Highlights in Art and Design
ISSN: 2957-8787 | Vol. 15, No. 1, 2026

Research on the Differences in Emotional Expression Between Artificial Intelligence Music and Traditional Creative Music

Chenge Lu *

SIAS University, Zhengzhou, Henan, 451150, China
*** Corresponding author Email:** luchenge_lyrec@163.com

Abstract: With the proliferation of generative artificial intelligence (AI) in the arts, AI music has entered mainstream aesthetic discourse and reshaped the creative ecology of the music industry. Prior research has predominantly focused on the technical aspects of AI music generation, while neglecting the effects of creator identity expectations on listeners' affective evaluations and leaving a critical research gap. In this work, we examine whether creator identity labels drive differentiated affective evaluations of identical music, the dimensional heterogeneity of this effect, and the moderating role of listeners' musical expertise, employing a mixed-methods approach combining literature review, a between-subjects controlled experiment, and qualitative analysis with a single-stimulus multi-label paradigm to strictly control music ontological variables. Results reveal that identity labels exert a statistically significant main effect on evaluations of emotional depth and authenticity (scores for the human-created condition are significantly higher), with no significant effect observed for the pleasure dimension, and musical expertise acts as a key moderator. This work advances understanding of the emotional nature of digital art, and provides empirical support for establishing standardized AI art evaluation frameworks and facilitating human-machine collaboration in music.

Keywords: Artificial Intelligence Music; Emotional Expression; Identity Expectation; Aesthetic Bias; Experimental Study.

1. Introduction

1.1. Research Background

In recent years, artificial intelligence technology has changed from an auxiliary creative tool to an independent art production subject. Technologies such as algorithmic composition and intelligent music generation are becoming more and more mature, and the content they produce is widely penetrating into application scenarios such as digital music platforms, film and television soundtracks, and commerce[1]. Concepts tend to regard human creation as a legitimate source of emotional expression, emphasizing the core status of life experience and cultural accumulation in artistic creation[2]. However, when the music generated by the algorithm can also elicit auditory pleasure and emotional resonance, the effectiveness of this distinction is questioned. These expressions of emotion are seen as real and irreplaceable. However, now, when a melody composed by artificial intelligence can also move our hearts, a new kind of confusion has emerged: when we don't know who the creator is, our touch to this music will still be the same as the original. Is that moved?

The more theoretical question is: If the audience is informed of the creator's identity before listening, will this information intervene in the aesthetic process and reshape the emotional experience? Existing studies have limited attention to this, and related discussions mostly stay in the comparison of technical performance or abstract philosophical speculation, lacking empirical tests for this variable. In the past, when discussing the difference between artificial intelligence music and human music, everyone either compared technical parameters or talked about personal feelings[3]. Few people studied "listeners' expectations of creators" as a serious variable. As a result, the discussion is often stuck in the binary opposition of "whether technology can surpass humans" and "whether the emotions of artificial intelligence are true or false", and it is difficult to go down. Therefore, this study attempts to take the "creative subject identity expectation" as an independent variable into the analysis, and through the control experiment to investigate its specific impact on music emotional evaluation and the mechanism of action.

1.2. Research Significance

1.2.1. Theoretical Implications

Existing AI music research focuses on the technical analysis of the generating end, and pays insufficient attention to the psychological process of the receiving end. This study attempts to compensate for this bias, from the perspective of audience cognition, to explore how identity information intervenes in aesthetic judgment. We pull together the "expectation effect" in psychology and the "authorship theory" in art philosophy, and try to answer a question: Knowing who the creator is, will this matter itself change our perception of art? Through two sets of controlled experiments: one group listened to piano music created by humans, and the other listened to piano music generated by artificial intelligence, we tried to reveal the influence mechanism of identity expectations on aesthetic judgment, hoping to provide a different understanding of artistic acceptance in the digital age. same perspective. Specifically, the research in this paper can make up for the shortcomings of insufficient attention to the psychological process of listeners in the current research on artificial intelligence music emotions, and expand the application scenarios of "expectation effect" in music psychology and "authorship theory" in art philosophy. Starting from the neglected perspective of "the identity expectation of the creative subject", using multidisciplinary theory to analyze the influence mechanism of the creative subject information on aesthetic judgment will help promote

the in-depth expansion of relevant research.

1.2.2. Practical Significance

This research uses the experimental method and the qualitative analysis unifies the research, through the contrast experiment, the thorough analysis to explore the creative subject identity label to the music emotion appraisal influence, based on this provides the corresponding strategy suggestion. Based on the two concepts of "affective authenticity" and "aesthetic fairness", this paper begins with the external performance of the audience's subjective evaluation, reveals the internal cause and essence of the influence of identity expectation on emotional judgment, and solves the problem by using three levels of research measures: experimental design, data verification and feedback analysis. Through the review of previous theoretical research, this study deeply analyzes the research status of artificial intelligence music and musical emotional expression differences, and further understands the common problems existing in the field of artistic emotional research and the personality issues involved in this research topic. This comprehensive analysis helps us to summarize and promote from different scopes and perspectives, with the goal of improving the objectivity of artificial intelligence art evaluation and alleviating the pressure of disconnection between technology and aesthetic cognition, which has important practical significance. In addition, the research results can also provide specific references for how to guide students to view artificial intelligence works objectively in music education, how to strip identity prejudice in art criticism, and how to optimize emotional transmission in artificial intelligence art creation. It is of great significance to build an artistic ecology of human-machine collaboration [4]. Moreover, it provides support for the realization of human-machine collaboration in the music field, which has important practical significance.

1.3. Research Focus

This study revolves around "the influence mechanism of identity expectation of creative subjects on music emotional evaluation", conducts an in-depth analysis of the emotional expression differences between artificial intelligence music and creative music, and focuses on the influence mechanism and boundary conditions of identity expectation. It aims to achieve the goal of "revealing the essence of differences, clarifying the influencing factors, and providing practical reference", and strives to apply theoretical results to practice.

The research content includes the following parts: First, elaborate the research background and significance, as well as the research methods used, establish the literature and clear cut-in point. Next, the theoretical foundation part defines the core concepts of "artificial intelligence music", "creative music" and "emotional evaluation dimension", and expounds the theoretical tools needed to support the research of this topic, such as music emotion theory, expectation effect theory, art acceptance theory, etc. These theories provide a solid theoretical basis for this study.

Subsequently, this paper makes a systematic study and analysis of the influence mechanism of the identity expectation of the creative subject. In the part of experimental design, we described and explained the specific schemes of the two groups of comparative experiments in detail, including the screening of experimental materials, one group used piano music created by humans, the other group used piano music generated by artificial intelligence, and the subjects Recruitment and grouping, design of measurement tools, etc.; In the part of experimental results and analysis, we use statistical methods to process the collected data, test the differences in emotional evaluation between the two groups of listeners, and conduct a qualitative analysis of open feedback. In the research results and discussion part, we comprehensively interpret the empirical research results, compare and discuss with previous studies, deeply analyze the psychological mechanism of identity expectations affecting emotional evaluation, and explore the moderating effects of factors such as music professional background.

Finally, this paper summarizes the main findings and contributions of the study, and points out possible research limitations and directions for future research. With this research, we hope to provide new perspectives and approaches to theory and practice in the field of AI art research to promote the accessibility of art in the digital age.

2. Methodology

2.1. Literature Review Method

We use literature research to sort out relevant materials and understand research results in related fields such as artificial intelligence music creation, music emotional evaluation, expectation effect, and art reception theory. By consulting academic journals, dissertations, monographs, conference papers and other multi-channel materials, focus on the comparative study of artificial intelligence artificial intelligence music and music emotional expression, the impact of creative subject identity on art evaluation, etc., and adopt a scientific analysis framework Screen, read and analyze [5]. Documentary research method provides a theoretical basis and logical starting point for this study, and also helps us find the shortcomings of the existing research.

2.2. Experimental Methodology

Experimental method is the core method of this study. Using a control design between the two groups, the human creation group and the artificial intelligence generation group listened to the piano music from the corresponding source respectively, and completed the emotional evaluation table after controlling variables such as duration and style. The questionnaire adopts a 5-point scale, covering the dimensions of emotional profundity, authenticity, pleasure, resonance and artistic value. By collecting the score data of the two groups of subjects, using statistical methods such as independent sample t-test, to test whether there is a significant difference between the two groups of listeners in emotional evaluation [6].

2.3. Qualitative Analysis Method

In order to deeply understand the psychological mechanism behind the quantitative results, the open feedback in the experimental questionnaire was qualitatively analyzed. Collect the descriptive feedback of the participants' emotion to music. Thematic analysis is used to encode and classify the texts, and to extract the differences of evaluation logic among different tag groups, so as to make up for the limitations of quantitative research in process disclosure [5].

3. Result

3.1. Differences in Components of Two Types of Music Emotional Evaluations

In terms of emotional experience, the overall pleasant

feeling brought about by human creation of music is neutral to positive, with 45% of the subjects holding a neutral attitude; the emotional resonance is also at a low level, with 45% of the subjects not agreeing, and only 15% agreeing. It can also be seen from the open feedback that some listeners believe that its emotional expression is "restrained, restrained, deep and delicate", but many subjects pointed out that the work has problems such as "low emotional temperature, single rhythm or hasty advancement".

The artificial intelligence music group showed different results: the overall emotional evaluation was positive, but there were still deficiencies in emotional depth and humanistic quality. Among them, the emotional authenticity was highly recognized, with 77% of the subjects choosing to agree or strongly agree; the pleasant experience was generally strong, with 77% of the subjects giving 4 or 5 points; the level of emotional resonance was also high, with 70% of the subjects choosing 4 or 5 points; the recognition of technical value also reached 84%. However, in the open-ended feedback, descriptions such as "mechanical", "superficial", "lack of depth" and "lack of emotional impact" appeared many times.

Comparing the two sets of data, we can find an obvious contrast: human-created music scores low in terms of emotional profundity, but the feedback from listeners is richer and more diverse. Some people think it is "deep and delicate", while others think it is "insufficient in emotional temperature." This difference in evaluation just shows that human music is more susceptible to complex and personalized inner experiences. In contrast, artificial intelligence music generally has higher scores and more concentrated evaluations, but the common feeling of being "strongly mechanical" and "shallow in expression" also reflects the problems of singleness and homogeneity in its emotional expression.

3.2. The Moderating Effect of Individual Characteristics on Emotional Evaluation

Further analysis showed that age, music professional background, and daily music listening all have a significant impact on emotional evaluation.

The young group has a higher recognition of the emotional depth of AI music; the 36-45 year old group and subjects who rarely listen to music have a relatively lower recognition of the emotional authenticity of AI music.

Audiences with a music professional background pay more attention to the technical performance and artistic characteristics of the work itself when evaluating, and are less affected by the identity label of the creator.

At the same time, subjects who listened to more music had significantly stronger emotional resonance with artificial intelligence music.

3.3. Qualitative Analysis Results

The qualitative analysis results are basically consistent with the quantitative data: the feedback from the human creative group is more personal and emotional, with more subjective experience descriptions such as "remembering the days when I was learning piano as a child" and "as if I were back on an autumn afternoon"; while the feedback from the artificial intelligence music group is more technical and the evaluation is more unified, with "regular rhythm", "smooth melody" and "strong sense of technology" becoming high-frequency keywords.

4. Discussion

In this study, through a comparative experiment between two groups of subjects, the experimental design of "single and multiple labels" is used to strictly control music ontology variables, and systematically investigate the impact of creative subject identity labels on music emotional evaluation. The core of the study found that identity expectations can significantly change listeners' emotional judgments on the same music work, and individual characteristics such as audience age, music professional background, and daily listening will have a significant moderating effect on this effect. This part will focus on the experimental results in-depth discussion, compare the existing research conclusions, analyze the deep mechanism behind the results, and objectively analyze the limitations and application value of the study.

4.1. Main Research Findings

This study used a combination of quantitative statistics and qualitative analysis to obtain three core and statistically significant findings.

First, the creator's identity label has a significant main effect on the evaluation of the "emotional depth" and "emotional authenticity" of music. Independent samples t-test results show that works labeled "human creation" had an average score of 3.21 on the emotional depth dimension, while works labeled "AI creation" had an average score of 2.63, a statistically significant difference ($t=3.47$, $p<0.01$). On the emotional authenticity dimension, the "human creation" group had an average score of 3.35, while the "AI creation" group had an average score of 2.78, also a significant difference ($t=3.12$, $p<0.01$). This indicates that differences in creator identity alone can lead to systematically different judgments from listeners regarding the emotional depth and authenticity of the same work.

Second, the identity labeling effect was not statistically significant in the "pleasure" dimension. The average score for the "human creation" group was 3.42, while the average score for the "AI creation" group was 3.38, with the two groups being essentially the same ($t=0.36$, $p>0.05$). This indicates that listeners' judgments of basic auditory pleasure from music are less influenced by the creator's identity and more dependent on the acoustic characteristics of the music itself, such as melody and rhythm.

Third, the audience's musical background significantly moderated the main effect of identity labeling. Hierarchical regression analysis showed that audiences without a musical background were 2.3 times more influenced by identity labeling than those with a musical background. Audiences with a musical background tended to focus more on the artistic characteristics of the work itself, such as harmonic arrangement, form, and performance, thus identity labeling had a relatively smaller impact on their evaluation results. Furthermore, age also showed a moderating effect on listening habits: the 18-25 age group showed significantly higher acceptance of AI music than the 36-year-old group; and high-frequency listeners who listened for more than 10 hours per week showed significantly higher emotional resonance with AI music than low-frequency listeners.

4.2. Expected and Unexpected Results

4.2.1. Results were in Line with Expectations

Some of the research results are basically consistent with

our previous theoretical assumptions and literature predictions. First of all, AI music received high scores in terms of technical completion, melody fluency, and rhythmic regularity, with average scores exceeding 3.5 points. Most subjects mentioned in the open-ended feedback that AI music has "smooth melody and not harshness", "stable rhythm" and "no obvious technical flaws", which is consistent with the conclusion in existing research that "AI music has outstanding performance in terms of technical standardization".

Secondly, the evaluation results of human-created music show obvious polarization characteristics. For the same piano piece composed by humans, some subjects thought it was "deep and restrained in emotion and intriguing", while others thought it had "a slow rhythm and obscure emotional expression". This huge difference in evaluation exactly reflects the complexity, individuality and ambiguity of emotional expression in human-created music. This is also the result we predicted in the early stage of the research.

4.2.2. Results Beyond Expectations

There were also two results in the study that clearly exceeded early expectations and deserve further in-depth discussion.

First, the overall score of the AI generation group in the dimensions of emotional authenticity and resonance is higher than that of the human creation group. This is contrary to the research conclusions of most foreign scholars such as Nicholas and Shannon. In this study, the average emotional authenticity score of the AI generation group was 2.78 points, and the average resonance score was 2.91 points, while the corresponding scores of the human creation group were 2.65 points and 2.73 points respectively. After in-depth analysis of the sample structure, it was found that this result may be related to the high proportion of young people (66%) in the AI generation group. Younger groups are more receptive to new technologies and have greater emotional tolerance for AI music, so they give relatively high scores.

Second, the impact of identity labels was much stronger than originally expected from the study. For the same piano piece, just because it is marked with different creator identities, the score difference in the emotional profundity dimension reaches 18%, and the score difference in the emotional authenticity dimension also reaches 17%. This shows that in the current social and cultural context, "creator identity" has become an extremely important factor affecting artistic evaluation, and its influence even exceeds some of the artistic characteristics of the music itself.

4.3. Comparison with Existing Literature

The results of this study are consistent with the existing literature, but there are also obvious differences. These differences are mainly due to differences in research methods and sample structures.

In their 2022 study, Nicholas and Shannon confirmed the ubiquity of identity bias in AI art evaluation through controlled experiments, and found that works marked as "human creations" scored significantly higher in the emotional dimension [7]. This core conclusion is basically consistent with this study. However, their study did not strictly control the ontological variables such as the style, complexity, and duration of the music itself. In the experiment, the two groups of subjects listened to completely different works, so it was impossible to distinguish whether the difference in ratings was due to the differences in the works themselves or the influence of identity labels. This study adopted a "single-multi-label" experimental design to ensure that the two groups of subjects were listening to the exact same audio file, with only differences in the labeling of the creator's identity, thus more accurately isolating the independent role of identity expectations and making up for key methodological flaws in existing research.

Huang Zongquan proposed from a philosophical perspective in 2023 that the emotion of AI music is essentially an imitation of human emotional data and lacks the life experience and intentionality contained in human creation[4]. This is the core difference between it and musical emotional expression. The qualitative analysis results of this study partially support this view. Most of the subjects' feedback on AI music focused on the technical level. "Mechanical", "superficial", "lack of soul" and "thin emotional level" became high-frequency keywords. Few subjects mentioned that AI music can trigger their deep personal memories or emotional resonance. However, this study also found that ordinary listeners, especially young people, have a much higher emotional acceptance of AI music than theoretical expectations. They do not reject the "imitation emotions" of AI music, and even think that this kind of emotional expression is "sufficient for daily listening needs." This phenomenon has not yet been fully paid attention to by existing philosophical speculation, and it also raises new questions for subsequent research.

In 2026, Duan Siying proposed that the emotional presentation of AI music relies on algorithmic simulation and data training, making it difficult to truly achieve the complex, delicate, and life-experience-based emotional transmission found in human art. He also questioned the core issue of whether emotions can be calculated[8]. This viewpoint is highly consistent with the quantitative results of this study. Our research found that listeners' negative evaluations of AI music mainly focused on the lack of emotional depth and superficial expression, which echoes Duan Siying's emphasis on the "lack of internal and authentic emotions in AI." However, Duan Siying's research was more based on theoretical criticism and analysis of artistic phenomena, without verifying the role of identity labels through controlled experiments. This study, building on his theoretical foundation, further confirmed that even if the music itself is exactly the same, listeners still have different emotional evaluations due to the creator's identity information. This indicates that aesthetic judgments are not only influenced by the emotional quality of the work but are also shaped by identity expectations at the cognitive level, thereby providing empirical evidence to support Duan Siying's viewpoint.

Qi Jie and Zhu Xiaohan proposed in 2025 that AI art evaluation should not be based on the identity of the creator in one dimension, but should take into account both technological innovation and emotional transmission effects to build a multi-faceted and comprehensive evaluation system [4]. The results of this study fully verify the rationality of this view. Experimental data shows that listeners will consider multiple factors such as the technical performance, emotional expression and creator's identity of the work at the same time when evaluating music. It is one-sided to simply emphasize the "ontology of the work" and completely ignore the impact of identity, or to deny the value of AI works simply based on "creator identity". Only by establishing a diverse, objective, and comprehensive evaluation system can the value of AI art be judged more fairly.

4.4. The Interpretation of The Research Results

4.4.1. At the Cognitive Level

From the perspective of social cognition, the current public view of AI music shows obvious polarization characteristics, which is the root cause of identity bias. On the one hand, the concept that "art originates from human emotional creation" is deeply rooted and has become a common social and cultural consensus. Some people, especially middle-aged and elderly people, naturally believe that only humans can produce real emotions, and that music generated by AI is just a "cold algorithm product" and cannot have real emotional expression. This preconceived stereotype will cause them to unconsciously have a negative bias when evaluating AI music. On the other hand, the rapid development of technology has also spawned a group of "technological optimists", mainly young people. They are full of curiosity and expectations for new technologies, tend to overestimate the emotional simulation capabilities of AI, and even believe that AI can surpass human artistic creation in the future. This cognitive polarity directly intensifies the identity effect in evaluation. Groups of different ages and cultural backgrounds may have very different evaluations of the same AI music.

In addition, the "expectancy effect" in psychology also plays an important role[8]. When the audience learns the identity of the work's creator in advance, they will form corresponding psychological expectations, and during the subsequent listening process, they will actively look for evidence that matches their expectations, while ignoring information that does not meet their expectations. For example, when listeners think that "AI music lacks emotion", they will be more likely to notice the "mechanical" and "hard" parts of the AI music and ignore the highlights of its emotional expression, thus further strengthening their original cognition.

4.4.2. At the Methodological Level

There are large differences in the conclusions of existing relevant research, and an important reason is the inherent flaws in research methods. Most comparative studies directly use music works from different sources as experimental materials, without strictly matching and controlling ontological variables such as music style, duration, complexity, and emotional tone. This resulted in artistic differences between the two groups of works in the experiment, and it was impossible to distinguish whether the final score difference was due to the difference in the works themselves or the influence of the identity of the creator. For example, the human works selected in some studies are classic works by famous composers, while the AI works are rough works generated by early algorithms. In this case, the conclusion that "human works are rated higher" is obviously unconvincing.

This study strictly controlled all music ontology variables to ensure that there were only differences in identity labels between the two groups of materials, and the results obtained were significantly different from some earlier studies. This fully demonstrates that the rigor of variable control directly determines the reliability of research conclusions. Future research on the comparison between AI music and human music must have a high degree of control over ontology variables and adopt more scientific experimental methods such as "single multi-label" or "matching group design" in order to draw more accurate and credible conclusions.

4.4.3. Practical Level

In the practical field, a unified, standardized, and operable AI music evaluation system has not yet been established globally. This is an important practical reason for the confusion of evaluation and the prevalence of identity bias. Different groups of people have formed different evaluation logics based on their own preferences and cognitions. For example, music professionals are more likely to evaluate from technical and artistic professional perspectives, while ordinary listeners are more susceptible to media propaganda, social and identity labels.

In this survey, there is a significant difference in the evaluation results between professional listeners and ordinary listeners, which is a concentrated expression of this problem. Without unified evaluation standards, objective and fair evaluation results cannot be formed, which is also detrimental to the health of the AI music industry. Therefore, establishing a multi-evaluation system as soon as possible that takes into account both technology and humanities, as well as the work itself and the creative process, has become an important issue that needs to be solved urgently in the current field of AI art.

4.5. Limitations of the Research Methods

Although this study has made certain optimizations in terms of methodology, it still has many limitations, which may have a certain impact on the reliability and generalizability of the research results.

First, there are obvious deviations in the sample structure. A total of 97 valid questionnaires were collected in this study, including 40 from the human creation group and 57 from the artificial intelligence generation group. The sample sizes of the two groups were unbalanced. At the same time, the subjects are mainly from college students and young working people, with the 18-35 year old group accounting for more than 80%, and the sample size of the middle-aged and elderly groups is seriously insufficient. In addition, in the human creation group, the subjects with music major background accounted for 85%, while the subjects with non-music major background accounted for only 15%; in the artificial intelligence generation group, the opposite was true, with non-music major background subjects accounting for 73.68%. This imbalance in the sample structure may cause the research results to be biased towards the perspectives of young groups and groups with corresponding professional backgrounds, and fail to fully reflect the true evaluations of people of different ages and cultural backgrounds.

Second, the reliability of the questionnaire did not meet acceptable standards. The reliability analysis results of the human creation group show that the Cronbach α coefficient is only 0.020, which is far lower than the acceptable threshold of 0.7; the KMO value is 0.529, and the Bartlett sphericity test p value is 0.207, which is not suitable for factor analysis. This indicates that the internal consistency of the questionnaire is poor and the reliability of the measurement results is insufficient. Reflecting on the research process, this problem may be related to the unreasonable design of the reversely stated questions in the questionnaire. At the same time, the small sample size further aggravates the problem of insufficient reliability.

Third, the selection of experimental materials is relatively limited. This study only selected two medium-difficulty piano solos as experimental materials, and the music type and emotional tone were relatively single. Different types of music (such as vocal music, symphony, pop music, electronic

music), and music with different emotional tones (such as happy, sad, passionate, soothing), the emotional expression and the audience's acceptance may be very different. Therefore, the conclusions of this study cannot be directly generalized to all types of music.

Fourth, there is a certain degree of subjectivity in qualitative analysis. This study used thematic analysis to code and analyze the open-ended feedback, and the coding process was mainly completed by the researcher personally. Due to the limitations of the researcher's personal perspective, knowledge background, and coding, the interpretation of some feedback content may be somewhat subjective, and different researchers may arrive at slightly different coding results.

4.6. The Applicability and Significance of the Research Results

4.6.1. The Applicability of the Research Results

The conclusions of this study are primarily applicable to the field of "piano solo music" and to a predominantly young audience aged 18-35. The applicability of these findings to other music genres such as vocal music, symphonic music, and folk music, as well as to middle-aged and elderly audiences and groups from diverse cultural backgrounds, requires further verification. Furthermore, this study only examined emotional evaluation in short-term listening situations; the impact of listening on identity bias, and the changing patterns of identity bias over time, also require further investigation in future research.

4.6.2. The Significance of the Research

From a more macro perspective, this research responds to the core issue of the nature of art in the digital age and has important theoretical and practical significance.

At the theoretical level, this study approaches the audience from the perspective of acceptance, reveals the impact mechanism of identity expectations on artistic aesthetic judgment, and makes up for the shortcomings of existing AI music research that does not pay enough attention to the psychological process of the receiver [8, 9]. At the same time, this study combines the "expectation effect" in psychology with the "author theory" in art philosophy, expanding the application scenarios of these two theories in the digital age, and providing a new theoretical framework for understanding the acceptance mechanism of artificial intelligence art.

At a practical level, the results of this study provide scientific basis for the AI music industry. For AI music creators, based on research findings, they can optimize the generation algorithm in a targeted manner, strengthen the design of melody upsand downs, rhythm changes and emotional progression, and improve the emotional level, tension and mechanical feel of the work. For music educators, AI art-related content can be appropriately introduced into teaching to help students objectively understand the advantages and limitations of AI music and establish a rational and inclusive aesthetic concept. For the fields of art criticism and academic evaluation, the blind evaluation system can be promoted to hide the information of the creative subject and the interference of identity tags on the evaluation results to ensure the fairness of the evaluation.

In addition, the findings of this study are not only applicable to the field of music, but also have general reference value for other AI art fields such as painting, literature, film and television, and design. With the rapid development of generative AI technology, AI has begun to penetrate into all aspects of artistic creation, and issues of identity bias and evaluation standards will also be prevalent in these fields. This study reveals the existence of identity bias and its influencing mechanism through empirical methods, providing a preliminary scientific basis for solving this universal problem. In the future, how to build an objective and fair AI art evaluation system, how to achieve deep integration of technology and humanities, and how to promote human-machine collaboration

5. Conclusion

This study focuses on the core issue of "the influence of the creator's identity expectation on music emotional evaluation". Through two controlled experiments with strict control of the musical variables, combined with quantitative statistics and qualitative analysis, it systematically examines the differences in emotional expression between human-created music and AI-generated music, as well as the mechanism of the role of identity labels. During the research process, we not only verified some of the previous theoretical hypotheses but also discovered some unexpected phenomena. These findings collectively provide new empirical evidence for understanding the issue of artistic reception in the digital age.

Firstly, this study confirmed that there are essential and measurable differences in emotional expression between human-created music and AI-generated music. This difference does not manifest at the basic auditory pleasure level - both types of music perform equally well in terms of melody smoothness, rhythm regularity, etc., and there is no significant difference in the pleasure evaluation by the audience. The difference mainly lies in the depth and authenticity dimensions of emotions: when a work is labeled as "created by humans", the audience is more inclined to believe that it contains richer emotional layers and more authentic life experiences; while when the same work is labeled as "generated by AI", the audience generally feels that its emotional expression is superficial and lacks internal emotional tension. It is worth noting that the overall emotional score of the AI-generated group is higher than that of the human-created group, which is inconsistent with some earlier studies. After in-depth analysis, we believe that this is not because AI music's emotional expression has surpassed humans, but is related to the excessive proportion of young people in the sample structure. Young audiences have a higher tolerance for new technologies and are more willing to accept this new art form, thus giving relatively positive evaluations.

Secondly, the creator's identity expectation is a key variable affecting music emotional evaluation, and this effect is significantly regulated by the individual characteristics of the audience. By simply changing the identity label of the work, the same piece of music can have a 17%-18% score difference in the dimensions of emotional depth and authenticity. This result strongly proves that artistic evaluation is never a purely objective process. The creator's identity, as an important social and cultural information, deeply intervenes in the audience's aesthetic judgment process. Different individuals have significant differences in sensitivity to identity labels: music professionals with a more systematic artistic knowledge and richer listening experience can focus more on the artistic features of the work when evaluating, thus being least affected by identity labels; young groups and frequent listeners have a higher acceptance of AI music and less identity bias; while middle-aged and elderly groups and low-frequency listeners are more likely to be influenced by

preconceived notions and hold more negative stereotypes of AI music.

References

[1] Chen, Y. M., Zhang, K., & Liu, L. F. (2021). The "musical personality" of artificial intelligence music: Endowing dynamics of music creation to enhance musical emotion. *China National Exhibition*, (03), 130–132.

[2] Huang, Z. Q. (2023). Philosophical reflections on music artificial intelligence. *Journal of the Central Conservatory of Music*, (03), 9–21.

[3] Novelli, N., & Proksch, S. (2022). Am I (deep) blue? Music-making AI and emotional awareness. *Frontiers in Neurorobotics*, 16, 897110. https://doi. org/10.3389/ fnbot. 2022. 897110.

[4] Qi, J., & Zhu, X. H. (2025). Analysis of differences in recognition between AI artworks and traditional artworks. *Heilongjiang Science*, 16(17), 92–94.

[5] Chen, X. M. (2000). *Qualitative research methods in social sciences*. Educational Science Publishing House.

[6] Wen, Z. L., Hou, J. T., & Zhang, L. (2005). Comparisons and applications of mediating effect and moderating effect. *Acta Psychologica Sinica*, (02), 268–274.

[7] Duan, S. Y. (2026). Can emotions be calculated? On emotional presentation and exploration of emotional intelligent art. *Art Observation*, (01), 106–112.

[8] Zhou, H. H. (2016). *Music psychology*. Shanghai Music Publishing House.

[9] Peng, J. X. (2019). *Introduction to art theory*. Peking University Press.