

Research On the Market and Willingness to Pay for Micro Short Drama Based on Machine Learning

Qi Wu¹, TianQi Cao^{1,*}, Zhen Wang¹, QiChun Tang¹, Lulu Yuan²

¹ School of International, Jiangxi University of Finance and Economics, Nanchang, China, 330013

² School of finance, Law&Business College of Hubei University of Economics, Wuhan, China, 430205

* Corresponding Author Email: 320600386@qq.com

Abstract. With the support and promotion of micro short dramas by the state and the continuous improvement of the basic network environment, China's micro short theater continues to develop vigorously. This paper takes residents of Nanchang as the research object, and analyzes the viewing status, payment intention and market development suggestions of micro short drama through questionnaire survey and in-depth interview. Stratified sampling and three-stage unequal probability sampling were used to collect data, and descriptive statistics, binary Logistic model, structural equation model and K-means clustering model were used for data analysis. The results of the survey show that: (1) The market potential of micro short drama is large, but the content quality and payment attraction need to be improved. (2) There are significant differences in the willingness of audiences of different ages, educational backgrounds and genders to pay for micro short drama. (3) Content quality, user experience and other factors significantly improve the willingness of micro short drama users to pay and promote payment. (4) The core of potential customers for micro short drama is young people. Based on the above research results, this paper puts forward several policy suggestions for consumers and platforms to promote the high-quality development of the micro short drama market.

Keywords: Micro short drama, Willingness to pay, Binary logistic model, Structural equation model, K-means clustering.

1. Introduction

With the rapid development of the Internet, various short video platforms led by Douyin gradually rise, and people gradually get used to short, flat and fast short video entertainment, and traditional long TV drama are gradually replaced by micro short drama [1]. At the same time, content creators are flocking to the micro short drama space. By taking advantage of the creation tools and distribution channels provided by the short video platform, creators can quickly transform their creativity into actual micro short drama. Micro short drama shows their advantages in content creativity, communication methods and audience interaction, and quickly occupy the online video market with their unique charm, showing huge development potential [2].

However, through market research and literature review, the micro short drama market still faces many challenges. On the one hand, there is a contradiction between the audience's evaluation of the micro short drama and the market performance. Although the market size of micro short drama continues to grow, the audience's criticism of its production level and quality is also endless. On the other hand, the micro short drama market also has problems of serious homogeneity and excessive piracy resources [3].

In order to deeply understand the development status and existing problems of the micro short drama market and explore the path of its sustainable development, this paper takes residents of Nanchang City as the research object, carries out a series of surveys and researches, and collects data on the current situation of the micro short drama market, willingness to pay and market development suggestions through questionnaires and in-depth interviews. Descriptive statistics, binary Logistic model, structural equation model and K-means clustering model were used to analyze the data. Through data analysis, conclusions are drawn, and policy recommendations are put forward to

promote the high-quality development of the micro short drama market, and achieve sustainable development of the industry.

2. Data

This paper takes residents of Nanchang City as the research object, mainly carried out a series of surveys through questionnaires, in-depth interviews and literature review, and collected data on the viewing status, payment willingness and development suggestions of micro short drama. The survey adopts the probability sampling method combining stratified sampling and three-stage sampling, in which the three-stage sampling is to extract administrative districts from Nanchang City, then communities from administrative districts, and finally permanent residents. Finally selected Hongcheng community, Guangming community, Hongming community, Yulin community, Chunhui community, Shajing community, Changdong Town, Machu Town, Wengang Town, Qili Township, Chang 'a community, Changfu community. We asked the residents in the specific communities we selected. After obtaining the consent, we conducted the survey in the form of questionnaires. We randomly stopped the relevant staff in the communities and conducted the survey by using the designed questionnaires. At the same time, the network questionnaire was issued for investigation to understand the relevant situation and collect sample data. In addition, the questionnaire content design focuses on four basic frameworks: personal basic information, skit understanding and viewing status, skit payment willingness, skit market suggestions and expectations to obtain the required data. In the preliminary survey, 219 questionnaires were distributed, and 195 were effectively recovered, with an effective recovery rate of 89%. In the formal investigation, 1014 questionnaires were issued, and 920 were effectively recovered, with an effective recovery rate of 90.7%. For the collected survey data, we use reliability test and validity test to conduct reliability and validity test to ensure the feasibility and accuracy of the survey.

Through the analysis of the data, we make the following descriptive statistics.

Analysis of the current situation of watching micro short drama: (1) Most users know and have watched micro short drama; (2) More than 50% of the audience think that the content of micro short drama is not up to expectations; (3) Micro short drama has become a new choice for most viewers to adapt to The Times.

Analysis of paying intention of micro short drama: (1) The audience's paying intention for micro short drama shows an almost 50/50 trend; (2) Non-paying resources become the fierce rival of paid micro short drama.

Analysis of audience's suggestions and expectations for the paid micro short drama market: (1) Content quality improvement and mechanism policy improvement have become the top priorities of the micro short drama market; (2) Faster and stable updates and platform mechanism optimization are most expected by the audience.

3. Research on consumption intention of micro short drama based on binary Logistic model

By establishing a binary selection model [4], this section quantitatively analyzes the influence of factors such as gender, age, education background, entertainment budget and frequency of watching micro short drama on whether they are willing to consume micro short drama, and then identifies the significant characteristics of consumers who are willing to consume micro short drama.

3.1. The choice of binary Logistic research model

Binary Logistic regression model can be used to analyze the dependent variable of "willing to pay for short drama". Through multivariate regression [5], this model explains the relationship between the audience's willingness to pay and the influencing factors, and limits the value of the dependent variable to the range of [0,1], which meets the analysis requirements of this study.

Logistic regression analysis model is:

$$\ln\left(\frac{p}{1-p}\right) = \beta_0 + \sum_{i=1}^n \beta_i x_i \quad (1)$$

Since the questionnaire data contains both ordered and disordered categorical variables, ordered variables (such as entertainment budget) should be coded directly, while disordered multi-categorical variables (such as age) should be coded by dummy variables to avoid the bias caused by simple coding, so as to more accurately analyze the influence of various factors on payment willingness.

3.2. Research on the characteristics of consumer groups of micro short drama

Logistic regression results, as shown in Table 1, show that gender, age, educational background, entertainment budget and viewing frequency significantly affect the willingness to consume micro short drama, among which men, young people, groups with higher entertainment budget and viewing frequency are more inclined to consume.

Table.1. Logistic regression result

Variable	Coefficient	Wald	P value
Constant term	1.215**	4.903	0.027
Gender((Male)	0.548***	3.234	0.004
Age((Under 8 years old)	1.952***	12.524	0.001
Age((18-25 years old)	2.658**	9.154	0.023
Age((26-30 years old)	2.859	19.564	0.241
Age((31-40 years old)	2.157	19.556	0.128
Age((41-50 years old)	-2.641***	7.325	0.006
Age((51-60 years old)	-1.254**	6.284	0.013
Age((Over 60 years old)	-1.741**	5.641	0.067
Education (junior high school and below)	0.609***	0.254	0.003
Education (High school/Technical School/Technical School)	0.456	0.487	0.685
Education (Junior College)	0.511**	0.824	0.014
Education (Bachelor degree	0.647**	0.745	0.025
Education (Master degree or above)	0.744	0.332	0.131
Entertainment budget	1.695***	9.892	0.000
Viewing frequency	2.465***	17.186	0.006

Note: ***, ** and * represent significance levels of 1%, 5% and 10% respectively.

The statistical significance of the model is good, which provides reliable support for user analysis and market strategy. Finally, a binary choice Logit model with the following formula is established:

$$y = 1.215 + 0.548\chi_{gender} + 1.952\chi_{age_{below18}} + 2.658\chi_{age_{18to25}} + 2.859\chi_{age_{26to30}} + 2.157\chi_{age_{31to40}} - 2.641\chi_{age_{41to50}} - 1.254\chi_{age_{51to60}} - 1.741\chi_{age_{above60}} + 0.609\chi_{MidSch} + 0.456\chi_{HighSch} + 0.511\chi_{JuniorCollege} + 0.647\chi_{College} + 0.744\chi_{Graduate} + 1.695\chi_{budget} + 2.465\chi_{frequency} \quad (2)$$

3.3. Model result analysis

The analysis of binary Logistic model shows that the main consumer group of micro short drama is young users, especially 18-25 years old, who prefer fast-paced and interactive content. Entertainment budget, viewing frequency and educational background have significant effects on consumption intention: users with higher budget and viewing frequency are more willing to consume; Among them, the groups of "junior high school and below", "college junior college" and "college undergraduate" have higher consumption intention, and the regression coefficients are 0.609, 0.511 and 0.647, respectively. Highly educated users tend to prefer in-depth content. The platform can attract users with different budgets, viewing frequencies and education levels through personalized

recommendations, preferential activities and other strategies to improve the overall payment conversion rate.

4. Analysis of influencing factors of willingness to pay for short drama based on structural equation model

4.1. Model selection

As a key index to predict consumer behavior, the willingness to pay has important guiding significance for the market strategy formulation of micro short drama enterprises. However, the factors influencing users' willingness to pay are complex and diverse, and their mechanisms need to be further explored. Based on the theoretical framework of TPB planned behavior, this study used structural equation model (SEM) to test the causal relationship among various influencing factors, with a view to revealing the formation mechanism of users' willingness to pay for micro short drama [6].

4.2. Structural equation principle

1. Basic structure of structural equation model

The structural equation model is divided into two parts: measurement model and structure model. The specific formulas and meanings are as follows:

$$x = \alpha\tau + \delta \quad (3)$$

$$y = \alpha\omega + \xi \quad (4)$$

The measurement equation of structural equation model consists of two parts: exogenous variable formula (3) and endogenous variable formula (4). In the exogenous variable equation, x is the exogenous observed variable vector, $\alpha\tau$ is the exogenous factor load matrix, τ is the exogenous latent variable vector, and δ is the residual term vector. In the endogenous variable equation, y is the endogenous observed variable vector, $\alpha\omega$ is the endogenous factor load matrix, ω is the endogenous latent variable vector, and ξ is the residual term vector [7].

Structural equation:

$$\omega = B\omega + v\tau + \sigma \quad (5)$$

In formula (5), B represents the relationship matrix between endogenous latent variables; v represents the relation matrix between exogenous and endogenous latent variables; σ represents the error term of the structural equation.

2. Theoretical model establishment

In this study, four dimensions under the framework of TPB planning behavior theory were used as latent variables, and the corresponding observation indexes were determined according to the screening results of the scale. After normative and accurate revision, the measurement index system including payment willingness and payment behavior was finally constructed, as shown in Table 2.

Table.2. Latent variable - observed variable correspondence table

Latent variable	Observed variable	Meaning of observed variable	Model variable
Content quality	Q24.1	Content likability	NR1
	Q24.2	Production level	NR2
	Q24.3	Innovativeness	NR3
User experience	Q25.1	Platform operation convenience	TY1
	Q25.2	The appropriateness of advertising	TY2
	Q25.3	Information service quality	TY3
Subjective norm	Q26.1	Influenced by family or friends who decide to buy micro short drama	ZG1
	Q26.2	Buying micro short drama will increase common topics with friends or family	ZG2
	Q26.3	Social and cultural identity	ZG3
	Q26.4	People are interested in buying a micro short drama because of the discussion	ZG4
Perceptual behavioral control	Q27.1	Have enough money to buy the micro short drama	ZJ1
	Q27.2	Have enough time and energy to watch paid micro short drama	ZJ2
	Q27.3	The process of buying micro short drama is simple and easy to operate	ZJ3
	Q27.4	Have the confidence not to indulge in paid micro short drama	ZJ4
Willingness to pay	Q28.1	Willing to pay for your favorite micro short drama	FF1
	Q28.2	Willing to continue to pay to watch micro short drama in the future	FF2
	Q28.3	Be willing to recommend paid micro short drama to friends or family	FF3
Payment behavior	Q17	Whether to pay for a micro short drama	JG1
	Q22	Recommend friends or family members to pay to watch the micro short drama	JG2

Based on the above four factors that affect the purchase intention of micro short drama, the following hypothesis summary table can be obtained, as shown in Table 3:

Table.3. Research hypothesis summary table

Hypothetical number	Research hypothesis
H1	Content quality has a positive and significant impact on the payment of micro short series
H2	User experience has a positive and significant impact on the payment of micro short series
H3	The subjective norm has a positive and significant influence on the payment of micro short drama
H4	Perceptual behavior control has a positive and significant effect on the payment of short drama
H5	The willingness to pay for micro short drama has a positive and significant effect on users' paying behavior

4.3. Analysis of influencing factors of payment willingness of micro short drama

1. SEM model fit test of influencing factors of payment willingness of micro short drama

According to the results of model fit test in Table 4, it can be seen that CMIN/DF (Chi-square freedom ratio) = 1.941, in the range of 1-3, RMSEA (mean square error) = 0.014, in the excellent range of less than 0.05. The test result of TLI was 0.853, which was a good level, while the test results of IFL and CFI both reached an excellent level above 0.9. Therefore, based on the analysis structure of this paper, it can be shown that the SEM model for the analysis of factors affecting the willingness to pay for short drama has a good fit [8].

Table.4. Model fit test

Index	Reference standard	Result of actual measurement
CMIN/DF	1-3 is excellent, 3-5 is good	1.941
RMSEA	<0.05 is excellent, <0.08 is good	0.014
IFL	>0.9 is excellent, >0.8 is good	0.938
TLI	>0.9 is excellent, >0.8 is good	0.853
CFI	>0.9 is excellent, >0.8 is good	0.921

2. Test results of the path relation hypothesis of SEM model on factors influencing the willingness to pay for short drama

In this paper, AMOS software was used to build the structural equation model shown in the figure. Maximum likelihood method was adopted for estimation, and reasonable modifications were made according to the model structure. The final model was shown in Figure 1.

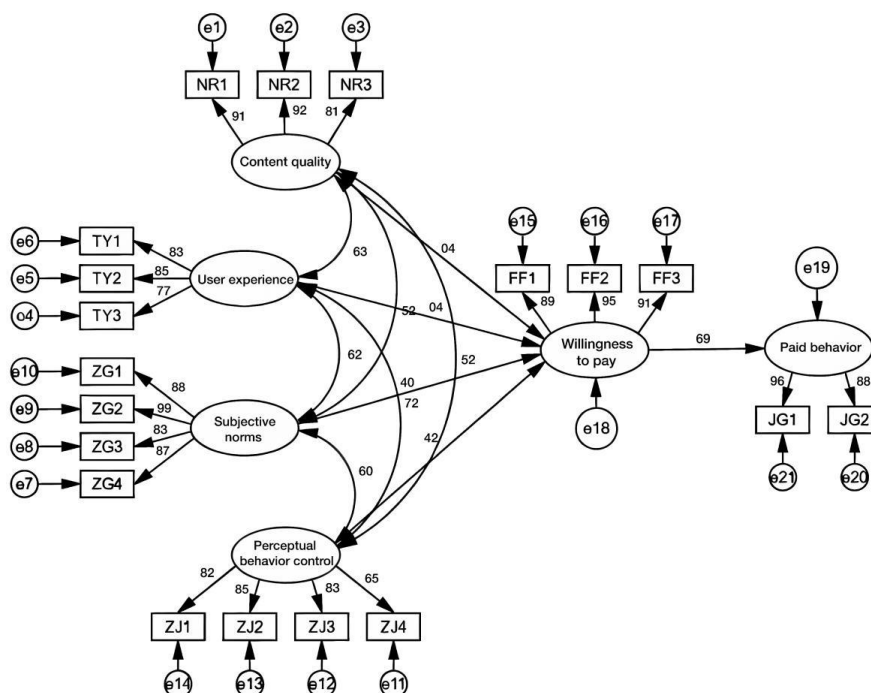


Figure. 1 SEM analysis model of influencing factors of willingness to pay for short drama

4.4. Interpretation and analysis of model results

Table.5. Path coefficient result

Path relationship			Estimate	S.E.	C.R.	P
Willingness to pay	<---	Subjective norm	0.397	0.111	4.077	***
Willingness to pay	<---	Perceptual behavioral control	0.424	0.175	3.472	***
Willingness to pay	<---	Content quality	0.041	0.021	3.446	***
Willingness to pay	<---	User experience	0.520	0.194	4.304	***
Payment behavior	<---	Willingness to pay	0.689	0.067	8.493	***

As shown in Table 5, the significance test results based on path coefficient show that the four dimensions of subjective norms, perceived behavior control, content quality and user experience all have significant positive impacts on users' willingness to pay for micro short drama. Further analysis shows that there is a significant positive correlation between users' willingness to pay and their actual paying behaviors, which indicates that when users form positive paying intentions, they are more likely to translate into actual paying behaviors. These findings confirm the rationality of hypothesis H1 to H5 proposed in this study. The specific analysis is as follows:

First, the significant impact of subjective norms on payment willingness verifies H1, indicating that users' payment decisions are influenced by social reference groups, which is consistent with the core view of planned behavior theory. Second, the positive effect of perceived behavioral control on payment intention supports H2, indicating that users' evaluation of their own resources and abilities directly affects their payment intention. Thirdly, content quality and user experience, as key factors at the product level, significantly affect users' willingness to pay, verifying H3 and H4 respectively, which reflects the important role of product attributes in users' decision-making process. Finally, H5 is supported by the significant predictive effect of payment intention on payment behavior, which further validates the applicability of attitudinal behavior theory in payment scenarios of micro short drama.

(1) Content quality

Table.6. Estimation of "content quality" of structural equation model

Path relationship	Path coefficient	Standardized path coefficient
NR1<--- Content quality	1.000	0.907
NR2<--- Content quality	1.104***	0.921
NR3<--- Content quality	1.043***	0.810

Note: "****" indicates that the path coefficient is significant at the 1% level

As shown in Table 6, the content quality of micro short drama has a significant positive impact on users' willingness to pay. This is mainly reflected in the three dimensions of high-quality script writing, professional production level and innovative content. The combined effect of these factors has effectively improved users' willingness to pay for viewing.

(2) User experience

Table.7. Structural equation model "user experience" estimation results

Path interpretation	Path coefficient	Standardized path coefficient
TY3<--- User experience	1.000	0.774
TY2<--- User experience	1.171***	0.851
TY1<--- User experience	1.000***	0.828

Note: "****" indicates that the path coefficient is significant at the 1% level

User experience has a significant positive impact on the willingness to pay for micro short drama. As shown in Table 7, platform interaction convenience, viewing fluency and after-sales service quality constitute the key factors affecting users' payment decision, which is consistent with the correlation between experience value and behavioral intention in user experience theory. A good user experience promotes willingness to pay by increasing user satisfaction and engagement.

(3) Subjective norms

Table.8. Structural equation model "subjective norm" estimation results

Path interpretation	Path coefficient	Standardized path coefficient
ZG4<--- Subjective norm	1.000	0.865
ZG3<--- Subjective norm	0.976***	0.825
ZG2<--- Subjective norm	1.105***	0.898
ZG1<--- Subjective norm	1.081***	0.876

Note: "****" indicates that the path coefficient is significant at the 1% level

As shown in Table 8, subjective norms have a significant positive impact on users' willingness to pay for micro short drama. As an important social influence factor, subjective norms affect users' payment decisions through two levels: the micro level is reflected in the recommendation and experience sharing of relatives and friends, and the macro level is reflected in the evaluation of mass media and the direction of public opinion. This finding is consistent with the discussion of the influence of social norms on behavioral intention in planned behavior theory.

(4) Perceptual behavior control

Table.9. Structural equation model "subjective norm" estimation results

Path interpretation	Path coefficient	Standardized path coefficient
ZJ4<--- Perceptual behavior control	1.000	0.654
ZJ3<--- Perceptual behavior control	1.205****	0.831
ZJ2<--- Perceptual behavior control	1.291****	0.848
ZJ1<--- Perceptual behavior control	1.350****	0.825

Note: "****" indicates that the path coefficient is significant at the 1% level

Table 9 shows that perceptual behavioral control has a significant positive impact on users' willingness to pay for micro short drama. Perceived behavioral control is mainly reflected in users' evaluation of their own economic ability, time resources and self-management ability. This finding is consistent with the discussion on the influence of perceived behavioral control on behavioral intention in planned behavior theory.

5. Potential consumer mining based on cluster analysis

5.1. Model building

1. Definition of potential customers

The purpose of this study is to identify the potential paying user groups of micro short drama and provide a basis for the content development and marketing strategy optimization of the platform. The study defines potential users as those who are not currently paying but have the potential to convert. Through the analysis of questionnaire data, it is found that some non-paying users show similar characteristics and attributes to existing paying users. Based on K-means cluster analysis, 313 non-paying users meeting certain criteria were selected from respondents who answered question 17 (willingness to pay) and identified as core potential customers.

2. Establishment of potential customer value model

After preliminary analysis of consumer characteristics of micro short drama, we established a potential customer value model and used K-means clustering to identify high-value potential consumer customers [9]. Through questionnaire data research, we selected five key indicators to quantify customer value, so as to identify the audience with the greatest consumption potential and provide support for the platform to formulate precise promotion strategies.

5.2. Model application and analysis

1. Selection of clustering factors

Before the cluster analysis, we selected five key factors from the questionnaire as the model indicators: age, occupation, education, viewing purpose and willingness to pay (according to the score of question 23, full score of 15) [10]. These indicators are denoted as "A, O, E, P, W". It is used to

group potential micro short drama users and identify the audience with the most consumption potential.

2. Construction of potential customer value model

(1) Potential customer clustering

With the help of SPSS software for cluster analysis, we classified potential customers into four categories, as shown in Table 10.

Table.10. Lead type clustering center table

Potential customer type	I	II	III	IV
Age	26-30 years old	18-25 years old	51-60 years old	31-40 years old
Occupation	Freelance	Students	Self-employed/contractor	business managers
Education	Junior College	Undergraduate	High school/technical school	Undergraduate
Viewing purpose	Meet the needs of fragmented time	release stress	pass time	keep social conversation
Willingness to pay	13.26	11.62	7.82	5.91

(2) Potential customer value analysis

According to the characteristics analysis of the clustering results, we divided the lead users into four categories: important leads, important development customers, minor leads, average and low value leads. The characteristics of each potential customer are as follows:

① Important Potential Users (Category I)

Mainly 26–30-year-old freelancers, college degree, the highest willingness to pay (13.26 points). They have a fast pace of life, meet the needs of fragmented time through micro short drama, have high requirements for content quality and diversity, and like to share the viewing experience. Platforms should focus on content innovation and user experience, leveraging social media promotion and personalized recommendations to enhance user engagement and maximize the consumption potential of this group.

② Important development Users (Category II)

It is mainly 18-25 years old students with a bachelor's degree and a willingness to pay is 11.62 points. This group is interested in micro short series, mainly through watching to release pressure, tend to choose light entertainment content. Platforms can expand their market by engaging them with personalized recommendations, social media ads and offers, and increasing their willingness to pay through increased engagement and content diversity.

③ Secondary Potential Users (Category III)

This group of users is mainly 51-60 years old, self-employed, with a high school education or below, and its willingness to pay is 7.82 points. This group does not have a strong demand for micro short drama consumption, and the platform can consider providing them with more content types that meet their interests, such as adding some easy and easy to understand micro short drama to stimulate their watching interest, so as to promote the transformation of potential consumption.

④ Average and low value potential users (Category IV)

This kind of users are mainly 31-40 years old business managers with a bachelor's degree, and their willingness to pay is 5.91 points, which shows that this group has a low value in the consumption of micro short drama.

6. Conclusions and policy recommendations

Based on the investigation and research of Nanchang residents' willingness to consume micro short drama, this paper draws relevant conclusions. The research points out that micro short drama has great market potential, but the content quality and payment attraction need to be improved. There are significant differences in the willingness to pay for micro short drama among audiences of different ages, educational backgrounds and genders. Content quality and user experience can promote users' willingness to pay, and the core potential customers of micro-short plays are young people.

At present, there are still some shortcomings in the development of the micro short drama market, such as the content quality, originality and copyright protection need to be improved, the update speed and user experience need to be optimized, and the willingness of different audience groups to pay varies greatly. Next, consumers should enhance copyright awareness, pay attention to cost performance and actively feedback to help platforms improve content; The micro short drama platform should improve the content quality, enrich the types of plays, optimize the payment model to improve the cost performance, and enhance the interaction and social attributes, so as to improve the overall quality and market acceptance of micro short drama, and better tap the market potential of micro short drama.

References

- [1] Sun M. Research on Consumption Characteristics and Marketing Strategy of Micro-short Drama from the Perspective of Self-Determination Theory[J]. Journal of Northeast Asian Arts and Cultural Management, 2024, 1(1): 1-11.
- [2] Aronsson K. Micro drama and morality-in-interaction: A commentary[J]. Text & Talk, 2020, 40(5): 695-703.
- [3] QinXinyuan.Cold thinking of micro short drama fever[J]. Acta YouthologicaSinica,2024(1):88-93. (inChinese)
- [4] Huang F L. Alternatives to logistic regression models when analyzing cluster randomized trials with binary outcomes[J]. Prevention Science, 2023, 24(3): 398-407.
- [5] de Rooij M, Groenen P J F. The melodic family for simultaneous binary logistic regression in a reduced space[M]//Facets of Behaviormetrics: The 50th Anniversary of the Behaviormetric Society. Singapore: Springer Nature Singapore, 2023: 67-97.
- [6] Cox K, Kelcey B, Bai F. Croon's bias-corrected estimation for multilevel structural equation models with latent interactions[J]. Structural Equation Modeling: A Multidisciplinary Journal, 2023, 30(3): 467-480.
- [7] Newsom J T, Hachem Z A, Granger A M, et al. Where did I go wrong with my model? Ten tips for getting results in SEM[J]. Structural Equation Modeling: A Multidisciplinary Journal, 2023, 30(3): 491-500.
- [8] Zyphur M J, Bonner C V, Tay L. Structural equation modeling in organizational research: The state of our science and some proposals for its future[J]. Annual Review of Organizational Psychology and Organizational Behavior, 2023, 10(1): 495-517.
- [9] Crowther D, Kim S, Lee J, et al. Methodological synthesis of cluster analysis in second language research[J]. Language Learning, 2021, 71(1): 99-130.
- [10] Vera J F, Macías R. On the behaviour of K-means clustering of a dissimilarity matrix by means of full multidimensional scaling[J]. psychometrika, 2021, 86(2): 489-513.