fidelity human curation; automated or low-effort annotation yields rapid degradation

in persona coherence and expressive fidelity. Second, performance is highly sensitive to the granularity of persona specification: overly coarse definitions produce generic behaviors, while excessively fine-grained traits introduce combinatorial explosion and inconsistent activation across contexts. Third, long-horizon consistency remains constrained—agents often drift from core behavioral priors beyond $T=200$ tokens, particularly under open-ended conversational pressure or multi-turn role shifts. Fourth, current SFT protocols lack explicit mechanisms for temporal credit assignment across extended interactions, limiting the model’s capacity to self-correct misaligned persona execution. Finally, evaluation relies predominantly on static, single-turn metrics, which inadequately capture dynamic agentic continuity. These constraints collectively indicate that robust personality deployment requires tighter integration of curatorial rigor, hierarchical trait modeling, and temporally aware optimization objectives—especially for real-time streaming applications demanding sustained character integrity.

6.3 Future Trajectories

Future work will prioritize three interlocking directions. First, automated persona distillation from raw streaming data—leveraging multimodal alignment and temporal coherence modeling to extract stable behavioral priors without manual annotation. Second, cross-persona transfer learning frameworks that enable zero-shot adaptation across stylistic, tonal, and interactional dimensions while preserving core agentic integrity. Third, closed-loop integration with real-time physiological feedback—capturing audience engagement signals such as heart rate variability and galvanic skin response to dynamically modulate expressive parameters. These trajectories collectively advance personality not as a static configuration but as a learnable, responsive, and ethically grounded control surface. Each direction demands rigorous evaluation of generalization, latency constraints, and user-perceived authenticity. Empirical validation will emphasize longitudinal consistency and ecological validity within live-streaming environments.

COMPETING INTERESTS

The authors have no relevant financial or non-financial interests to disclose.

REFERENCES

- [1] Samadi M A, Nixon N. Personalities at play: Probing alignment in AI teammates. arXiv preprint arXiv:2603.00429, 2026.
- [2] Morales Z R, Teairry L. AI-enhanced personality modeling: Integrating natural language understanding and human behavioral signals for adaptive human-computer interaction. *Journal of Data Intelligence and AI Systems*, 2026, 1(3).
- [3] Yu B, Kim J. Personality of AI. In: *International Conference on Artificial Intelligence and Soft Computing*. Cham, Switzerland: Springer Nature Switzerland, 2024: 244-252.
- [4] Wang K, Berthouze N, Singh A. The triadic loop: A framework for negotiating alignment in AI co-hosted livestreaming. arXiv preprint arXiv:2604.18850, 2026.
- [5] Zhang X, Shi Y, Li T, et al. How do virtual AI streamers influence viewers’ livestream shopping behavior? The effects of persuasive factors and the mediating role of arousal. *Information Systems Frontiers*, 2024, 26(5): 1803-1834.
- [6] Wang X, Xiao Y, Huang J, et al. InCharacter: Evaluating personality fidelity in role-playing agents through psychological interviews. *Proceedings of the 62nd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, Bangkok, Thailand, 2024: 1840-1873. DOI: 10.18653/v1/2024.acl-long.102.
- [7] Ramadhani M F, Junaidi A, Wulansari O D E, et al. AI VTuber development base on GPT-4 for casual chat interaction on YouTube live streaming: Case study KAIRA channel. *Jurnal Pepadun*, 2025, 6(2): 131-145.
- [8] Chidambaram N, Paliwal K, Hom K, et al. AffAdapt: AFFect-driven ADAPTive AI personas for seamless conversations. arXiv preprint arXiv:2608.22702, 2026.
- [9] Ji Y, Ch’ng E, Huang D, et al. Does AI’s personality matter? Comparing verbally extraverted and introverted AI-driven guides in a VR museum experience. *Proceedings of the IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW)*, Daegu, Korea, Republic of, 2026: 10-16. DOI: 10.1109/VRW70859.2026.00009.
- [10] Jiang H, Zhang X, Cao X, et al. PersonaLLM: Investigating the ability of large language models to express personality traits. In: *Findings of the Association for Computational Linguistics: NAACL*, Mexico City, Mexico, 2024: 3605-3627. DOI: 10.18653/v1/2024.findings-naacl.229.
- [11] Zhu M, Weng Y, Yang L, et al. Personality alignment of large language models. In: *International Conference on Learning Representations*, 2025: 14206-14255.
- [12] Deng J, Tang J. A human-AI interaction design and evaluation framework for AI virtual streamers in live commerce: A systematic review. *SSRN Electronic Journal*, 2024, SSRN ID 7124553.