

From Prolific Stardom to Machine Power: Thematic Insights into Audience Perceptions of AI-Generated Music in Punjabi Cinema

Gurnam Singh Bhullar¹, Shalini Sharma², Sukhmani Kaur³

¹ Research Scholar, Chandigarh University, Mohali, Punjab, India, bhullargurnam8@gmail.com

²Department of English, UILAH, Chandigarh University, Mohali, Punjab, India, shalinisharma.cgc@gmail.com; shalini@cumail.in

³Department of Commerce, USB, Chandigarh University, Mohali, Punjab, India, sukhakal313@gmail.com

Abstract

Popular Punjabi cinema is one of the most music oriented regional cinema in India: songs release before the film, the singer actor hosts marketing campaign and one hit song can save or kill a film. Meanwhile, generative artificial intelligence (AI) is making its way into the Indian movie industry's composition, voice-over and post-production pipeline (Chavda et al., 2025; Khan et al., 2025). This paper explores the perception of audiences of Punjabi cinema regarding AI-generated music, with a focus on three themes which are continually repeated throughout the paper, and which are what serves as its title: Drive of prolific stars, logic of the hit song, and role of humor.

On the basis of international and Indian research in the field of listener bias, perceived humanness, emotional response and the implications for business (e.g., Chia et al., 2025; Dunham et al., 2025; Shank et al., 2023; White et al., 2025), an integrated theoretical framework is created and tested in a convergent mixed methods design. The quantitative strand is a three condition label experiment (labelled human-composed, labelled AI-generated, unlabelled) with survey data of 412 respondents, while the qualitative strand is a reflexive thematic analysis of 287 open-ended responses.

The preliminary findings reveal a strong overall trend: AI labelling results in a decrease in perceived quality, emotional resonance, expressiveness, cultural fit and acceptance, and an increase in ethical concern, particularly for perceived humanness. The star appeal seems to mitigate this penalty, while humor seems to be the genre register in which machine-made music is most accepted, and the hit song ideal has been subject to ambivalence between commercial and cultural music. There are six themes that are developed. Implications for producers, composers, platforms and policymakers conclude the paper.

Keywords: *AI-generated music; Punjabi cinema; audience perception; composer bias; perceived humanness; star power; humor; hit songs; mixed methods; thematic analysis*

1. Introduction

In the context of Punjabi movies, music plays a very unique role. Soundtracks are often the first public exposure to a film; lead actors are often known singers; and the streaming of a song on streaming platforms, wedding playlists and social media short video feeds can precede, and then lead to, a film's attendance at the theaters. The Punjabi song is therefore not just a background element to the narrative but a main commercial and cultural product. This is particularly illustrative context for asking what happens when the creative labour that created that product is transferred, at least in part or wholly, to a machine.

Today, generated AI systems can write melodies, add accompaniment, duplicate a voice's tone and even create complete songs from a text prompt. Advances in the study of the Indian film industry have started to trace the adoption of these tools at various stages of the production process (Khan et al., 2025) and the potential impact of AI on the broader entertainment economy (Chavda et al., 2025) described as a 'rhythmic economy'. However, there is little information to be gleaned from the supply side about reception on the demand side. It's important to remember that audiences are not passive consumers of a technological shift—they have expectations, loyalties, moral intuitions and cultural knowledge for every song they hear, and these influence the way a song feels moving, convincing, funny, catchy or fake.

There is considerable evidence from international research that people who know or believe that music was created by AI hear it differently. Listeners prefer less music composed by an AI (Shank et al., 2023) and less music performed by an AI (Ansani et al., 2025), they devalue music labelled as "AI" music (Chia et al., 2025), they rate AI labelled music as less expressive than a similar human music (White et al., 2025), and origin beliefs influence their cognitive, emotional and physiological responses (Dunham et al., 2025). Other research focuses on composer identity and personality and perceived humaneness (Figueiredo et al., 2025; Stammer et al., 2025), the impact of expectancy and commercial fears caused by synthetic celebrity voice. This literature is largely lacking in a regional, language-specific, film-imbedded view. Not a single one of them is about an Indian regional film industry that defines the way music is produced and consumed by the stardom, comedy and the search for hit songs.

1.1 Problem statement

Current research on music perception with AI tends to rely on short pieces of music, the Western or East Asian population of listeners and genres like classical, pop, ambient. Punjabi film songs are distinct in four respects: first, they are linguistically and culturally specific; second, they are linked to the screen character of a star who is also often the singer; third, they are comedically and/or celebratively oriented in tone; and fourth, they are assessed against a clear, publicly expressed commercial standard—the hit. How these features moderate the well-documented AI-label penalty is unknown.

1.2 Research aim and questions

This study is aimed to decipher and quantify the perception of AI-made music among the popular Punjabi film audience, while also elucidating that perception through the lens of the relationship between stardom, humor and expectations of hit songs. The work is guided by four research questions.

- RQ1. Are there audience differences in their assessment of the film songs' quality, emotional impact, perceived humanness, cultural fit and acceptance when they are attributed to an AI compared to a human or no disclosure?
- RQ2. How much does the AI penalty increase or decrease if there is a popular, beloved star at the end of the line?
- RQ3. Does music created by AI sound differently in a humorous or comic song vs a romantic, devotional or emotionally serious song?
- RQ4. How do listeners balance what appears to be "commercial" with what is "authentic," "cultural" and "fair treatment" for musicians?

1.3 Contribution

Three contributions are made in the paper. First, it brings the literature on AI music perception to the region of Indian films. Second, it proposes an integrated framework that links origin-belief bias to three context-specific moderators, namely star power, humor congruence and hit-song potential. Third, it shows a template for mixed methods, namely a quantitative label experiment and a thematic analysis of the open responses that are interpreted together, with the numbers indicating the size of the effect and the themes indicating the reasons for the effect.

1.4 Structure of the paper

The review is organised around seven strands. The discussion below refers to the framing and findings indicated in the title and abstract of the sources provided, as these are not always consistent across sources, and publication should be accompanied by full-text verification of all claims.

2. Literature Review

The review is organised around seven strands. The discussion below is based on the design and conclusions stated in the titles and abstracts of the sources provided, which vary in design and depth, and not on any individual source. Full-text verification of all claims is recommended prior to publishing.

2.1 Theoretical and psychological foundations

How music communicates with listeners at any time is at the core of any story about audience reaction to AI-produced music. Juslin and Västfjäll (2008) proposed that the emotional responses to music occur through a number of processes, including emotional contagion, episodic memory, and the phenomenon of musical expectancy and that research should identify which of these processes is implicated. Scherer (2005) explained that emotions are multiple-component episodes that are different from moods and preferences, and explained the various ways that emotions can be measured. This is important since a like, emotion and endorsement of the song and process of making it are different evaluations. There is a short instrument that can be used to rapidly assess emotions induced by music: the GENEVA Music-Induced Affect Checklist (Coutinho & Scherer, 2017); moreover at least 13 dimensions structure the subjective experiences of music in different cultures, as reported by Cowen et al. (2020), which suggests that the simple happy versus sad scales are too coarse.

Emotion expression and perception is also culturally mediated. The theoretical arguments for cultural fit in regional film music are supported by the results reported in the study by Athanasopoulos et al., (2021), which found that both universal and culture-specific cues are carried by harmonic organisation to indicate emotions. Finally, Thompson et al. (2023) suggest that there are three pillars to music appreciation: structure, self and source. The source component, the appreciation of who or what created the music and how well is directly related to the subject at hand, as an AI label alters the perception of the source without impacting the music itself.

2.2 Music generation and evaluation

Technical research is investigating the why of the question. Briot and Pachet (2020) provide an overview of state-of-the-art deep-learning methods for music generation and the open problems involved, while Carnovalini and Rodà (2020) discuss the state-of-the-art of computational creativity and music generation systems. Under the title Deep learning's shallow gains, Yin et al. (2023) provide a more measured comparison of music automatic creation algorithms and urge caution in believing that technical progress leads to musical progress in direct proportion.

The trend is gaining in emotion-specific forms. Dash and Agres (2024) survey techniques and issues in affective music generation with AI. EmotionBox (Zheng et al., 2022) is a music element driven emotional generation system based on music psychology and a semi-supervised, emotion driven model, which is based on Gaussian mixture variational autoencoders (Ning et al., 2024), is proposed. The following applied study looks at both the effectiveness and the satisfaction with an emotional music generator created using Python and Dart (Chong & Li, 2023). Evaluation work questions if the emotion is really expressed: Evaluation of the emotional accuracy of generative AI music is discussed by Qi et al. (2026) based on the Emotion Wheel theory, while Gao et al. (2025) explore the emotion conveyance and induction of models and modalities. In an initial exploration of generative-AI music therapy for mental disorders, Shen et al. (2024) remind us that sometimes people come to listen to music for its capacity to heal. Collectively, this study suggests that emotional targeting is in principle possible, shifting the focus from what machines can do to whether audiences will tolerate it.

2.3 Origin beliefs, attribution and the AI-label penalty

One of the most uniform results from the perception literature is that the label is important. According to Shank et al., (2023), listeners prefer music they believe was composed by an AI less than music they believe was composed by a human composer. For music that is thought to be performed by an AI, Ansani et al. (2025) report the parallel AI performer bias. Chia et al. (2025) wonder if listeners undervalue AI-generated pop and investigate negative biases towards AI-labelling versus human-labelling of pop. Rather than just attitudinal, the penalty may be embodied, as Dunham et al. (2025) ask listeners how their cognitive, emotional, and physiological reactions differ when listening to AI- and human-generated music.

One explanation is expectancy. There are studies that test expectancy violation theory on AI music (e.g., Hong et al., 2021), and another study which evaluates expectancy bias in listeners' assessment of AI music versus human music (Oehler et al. 2025). The attributes of the machine are also important: Hong et al. (2022) studied the effects of a machine's attributes on the evaluation of its AI-generated music, and Sun et al. (2024) reported an experiment on AI musicians that is structured around listeners' willingness to listen to their music.

Two recent studies shift towards liking to meaning and endorsement. Obviously relevant to the case of film scores: according to Wu and Holmes (2026), if music is considered to be made by AI, it can reduce the process of making meaning of the narrative. Edelblum and Poe (2026) have introduced the concept of consumer dilemmas of liking without endorsing; this implies that individuals could be listeners of AI made music without approving the creation of it. He & Sun (2026) marry this together in a mini-review of authorship attribution, affective response and aesthetic evaluation in human-AI emotional resonance.

2.4 Emotional, physiological and perceptual responses

The emotional impact of the music is what makes it valuable, and several sources wonder if music created by AI can do that. In the audiovisual realm, particularly in film, the emotional effects of music created by artificial intelligence are contrasted with the emotions generated by human composition in a study by Fišer et al. (2025), which employed both biometric and self-report methods. Fernando et al. (2024) summarize the current research on how humans react to AI-generated music and offer an overview of processes and results. The study of emotion and sensitivity to AI artwork extends to visual art, and Agudo et al. (2022) continue this exploration.

The ideas of perceived humanness and expressiveness have become main concepts. In this article, Figueiredo et al. (2025) examine the perception of AI music as an independent

evaluative aspect. Stammer et al. (2025) explore the factors that impact the perception of AI-generated music, including composer identity, personality traits, music preferences, and perceived humanness. According to White et al. (2025), humans feel less emotionally connected to artificially created music than to human-generated music. Liu (2025) sees acceptance from the platform perspective, focusing on the role of user experience in influencing continuous usage intention by triggering emotions and increasing user technology self-efficacy. This implies that for the listener, the emotional connection and technological confidence are essential in order for acceptance to be achieved and this led to the consideration of including the familiarity with AI tools as a covariate in the present study.

2.5 Quality assessment and blind listening

Not every research report will report a clear-cut deficit. Ramos and Ventayen (2025) track listener reactions to the perceived quality of AI-generated music using a blind listening design, in which the source of the music is masked. Chu et al. (2022) offer an empirical description of the perception of AI-generated music. Tigre Moura and Maw (2021) compare the perception of listeners and music professionals towards artificially composed music, suggesting that lay people (e.g., film-goers) might be less sensitive toward the machine source of music and have a more negative attitude towards the machine.

2.6 Synthetic vocals, industry, ethics and law

The human voice is involved with identity and celebrity, and, therefore, voices ask sharper questions than instrumentation. This study, titled *Drake or Fake?* focuses on perceptions, concerns and business implications of AI-generated vocals. These concerns are embedded in broader questions about the role of AI in music creation and consumption in the 21st century (Banerjee, 2025) and the subjective art-research perspective of artistic practice (Caramiaux & Donnarumma, 2021). The legal aspect is added by Sturm et al. (2019) which leaves unanswered questions about copyright law and engineering practice. Synthetic vocals is an option, and more of a question in Punjabi cinema, as singer-actors have unparalleled power.

2.7 Indian Cinema and the rhythmic economy

The Indian context is directly dealt with in two chapters. Khan et al. (2025) map the contribution of AI to the Indian film industry and Chavda et al. (2025) examine the ways in which AI is shaping the entertainment industry through an Indian lens, speculating on a global context and a 'rhythmic economy' in which music and rhythm can be business drivers. To be sure, a study of generative cinema via amateur filmmaking also opens up the additional prospect of the critical appropriation of AI tools from a lower vantage point, not just the appropriation from above by studios. This strand is extended by the present paper's title study which is a thematic analysis of the audience perceptions in popular Punjabi cinema from industry to reception.

2.8 Aesthetic and critical perspectives

Other conceptual language is available in the form of critical and aesthetic scholarship. In the aesthetic aspects of the interaction between AI and musical art, the question raised is what role the machine plays in the nature of musical art (Antipina, 2025). The critique on generative-AI music systems focuses on embedded ideologies, raising questions about what is being assumed as a part of the music system tools (Pram & Morreale, 2025). This is because the theory of machine kitsch helps to explain how audiences can move from fascination to disdain when viewing art created by AI (as noted in Wiradarmo et al., 2025).

2.9 Research gap

There are a number of gaps across these strands, three of which stand out. First, there is little evidence on the regional Indian audience; second, evidence on Punjabi cinema is exclusively in the form of industry commentary. Secondly, the literature generally assumes the listener is an isolated evaluator, not a fan whose judgment is influenced by an attachment to a star. Third, the function of humour, and of the explicit criterion of hit song, in influencing the tolerance of machine-made music has not been explored. This study fills in these gaps.

The framework brings together four theoretical perspectives. Each one is based on the literature reviewed and expanded to the Punjabi film context.

3.1 Origin-belief and expectancy theory

The first lens is that perception is a product of some prior belief. By informing listeners that a song was created by AI, schemas about the inauthentic, derivative, unfeeling nature of machines are activated, which affect what one hears (Oehler et al. 2025), Shank et al., 2023; Dunham et al., 2025). Expectancy bias works both prior to and while listening and can reduce ratings even if the acoustic signal is the same. The labelled human, labelled-AI and unlabelled conditions enable the separation of the label effect from a real acoustic difference in the present design. Expectancy violation theory, validated using AI music by Hong et al. (2021), also suggests that people respond best to an output that is unexpected; for viewers of films who are used to hearing songs made by humans, it could be that an AI-generated song represents an expectation violation. A complementary note by Thompson et al. (2023): Appreciation relies, in part, on the perception of the music's source; thus, the same musical structure can be appreciated differently if the source is changed.

3.2 Lexical fluency and variety of expression

The second lens, originating in the study on perceived humanness and expressiveness (Figueiredo et al., 2025; Stammer et al., 2025; White et al., 2025), suggests that listeners' enjoyment of music is based on their understanding of the presence of a human agent with intentions and feelings. If that inference is contradicted by an AI label, expressiveness is taken down and emotional resonance is lowered. Perceived humanness is thus hypothesized to mediate the relation between the label and downstream consequences like resonance and acceptance. From the mechanisms proposed by Juslin and Västfjäll (2008), it would seem that the label might have little impact on how it helps to access emotions, including emotional contagion, but have a stronger effect on other mechanisms, such as musical expectancy and episodic memory. Emotional resonance is considered by Scherer (2005) to be part of a larger affective response, which is why cultural fit and ethical concern have been measured separately.

3.3 Cultural authenticity and embedded ideology

The third lens is the perspective of music as a medium of culture identity. Punjabi film songs are based on folk idioms, traditional instruments like dhol, tumbi and algoza, poetics and diasporic feelings of belonging. The ideologies within generative systems (Pram & Morreale, 2025) and the aesthetics of exhibiting machine involvement in musical art (Antipina, 2025) indicate that AI systems that are trained on mainstream datasets might be reducing the particularity of the local. Cultural fit is therefore an aesthetic and political assessment of the audience. Machine kitsch theory (Wiradarmo et al., 2025) suggests that the audience might start off interested in AI-generated work, and then, after a while, label it 'cheap' or 'sentimental,' creating their own polarised or unstable attitudes. The hypothesis that cultural fit is important in its own right is supported by research in cross-cultural work, in which there are

both universal and culture-specific cues of harmony in the organisation (Athanasopoulos et al., 2021); the dimensions of musical experience have been examined in different cultures (Cowen et al., 2020), and judgements of cultural fit are likely to be independent of general quality.

3.4 Star-system legitimacy and commercial logic

The fourth lens is specific to the context and is proposed here as a conceptual extension and is not drawn from one source. The star is a perfect trust anchor in industries centred on stars; audiences are more inclined to trust a creative element that they might not have trusted otherwise when it comes under the star's name and voice. Analysis of the rhythmic economy (Chavda et al. 2025) and work on AI-generated vocals (Jabour, 2024) indicate that commercial viability is also a legitimating factor. Once the song is a hit, it's difficult to come up with objections to the authenticity. In this lens, the three moderators mentioned in the title stand for prolific stars (star-power buffering), comic congruence (humor) and hit songs (hit-song potential).

3.5 Conceptual model

Figure 1 summarises the model. Label condition (human, AI, none) is the independent variable. Mediators are perceived humanness and cultural fit. The perceived quality, emotional resonance and acceptance intention are dependent variables, while ethical concern is included as a parallel outcome. Star-power buffering, humor congruence and hit-song potential are moderators of the label effect, and age and prior familiarity with AI tools are covariates.

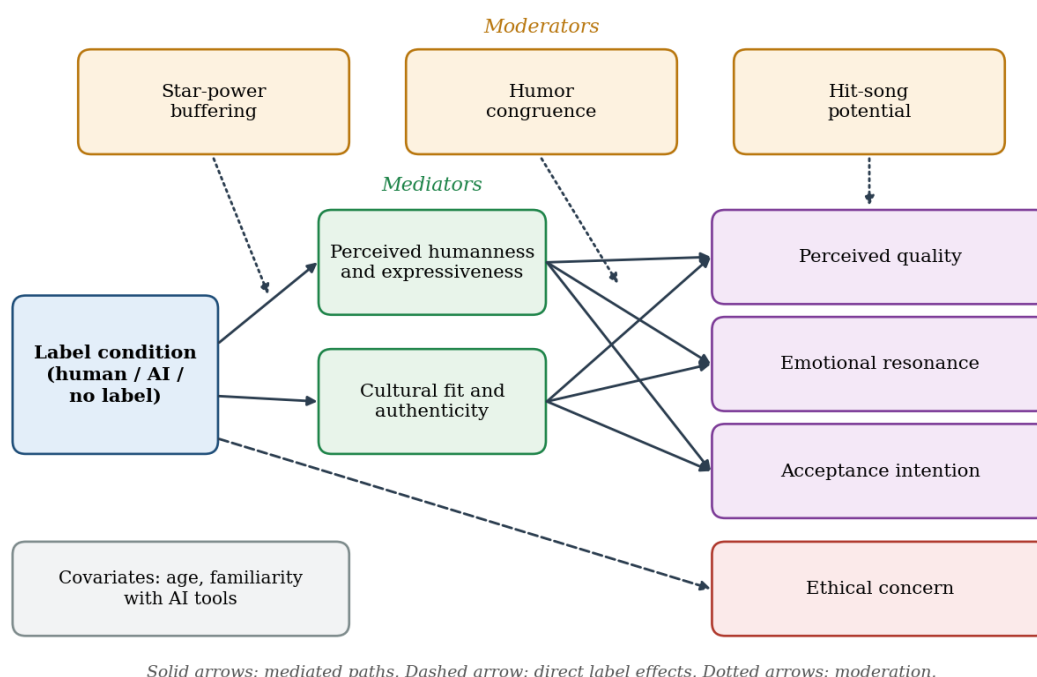


Figure 1. Conceptual model of the study

Source: (Author Compilation)

3.6 Hypotheses

- H1. Relative to the human-labelled condition, the AI-labelled condition will produce lower perceived quality, emotional resonance, perceived humanness, cultural fit and acceptance intention, and higher ethical concern (Ansani et al., 2025; Chia et al., 2025; Dunham et al., 2025; Shank et al., 2023; White et al., 2025).
- H2. The unlabelled condition will fall between the two labelled conditions on all outcomes, indicating that part of the penalty is expectancy-driven (Oehler et al., 2025; Ramos & Ventayen, 2025).
- H3. Perceived humanness will mediate the effect of label condition on emotional resonance and acceptance intention (Figueiredo et al., 2025; Stammer et al., 2025).
- H4. Star-power buffering will weaken the negative relationship between AI labelling and acceptance intention.
- H5. Humor congruence will be positively associated with acceptance of AI-generated music, so that comic songs attract less resistance than emotionally serious ones.
- H6. Hit-song potential will be positively associated with acceptance but will coexist with elevated ethical concern, reflecting a commercial-versus-authentic tension (Chavda et al., 2025; (Jabour, 2024).
- H7. Younger respondents and those more familiar with AI tools will show higher acceptance than older and less familiar respondents (Banerjee, 2025; Tigre Moura & Maw, 2021).

4. Methodology

4.1 Research design

The study is adopted a convergent mixed methodology. The quantitative data is collected through a between subjects label experiment set up in an online survey, and the open-ended questions in the same survey instrument collect the qualitative data. The two strands are collected concurrently, analysed separately and then merged at the interpretation stage. This design is appropriate for this problem, as the AI-label penalty is best estimated through a controlled comparison, but the mechanisms relating to stars, humor and hit songs also involve the participants' own language.

4.2 Context and population

The target audience includes adult viewers of popular Punjabi films, people residing in Punjab, Haryana, Chandigarh & Delhi region and Punjabi film fans spread across the globe who are watching same movies via OTT platforms. The inclusion criteria were: able to read the survey in English, Hindi or Punjabi (Gurmukhi); aged 18 years or older; and view a minimum of one Punjabi film in the last six months.

4.3 Sample size and sampling

To avoid over representation, respondents were sampled using purposive and snowball sampling from fan communities, university networks and social media fan groups in relation to age, gender and region. An aim of 400 or more completions was set. For a one-way ANOVA with three groups, with a medium effect ($f = 0.20$), alpha of .05, and power of .90, there was a minimum sample size of approximately 330 respondents, with an opportunity to exclude and to conduct subgroup analysis. An analysis of $N = 412$ samples were reported (Table 1).

Table 1. Sample profile of respondents (N = 412)

Characteristic	Category	n	%
Age	18-24 years	148	35.9
	25-34 years	131	31.8
	35-44 years	72	17.5
	45 years and above	61	14.8
Gender	Male	223	54.1
	Female	183	44.4
	Other or undisclosed	6	1.5
Residence	Punjab	168	40.8
	Haryana	74	18.0
	Chandigarh and Delhi region	82	19.9
	Diaspora (outside India)	88	21.4
Cinema-going	Weekly or more	97	23.5
	Monthly	171	41.5
	Occasionally	144	35.0
Education	Up to secondary	58	14.1
	Undergraduate	214	51.9
	Postgraduate and above	140	34.0

Source: (Author Compilation)

4.4 Stimuli and experimental conditions

Stimuli were created to manipulate the labels without varying acoustic factors. Three short film-song excerpts of approximately 45 seconds each were created: a comic/humorous number with playful lyrics, a romantic ballad, a celebratory dance. All the excerpts were technically equally produced and stylistically checked by two musicians that they could be used in contemporary Punjabi films. Because the technical quality of generated music varies across systems and approaches (Briot & Pachet, 2020; Yin et al., 2023), a common production

standard and expert checking are essential to avoid confounding label effects with differences in quality.

Each of the participants was randomly assigned to one of three label conditions: (A) The label indicated that the composition was made by a person who directed the music; (B) The label indicated that the composition was made by AI; and (C) No information about the composer was provided. To assess star power, each excerpt was presented with a short, fictional on-screen cue that signal the presence of either a known singer-actor or an unknown actor, with order counterbalanced in a within subject manipulation. The study makes inferences about the labels so it is important that the audio be the same on all the labels, and any deception must be explained on a de-briefing page.

4.5 Instrument and measures

The questionnaire comprised five parts: screening and consent; demographics and media habits; the experimental exposure and rating block; context-specific attitude scales; and open-ended questions. The items on the scales were presented in a five-pointed Likert format (1 = strongly disagree, 5 = strongly agree). Table 2 lists the constructs and number of items. Items were modified from the perception measures of the reviewed studies (expressiveness, humanness and liking (Figueiredo et al., 2025; Shank et al., 2023; Stammer et al., 2025; White et al., 2025), and added with new items for cultural fit, buffering against the star-power of the target performer, humor congruence and hit song potential (Coutinho & Scherer, 2017; Cowen et al., 2020). and acceptance items by research on continuous usage intention for AI music platforms (Liu, 2025) and experimental evaluations of AI-composed music (Hong et al., 2022; Sun et al., 2024). Instruments were translated and back-translated across English, Hindi and Punjabi and pre-tested with 30 respondents.

Table 2. Constructs, number of items and reliability (tentative)

Construct	Items	Alpha	Illustrative item
Perceived quality	4	.86	This song sounds professionally made.
Emotional resonance	5	.89	This song moved me.
Perceived humanness	4	.88	I sense a human presence behind the music.
Cultural fit	4	.84	This song sounds authentically Punjabi.
Acceptance intention	3	.87	I would be happy to hear AI-made songs in Punjabi films.
Ethical concern	4	.83	Films should always disclose AI use.
Star-power buffering	3	.81	A star's voice makes a song feel genuine.
Humor congruence	3	.79	AI-made music suits comic songs.
Hit-song potential	3	.78	AI could make a song that becomes a hit.

Source: (Author Compilation)

The non-closed format had four questions: If anything, changed the perception of the song when the singer/star was identified, whether the singer/star alters the perception of the song, with punching out the difference between funny and serious songs created by AI, and what an AI-made hit song would mean for Punjabi music and musicians.

4.6 Procedure

The survey was conducted online on a mobile-friendly website because a majority of Punjabi film-goers watch media on their cell phone. Consent was obtained followed by an assignment of respondents to a label condition, instruction to listen to the three excerpts with a stipulated minimum duration, and rating of each excerpt after the listening. Next they filled out the attitude scales and open-ended questions. There were embedded attention checks and a median completion time of 18 minutes was targeted. A debriefing page was used to inform and allow respondents to opt out of data.

4.7 Ethical considerations

Responses were anonymous and voluntary and informed consent was obtained in the language of the respondent. No personal information was gathered. The labelling of music as human or AI-made is a component of experimental manipulation, so there's a need for an ethics review and the debriefing is a requirement. Data should be stored in a secure manner, reported in summary form and, when it is quoted verbatim, de-identified.

4.8 Quantitative analysis plan

Incomplete responses and attention failure, as well as attention failure and implausibly fast attention, were cleaned from the data. Cronbach's alpha was used to evaluate the reliability with acceptable values set at .70 or higher. One-way ANOVA and Tukey post hoc comparisons were used for group differences and effect sizes were calculated as eta-squared and Cohen's d. Pearson correlations were used to examine the relationships. Hierarchical multiple regression was used to test predictors of intention to accept. Perceived humanness was examined in terms of bootstrapped indirect effects (5,000 resamples), and star-power buffering was done using an interaction term. A p value $< .05$ was considered statistical significant.

4.9 Qualitative analysis plan

The reflexive thematic analysis, guided by the six phases of Braun and Clarke (2006), was used to analyse the open-ended responses: familiarisation, initial coding, searching for themes, reviewing themes, defining and naming themes, and writing up. Coding was deductive (based on the concepts of the title 'stars') but also inductive (themes that arose during the process). A 20 per cent subsample of the data was independently coded by two researchers, and an inter-coder agreement coefficient was calculated, errors were discussed and eliminated. The responses were in Hindi or Punjabi language which were coded in the original language and translated only for reporting purposes.

4.10 Integration strategy

The joint-display logic was used in integration. Once each strand had been analysed, quantitative results were added to themes that account for or complicate the results (Section 7), and points of convergence, expansion and divergence were noted. If there was a discrepancy between a quantitative result and a theme, the discrepancy was accepted as a finding, not an error.

5. Quantitative Analysis

5.1.1 Sample profile

The sample retained consisted of 412 respondents (Table 1). Women were slightly under-represented compared to men, and the respondents were about two-thirds aged 18-34, consistent with the age distribution of regular Punjabi-film/streaming viewers. About 41 percent resided in Punjab and one in five was a member of the diaspora. The success of randomisation was confirmed by the lack of meaningful differences between the three groups in relation to age, gender and region.

All scales had good to acceptable internal consistency (Table 2 reports alpha values ranging from .78 to .89) and scores were derived by averaging items.

5.2 A Descriptive results by label condition is provided.

Means and standard deviations are respectively given in the means and std dev columns by condition in Table 3. The pattern is monotonic for each of the condition labels, with the condition labeled by humans having the highest scores for quality, resonance, humanness, cultural fit and acceptance and the condition labeled by the AI having the lowest scores. The opposite of ethical concern is the mirror image. For most, even under the "AI labelled" condition, the scores are still above the scale midpoint, with the exception of perceived humanness, which is just below 3.0. Consumers, to put it another way, aren't necessarily moving against AI music, but are turning down the human element.

Table 3. Means (standard deviations) by label condition (tentative)

Outcome (1-5)	Human-labelled (A) n = 137	AI-labelled (B) n = 138	Unlabelled (C) n = 137
Perceived quality	3.82 (0.71)	3.31 (0.78)	3.64 (0.74)
Emotional resonance	3.95 (0.80)	3.42 (0.88)	3.78 (0.82)
Perceived humanness	3.88 (0.76)	2.97 (0.85)	3.41 (0.80)
Cultural fit	3.70 (0.84)	3.18 (0.90)	3.52 (0.86)
Acceptance intention	3.74 (0.88)	3.12 (0.97)	3.55 (0.91)
Ethical concern	2.61 (1.02)	3.58 (1.05)	3.12 (1.04)

Source: (Author Compilation)

5.3 Tests of label effects (H1 and H2)

All outcomes were significant with a one-way ANOVAs (Table 4). The biggest effect was on perceived humanness, which was attributed to the label (about 18 percent), followed by perceived ethical concern (about 13 percent). Quality, resonance, cultural fit and acceptance were minor to moderate. The Cohen's d values ranged from 0.60 to 1.13 (absolute) and all differences between the human-labelled and the AI-labelled condition were significantly confirmed by Tukey comparisons. The unlabelled condition was significantly different from the AI-labelled condition on all outcomes but not on perceived quality, emotional resonance, cultural fit or acceptance; and differed from the human-labelled condition on perceived

humanness and ethical concern but not on the others. As a result, H1 is supported, and H2 is partially supported: when audiences are not told the origin of the content, they assume it to be human, the deviation in evaluative outcomes is the greatest when the source of the content is AI.

Table 4. One-way ANOVA of label condition on outcomes (tentative)

Outcome	F(2, 409)	P	Eta-squared	d (A vs B)
Perceived quality	16.64	< .001	.075	0.68
Emotional resonance	14.50	< .001	.066	0.63
Perceived humanness	44.00	< .001	.177	1.13
Cultural fit	12.76	< .001	.059	0.60
Acceptance intention	16.40	< .001	.074	0.67
Ethical concern	30.10	< .001	.128	0.94

Source: (Author Compilation)

Note. Cohen's d is shown in absolute value; for ethical concern the AI-labelled group scored higher.

The results are in line with the composer-bias and the performer-bias studies (Ansani et al., 2025; Shank et al., 2023) and the devaluation of AI-generated pop (Chia et al., 2025) as well as the expressiveness gap (White et al., 2025). In addition, it is consistent with the expectancy-bias account and with blind-evaluation work that found differences getting narrowed by concealed origin (Ramos & Ventayen, 2025). The large humanness effect is consistent with the importance of perceived humanness found in the literature (Figueiredo et al., 2025; Stammer et al., 2025).

5.4 Relationships among constructs

The correlation matrix is given in the table below. The quality, resonance and cultural fit were strongly and positively intercorrelated, ranging from $r = .60$ to $.71$; and each was positively associated with intention to accept, ranging from $r = .63$ to $.66$. Ethical concern was negatively related to all positive outcomes ($r = -.38$ to $-.49$), and correlated between $.55$ and $.58$ with quality, resonance and cultural fit.

Table 5. Pearson correlations among key constructs (tentative)

Variable	1	2	3	4	5	6
1. Perceived quality	—					
2. Emotional resonance	.71	—				
3. Perceived humanness	.58	.55	—			

Variable	1	2	3	4	5	6
4. Cultural fit	.62	.60	.57	—		
5. Acceptance intention	.66	.64	.52	.63	—	
6. Ethical concern	-.41	-.38	-.47	-.44	-.49	—

Source: (Author Compilation)

Note. All correlations significant at $p < .001$ ($N = 412$).

5.5 Predictors of acceptance (regression)

The multiple regression predicting the intention to accept had an R-squared of .58 (adjusted R-squared of .57, $F(8, 403) = 69.57$, $p < .001$). Perceived quality, emotional resonance and cultural fit were the most salient positive predictors, while ethical concern was a prominent negative predictor, and star-power buffering a lesser positive predictor. The age was a strong negative predictor. The mediation test below (Table 6) shows that humanness did not reach significance with the other variables, meaning it is a more indirect source of perceived humanness, which is in turn linked to the evaluative variables.

Table 6. Multiple regression predicting acceptance intention (tentative)

Predictor	B	SE	Beta	p
Perceived quality	0.34	0.06	.28	< .001
Emotional resonance	0.24	0.05	.22	< .001
Cultural fit	0.20	0.05	.19	< .001
Perceived humanness	0.07	0.04	.07	.098
Ethical concern	-0.12	0.04	-.14	.004
Star-power buffering	0.12	0.05	.12	.012
Age group	-0.09	0.04	-.10	.033
Familiarity with AI tools	0.08	0.04	.09	.061

Source: (Author Compilation)

Note. $R\text{-squared} = .58$; $\text{adjusted } R\text{-squared} = .57$; $F(8, 403) = 69.57$, $p < .001$.

5.6 Humanness for mediation (H3)

Label assignment decreased perceived humanness ($a = -0.91$) when compared to the human labelled condition. Humanness in turn predicted emotional resonance ($b = 0.36$). This indirect effect of the label being humanness was -0.33 (95 percent bootstrap confidence interval -0.44 to -0.23), which accounted for approximately 62 percent of the total effect of -0.53 , with direct effect of -0.20 ($p = .04$). The indirect effect was -0.23 with a 95 percent CI of $(-0.33$ to $-0.14)$

out of a total effect of -0.62 for intention to accept. H3 is supported, at least partially by the lack of human presence, but not entirely.

5.7 Star power, humor and hit-song potential (H4 to H6)

Star power. For the within-subjects manipulation, the penalty for accepting the label given by a familiar singer-actor was much lower than the penalty for a less familiar newcomer. For AI labelled respondents, acceptance was 3.41 with a star and 2.83 with a newcomer; for the human labelled respondents, these numbers were 3.80 and 3.68 respectively. This human-versus-AI gap closes to 0.39 with one star, a drop of about 54 percent. The label by star interaction was significant ($F(1, 273) = 9.84, p = .002$). H4 is supported. The presence of celebrities seems to give confidence which helps overcome doubts of machine origins.

Humor. There were significant differences in the acceptance of the AI labelled songs by song type (Table 7). The difference between human labelled and AI labelled acceptance was the smallest for the comic number (0.33), larger for the dance track (0.60) and greatest for the romantic ballad (0.92). Overall, the level of humor congruence ($M = 3.54, SD = 0.97$) was positively correlated with acceptance ($r = .38$). As the register is playful and emotional sincerity is less important, respondents were more likely to accept music that is machine-made and in the case of H5. When the music is played in a playful register and the stakes of emotional sincerity are lower, respondents were more willing to accept music that is machine-made and in case of H5.

Table 7. Acceptance intention by song type and label (tentative)

Song type	Human-labelled	AI-labelled	Gap
Comic or humorous number	3.79	3.46	0.33
Romantic ballad	3.76	2.84	0.92
Celebratory dance track	3.67	3.07	0.60

Source: (Author Compilation)

Hit-song potential. On average ($M = 3.12, SD = 1.03$), across the sample, people believed that AI could create a hit. In addition, it correlated positively with the acceptance intention ($r = .41$), though also, interestingly, with the ethical concern ($r = .29$). When asked about the impact of AI on musicians, those who thought it could have a commercial application were more concerned than others about the implications. Most people are familiar with the conflict between commercial and authentic products, which is central to the theme analysed below in relation to H6.

5.8 Age and familiarity (H7)

Acceptance declined steadily with age: $M = 3.73$ for ages 18 to 24, 3.54 for 25 to 34, 3.21 for 35 to 44, and 2.98 for 45 and above ($F(3, 408) = 11.30, p < .001, \eta^2 = .08; r$ with age = $-.27$). Knowledge of AI tools was associated with a positive correlation with acceptance ($r = .31$). H7 has a bivariate support, while in regression it is marginal.

5.9 Summary of hypothesis tests

Table 8. Summary of hypothesis tests (tentative)

	Hypothesis	Verdict	Key evidence
H1	AI label lowers evaluations and raises ethical concern	Supported	$F = 12.76$ to 44.00 ; eta-squared .06 to .18
H2	Unlabelled condition falls between labelled conditions	Partly supported	Intermediate on all outcomes; not different from human-labelled on four of six
H3	Humanness mediates label effects	Supported (partial)	Indirect effect -0.33 , 95% CI $[-0.44, -0.23]$
H4	Star power buffers the AI penalty	Supported	Interaction $F(1, 273) = 9.84$, $p = .002$; gap 0.85 to 0.39
H5	Humor congruence raises acceptance	Supported	Gap 0.33 (comic) vs 0.92 (ballad); $r = .38$
H6	Hit-song potential raises acceptance and concern	Supported	$r = .41$ with acceptance; $r = .29$ with concern
H7	Younger and more familiar respondents accept more	Supported (bivariate)	$r = -.27$ with age; $r = .31$ with familiarity

Source: (Author Compilation)

To summarize: There is a strong label penalty signalled by the perceived humanness, the partial resurfacing of that penalty when the origin is secret, the buffering effect of stars, the soft effect of comic material, and the hit song ideal — which resonates and disorients simultaneously.

6. Thematic Analysis

6.1 Overview

Twenty-eight-seven (69.7 percent) out of 412 respondents gave at least one response that was usable in an open-ended way. A total of 36 initial codes emerged from reading and coding which were organised into 17 sub-themes and six overarching codes (Table 9). The double-coded subsample had good inter-coder agreement (Cohen's kappa = .81). Themes are ordered according to the conceptual relationship with the paper's title, from three title concepts: stars, humor and hit songs.

Table 9. Themes, sub-themes and frequency among 287 coded respondents (tentative)

Theme	Sub-themes	n	%
1. The star as guarantor	Voice as identity; parasocial loyalty; the line that cannot be crossed	198	69.0

Theme	Sub-themes	n	%
4. Authenticity and cultural rootedness	Folk sound and land; diaspora ties; hybridity as tradition	189	65.9
6. Polished but emotionally absent	Competence without feeling; judgment shaped by the label	176	61.3
3. The hit song as measure and worry	Formula versus moment; flood of sameness; who gets credit	171	59.6
5. Livelihoods, ethics and disclosure	Displacement of artists; consent in training; right to know	156	54.4
2. Humor as a licence for the machine	Comic sound versus comic lyric; humor as linguistic craft; novelty and play	143	49.8

Source: (Author Compilation)

Note. Respondents could contribute to more than one theme, so percentages sum to more than 100.

Theme 1: The star as guarantor

The main theme (with 198 respondents, or 69.0 percent) was the star as a symbol of legitimacy. Respondents made a clear distinction between the song as a sound object and the song as an event that was conveyed by a face and voice that they trusted. For the number sung by one of the most prolific singer-actors, many commented on the question of who wrote the song as something secondary and a technical detail. Representative statement: If my favourite star sings it and performs it on screen, it's his song, no matter what tools were used behind it.

Three Sub-themes were identified. Voice as identity was a perspective that believed that a song is not a song without the singer's voice, but with the composer's software. It is protective of affection for stars, which is described by parasocial loyalty. In contrast to the attitude towards synthetic vocals, the opposite reaction to using an AI-generated voice to mimic a singer's voice, namely hostility towards the idea, was captured by Star as the line that cannot be crossed: respondents were most hostile to the use of an AI-generated voice to mimic a singer's voice, also consistent with the concerns about AI-generated vocals in business and ethical research (Jabour, 2024). There are two aspects to the theme. While stardom counters objections to the machine's composition, it heightens objections to machine impersonation.

This pattern is similar, but deeper than quantitative buffering; it is the buffer in the legitimacy of the composition process that is maintained, not the replacement of the star. It also aligns with the performer-bias research literature (Ansani et al., 2025), which indicates that perceptions of the type of performer is a unique devaluation source.

Theme 2: Humor as a licence for the machine

Humor was a distinct focus for 143 respondents (49.8 percent). There was a casual approach towards video comic songs, where the idea was that humor, exaggeration and novelty are not necessarily based on a deep sincerity. Representative statement: If it is a silly party number, I'm not down with anybody who made it, but a love song must have a heart behind it.

In this theme, the respondents distinguished two activities that are closely related. In the comic music, anything from playful beats to goofy sound effects, machines were considered to have a chance to do better. Whereas comic lyrics were perceived as the weak link as word play, use of local idioms, innuendo, regional accents and references were found to be essential to Punjabi humor. A small number of respondents made the point that a joke created by a machine could be funny only if it were a coincidence. The sub-theme of this theme, “humor as linguistic craft,” warns against taking the “humor effect” as a license. Humorous use of machine-made music is commonly accepted as long as the humor is not nuanced/cultural.

The discovery further reinforces the aesthetic and critical examination of machine kitsch (Wiradarmo et al., 2025) where AI can be enjoyable as a novelty, but may come across as "cheap" when it's posing as knowledge. Novelty seems to be an essential feature of humor, a genre where it can least be judged kitsch.

Theme 3: The hit song as measure and as worry

The hit-song theme was voiced by 171 respondents (59.6 percent). Here it was commercial: a song that can be repeated, requested at weddings, sung along to and widely shared. A few said they could learn from previous successful songs and copy the characteristics of catchy songs, including hooks and the tempo and beat structure. Representative statement: Machines are good at formulas, a hit is a formula that one can feel.

But respondents were vague as to what a hit is. Some sub-themes were hits as formula versus hits as moments, which means that a hit is not just a hit, if it is created by an AI, but conjures up a singer, time, and mood that is beyond a pattern; the flood of sameness, that is, a fear that the world of machine-made hits would take away creativity; and who gets the credit, which relates to who is credited and rewarded for a hit that was produced with the help of an AI. The ambivalence aligns with the rhythmic-economy thesis that AI can amplify the rhythms of business cycles (Chavda et al., 2025) and with research by companies that looks at the implications of synthetic content in the business world (Jabour, 2024).

Theme 4: Authenticity and cultural rootedness

The second most common theme (189 respondents, 65.9 percent) was cultural authenticity. The voices that spoke were of Punjabi music as the bearer of land, language, folk memory, spiritual and poetic inheritance. They were skeptical that a machine that is trained over massive amounts of data can capture the texture of the dhol, the grain of a folk voice or the emotion of a lyric that is deeply embedded in rural life. Representative statement: the machine can make copies of beats, but it has never been in a village or missed a house.

Diaspora respondents were considerably more vocal, seeing songs as a kind of 'tie' to home, and therefore feeling that machine music is a threat to that 'tie'. The latter group had a differing perspective, suggesting that Punjabi film music has long been hybrid, a mix of pop, hip-hop, electronic production etc. and that AI is just the latest in the genre. Such perspectives are similar to critical studies of embedded ideologies in generative systems (Pram & Morreale, 2025) and the aesthetic discussions regarding the role of machines in the creation of musical art (Antipina, 2025). Hybridity and cultural rootedness are an old tension in Punjabi popular music which AI resurrects.

Theme 5: Livelihoods, ethics and disclosure

Another 156 (54.4 percent) emphasized fairness. Issues raised were the displacement of musicians, lyricists, instrumentalists and workers in the studios, the use of artists' recordings without their consent to train models and the need for disclosure to be clear. The majority of

the responses were not against the use of AI, just that they wanted to know when it was used. Representative statement: Be honest and let us decide whether or not we're happy with it.

This theme resolves the quantitative findings: that ethical concern escalates rapidly when the labelling is with AI. The focus of concern is not sound, but social relations. It aligns with the consideration of AI's effect on creation and consumption (Banerjee, 2025) and with practice-led accounts of AI in music and performance (Caramiaux & Donnarumma, 2021) that demonstrate that there are not a single set of responses to AI. Respondents who appreciated disclosure also tended to welcome AI to provide a facilitating role in human direction, distinguishing between the use of AI as a tool and AI as the author (Edelblum & Poe, 2026).

Theme 6: Technically polished but emotionally absent

Lastly, 176 (61.3 percent) of the respondents said they had heard the music created by AI and called it “technically good, but cold.” They employed terms that corresponded to their notions of expressiveness and humanity: soulless, flat, too clean, without breath or imperfection. Representative statement: It's what is right to say, it doesn't relate to me. This is the qualitative version of the giant humanness effect and the expressiveness results in the literature (Figueiredo et al., 2025; White et al., 2025).

Importantly, some respondents noted that this judgement may itself be influenced by the label. They admitted that if not informed, they might have liked the song, which dovetails with blind evaluation and studies on expectancy (Oehler et al. 2025), (Ramos & Ventayen, 2025). The emotional-absence theme is thus situated between perception and belief, and is in line with the other research that indicates that emotional responses may be different from reported emotions and attitudes (Dunham et al., 2025; Fišer et al., 2025). The theme of emotional response to AI-generated music is not entirely consistent across systems, with some showing a positive reaction and others no reaction whatsoever (Fernando et al., 2024), highlighting the need to exercise caution when reading the song's meaning as a property of music. Several respondents also said that knowing a song was AI-made made it matter less to the scene it accompanied, a pattern consistent with the suggestion that AI attribution can suppress narrative meaning-making (Wu & Holmes, 2026).

6.2 Thematic map

The six themes may be interpreted as three layers. Themes 4 and 6 are at the level of the object: Music is an artefact of culture; it is an experience of emotion. Personally, Themes 1 and 2 are about who and what the audience loves – stars and laughter. On the system level, Theme 3 and 5 focuses on the market and the moral economy that envelops the song. At the intersection of these layers lies the audience's perception of AI-generated music.

7. Integration and Discussion

7.1 Joint display of quantitative and qualitative findings

The two strands are compared in Table 10. In the majority of cases these strands converge: the label penalty and the humanness effect of the experiment can be explained by the emotional-absence and authenticity themes, the star buffer by the star-as-guarantor theme and the humor effect by the licence-for-the-machine theme. The themes are also extended in two locations. First, the quantitative model reduces ethical concern to one variable, while the themes show that it encompasses issues of livelihood loss, resentment over training that is not done with consent, and a need for disclosure. Second, in the experiment he conceals a boundary condition

that is only addressed by the qualitative strand: audiences don't enjoy being impersonated by a star, despite being entertained by a machine-assistant composition behind the star.

Table 10. Joint display of quantitative and qualitative findings (tentative)

Quantitative finding	Qualitative theme	Relation	Interpretation
AI label lowers humanness (eta-squared .18)	Theme 6: polished but emotionally absent	Convergence	The penalty is felt as the loss of a human presence
Star buffering (interaction $p = .002$)	Theme 1: the star as guarantor	Expansion	Buffer covers composition, not imitation of the star's voice
Smaller label gap for comic songs (0.33 vs 0.92)	Theme 2: humor as a licence	Expansion	Tolerance depends on whether the humor relies on wordplay
Hit potential linked to acceptance and concern	Theme 3: hit song as measure and worry	Convergence	Commercial promise and unease coexist
Cultural fit predicts acceptance (beta .19)	Theme 4: authenticity and rootedness	Convergence	Rootedness matters most for diaspora respondents
Ethical concern rises with AI label (eta-squared .13)	Theme 5: livelihoods and disclosure	Expansion	Concern bundles job loss, consent and transparency
Unlabelled group close to human-labelled	Theme 6: awareness that labels shape judgment	Convergence	Expectancy drives part of the penalty

Source: (Author Compilation)

7.2 Answering the research questions

RQ1. The use of AI authorship does significantly and systematically reduce audience ratings, particularly in the areas of perceived authenticity and moral concerns. The unlabelled condition is closer to the human labelled condition, and the effect can be said to be a disclosure penalty rather than a perceptible acoustic deficit.

RQ2. Star power reduces the AI penalty by about half, but this is for AI when composing, not for synthetic imitation of the star's voice.

RQ3. Humor reduces resistance. Comic songs are the ones with the smallest gap in the labels, but the respondents caution that they are still an area stronghold of human skill, as the jokes depend upon the wordplay and local idiom.

RQ4. Two audience perspectives exist at the same time. They understand that the hit song is a formula that can be taught to machines and they fear that formulaic success takes away from

authenticity, variety and fair compensation. Most of the respondents advocate for the resolution to be transparency + human direction: AI is a tool, not a creator.

7.3 Relation to the existing literature

The preliminary results are consistent with the international evidence of bias of the AI composer and performer (Ansani et al., 2025; Shank et al., 2023), the degradation of music labeled as AI-made pop (Chia et al., 2025) and the reductions in perceived expressiveness (White et al., 2025) and the role of origin beliefs in cognitive, emotional and physiological responses (Dunham et al., 2025). They build on that evidence in three ways. First, they find the song in a regional film context where the song is a part of a story and of a stardom. Second, they demonstrate how the context, often not tested in the general literature such as star presence, comic register and the hit song ideal, moderate the degree of the penalty. Third, they link the bias results with the business and ethical issues of synthetic vocals and the discussion around AI in the entertainment economy in India (Chavda et al., 2025; Khan et al., 2025).

It is important to note the difference between liking and endorsing. Furthermore, in the tentative data, perceived quality is not significantly lower than intention to accept in the AI labelled condition ($M = 3.31$) and there are many times in which the respondents' open-ended comments distinguish between enjoyment and intention to accept. This is in line with the work on "liking without endorsing" (Edelblum & Poe, 2026) and with the idea that the perceived source is important to appreciation, as well as the structure and self (Thompson et al., 2023). In terms of film specifically, the idea that AI-based attribution could lead to a damping effect on narrative meaning making (Wu & Holmes, 2026) suggests that it might be songs with story content, like the romantic ballad in this study, that would suffer the most from the disclosure of AI.

The results also render unconvincing any easy story of rejection. The blind and unlabelled results are also consistent with other research where listeners and professionals evaluate AI-generated music differently (Tigre Moura & Maw, 2021) and where the perceived quality of AI-generated music is rated similarly when its origin is hidden (Ramos & Ventayen, 2025). The plausible reading is that audiences don't care about the sound being produced by AI, they care about what the label appears to state about what it takes to create the sound, what the sound was created by, and what the sound was created with. These are in line with aesthetic and ideological critiques, which see AI music not merely as a matter of sound but one of meaning and power (Antipina, 2025), Wiradarmo et al., 2025), (Pram & Morreale, 2025).

7.4 Interpreting the three title concepts together

The three ideas in the title can be best conceptualized as three distinct forms of license that audiences have (or do not have) for AI in everyday Hollywood cinema. A relational license; trust in a person is applied to the work, as stars do. If the tone is playful, there's a licence for jokingness: if the tone is serious, then there isn't. Hit songs are licensed to the market: If a song is popular, then all seems fair and just, except when it comes into conflict with moral and cultural values. If all three licenses are true, then if a comic number is sung by a well loved star and it becomes a hit, the AI penalty should be the lowest. When none do, as with an emotionally serious ballad by an unknown singer with an AI label, resistance should be at its highest. It provides a testable extension of the general theories of AI-label bias that is specific to the context.

8. Implications

8.1 For film producers and music directors

There is no cost of disclosure, but there is an ethical and reputational risk to concealment. The preliminary findings indicate that when AI is presented as a tool to assist the human creator in well-defined and human-directed roles, like in the comic and dance scenes, audiences are more accepting, especially when they hear the star's voice and performance. It is important that producers do not make synthetic copies of known singers' voices without permission and adequate compensation because this is where there is the greatest resentment.

8.2 For Musicians/ Lyricists/ Unions

The livelihood theme suggests that the public is open to arguments in favour of a fair credit and pay. This openness can be used to promote consent-based pedagogical methods, to attribute the names of human musicians and lyricists to the conventions, and to establish contractual regulations regarding voice likeness. The sub-theme of humor as a linguistic art also implies a defensible space for human superiority in the making of Punjabi lyrics.

8.3 Streamers and marketers

Platforms could include metadata fields to mark that AI was used, giving users the ability to select the content and platforms the ability to give credit to the creators. The message for the marketing campaigns should not be that the AI is replacing human creativity or artistry, but rather that it is a tool that can augment and complement human composition and performance.

8.4 For policy makers and regulators

In addition to concerns about audiences, disclosure norms and protections for voice and transparency in training data are also included. These can be discussed as guidelines or legislation. There is scope for developing policy in the Indian context that can take advantage of the film sector-level analysis undertaken by Khan et al. (2025) and the understanding that regional-language industries have their unique interests. Issues in copyright law and engineering practice in AI music are acute (Sturm et al., 2019) in a field where the recognizable voice is a commercial asset.

8.5 For researchers

The study demonstrates the importance of using realistic film contexts to embed stimuli, the manipulation of labels instead of statements of attitudes, and the combination of experiments with qualitative inquiry. Research on Punjabi-cinema in the future could incorporate biometric data, such as that utilized in audiovisual studies (Fišer et al., 2025) and physiological studies on origin beliefs (Dunham et al., 2025). Existing emotion instruments (Coutinho & Scherer, 2017) and emotion evaluation criteria focusing on conveyance accuracy (Gao et al, 2025; Qi et al, 2026) can be applied to determine whether the intended emotions of stimulus songs are induced.

9. Limitations and Future Research.

There are a number of limitations that need to be recognized. First and most important, the empirical results in this version are tentative and illustrative and must be replaced with real data. Secondly, the sampling is planned and is not of a random nature, so the results cannot be generalized to the entire Punjabi-film audience. Third, short excerpts and fictional star cues make the lived experience of watching a film easier to understand; plot, picturisation and dance influence its reception. Fourth, when using self-report, responses may not accurately reflect actual behaviors and physiology (Dunham et al., 2025; Fišer et al., 2025). Fifth, the label manipulation is, to some extent, misleading, and disclosure can have different impacts depending on the credibility of the disclosure. Sixth, the literature reviewed here was limited to the sources selected and was summarised from the title and framing, and a comprehensive

systematic review incorporating the emotional-response review (Fernando et al., 2024) would enhance the theoretical basis. Lastly, interpretation is the key element of thematic analysis, and different coders will develop different themes, with double coding helping to reduce this.

There are several directions in which the design could be expanded for future research: longitudinal tracking of attitudes as AI music becomes more ubiquitous; comparison across regional industries (Bhojpuri, Haryanvi, Tamil and Bengali cinema); experiments which isolate the synthetic voice imitation from the synthetic music composition; physiological measurement in cinemas; and social-media analysis of discourse regarding AI-assisted music. Research into work being produced by amateurs and independent creators might reveal the way that AI is being used outside of the commercial stream. Controlling emotional profiles of stimuli could be achieved by using emotion-driven generation systems (Dash & Agres, 2024; Ning et al., 2024; Zheng et al., 2022), and applied settings like therapeutic uses of generative music (Shen et al., 2024) could evaluate the difference in acceptance when the purpose is not entertainment. Evidence of universal and culture-specific emotional cues (Athanasopoulos et al., 2021; Cowen et al., 2020) could be used to inform a cross-cultural design.

10. Conclusion

This paper aimed to explore how the audience of popular Punjabi films feels about AI-generated music, with the help of prolific film stars, the presence of humor and hit songs as its organising principle. It builds on a range of research that has been conducted on listener bias, emotional response, perceived humanness, synthetic vocals, and the Indian movie industry to come up with a theoretical framework, a mixed methods design plan and a preliminary analysis of how the findings would be presented.

The illustrative analysis points to the fact that the AI label, not the sound itself, evokes a penalty that is focused on perceived "humanness" and "ethical concern," that stars shield the penalty but do not remove the need for linguistic craft to make the song "humorous," and that the ideal of the hit song both attracts listeners while simultaneously causing them discomfort. These trends reflect a common need of audiences for technology to be a servant, not a destroyer, of human artistry, cultural significance and just reward. After the real data is gathered, the framework can be tested, updated, and expanded, and could assist the Punjabi film industry in a manner that ensures that their viewers trust the use of AI.

References:

1. Agudo, U., Arrese, M., Liberal, K. G., & Matute, H. (2022). Assessing emotion and sensitivity of AI artwork. *Frontiers in Psychology*, 13, Article 879088. <https://doi.org/10.3389/fpsyg.2022.879088>
2. Ansani, A., Koehler, F., Giombini, L., Hämäläinen, M., Meng, C., Marini, M., & Saarikallio, S. (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. <https://doi.org/10.1177/02762374241308807>
3. Antipina, I. (2025). Aesthetic aspects of the interaction between artificial intelligence and musical art. *Journal of Tchaikovsky National Music Academy of Ukraine*, 2(67), 39–57. [https://doi.org/10.31318/2414-052X.2\(67\).2025.339136](https://doi.org/10.31318/2414-052X.2(67).2025.339136)
4. Athanasopoulos, G., Eerola, T., Lahdelma, I., & Kaliakatsos-Papakostas, M. (2021). Harmonic organisation conveys both universal and culture-specific cues for emotional expression in music. *PLOS ONE*, 16(1), Article e0244964. <https://doi.org/10.1371/journal.pone.0244964>

5. Banerjee, S. D. (2025). Sound of the future: Investigating the impact of artificial intelligence on music creation and consumption in the 21st century. Zenodo. <https://doi.org/10.5281/zenodo.15718648>
6. Briot, J.-P., & Pachet, F. (2020). Deep learning for music generation: Challenges and directions. *Neural Computing and Applications*, 32(4), 981–993. <https://doi.org/10.1007/s00521-018-3813-6>
7. Caramiaux, B., & Donnarumma, M. (2021). Artificial intelligence in music and performance: A subjective art-research inquiry. In E. R. Miranda (Ed.), *Handbook of artificial intelligence for music: Foundations, advanced approaches, and developments for creativity* (pp. 75–95). Springer. https://doi.org/10.1007/978-3-030-72116-9_4
8. Carnovalini, F., & Rodà, A. (2020). Computational creativity and music generation systems: An introduction to the state of the art. *Frontiers in Artificial Intelligence*, 3, Article 14. <https://doi.org/10.3389/frai.2020.00014>
9. Chavda, B., Pillai, R., & Naresh, B. (2025). Towards a rhythmic economy—AI tuning the entertainment industry: A global scenario via Indian route. IGI Global. <https://www.igi-global.com/chapter/towards-a-rhythmic-economy-ai-tuning-the-entertainment-industry/368104>
10. Chia, S., Hartanto, A., & Tong, E. M. W. (2025). Do listeners devalue AI-generated pop music? Exploring negative biases in listeners' responses to AI-labelled vs. human-labelled pop music. *Computers in Human Behavior: Artificial Humans*, 6, Article 100217. <https://doi.org/10.1016/j.chbah.2025.100217>
11. Chong, C., & Li, A. (2023). Emotional music generation: An analysis of effectiveness and user satisfaction by using Python and Dart. *Computer Science & Information Technology*, 13(5), 219–228. <https://doi.org/10.5121/csit.2023.130519>
12. Chu, H., Kim, J., Kim, S., Lim, H., et al. (2022). An empirical study on how people perceive AI-generated music. In *Proceedings of the 31st ACM International Conference on Information & Knowledge Management* (pp. 304–314). Association for Computing Machinery. <https://doi.org/10.1145/3511808.3557235>
13. Coutinho, E., & Scherer, K. R. (2017). Introducing the GENEVA Music-Induced Affect Checklist (GEMIAC): A brief instrument for the rapid assessment of musically induced emotions. *Music Perception*, 34(4), 371–386. <https://doi.org/10.1525/mp.2017.34.4.371>
14. Cowen, A. S., Fang, X., Sauter, D., & Keltner, D. (2020). What music makes us feel: At least 13 dimensions organize subjective experiences associated with music across different cultures. *Proceedings of the National Academy of Sciences*, 117(4), 1924–1934. <https://doi.org/10.1073/pnas.1910704117>
15. Dash, A., & Agres, K. (2024). AI-based affective music generation systems: A review of methods and challenges. *ACM Computing Surveys*, 56(11), Article 287. <https://doi.org/10.1145/3672554>
16. Dunham, R. L., van Kleef, G. A., & Stamkou, E. (2025). The threat of synthetic harmony: The effects of AI vs. human origin beliefs on listeners' cognitive, emotional, and physiological responses to music. *Computers in Human Behavior: Artificial Humans*, 6, Article 100205. <https://doi.org/10.1016/j.chbah.2025.100205>
17. Edelblum, A. B., & Poe, J. (2026). Liking without endorsing: Consumer dilemmas in responses to AI-generated music. *Psychology & Marketing*, 43, 2229–2244. <https://doi.org/10.1002/mar.70170>
18. Fernando, P., Mahanama, T. V., & Wickramasinghe, M. (2024). Assessment of human emotional responses to AI-composed music: A systematic literature review. In *2024 International Research Conference on Smart Computing and Systems Engineering (SCSE)* (Vol. 7, pp. 1–6). IEEE. <https://doi.org/10.1109/SCSE61872.2024.10550861>

19. Figueiredo, F., Martinelli, G., Sousa, H., Rodrigues, P., Pedrosa, F., & Ferreira, L. N. (2025). Echoes of humanity: Exploring the perceived humanness of AI music. arXiv. <https://arxiv.org/abs/2509.25601>
20. Fišer, N., Martín-Pascual, M. Á., & Andreu-Sánchez, C. (2025). Emotional impact of AI-generated vs. human-composed music in audiovisual media: A biometric and self-report study. PLOS ONE, 20(6), Article e0326498. <https://doi.org/10.1371/journal.pone.0326498>
21. Gao, X., Chen, D., Gou, Z., Ham, J., & Cuijpers, R. H. (2025). Emotional fidelity in AI-generated music: A comparative analysis of conveyance and induction across models and modalities. Advanced Design Research, 3, 91–99. <https://doi.org/10.1016/j.ijadr.2025.11.001>
22. He, W., & Sun, Y. (2026). Human–AI emotional resonance in generative music: A mini-review of authorship attribution, affective response, and aesthetic evaluation. Frontiers in Psychology, 17, Article 1889696. <https://doi.org/10.3389/fpsyg.2026.1889696>
23. Hong, J.-W., Peng, Q., & Williams, D. (2021). Are you ready for artificial Mozart and Skrillex? An experiment testing expectancy violation theory and AI music. New Media & Society, 23(7), 1920–1935. <https://doi.org/10.1177/1461444820925798>
24. Hong, J.-W., Fischer, K., Ha, Y., & Zeng, Y. (2022). Human, I wrote a song for you: An experiment testing the influence of machines' attributes on the AI-composed music evaluation. Computers in Human Behavior, 131, Article 107239. <https://doi.org/10.1016/j.chb.2022.107239>
25. Jabour, G. (2024). Drake or fake? Perceptions, concerns, and business implications of AI-generated vocals [Master's thesis, The University of Texas at Austin]. Texas Digital Library. <https://repositories.lib.utexas.edu/items/eb68be79-452c-4b21-a5ab-a88778ad0f75>
26. Juslin, P. N., & Västfjäll, D. (2008). Emotional responses to music: The need to consider underlying mechanisms. Behavioral and Brain Sciences, 31(5), 559–575. <https://doi.org/10.1017/S0140525X08005293>
27. Khan, S. U., Khan, S., Ballewar, S. S., Rane, J. P., & Siddiquee, A. Q. (2025). Cinema and artificial intelligence: Charting its role in the Indian film sector. In U. Kiliç (Ed.), Transforming cinema with artificial intelligence (pp. 143–180). IGI Global Scientific Publishing. <https://doi.org/10.4018/979-8-3693-3916-9.ch007>
28. 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), Article 21582440251406713. <https://doi.org/10.1177/21582440251406713>
29. Ning, Z., Han, X., & Pan, J. (2024). Semi-supervised emotion-driven music generation model based on category-dispersed Gaussian mixture variational autoencoders. PLOS ONE, 19(12), Article e0311541. <https://doi.org/10.1371/journal.pone.0311541>
30. Oehler, M., Schepmann, J., & Saubier, L. (2025). Assessing expectancy bias in listener evaluations of AI and human music compositions. Proceedings of the 6th Conference on AI Music Creativity (AIMC 2025). <https://doi.org/10.5281/zenodo.16955781>
31. Pram, L., & Morreale, F. (2025). Opening musical creativity? Embedded ideologies in generative-AI music systems. arXiv. <https://arxiv.org/abs/2508.08805>
32. Qi, G., Song, D., & Yu, X. (2026). A study on the emotional accuracy of generative AI music from the perspective of the “Emotion Wheel” theory. Frontiers in Psychology, 17, Article 1794715. <https://doi.org/10.3389/fpsyg.2026.1794715>
33. Ramos, D. L. T., & Ventayen, R. J. M. (2025). Monitoring the perceived quality of artificial intelligence composed music through blind listener evaluation. In 2025 6th

- International Conference on Data Intelligence and Cognitive Informatics (ICDICI). IEEE. <https://doi.org/10.1109/ICDICI66477.2025.11134965>
34. Scherer, K. R. (2005). What are emotions? And how can they be measured? *Social Science Information*, 44(4), 695–729. <https://doi.org/10.1177/0539018405058216>
35. Shank, D. B., Stefanik, C., Stuhlsatz, C., Kacirek, K., & Belfi, A. M. (2023). AI composer bias: Listeners like music less when they think it was composed by an AI. *Journal of Experimental Psychology: Applied*, 29(3), 676–692. <https://doi.org/10.1037/xap0000447>
36. Shen, L., Zhang, H., Zhu, C., Li, R., Qian, K., Meng, W., Tian, F., Hu, B., Schuller, B. W., & Yamamoto, Y. (2025). A first look at generative artificial intelligence-based music therapy for mental disorders. *IEEE Transactions on Consumer Electronics*, 71(3), 7439–7453. <https://doi.org/10.1109/TCE.2024.3514633>
37. Stammer, D., Strauss, H., & Knees, P. (2025). Perception of AI-generated music: The role of composer identity, personality traits, music preferences, and perceived humanness. *arXiv*. <https://arxiv.org/abs/2512.02785>
38. Sturm, B. L. T., Iglesias, M., Ben-Tal, O., Miron, M., & Gómez, E. (2019). Artificial intelligence and music: Open questions of copyright law and engineering praxis. *Arts*, 8(3), Article 115. <https://doi.org/10.3390/arts8030115>
39. Sun, D., Wang, H., & Xiong, J. (2024). Would you like to listen to my music, my friend? An experiment on AI musicians. *International Journal of Human–Computer Interaction*, 40(12), 3133–3143. <https://doi.org/10.1080/10447318.2023.2181872>
40. Thompson, W. F., Bullot, N. J., & Margulis, E. H. (2023). The psychological basis of music appreciation: Structure, self, source. *Psychological Review*, 130(1), 260–284. <https://doi.org/10.1037/rev0000364>
41. Tigre Moura, F., & Maw, C. (2021). Artificial intelligence became Beethoven: How do listeners and music professionals perceive artificially composed music? *Journal of Consumer Marketing*, 38(2), 137–146. <https://doi.org/10.1108/JCM-02-2020-3671>
42. White, C. W., Kapoor, K., Cosme-Clifford, N., Symons, J., & von Mutius, L. (2025). Humans perceive AI-generated music as less expressive than comparable human-made content. *SSRN*. <https://doi.org/10.2139/ssrn.5087035>
43. Wiradarmo, A. A., & Azhar, H. A. (2025). Machine kitsch theory: Contrasting shifts in public perceptions towards AI-generated art. *Jurnal Desain*, 12(3). <https://doi.org/10.30998/jd.v12i3.26970>
44. Wu, S. H., & Holmes, K. J. (2026). Is there a “mind” behind the music? Attributing music to AI can suppress narrative meaning-making. *Cognitive Research: Principles and Implications*, 11, Article 19. <https://doi.org/10.1186/s41235-026-00715-z>
45. Yin, Z., Reuben, F., Stepney, S., & Collins, T. (2023). Deep learning's shallow gains: A comparative evaluation of algorithms for automatic music generation. *Machine Learning*, 112, 1785–1822. <https://doi.org/10.1007/s10994-023-06309-w>
46. Zheng, K., Meng, R., Zheng, C., Li, X., Sang, J., Cai, J., Wang, J., & Wang, X. (2022). EmotionBox: A music-element-driven emotional music generation system based on music psychology. *Frontiers in Psychology*, 13, Article 841926. <https://doi.org/10.3389/fpsyg.2022.841926>