Journal of Computers Vol. 36 No. 3, June 2025, pp. 335-348
https://doi.org/10.63367/199115992025063603022

The Impact Mechanism of Algorithmic Transparency on User Trust in Intelligent Recommendation Systems of Internet Platforms

Yan-Xue Guo[1*] and Xiang Zhang[2]

[1] Department of Economics and Management, Communication University of China,
Chaoyang District, Beijing, 100024, China
1205362827@qq.com

[2] Department of Mechanical Engineering, University of Adelaide,
North Terrace, Adelaide, 5005, Australia
1832394417@qq.com

Received 20 May 2025; Revised 2 June 2025; Accepted 9 June 2025

Abstract. Whether intelligent recommendation algorithms based on computational systems can influence user trust in internet platforms has become a key issue in today's digital society. However, relevant research remains limited, particularly in terms of systematic empirical investigation into the underlying mechanisms. Drawing on Fairness Heuristic Theory, this study introduces algorithmic fairness perception as a mediating variable and algorithmic literacy as a moderating variable to construct a conceptual model linking algorithmic transparency and user trust. A scenario-based experimental survey was conducted to systematically investigate the psychological mechanism through which algorithmic transparency affects user trust. The results demonstrate that overall, algorithmic transparency significantly enhances users' trust in both the platform and its recommendation system. Furthermore, algorithmic fairness perception mediates this relationship, while algorithmic literacy positively moderates the effect of algorithmic transparency on fairness perception. These findings not only provide empirical support for understanding the relationship between algorithmic transparency and user trust but also offer practical insights for platform governance. Specifically, designers of recommendation algorithms should establish more standardized and effective transparency mechanisms and promote users' algorithmic literacy to enhance their understanding of algorithmic logic, thereby strengthening trust at the cognitive level.

Keywords: algorithmic transparency, user trust, fairness heuristic theory, algorithmic literacy

1 Introduction

To enhance user stickiness and improve competitive advantage, major internet platforms have widely adopted intelligent algorithm-based personalized recommendation systems, with their application rate exceeding 70% [1]. However, due to the inherent non-transparency of such algorithms, users are often placed in a passive position, subject to the influence of algorithmic logic in a highly controlled digital environment. Professor Frank Pasquale has critically examined the pervasive application of algorithmic technologies across various domains and argued that humans are already living in an algorithm dominated world, yet they frequently face a reality in which the rules are unknown and the reasons are unclear, a phenomenon he terms the algorithmic black box [2]. The existence of such black-box algorithms has facilitated the abuse of algorithmic power, including the manipulation of information flows, the shaping of public discourse, and the emergence of phenomena such as traffic fraud and information hijacking, as well as more subtle interventions in content presentation [3]. As a result, algorithmic transparency has long been regarded as a fundamental solution to the black box problem. With the introduction of AI-specific legislation, the regulation of controversial opaque algorithms has gained clearer legal foundations [4]. In March 2024, the European Union passed its first dedicated legislation on artificial intelligence called the Artificial Intelligence Act. This act defines an "AI system" as a machine-based system designed to operate with varying levels of autonomy, capable of adapting after deployment, and able to infer how to generate outputs (e.g., predictions, content, suggestions, or decisions) that influence physical or virtual environments based on received input. This definition closely overlaps with the characteristics of intelligent recommendation algo-

* Corresponding Author

rithms widely used on internet platforms, thereby subjecting them to legally mandated transparency obligations. Simultaneously, China has also advanced its AI legislative efforts, releasing the AI Law (Scholarly Draft) and the AI Model Law 2.0 (Expert Draft) in 2024 [5]. Both documents identify transparency as a foundational principle of AI systems and explicitly include intelligent recommendation algorithms within the scope of legal regulation. Scholars generally agree that enhancing the transparency of intelligent recommendation algorithms is a fundamental approach to addressing the problem of algorithmic black boxes. In the academic domain, substantial progress has been made in exploring the theoretical underpinnings, functional mechanisms, and practical strategies for implementing algorithmic transparency. In parallel, at the policy and regulatory level, a series of legislative measures have been introduced to address the challenges posed by opaque algorithmic practices, aiming to institutionalize transparency as a core principle in algorithm governance. However, the disconnection between algorithmic transparency and user trust on internet platforms poses challenges for the implementation of transparency principles. This "metaphysical disconnect" makes it difficult to align formal algorithmic disclosure with users' perceptual understanding, ultimately undermining trust and engagement. At the governmental level [6], there has been a growing body of research exploring the relationship between algorithmic transparency in computational systems [7] and public trust [8], but most of this research focuses on government use of algorithms [9]. In contrast, studies examining how algorithmic transparency on internet platforms influences user trust remain limited. Guo [10] suggests that platforms should regularly publish algorithmic transparency reports to clearly inform users about the recommendation mechanisms, decision processes, and underlying value orientations, thereby fostering user trust and creating space for social oversight. Thus, further research is needed to understand and reveal how algorithmic transparency in intelligent recommendation systems affects user trust on digital platforms.

This study introduces two critical variables, algorithmic fairness perception and algorithmic literacy, to explore this issue. Innovatively drawing on Fairness Heuristic Theory [11], it constructs a theoretical framework in which algorithmic fairness perception mediates the relationship between algorithmic transparency and user trust, thereby revealing how users form trust in situations of information asymmetry. Transferring this theory from organizational management to the digital platform context suggests that users, when interacting with algorithmic recommendation systems, are also embedded in a "fundamental social dilemma": they desire to rely on platform algorithms for information efficiency while remaining wary of potential manipulation and unfairness. In such contexts, users tend to rely on external fairness cues provided by transparent algorithms to make trust judgments [12]. When users in a high transparency environment believe that they are being treated fairly by the recommendation algorithm, this algorithmic fairness perception serves as a social heuristic, a psychological shortcut based on fairness impressions, which mediates the effect of algorithmic transparency on user trust. Meanwhile, algorithmic literacy has attracted widespread academic attention [13], with scholars generally recognizing its role in helping the public better understand algorithmic systems [14] and, on that basis, serving as an effective complement to algorithmic transparency [15]. However, studies exploring whether algorithmic literacy affects algorithmic fairness perception remain relatively scarce and have not yet been adequately examined.

As algorithmic governance on internet platforms has become a key issue of contemporary concern, empirical research examining the relationship between algorithmic transparency in intelligent recommendation systems and user trust remains insufficient. To more effectively explore and validate this relationship, there is an urgent need for the adoption of diverse methodological approaches. Against this backdrop, the present study constructs an analytical framework based on Fairness Heuristic Theory and employs a scenario-based experimental method to conduct empirical analysis. It seeks to address the following research questions:

1) Does algorithmic transparency on internet platforms influence users' trust in the platform, and if so, through what processes and mechanisms?

2) To what extent does algorithmic fairness perception mediate the relationship between algorithmic transparency and user trust?

3) To what extent does algorithmic literacy moderate the relationship between algorithmic transparency and algorithmic fairness perception? Addressing these questions will contribute to a deeper understanding of the relationship between algorithmic transparency and user trust, as well as the underlying mechanisms through which this relationship operates.

2 Literature Review and Research Hypotheses

This section develops a theoretical model and proposes several hypotheses. It explores how algorithmic transparency affects user trust directly and indirectly through algorithmic fairness perception, with algorithmic literacy moderating the effect on fairness perception. The model seeks to explain how these factors interact to shape user trust.

2.1 Conceptual Analysis of Algorithmic Transparency

Originally, algorithms were a mathematical concept primarily used to solve mathematical problems in daily life, defined as finite, ordered, and computable sequences of steps or procedures [16]. With the development of computer technology, the definition of algorithms expanded to "a finite, ordered, stable, and efficient computer program for solving problems," [17] which refers to logical instructions executed by intelligent machines to analyze, process, and solve problems. As algorithms gradually embedded themselves into various aspects of life, they became consciously designed equations, formulas, or codes created by people to address specific problems. From this point on, algorithms were no longer confined to a narrow technical concept but evolved into reasoning rules with cultural and social significance, exerting broad dominance and influence over people and things in the real world. Decision-making systems centered on algorithmic technology strongly intervene in and shape human behavior, forming what is called "algorithmic power," continuously constructing and maintaining relationships between people and between people and things [18]. Algorithmic power often shapes human cognitive patterns and regulates behavioral capabilities in implicit, invisible, and hard-to-detect ways.

To unravel the mystery of algorithmic power, establishing a sound system of algorithmic transparency is especially necessary to achieve digital justice in the algorithmic era. Algorithmic transparency refers to the clear visibility to users and stakeholders of all aspects of the algorithm, including training data, input logic, parameter selection, computational process, and output results [19]. Its core objectives include helping people understand the operational mechanisms of algorithms, promoting algorithmic improvements, and providing grounds for questioning and feedback on predictions [20]. Functionally, algorithmic transparency not only holds value in improving and validating algorithms but also enhances social communication and rationality around algorithms, increases public trust, and effectively supervises the exercise of algorithmic power, thereby promoting individual freedom and safeguarding human dignity.

Therefore, this study focuses on algorithmic transparency in internet platform recommendation algorithms, with particular attention to its manifestations in the dimensions of training data transparency, algorithmic logic transparency, and result feedback transparency.

2.2 Research Hypotheses

In recent years, algorithmic transparency mechanisms aimed at enhancing user trust have become core requirements and guiding principles in the global development of computer algorithms. The European Commission's High-Level Expert Group has successively issued the "Ethics Guidelines for Trustworthy AI," the "White Paper on Artificial Intelligence: A European Approach to Excellence and Trust," and the "AI Act," repeatedly emphasizing the importance of building Trustworthy AI. In September 2021, China's New Generation Artificial Intelligence Governance Professional Committee released the "Ethical Norms for the New Generation Artificial Intelligence," establishing trustworthiness and controllability as fundamental ethical standards for AI activities. In March 2024, the United Nations General Assembly adopted the first global resolution on artificial intelligence, demonstrating a global commitment to developing safe, reliable, and trustworthy AI.

Meanwhile, multiple studies have pointed out that incorporating "trust" as a key evaluative factor in algorithm governance represents a viewpoint centered on both users and the public. Algorithmic trust refers to the willingness of the public to accept and use algorithms despite certain risks and uncertainties, because they believe the algorithms will function as expected. This trust is vital, as "any future development, implementation, and application of artificial intelligence is closely tied to public trust and support." [21] Algorithmic trust does not demand zero risk but aims to keep risks within levels acceptable to the public, a tolerance that fosters technological advancement and innovation. Based on this, the following research hypothesis is proposed:

H1: Algorithmic transparency is positively associated with user trust.

This study posits that the relationship between algorithmic transparency and user trust is complex and influenced by users' perceptions of algorithmic fairness. Drawing on prior discussion, fairness heuristic theory confirms that perceptions of fairness in algorithmic recommendations not only serve as psychological triggers for trust formation but also act as key channels through which algorithmic transparency exerts its effect. Therefore, transparency in intelligent recommendation algorithms can, to some extent, improve users' perceptions of fairness in algorithmic recommendations, thereby enhancing their trust in the algorithms. Accordingly, the following hypothesis is proposed:

H2: Algorithmic fairness perception mediates the relationship between algorithmic transparency and user trust.

However, users' varying levels of algorithmic literacy lead to differences in how they perceive fairness in algorithmic recommendations. Although a unified definition of algorithmic literacy has yet to be established domestically or internationally, it is widely recognized as a complex concept system encompassing algorithm awareness, skills, strategies, and resistance [22]. Some scholars have concretized algorithmic literacy as news literacy in the field of journalism, noting that users' cognition of algorithms may influence their acceptance, understanding, and behavioral responses to fake news. For example, users' perceptions of algorithmic transparency and information filtering, as well as their feedback on platform information, may strengthen their ability to identify and report fake news [23]. Therefore, enhancing algorithmic literacy helps promote user attention to and discussion of fairness, transparency, interpretability, and accountability of algorithms in platform recommendation environments. Based on this, this study proposes the following hypothesis:

H3: Algorithmic literacy positively moderates the effect of algorithmic transparency on algorithmic fairness perception.

Accordingly, a theoretical model is constructed to illustrate the relationships among the core variables (see Fig. 1). In this model, algorithmic transparency is the independent variable, user trust is the dependent variable, algorithmic fairness perception serves as the mediating variable, and algorithmic literacy functions as the moderating variable. Specifically, algorithmic transparency is expected to have a direct impact on user trust, while algorithmic fairness perception mediates this relationship. Meanwhile, algorithmic literacy moderates the effect of algorithmic transparency on algorithmic fairness perception. The model aims to uncover the interactions among these variables and explore how they jointly contribute to the formation of user trust.

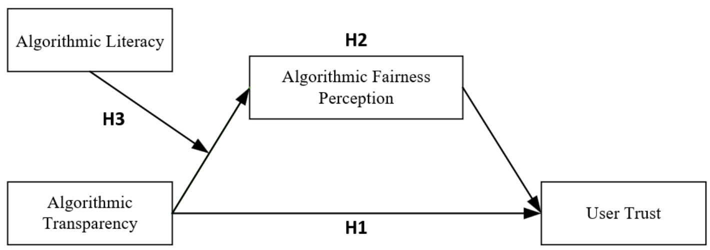

Fig. 1. Theoretical model

3 Variable Measurement and Data Collection

Data for this study were collected through the online survey platform Credoma. The questionnaire was divided into four parts. The first part consisted of instructions, providing participants with background information and guidance on how to complete the survey. The second part measured algorithmic literacy and platform activity level, and these two sections were identical for all participants without any distinction. The third part constituted the core of the questionnaire, including the experimental manipulation and measurement items. Two experimental scenarios were designed, one with high transparency and one with low transparency, to simulate users' interaction with internet platform algorithms under varying degrees of algorithmic transparency. Participants were randomly assigned to read one of the scenarios and then responded to items measuring algorithmic transparency perception, user trust, and algorithmic fairness perception. The fourth part collected participants' evaluations of algorithmic recommendation systems and basic demographic information, including gender, age, education level, income, and years of internet usage.

All variables were measured using established scales with a 7-point Likert response format. The scales for algorithmic literacy and algorithmic transparency perception were adapted from Liu et al., each covering three dimensions. For algorithmic literacy, these dimensions are data sources, technical familiarity, and application domains. For algorithmic transparency perception, the dimensions include accessibility, explainability, and interpretability. The user trust scale was adapted from Molina and Sundar [24] and comprises two higher-order dimensions, attitudinal trust and behavioral trust, which are further divided into four subdimensions: process trust, outcome trust, application trust, and recommendation trust. The scale for algorithmic fairness perception was adapted from fairness characteristics proposed by Thibaut [25] and Leventhal [26], including the dimensions of consistency, diversity, correctability, morality, and impact.

A total of 275 responses were collected via random sampling nationwide. One response was excluded for failing the attention check, resulting in 274 valid responses. The sample was divided into two groups: low transparency (transparency = 0, n = 136) and high transparency (transparency = 1, n = 138), with comparable group sizes ensuring a good level of comparability.

In terms of gender distribution, the average for the low transparency group was 0.30, while the high transparency group averaged 0.33, showing minimal gender differences. Regarding age, the average age was 30.81 years in the low transparency group and 29.96 years in the high transparency group, indicating little variation and an overall concentration around age 30. For education level, the low transparency group had an SD of 0.46 and variance of 0.209, while the high transparency group had an SD of 0.43 and variance of 0.186. These results suggest both groups had similarly high educational backgrounds, with most participants holding at least a bachelor's degree. Regarding income, the average income level for the low transparency group was 2.50, while for the high transparency group, it was 2.55. Based on the income classification standard, both groups' income levels fall within the "10,001-15,000 RMB" range, indicating that the majority of respondents have similar income levels, with minimal income differences between the two groups. Regarding internet usage years, the average level for the low transparency group was 2.85, and for the high transparency group, it was 2.86. Both groups fell within the "11-15 years" range, indicating minimal differences between the two groups in terms of internet usage experience. An overview of the demographic characteristics of participants is provided in Table 1. Based on the data presented in Table 1, the variance comparison of different variables between the low and high transparency groups is shown in Fig. 2. The differences in variance are minimal, indicating that the variances of each variable are approximately consistent across the different transparency groups.

Table 1. Demographic characteristics of participants

Transparency group		N	Min	Max	Mean	Standard deviation	Variance
Low	Gender	136	0	1	0.30	0.460	0.211
	Age	136	20	53	30.81	6.565	43.104
	Education	136	2	5	3.14	0.457	0.209
	Income	136	1	5	2.50	1.044	1.090
	Internet usage years	136	1	4	2.85	0.818	0.670
High	Gender	138	0	1	0.33	0.470	0.221
	Age	138	19	58	29.96	6.859	47.042
	Education	138	1	5	3.06	0.432	0.186
	Income	138	1	5	2.35	1.092	1.192
	Internet usage years	138	1	4	2.86	0.720	0.519

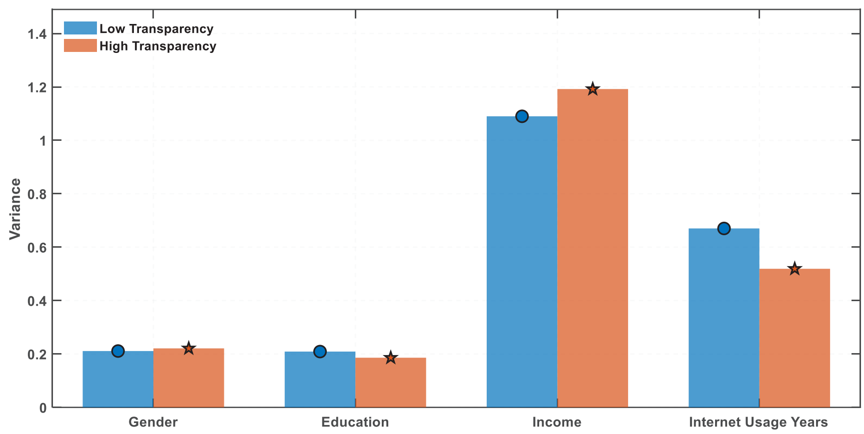

Fig. 2. Variance comparison between low and high transparency groups (Partial)

4 Model Specification

To examine the direct effect of algorithm transparency on user trust (Equation 1) and its indirect effect through Fairness Perception (Equations 2, 3, and 4), as well as to further investigate the moderating effect of Algorithm Literacy (Equation 5), the model is specified as follows.

4.1 Ordinary Least Squares (OLS) Model

To accurately examine the impact of algorithmic transparency on user trust, this study draws on the research of Wang Wenbin [27] and constructs the following baseline Ordinary Least Squares (OLS) regression model:

$$UT = \alpha_0 + \alpha_1 AT + \gamma Controls + \varepsilon_1 \tag{1}$$

In the equation, the dependent variable is User Trust (UT), while the key independent variable is Algorithmic Transparency (AT). The term Controls represents a set of control variables, which include user activity level, user satisfaction, gender, age, education level, income, and years of internet usage. The coefficient α_0 denotes the constant term in the model. The parameters α_1 and γ are coefficients to be estimated, where α_1 is the primary coefficient of interest, representing the magnitude and direction of the effect of algorithmic transparency on user trust. Finally, ε_1 denotes the stochastic error term capturing unobserved factors and random disturbances in the model.

4.2 Mediation Effect Model

To investigate through which pathways algorithmic transparency influences user trust, this study follows the causal stepwise regression approach proposed by Wen et al. [28]. The relationship between algorithmic transparency and user trust, mediated by perceived fairness of algorithmic recommendations, is examined. The mediation effect models are defined as follows:

$$AFP = \alpha_0 + \alpha_1 AT + \gamma Controls + \varepsilon_2 \tag{2}$$

$$UT = \beta_0 + \beta_1 AT + \beta_2 AFP + \gamma Controls + \varepsilon_3 \tag{3}$$

In the equations, AFP represents Algorithmic Fairness Perception, AT denotes Algorithmic Transparency, and UT stands for User Trust. The terms α_0 and β_0 indicate the constant terms in the respective models. Controls refers to the set of control variables. The coefficients α_1, α_1, β_1, β_2 are parameters to be estimated corresponding to their respective dependent variables. The terms ε_2 and ε_3 represent the stochastic error terms accounting for unobserved influences and random disturbances in the models.

This study employs a stepwise regression method to test the mediation effect. Based on examining the main effect of algorithm transparency on user trust, the mediation variable, perceived fairness of algorithmic recommendations, is introduced to test the mediation effect. The stepwise regression procedure consists of two main steps: first, regressing the independent variable on the mediator; subsequently, including the mediator in the regression of the independent variable on the dependent variable to examine the regression coefficients. The specific paths are illustrated in Fig. 3.

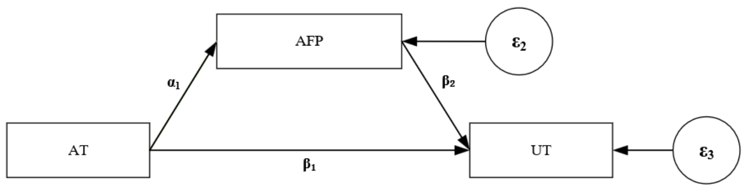

Fig. 3. Mediation effect testing model path

4.3 Moderation Effect Model

To investigate the moderating role of algorithm literacy between algorithm transparency and users' perceived fairness of algorithmic recommendations, this study adopts the research framework of Zhang [29] and establishes the following moderation effect testing model:

$$AFP = \delta_0 + \delta_1 AT + \delta_2 AL + \gamma Controls + \varepsilon_4 \quad \textbf{(4)}$$

$$AFP = \eta_0 + \eta_1 AT + \eta_2 AL + \eta_3 AT * AL + \gamma Controls + \varepsilon_5 \quad \textbf{(5)}$$

In the equation, AT represents algorithm transparency, Controls are the control variables, AL is the moderator (algorithm literacy), AT*AL is the interaction term between algorithm transparency and algorithm literacy, and AFP stands for perceived fairness of algorithmic recommendations. δ_0 and η_0 are the constant terms, δ_1, δ_2, γ, η_1, η_2, and η_3 are the coefficients to be estimated for the respective dependent variables, and ε_4 and ε_5 are the model's error terms.

5 Empirical Results

This section presents the empirical analysis through several key components: Reliability, Validity, and Manipulation Checks, Descriptive Statistical Analysis, Main Effect Analysis of Variance, Regression Analysis, and Mediation Effect Test.

5.1 Reliability, Validity, and Manipulation Checks

In this study, SPSS software was used to conduct validity and reliability checks on the questionnaire data. For validity analysis, factor analysis was first performed on the scales in the questionnaire using the KMO (Kaiser-Meyer-Olkin) test and Bartlett's Test of Sphericity. The calculation formula for the KMO statistic is presented in Equation 6.

$$KMO = \frac{\sum\sum_{i \neq j} r_{ij}^2}{\sum\sum_{i \neq j} r_{ij}^2 + \sum\sum_{i \neq j} \alpha_{ij}^2} \quad \textbf{(6)}$$

In the formula, KMO values range from 0 to 1, with 0.6 commonly used as the threshold. When KMO ≥ 0.6, it indicates that the data is suitable for principal component analysis r_{ij}^2 represents the correlation between the i-th and j-th variables, and α_{ij}^2 represents the partial correlation between the i-th and j-th variables. The test results show that the KMO values of the scales in the questionnaire are all greater than 0.6, and Bartlett's Test of Sphericity is significant (P = 0.000), indicating that there is a correlation between the selected indicators, making them suitable for factor extraction [30].

For reliability analysis, internal consistency reliability of the scale was tested using 274 data points, and Cronbach's α coefficient was calculated. The calculation formula is shown in Equation 7.

$$\alpha = k(1 - \sum s_i^2) / ((k-1)s^2) \quad \textbf{(7)}$$

In this formula, k represents the number of items, s_i^2 is the variance of the i-th item, and s^2 is the variance of the total score [31]. Based on the reliability analysis, the Cronbach's α coefficients were found to be greater than 0.8, indicating that the scales in the questionnaire have high internal consistency and good reliability.

Additionally, an independent samples t-test was conducted to examine the effectiveness of the experimental manipulation of algorithmic transparency in the questionnaire design. The test results (see Table 2) show that the three dimensions of algorithmic transparency (accessibility, explainability, and interpretability) exhibited significant differences, with the mean of the low transparency group being significantly lower than that of the high transparency group, and $p < 0.001$. This indicates that the experimental manipulation of algorithmic transparency in the questionnaire was effective.

Table 2. Independent samples t-test

Algorithmic transparency	Group 1 = High transparency	Mean	Standard deviation	T-value
Accessibility	1	4.94	1.382	11.157***
	0	3.04	1.439	
Explainability	1	5.70	0.964	12.964***
	0	3.52	1.720	
Interpretability	1	5.52	0.991	11.988***
	0	3.63	1.567	

Note: ****p<0.01, **p<0.05, *p<0.1

5.2 Descriptive Statistical Analysis

Table 3 reports descriptive statistical analysis results of the main variables in this study, and the correlation matrix of the main variables is provided in the appendix. Based on the contents reported in the above tables, it can be seen that there are significant correlations between algorithmic transparency, algorithmic fairness perception, algorithmic literacy, and user trust, as well as between user trust and user activity level and user satisfaction. Therefore, these variables will be included as control variables in the subsequent analysis. Based on the factor analysis results of the core variables in this study, the following conclusions can be drawn:

Factor analysis of accessibility, explainability, and interpretability extracted one main factor as algorithmic transparency. Factor analysis of process trust, outcome trust, application trust, and recommendation trust extracted one main factor as user trust. Factor analysis of the five indicators of algorithmic fairness perception, including consistency, diversity, correctability, morality, and impact, identified a single primary factor representing algorithmic fairness perception. Factor analysis of technical familiarity, data sources, and application domains extracted one main factor as algorithmic literacy.

Table 3. Descriptive statistical analysis of key variables

Variable		N	Min	Max	Mean	Standard deviation	Variance
Algorithmic literacy	Data sources	274	1	7	5.07	1.051	1.105
	Technical familiarity	274	1	7	4.56	1.323	1.751
	Algorithm application domains	274	1	7	4.86	1.213	1.473
Algorithmic transparency	Accessibility	274	1	7	4.00	1.699	2.887
	Explainability	274	1	7	4.61	1.767	3.122
	Interpretability	274	1	7	4.58	1.615	2.610
Algorithmic fairness perception	Consistency	274	1	7	5.45	1.241	1.541
	Diversity	274	1	7	4.48	1.399	1.958
	Correctability	274	1	7	5.21	1.419	2.014
	Morality	274	1	7	5.07	1.453	2.112
	Impact	274	1	7	5.52	1.268	1.608
User trust	Process trust	274	1	7	4.8	1.506	2.268
	Outcome trust	274	1	7	4.78	1.528	2.334
	Application trust	274	1	7	5.63	1.606	2.578
	Recommendation trust	274	1	7	5.04	1.472	2.167
User activity level		274	-3.607	1.365	0	1	1
User satisfaction		274	2	7	5.48	0.856	0.732
Gender		274	0	1	0.31	0.464	0.216
Age		274	19	58	30.38	6.716	45.102
Education level		274	1	5	3.10	0.445	0.198
Income		274	1	5	2.42	1.069	1.143
Internet usage experience		274	1	4	2.85	0.769	0.592

5.3 Main Effect Analysis of Variance

This study uses a multifactor analysis of variance to test the impact of algorithmic transparency on algorithmic fairness perception and the moderating role of algorithmic literacy. Specifically, as shown in Table 4, algorithmic transparency was used as the independent variable and algorithmic fairness perception as the dependent variable. Both algorithmic transparency and algorithmic literacy show significant positive correlations with algorithmic fairness perception. The interaction effect between algorithmic transparency and algorithmic literacy on algorithmic fairness perception is also significant ($P < 0.01$). This indicates that algorithmic literacy plays a positive moderating role in the relationship between algorithmic transparency and algorithmic fairness perception. Therefore, Hypotheses 1 and 3 are supported. Fig. 4 illustrates the main effect strengths of three independent variables: algorithm transparency, algorithm literacy, and their interaction on perceived fairness of algorithmic recommendations, using the negative logarithm of the p value (negative log base 10 of p) as the effect measure. The figure shows that both algorithm transparency and its interaction with algorithm literacy have significant and positive impacts on perceived fairness, while the standalone effect of algorithm literacy, though weaker, remains statistically significant. This indicates that algorithm literacy plays a positive moderating role between algorithm transparency and perceived fairness of algorithmic recommendations.

Table 4. Main effect of algorithmic transparency and moderating effect of algorithmic literacy

Independent variable	Dependent variable	Mean square	F	P	Ajusted R^2
Algorithmic transparency	Algorithmic fairness perception	2.302	4.871	0.000	0.527
Algorithmic literacy		1.318	1.438	0.034	0.084
Algorithmic transparency*algorithmic literacy		1.120	2.576	0.000	0.565

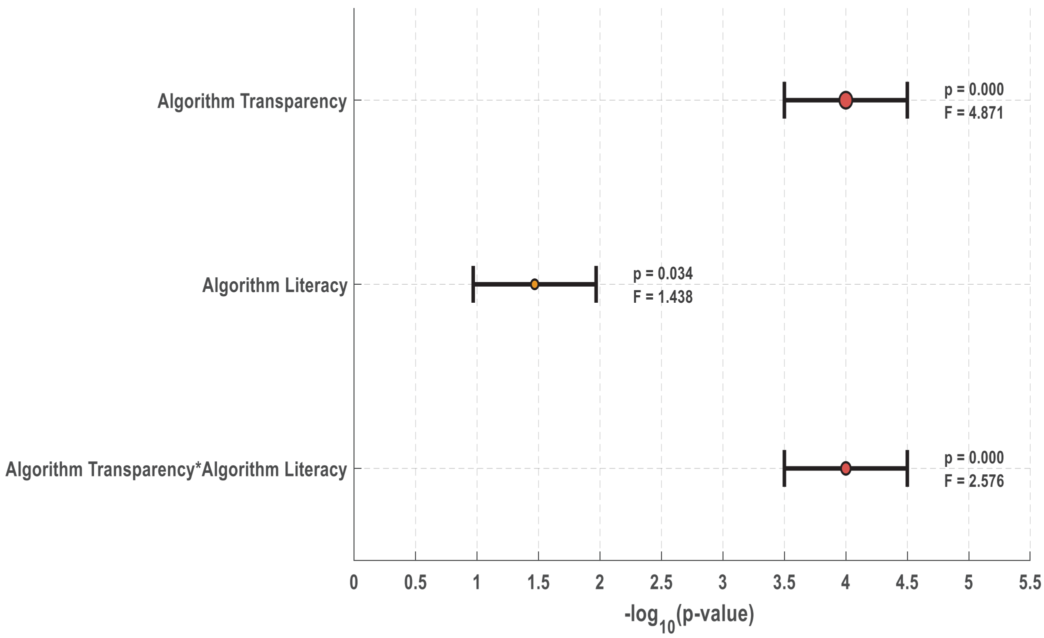

Fig. 4. Enhanced forest plot of algorithm factors

5.4 Regression Analysis

Table 5 presents the results of the OLS regression analysis and mediation effects. First, the three dimensions of algorithmic transparency (accessibility, explainability, and interpretability) were used as independent variables, and the four dimensions of user trust (process trust, outcome trust, application trust, and recommendation trust) were used as dependent variables in the regression analysis. Models 1 and 3 report the effects of algorithmic transparency on user trust without including mediating variables. The results show that accessibility, explainability, and interpretability have a significant positive relationship with process trust and outcome trust. User activity level is negatively related to attitude trust (process trust, outcome trust), and user satisfaction is positively correlated with outcome trust. Model 5 shows a significant positive relationship between explainability, interpretability, and application trust, while accessibility of the intelligent recommendation algorithm does not affect users' application trust toward the platform. User activity level still has a significant negative relationship with application trust. Notably, males are more likely to develop application trust. Model 7 reveals a significant positive relationship between explainability, interpretability, and recommendation trust, while accessibility is negatively correlated with recommendation trust. Thus, the higher the algorithm's explainability and interpretability, the higher the level of trust users have in the platform and the intelligent recommendation algorithm. Additionally, the higher the user activity level, the lower the levels of process trust, outcome trust, and application trust in the platform and the intelligent recommendation algorithm. The negative correlation be-tween user activity level and trust can be understood as active users having more cognitive awareness, expectations, and critical feedback about the platform's recommendation algorithm. This phenomenon may stem from factors such as information overload, personalized recommendation mismatches, concerns about transparency and privacy, and others [32].

Furthermore, the introduction of the mediating variable algorithmic fairness perception into the model shows that accessibility, explainability, interpretability, and algorithmic fairness perception are all positively correlated with attitude trust (Models 2 and 4), while user activity level is negatively correlated with attitude trust. Explainability, interpretability, and algorithmic fairness perception all positively affect behavioral trust, while accessibility negatively influences behavioral trust. Additionally, user satisfaction and user activity level negatively impact application trust, which is consistent with the baseline regression model. Again, males are more likely to develop application trust.

Next, Model 9 reports the effect of algorithmic transparency on algorithmic fairness perception. Both explainability and interpretability pass significance tests, indicating that algorithmic fairness perception plays a significant mediating role in the relationship between algorithmic transparency and user trust.

Table 5. OLS regression analysis (Without merged variables)

Variable		User trust								Algorithmic fairness perception
		Attitude trust				Behavioral trust				
		Process trust		Outcome trust		Application trust		Recommendation trust		
		Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model7	Model8	Model 9
Algorithmic transparency	Accessibility	0.173**	0.153**	0.149**	0.124**	-0.081	-0.111*	-0.167**	-0.194***	0.056
	Explainability	0.382***	0.235***	0.321***	0.137*	0.534***	0.315***	0.634***	0.441***	0.409***
	Interpretability	0.249***	0.157*	0.296***	0.18**	0.263***	0.126	0.218**	0.097	0.255***
Algorithmic fairness perception			0.36***		0.451***		0.535***		0.473***	
User satisfaction		0.053	0.016	0.086*	0.04	-0.046	-0.1**	0.030	-0.018	0.101**
Gender		0.049	0.025	0.05	0.02	0.131***	0.095**	0.054	0.023	0.066
Age		0.043	0.05	0.067	0.076	-0.024	-0.013	-0.005	0.004	-0.020
Education level		0.029	0.03	0.044	0.046	0.031	0.034	0.046	0.049	-0.005
Income		0.01	0.033	-0.01	0.019	-0.003	0.032	0.009	0.040	-0.065
Internet usage experience		0.038	0.016	0.012	-0.015	0.038	0.006	-0.034	-0.063	0.060
User activity level		-0.123**	-0.116**	-0.124**	-0.115**	-0.093*	-0.083*	0.005	0.015	-0.020
R-squared		0.562	0.625	0.523	0.623	0.504	0.644	0.514	0.624	0.509
Adjusted R-squared		0.545	0.609	0.505	0.607	0.485	0.629	0.496	0.608	0.491
F-value		33.681	39.733	28.837	39.342	26.740	43.151	27.843	39.520	27.297

To further analyze the overall impact of algorithmic transparency on user trust and the mediating role of algorithmic fairness perception, linear regression was conducted with algorithmic transparency as the independent variable and user trust as the dependent variable. The results show (see Table 6) that, overall, there is a significant positive correlation between algorithmic transparency and user trust, supporting Hypothesis 1. User activity level is negatively correlated with user trust, and males are more likely to develop user trust. After introducing algorithmic fairness perception as a mediating variable, the effect of algorithmic transparency on user trust remains significant, with the coefficient decreasing from 0.770 to 0.415. This suggests that algorithmic fairness perception plays a significant mediating role in this relationship.

Table 6. OLS regression analysis (Merged variables)

Variable	User trust	User trust	Algorithmic fairness perception
Algorithmic transparency	0.770***	0.415***	0.681***
Algorithmic fairness perception		0.520***	
User satisfaction	0.036	-0.017	0.103**
Gender	0.075*	0.042	0.064
Age	0.030	0.037	-0.015
Education level	0.048	0.047	0.000
Income	0.009	0.041	-0.060
Internet usage experience	0.001	-0.025	0.050
User activity level	-0.143***	-0.114***	-0.056
R-squared	0.580	0.717	0.495
Adjusted R-squared	0.568	0.707	0.480
F-value	45.789	74.297	32.522

Table 7 reports the results of the OLS regression analysis with user trust as the dependent variable under different levels of algorithmic transparency. The results show that, regardless of whether the transparency is low or high, algorithmic transparency has a significant positive effect on user trust, indicating that users' perception of algorithmic transparency is an important river of their trust. Moreover, the study finds that, within the higher transparency group, the impact of algorithmic transparency on user trust is even more significant, with a higher standardized regression coefficient (Beta). This further confirms the central role of algorithmic transparency in the development of user trust. This suggests that the higher the platform's transparency, the higher the level of user trust in the algorithm.

Table 7. OLS regression analysis results under different levels of algorithmic transparency

Variable	Low algorithmic transparency group (User trust)	High algorithmic transparency group (User trust)
Algorithmic Transparency	0.318***	0.494***
Algorithmic Fairness Perception	0.612***	0.437***
User satisfaction	0.021	-0.036
Gender	0.02	0.057
Age	0.073	0.013
Education level	0.048	0.046
Income	-0.009	0.079
Internet usage Experience	-0.044	-0.009
User activity level	-0.116**	-0.123**
R-squared	0.714	0.737
Adjusted R-squared	0.694	0.719
F-value	35.309	39.626

5.5 Mediation Effect Test

To further investigate whether algorithmic fairness perception mediates the relationship between algorithmic transparency and user trust, this study used the Bootstrap method for testing. The mediation effect test results show (see Table 8) that when algorithmic fairness perception is used as the mediating variable, the Bootstrap 95% confidence interval is [0.2653, 0.4625], which does not include 0, indicating that the mediation effect of algorithmic fairness perception is significant. Further, the direct effect's 95% confidence interval is [0.3245, 0.5087], which also does not include 0, showing that algorithmic fairness perception plays a mediating role in the relationship between algorithmic transparency and user trust. Fig. 5 illustrates the 95% confidence intervals for the total effect, direct effect, and mediation effect.

Table 8. Mediation effect test of algorithmic fairness perception

Variable	Effect value		Bootstrapped standard error	95% confidence interv	
				Lower bound	Upper bound
Algorithmic fairness perception	Total effect	0.7797	0.0415	0.6980	0.8614
	Direct Effect	0.4166	0.0468	0.3245	0.5087
	Mediation Effect	0.3631	0.0509	0.2653	0.4625

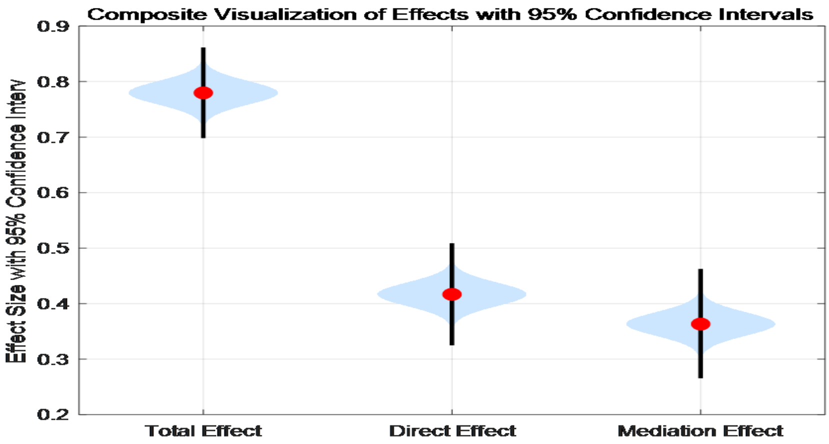

Fig. 5. Composite visualization of effects with 95% confidence intervals

6 Conclusion and Discussion

This study employs a scenario-based experimental approach to systematically analyze the mechanism by which algorithmic transparency affects user trust, leading to three major conclusions. These conclusions provide new empirical evidence for understanding the relationship between algorithmic transparency in intelligent recommendation algorithms and user trust. Overall, the study finds that algorithmic transparency significantly enhances users' trust in the platform and its recommendation algorithms. Specifically, the three core dimensions of algorithmic transparency, which are accessibility, explainability, and interpretability, are significantly positively correlated with user trust. This indicates that when users can more easily access, understand, and explain the operational logic and decision-making rules of the algorithm, their trust in the platform and algorithm system correspondingly increases.

Furthermore, the subgroup analysis under different levels of transparency shows that, whether in low or high transparency contexts, algorithmic transparency has a significant positive impact on user trust. This suggests that users' perception of transparency is a crucial driving factor in building their trust. Additionally, within the higher transparency group, this effect is more significant, indicating that users tend to trust the platform's algorithm more when transparency is higher, further highlighting the core role of algorithmic transparency in trust-building.

Moreover, the study uncovers some thought-provoking results. Among user characteristics, user activity level is negatively correlated with algorithmic trust, meaning that users with higher inter-action frequency on the platform tend to exhibit lower trust in the recommendation algorithm. This may be because active users encounter recommendation content more frequently, are more likely to experience information overload and personalized recommendation mismatches, and become more sensitive to data handling and privacy protection issues, leading to higher cognitive demands and critical feedback on the recommendation system. Additionally, gender differences in application trust were found, with male users being more likely than female users to develop higher application trust, showing a stronger dependency and willingness to use the platform. These findings not only expand our understanding of the formation of user trust but also provide empirical evidence for platform design in user segmentation governance and the differentiated design of transparency strategies.

Regarding Hypothesis 2, this study finds that algorithmic fairness perception plays a significant mediating role in the relationship between algorithmic transparency and user trust. This finding provides empirical support for the path of Fairness Heuristic Theory. The theory aims to explain “how fairness perceptions arise” and “how fairness perceptions influence subsequent behaviors,” and this study verifies its applicability in the context of intelligent recommendation algorithms. Users, who are in the “basic social dilemma” of algorithm-driven environments, often struggle to fully understand the platform’s operational mechanisms and thus tend to rely on explicit fairness cues provided by algorithmic transparency to assess the fairness of their environment. When users perceive that they are treated fairly by the algorithmic system in a high transparency context, they are more likely to form trust judgments based on this fairness perception, making “shortcut” psychological decisions, and thus increasing their overall trust in the platform. Based on this, recommendation algorithm designers should pay attention to the procedural fairness of the recommendation process and the rationality of decision rules. This means that not only should transparency be improved, but the algorithmic operational logic should also embed respect and protection for user interests. Only by doing so can the platform strengthen user trust and, in turn, enhance user engagement and long-term willingness to use the platform.

Hypothesis 3 is also empirically supported in this study, which found that algorithmic literacy plays a positive moderating role in the relationship between algorithmic transparency and algorithmic fairness perception. As users’ algorithmic literacy improves, their ability to understand and judge algorithmic systems strengthens, allowing them to more effectively identify and interpret the fairness information conveyed by algorithmic transparency.

The findings and conclusions of this study not only provide valuable insights for stakeholders in internet platforms to deeply understand the relationship between algorithmic transparency and user trust, but also offer important references for global digital governance. Currently, internet platforms have gradually evolved into essential infrastructures in the digital society. Their intelligent recommendation algorithms not only profoundly impact user behavior and information acquisition methods but also play a fundamental role in key areas such as public opinion guidance, public services, and social governance. In this context, building a governance mechanism centered on algorithmic transparency is not only a micro-path to enhancing user trust but also a crucial pillar for achieving effective global digital governance, promoting fairness and justice, and fostering social stability. Therefore, recommendation algorithm designers should work towards establishing more standardized and effective algorithmic governance mechanisms, always adhering to the principle of transparency. At the same time, attention should be paid to improving the public’s algorithmic literacy, helping users better understand the operational logic of algorithms, thus enhancing their trust in the platform and its recommendation systems at the cognitive level.

References

[1] M.-Y. Liao, J.-M. Fang, J.-J. Yang, M.-A. Hossin, The Effect of Recommendation Content Similarity on Sustained App Use from the Perspective of User Coping, Nankai Business Review 26(3)(2023) 178-190.

[2] W. Dai, X.-X. Wang, Enhancing Algorithmic Transparency: An Analysis on Limitations of Transparency Mechanism and Its Reconstruction, ECUPL Journal 28(1)(2025) 33-44.

[3] W.-D. Li, L.-L. Yang, Research on the Stakeholders of Transparent Mechanisms of Information Recommendation Algorithms, New Media and Society (4)(2024) 13-28+413.

[4] T. Liu, Y.-P. Zheng, M.-B. Cao, Algorithmic Administration: Literature Review and Research Prospects, Journal of Public Administration 17(1)(2024) 84-104+198.

[5] H. Zhang, Y. Yue, Regulatory Design for the Dynamic Balance between AI Innovation and Supervision : Based on the Perspective of Regulatory Sandbox, Zhejiang Academic Journal (1)(2025) 76-87.

[6] S. Grimmelikhuijsen, Explaining Why the Computer Says No: Algorithmic Transparency Affects the Perceived Trustworthiness of Automated Decision-Making, Public Administration Review 83(2)(2023) 241-262.

[7] D.-S. Schiff, K.-J. Schiff, P. Pierson, Assessing public value failure in government adoption of artificial intelligence, Public Administration 100(3)(2021) 653-673.

[8] D. Shin, Y.-J. Park, Role of fairness, accountability, and transparency in algorithmic affordance, Computers in Human Behavior 98(2019) 277-284.

[9] R. Gorwa, R. Binns, C. Katzenbach, Algorithmic content moderation: Technical and political challenges in the automation of platform governance, Big Data & Society 7(1)(2020) 1-15.

[10] X.-A. Guo, The Disorderly Development of Mobile Short Videos and Strategies for Collaborative Governance, Frontiers (6)(2025) 77-86.

[11] S.-L. Jordan, J.-C. Palmer, S.-R. Daniels, W.-A. Hochwarter, P.-L. Perrewé, G.-R. Ferris, Subjectivity in fairness per-

ceptions: How heuristics and self-efficacy shape the fairness expectations and perceptions of organisational newcomers, Applied Psychology 71(1)(2022) 103-128.
[12] Y. Zhou, X. Huang, W.-J. Xie, Does Perceived Overqualification Inspire Employee Voice? —Based on the Lens of Fairness Heuristic, Management Review 32(12)(2020) 192-203.
[13] K. Kim, S.-I. Moon, When Algorithmic Transparency Failed: Controversies Over Algorithm-Driven Content Curation in the South Korean Digital Environment, American Behavioral Scientist 65(6)(2021) 847-862.
[14] D. Shin, A. Rasul, A. Fotiadis, Why am I seeing this? Deconstructing algorithm literacy through the lens of users, Internet Research 32(4)(2022) 1214-1234.
[15] L. Dogruel, P. Masur, S. Joeckel, Development and Validation of an Algorithm Literacy Scale for Internet Users, Communication Methods and Measures 16(2)(2022) 115-133.
[16] M.-H. Zhou, A New Discussion on the Method of Proof Evaluation—An Algorithm-Based Perspective, Science of Law (Journal of Northwest University of Political Science and Law) 38(1)(2020) 46-53.
[17] K. Long, L.-H. Wen, The Rise of Algorithmic Warfare and the Construction of Global Governance Mechanism for Military AI, Foreign Affairs Review 42(2)(2025) 61-91+6.
[18] X.-X. He, The Justification and Limitations of the Algorithm Interpretation Rights, Journal of Dalian University of Technology (Social Sciences) 44(3)(2023) 65-73.
[19] F. Yang, F.-S. Tang, K.-L. Zheng, Empirical Analysis of Transparency of Personalized Recommendation Algorithm Used in Digital Reading APPs in China, Journal of the National Library of China 33(3)(2024) 37-48.
[20] H.-P. Olsen, T.-T. Hildebrandt, C. Wiesener, M.-S. Larsen, A.-W.-A. Flügge, The Right to Transparency in Public Governance: Freedom of Information and the Use of Artificial Intelligence by Public Agencies, Digital Government: Research and Practice 5(1)(2024) 8.11-8.15.
[21] S. Afroogh, A probabilistic theory of trust concerning artificial intelligence: can intelligent robots trust humans?, AI and Ethics 3(2023) 469-484.
[22] Y. Wang, P. Wang, J.-H. Hong, “Intermediary ” or “Axes”: The Influence of Algorithms Visibility Scenes on RedNote Amateur Bloggers’ Content Production and Algorithmic Imagination, Journalism Research (2)(2025) 62-78+119-120.
[23] S. Wei, H.-Q. Ma, Research on the Realistic Logic and Implementation Path of Digital Intelligence Enabling Library Transformation, Library Work and Study (11)(2022) 27-36.
[24] M.-D. Molina, S.-S. Sundar, Does distrust in humans predict greater trust in AI? Role of individual differences in user responses to content moderation, New Media & Society 26(6)(2024) 3638-3656.
[25] R.-H. Zhu, Y.-X. Wang. F.-F. Shi, S.-H. He, Does Talking about Money Really Hurt Feelings? The Impact of Entrepreneurial Team Contractual Governance on Team Closeness, Foreign Economics & Management 45(2)(2023) 118-133.
[26] Z.-W. Huang, G. Fedeli, M.-M. Cheng, Trouble in paradise? Collaboration behavior and ethics of micro-influencers in the hospitality industry, Journal of Hospitality and Tourism Management 59(2024) 25-35.
[27] W.-B. Wang, Research on the Impact of Residents’ Participation in Pension Insurance on Fertility Intention, Price: Theory & Practice (8)(2023) 164-167+210.
[28] Z.-L. Wen, L. Zhang, J.-T. Hou, H.-Y. Liu, Testing and Application of the Mediating Effects, Acta Psychologica Sinica 36(5)(2004) 614-620.
[29] Y. Zhang, Study on the impact of blockchain traceability labeling of agricultural products on consumers’ willingness to pay premium—Taking the apple as the example, [dissertation] Yangling: Northwest A&F University, 2024.
[30] D.-H. Sun, X.-H. Qian, The Influencing Factors of Medical Assistance Satisfaction of the Low-income Population: A Case Study of Guangzhou, Medicine & Philosophy 42(3)(2021) 47-50.
[31] P. Wang, J.-C. Ma, F.-X Zheng, F. Geng, Y.-X. Zheng, M.-Z. Zhang, Q.-L. Xing, The Development of the Parent Rating Scale for Young Children’s Drawing Potential Based on Multidimensional Item Response Theory, Journal of Jiangxi Normal University (Natural Science Edition) 48(5)(2024) 441-448.
[32] B.-Y. Huangfu, “Algorithms Offend Me”: Emotional User-algorithm Practices and User Agency, Journalism Research (2)(2023) 16-27+117-118.