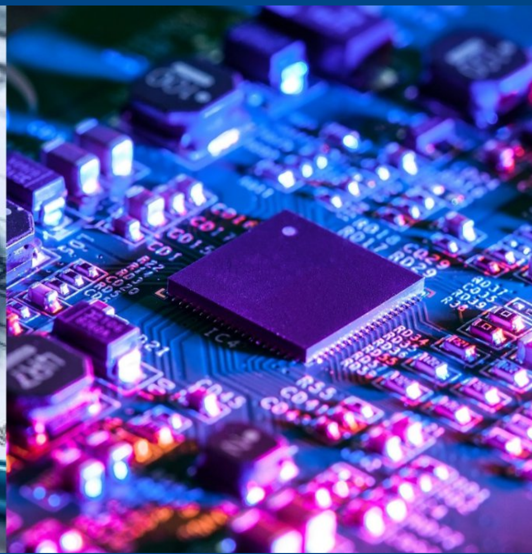
ISSN: 2582-7219

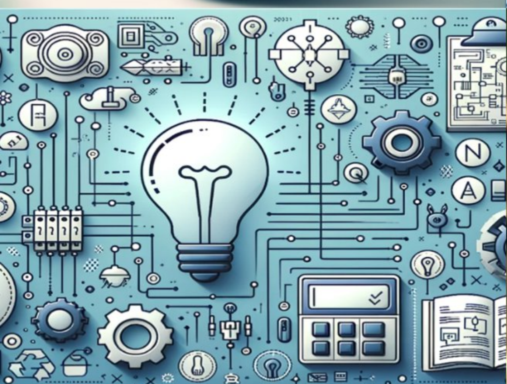




International Journal of Multidisciplinary Research in Science, Engineering and Technology

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)



Impact Factor: 8.206

Volume 9, Issue 4, April 2026

www.ijmrset.com ijmrset@gmail.com +91 6381 907 438 +91 63819 07438

© 2026 IJMRSET | Volume 9, Issue 4, April 2026 | DOI:10.15680/IJMRSET.2026.0904220

ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 8.206| ESTD Year: 2018|

International Journal of Multidisciplinary Research in Science, Engineering and Technology (IJMRSET)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

Assessing the Narrative on AI Music / Music Tools Using Youtube Content and Comments

Vardhan A. Shetty, Dr. Avinash Rana,

MBA, Business Analytics, CMS Business School,JAIN (Deemed-to-be University), Bengaluru, Karnataka, India

Associate Professor, Business Analytics, CMS Business School,JAIN (Deemed-to-be University), Bengaluru, Karnataka, India

ABSTRACT: This paper examines the impact of AI music generating tools on audiences and creators, and their implications for the broader music industry. The music industry currently faces challenges in the acceptance of AI integration. This investigation gathers textual data from 84 YouTube videos on the topic, comprising 37,157 viewer comments and 84 video transcripts. The analysis employs a multi-method NLP pipeline—VADER sentiment analysis, LDA topic modelling, transformer-based emotion classification, and aspect-based sentiment analysis (ABSA). Results reveal that AI music tools receive predominantly positive feedback (60.97%), yet a substantial critical minority (22.58%) expresses concerns around artistic authenticity, musician livelihoods, and copyright ethics. Content framing effects are significant: tutorial content elicits the highest positivity (67.19%), while comparison and review content produces the most scepticism (25.71% negative). The paper concludes that AI music tools show potential for mainstream acceptance, but industry stakeholders must address ethical and authenticity concerns proactively.

KEYWORDS: Music generating tools, AI integration, user sentiment, topic modelling, sentiment analysis, YouTube, VADER, NLP

I. INTRODUCTION

The global recorded music industry reached $31.7 billion in revenues in 2025—the eleventh consecutive year of expansion—driven primarily by digital streaming, which accounts for 69.6% of total revenues and 837 million paid subscribers worldwide (IFPI, 2026). Against this backdrop of remarkable resilience, a disruptive technological force has emerged: AI-powered music generation tools capable of producing full-length tracks from text prompts in seconds.

The AI-in-music market was valued at approximately $4.48 billion in 2025 and is projected to reach $5.55 billion by 2026 at a CAGR of 23.7%, with longer-term forecasts reaching $12.86 billion by 2035. Leading platforms—Suno, Udio, AIVA, and Soundraw—have democratised production, enabling creators without formal training to generate professional-quality music affordably. Yet this democratisation provokes a parallel discourse around copyright infringement, artistic authenticity, and the economic displacement of professional musicians, with projections suggesting AI may threaten up to 24% of traditional creators' income by 2028.

Despite significant scholarly interest, a gap exists in research on naturalistic, large-scale public sentiment about AI music tools. Most prior work relies on controlled surveys or experimental designs, overlooking the rich, unfiltered commentary found on platforms like YouTube. This study addresses that gap by systematically mining 37,157 YouTube comments across 84 videos to assess viewer sentiment, identify key acceptance factors, and derive actionable implications for developers, creators, and industry stakeholders.

1.1 Research Problem and Objectives

The fundamental research problem concerns understanding how YouTube audiences perceive and respond to AI-integrated music generating tools. Despite abundant viewer commentary, organisations and researchers lack systematic insight into public mood regarding adoption and acceptance of these technologies. This study's primary objective is to classify and interpret sentiment patterns in YouTube comment data. Secondary objectives include: (1) building and validating a sentiment classification pipeline; and (2) identifying key acceptance factors—perceived utility, ethical concerns, job displacement anxieties, ease of use, and authenticity.

ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 8.206| ESTD Year: 2018|

International Journal of Multidisciplinary Research in Science, Engineering and Technology (IJMRSET)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

1.2 Research Hypotheses

H1 (Sentiment Polarity): Comments display a positively skewed distribution rather than an equal split across sentiment categories.
H2 (Content Type): Promotional/tutorial content attracts higher positivity than evaluative/comparison content.
H3 (Topic–Sentiment): Topics centring on copyright, job loss, or authenticity correlate with more negative sentiment than tool-usage topics.
H4 (Emotional Dimension): Happiness is the modal discrete emotion, with fear and anger more prevalent in negative comments.

1.3 Review of Literature

Early AI music systems, built on Digital Audio Workstation (DAW) infrastructures, treated music as controllable processes. LSTM-based generative models trained on thousands of iterations showed superior convergence to classical RBM algorithms, establishing proof-of-concept for autonomous generation (Pan et al., 2022). Commercially, these foundations now underpin platforms such as Suno and Udio.

In educational settings, AI-driven music platforms using RBF algorithms and MVC frameworks yield significantly higher learner satisfaction than conventional methods (Chen, 2022; Ma, 2021). Self-Determination Theory analysis confirms that techno-philia and positive experience predict flow and perceived learning efficiency (Wang & Liu, 2025). UTAUT-based studies further show that social influence and personal innovativeness—not hedonic motivation alone—drive adoption of AI music materials in classrooms (Weng et al., 2025).

On the perceptual side, sentiment analysis of YouTube comments about Queen's AI-remixed track revealed a sharp divide: professional legacy works elicit negativity (threats to authenticity, artistic integrity) while amateur outputs are embraced as democratising preservation (Pfeiffer, 2025). Experimentally, identical performances rated lower in likability and emotional valence when labelled "AI-generated" confirm perceptual bias independent of musical expertise (Ansani, 2025). Industry discourse analysis at trade conferences reveals conflicting stances—protectionist calls for "responsible AI," liberalising cost-cutting pushes, and union compromises—exposing governance tensions (Campos Valverde, 2026).Gaps persist: most evidence is cross-sectional and geographically confined to China, Europe, or North America; longitudinal studies and cross-genre evaluations are scarce; and few works integrate technical, educational, perceptual, and fairness analyses into a single framework. This study directly addresses the gap in naturalistic, large-scale sentiment analysis of YouTube discourse.

II. RESEARCH METHODOLOGY

2.1 Research Design

This study adopts a mixed-methods sequential explanatory design, combining large-scale quantitative sentiment analysis with qualitative content analysis. A pragmatic research paradigm acknowledges that sentiment in comments represents authentic user experience that neither purely quantitative nor purely qualitative analysis can fully capture in isolation.Phase 1 (Quantitative) deployed automated sentiment analysis using NLP and machine learning across the full 37,157-comment corpus. Phase 2 (Qualitative) examined representative comment subsets to contextualise quantitative patterns, identifying nuanced concerns and acceptance drivers.

2.2 Data Collection

YouTube served as the primary data source. Videos were identified via systematic search terms (e.g., "AI music generating tools," "Suno review," "AI-integrated music studio") and selected using inclusion criteria: minimum 5-minute duration, at least 50 comments, primary emphasis on AI music tools, English-language content. Purposive sampling ensured diversity across content types: tutorials, product reviews, comparisons, and ethical discussions.
Comments and metadata were extracted using ytdt.digitalmethods.net, yielding 37,157 comments across 84 videos from 63 distinct channels. Comments under two characters, deleted entries, and spam-flagged items were excluded. Extracted variables included comment text, author metadata, timestamps, engagement metrics (likes, reply counts), and video metadata (title, view count, classification category).

2.3 Analytical Pipeline

Text preprocessing followed a rigorous pipeline: cleaning (removal of URLs, special characters, promotional noise), tokenisation via spaCy, lemmatisation, part-of-speech tagging, and dependency parsing. This linguistic preparation underpinned four analytical stages:

ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 8.206| ESTD Year: 2018|

International Journal of Multidisciplinary Research in Science, Engineering and Technology (IJMRSET)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

- VADER Sentiment Analysis (H1, H2): Each comment received a compound score classifying it as positive (> 0.05), negative (< -0.05), or neutral. Results were cross-tabulated against video classification labels.
- LDA Topic Modelling (H3): Applied to a 17,261-term vocabulary across 37,157 documents (comments) and separately to 884 transcript documents. Five coherent latent topics were extracted per corpus.
- Emotion Classification (H4): The j-hartmann/emotion-english-distilroberta-base transformer model assigned discrete emotion scores (happiness, sadness, surprise, anger, fear) to each comment.
- Aspect-Based Sentiment Analysis (ABSA): spaCy noun-chunk extraction identified 200,910 aspect-sentiment pairs across the comment corpus, enabling aspect-level sentiment profiling.

Table 1: Hypothesis-to-Method Mapping

Hypothesis	Method	Tool/Model	Data Source
H1	Sentiment Analysis (VADER)	NLTK Sentiment Intensity Analyser	Comments (n = 37,157)
H2	Cross-tabulation: Sentiment x Classification	Pandas groupby + value_counts	Comments with Classification labels
H3	LDA Topic Modelling + Sentiment Overlay	Sklearn LDA; CountVectorizer	Comments + Transcripts
H4	Emotion Classification	j-hartmann/emotion-english-distilroberta-base	Comments (n = 37,157)

2.4 Validity and Limitations

Corpus validity is supported by coverage of 84 videos across seven classification categories, ensuring breadth across the AI music content niche. VADER's rule-based architecture guarantees full replicability. Limitations include corpus self-selection (only active commenters are represented), VADER's known difficulty with sarcasm and irony, single-label emotion assignment, temporal scope confined to a rapidly evolving domain, and English-only comment inclusion.

III. DATA ANALYSIS AND RESULTS

3.1 Corpus Overview

The comment corpus spans 63 distinct YouTube channels. One channel contributed 21,224 of 37,157 total comments, and the most-commented video attracted 8,509 comments—reflecting platform economics' characteristic attention skew. Videos spanned seven categories: Tutorial/How-To (31 videos), Comparison and Review (19), Product Review (17), Analysis and Commentary (9), Product Deep Dive (4), and two smaller categories.Video engagement ranged from 3,590 to 2.47 million views (mean = 196,223), with comment counts from 8 to 13,929 (mean = 922). The Pearson correlation between views and comments was r = 0.85, indicating a very strong co-variation between viewership and participation. Video duration showed no significant relationship with views or comment counts, suggesting content quality and relevance drive engagement rather than length. Transcript lengths averaged 976 words (SD = 508.9), reflecting a moderately right-skewed distribution dominated by standard tutorials supplemented by a minority of detailed long-form reviews.

3.2 Sentiment Analysis — Testing H1

Table 2: Overall Sentiment Distribution Across Comment Corpus

Sentiment	Count	Percentage (%)
Positive	22,665	60.97
Negative	8,391	22.58
Neutral	6,101	16.45

ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 8.206| ESTD Year: 2018|

International Journal of Multidisciplinary Research in Science, Engineering and Technology (IJMRSET)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

Sentiment	Count	Percentage (%)
Total	37,157	100.00

H1 is strongly supported. Positive sentiment (60.97%) substantially outweighed negative (22.58%) and neutral (16.44%) responses—approximately three times the negative proportion. This positive skew reflects audience self-selection: viewers actively seeking AI music content are predisposed toward interest in these technologies.

The substantial negative minority clusters around three core concerns identified through qualitative examination: (1) perceived threats to musicians' livelihoods and artistic identity; (2) debate over whether AI-generated music constitutes genuine artistic expression; and (3) legal and ethical concerns about training data, copyright, and platform accountability. Creator transcripts showed near-uniformly positive VADER scores across all 84 documents, demonstrating that audience sentiment heterogeneity is not reflected in creator discourse, which maintains a professionally neutral register.

3.3 Sentiment by Content Classification — Testing H2

Table 3: Sentiment Percentage Distribution by Video Classification

Content Classification	% Positive	% Negative	% Neutral
Tutorial / How-To	67.19	17.77	15.04
Primary Classification	66.53	15.10	18.37
Product Deep Dive	66.34	19.14	14.52
Analysis and Commentary	63.25	21.39	15.36
Comparison and Review	59.30	25.71	14.98
Product Review	57.90	22.44	19.66

H2 is supported. Tutorial/How-To content attracted the highest positive sentiment (67.19%) and lowest negative (17.77%), while Comparison and Review content generated the most negativity (25.71%), a gap of 7.94 percentage points. Instructional content frames AI tools as practical creative resources, producing an enabling mindset in audiences; evaluative/comparative formats structurally invite critical assessment. The transcript corpus showed no classification-based sentiment gradient, confirming that the observed audience sentiment variation is not creator-driven but audience-driven.

3.4 Topic Modelling — Testing H3

Table 4: LDA Topic Keywords and Labels (Comment Corpus)

#	Topic Label	Key Terms
1	AI vs. Human Music	music, ai, people, real, musician, human, artist, band
2	Copyright & Ethics	copyright, claim, law, right, problem, generated, youtube
3	Tool Usage & Access	use, make, free, paid, need, using, don't
4	AI Creativity & Art	art, creative, create, way, think, tool, work
5	Song & Audio Production	suno, lyric, vocal, voice, track, sound, song

H3 receives partial support. Topic 2 (Copyright and Ethics) directionally aligns with elevated negative sentiment in Comparison/Review and Analysis/Commentary content. Topics 3 (Tool Usage) and 5 (Song Production) align with higher positivity in Tutorial content. Topic 4 (AI Creativity and Art) exhibits the most mixed sentiment, reflecting

© 2026 IJMRSET | Volume 9, Issue 4, April 2026 | DOI:10.15680/IJMRSET.2026.0904220

ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 8.206| ESTD Year: 2018|

International Journal of Multidisciplinary Research in Science, Engineering and Technology (IJMRSET)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

genuine philosophical ambivalence. The transcript LDA identified functionally distinct topics—vocal generation, lyric creation, stem editing, model prompting—with no equivalent ethical or normative debate themes, reinforcing the conclusion that critical discourse is audience-generated.

3.5 Emotion Classification and ABSA — Testing H4

The DistilRoBERTa model produced the following mean emotion scores: happiness (0.154), sadness (0.108), surprise (0.107), anger (0.075), fear (0.045). H4 is partially supported. Happiness is the modal emotion, consistent with majority positive sentiment, but its narrow margin over sadness and surprise indicates emotional complexity. Sadness likely reflects apprehension about musical authenticity and musician livelihoods; surprise captures audience reaction to AI's unexpectedly capable outputs. The low fear score, despite widespread discourse about existential risk to musicians, suggests YouTube commenters engage in more measured affective expression than headlines imply.

Table 5: Aspect-Level Sentiment Proportions — Selected Top Aspects (Comments)

Aspect	% Positive	% Negative
Suno	82.5	10.4
Sound	79.3	15.6
Song	76.5	16.2
Tool	74.3	13.9
Musician	70.5	19.5
Music	68.2	20.4
Art	67.4	25.8
AI	64.5	23.4
Copyright	63.7	23.1

ABSA produced 200,910 comment aspect-sentiment pairs. Suno achieved the highest positive proportion (82.5%), followed by sound quality (79.3%) and song (76.5%). Notably, "musician" (70.5%) and "music" (68.2%) outperformed "AI" (64.5%) and "copyright" (63.7%) in positive proportions, indicating that audiences retain strong positive orientation toward music as a cultural practice and musicians as creative forces—even amid ambivalence about AI as a disruptive technology.

IV. FINDINGS, IMPLICATIONS, AND CONCLUSION

4.1 Key Findings

Across all four analytical dimensions, the findings converge on a coherent picture: AI music tools are received with dominant enthusiasm co-existing with a substantive, thematically structured critical undercurrent. The positive majority (60.97%) reflects curiosity, creative interest, and practical engagement with tools that genuinely lower production barriers. The critical minority (22.58%) is not homogenously negative but clusters around copyright ethics, artistic authenticity, and musician displacement—issues that constitute a coherent discourse strand across video genres, dates, and channel communities.

Content framing exerts a measurable and directed influence on audience affect. Instructional formats produce enabling emotional states; evaluative formats invite scepticism—a finding with direct practical implications for content strategy on YouTube. Topic modelling confirms that copyright and ethics topics drive the most negative commentary, while tool-usage topics drive the most positive, with AI creativity occupying philosophically contested middle ground.

4.2 Theoretical Implications

Framing Theory is confirmed at the scale of YouTube video categorisation: content framing systematically shapes viewer emotional response, extending framing effects to user-generated platform contexts. Uses and Gratifications

© 2026 IJMRSET | Volume 9, Issue 4, April 2026 | DOI:10.15680/IJMRSET.2026.0904220

ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 8.206| ESTD Year: 2018|

International Journal of Multidisciplinary Research in Science, Engineering and Technology (IJMRSET)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

Theory is supported by the divergence between tutorial audiences seeking practical skill acquisition and review audiences seeking critical evaluation, with different motivational orientations producing fundamentally different affective outputs. Technology Acceptance and Resistance frameworks are also affirmed: the dominant happiness emotion coexisting with elevated sadness and muted anger suggests that AI music tool resistance reflects cultural mourning—grief at perceived loss of musical authenticity—rather than technology-directed fear. Finally, the multi-method NLP pipeline (VADER + LDA + DistilRoBERTa + ABSA) demonstrates robust methodology for platform audience studies at scale.

4.3 Managerial Implications

For AI music tool developers: Suno's dominance in positive ABSA scores (82.5%) confirms that platform reputation is shaped through YouTube discourse and that community engagement investment yields measurable returns. The persistently low positive proportions for copyright (63.7%) indicate that unaddressed training data ethics and intellectual property concerns remain active reputation risks; transparent licensing and artist compensation frameworks are strategic priorities.

For YouTube content creators: the empirically demonstrated sentiment gradient across content formats provides an evidence base for strategic decisions. Tutorial formats maximise positive engagement; creators employing comparison formats should actively moderate comment sections and frame evaluations constructively to mitigate polarisation.

For music industry stakeholders: the finding that "musician" and "music" consistently outperform "AI" in positive sentiment proportions reveals latent goodwill that industry bodies can leverage. Advocacy framing around protecting musicians and musical heritage—rather than opposing AI—is better aligned with demonstrated audience values. The near-parity of sadness and happiness scores signals that audiences are emotionally invested in the cultural stakes of AI's impact and are receptive to campaigns that take those stakes seriously.

4.4 Limitations

Several limitations qualify interpretation. First, corpus self-selection: only active commenters are represented, likely skewing positive relative to passive viewers. Second, VADER's known sensitivity to sarcasm and context-dependent language may inflate positive classifications. Third, single-label emotion assignment by DistilRoBERTa may not adequately capture the multi-emotional complexity of longer comments. Fourth, the temporal scope is limited to a rapidly evolving domain; sentiment patterns will shift as technology matures and regulatory frameworks develop. Fifth, noun-phrase ABSA links aspects to comment-level rather than aspect-level sentiment, limiting granularity.

4.5 Directions for Future Research

Longitudinal analysis would capture how sentiment and topic distributions evolve in response to tool maturation, regulatory interventions, and industry actions. Cross-platform comparison—extending to Reddit, X (Twitter), and TikTok—would establish whether the observed patterns are YouTube-specific or generalisable across social media discourse. Demographic disaggregation, distinguishing professional musicians from hobbyist producers and general listeners, would illuminate whether different audience segments hold systematically different views. Experimental designs manipulating content framing could provide causal evidence of the classification-based sentiment gradient observed here. Finally, multilingual corpus expansion would broaden generalisability and illuminate cultural universality versus regional specificity in AI music discourse.

4.6 Conclusion

This study provides a rigorous empirical foundation for understanding public perception of AI music generation tools on YouTube. Across 37,157 comments and 84 video transcripts, a convergent multi-method NLP pipeline demonstrates that audience discourse is characterised by dominant positive engagement—curiosity, enthusiasm, and creative interest—co-existing with a substantive critical undercurrent concentrated around copyright, artistic authenticity, and the implications of AI for professional musicians. Content framing effects, topic-sentiment alignments, and aspect-level profiles all reinforce this dual character of audience reception.

Critically, the dual corpus approach reveals that evaluative discourse about AI and music is an audience-generated, not creator-generated, phenomenon: creator transcripts maintain uniformly neutral emotional registers while audience comments carry the full affective complexity of public response. This distinction has methodological implications for future platform studies, which should treat creator and audience content as fundamentally different data types.

ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 8.206| ESTD Year: 2018|

International Journal of Multidisciplinary Research in Science, Engineering and Technology (IJMRSET)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

The findings carry direct implications for AI tool developers, content creators, and music industry stakeholders seeking to navigate the challenges and opportunities of AI integration. As generative music tools proliferate, sustaining the balance between technological democratisation and the protection of human creative agency will be the defining challenge for the industry—and understanding where public sentiment stands is an essential first step.

REFERENCES

1. Ansani, A. K. (2025). AI performer bias: Listeners like music less when they think it was performed by an AI. Empirical Studies of the Arts, 43(2), 1137–1161.
2. Bell, A. P. (2026). Pedagogy of the producer: DAWs, YouTube, and music education. International Journal of Music Education.
3. Campos Valverde, R., & colleagues (2026). "A safe, responsible, and profitable ecosystem of music": Analyzing perceptions and implementation of generative AI in the music industry. New Media & Society, 0(0).
4. Chen, W. (2022). Design of music teaching system based on artificial intelligence. Mathematical Problems in Engineering.
5. Cheng, L. M. (2024). Digital musicianship in post-pandemic popular music education. International Journal of Music Education.
6. Henry, A., Wiratama, V., Afilipoaie, A., & Ranaivoson, H. (2024). Impacts of AI on music consumption and fairness. Emerging Media.
7. IFPI. (2026). Global music report 2026. International Federation of the Phonographic Industry.
8. Liu, Z. (2025). The impact of user experience on continuous usage intention for AI-generated digital music platforms: Examining the mediating roles of emotional attachment and technology self-efficacy. Sage Open, 15(4).
9. Ma, X. (2021). Analysis on the application of multimedia-assisted music teaching based on AI technology. Advances in Multimedia.
10. Pan, J., Y., & colleagues (2022). The generation of piano music using deep learning aided by robotic technology. Computational Intelligence and Neuroscience.
11. Pfeiffer, A., & colleagues (2025). AI in music reconstruction: Sentiment in pro versus amateur contexts. Journal of Cultural Management and Cultural Policy, 11(2), 256–268.
12. Qiusi, M. (2022). Research on the improvement method of music education level under the background of AI technology. Mobile Information Systems.
13. Touchette, J. (2025). Well DAW! That's why I don't sound like the recording: Teacher perspectives on music production in elementary schools. International Journal of Music Education.
14. Wang, P., & Liu, X. (2025). The role of technophilia, positive experience and flow in enhancing learning efficiency in AI-based music education: A self-determination theory approach. European Journal of Education, 60.
15. Weng, T.-S., & colleagues (2025). Exploring the impact of AI-generated music on learning outcomes. Management Science and Industrial Engineering: Proceedings of the 7th International Conference (MSIE 2025).
16. Yihan, L. C. (2025). The use and effectiveness of digital tools in elementary music education: A systematic review. Music & Science.
17. Zhao, S. (2024). Rethinking materiality of music: The transformative practices of music streaming. Cultural Sociology, 0(0).
18. Zhao, S. (2026). Authorial positioning and creative agency in AI-mediated music. Global Media and China.