Proceedings of ICFTBA 2026 Symposium: Driving Organizational Strategic Change for Financial Gains
DOI: 10.54254/2754-1169/2026.BH37330

Public Opinion and Audience Perception of Cross-Cultural Sports Stars Endorsements: A Text Mining Study Based on Haaland Endorsement Wang Laoji's Weibo Comment Section

Siheng Jing

Massey Institute, Nanjing University of Finance & Economics, Nanjing, China
jingsiheng4@gmail.com

Abstract. Under the background of globalization, it is increasingly common for domestic time-honored brands to jointly sign western sports stars to carry out cross-cultural endorsement, but the cultural differences between China and the west can easily lead to polarized audience feedback on social platforms. Most of the existing studies evaluate the marketing effectiveness from the macro perspective of enterprises, and lack the analysis of the discussion content at the social media comment level. Taking the case of Haaland's endorsement of Wanglaoji as an example, this study selected Weibo as the research platform, supported by the theory of source credibility and matching hypothesis, and used two-stage LDA topic modeling based on algorithm modeling and artificial semantics to carry out empirical analysis. The comments were refined into nine sub topics and summarized into three core dimensions: star emotional identity, cross-cultural advertising conflict, and brand international communication. The study found that there was a clear stratification of public opinion, with brand spread to the sea as the mainstream discussion, followed by fan circle evaluation, and cultural adaptation disputes in the minority. This study complements the micro audience demonstration of cross-cultural sports endorsement with strong contrast, and provides reference for domestic time-honored brands to optimize overseas endorsement strategies, advertising creativity and network public opinion management.

Keywords: Cross-cultural sports endorsement, LDA topic modeling, Social media opinion, Brand internationalization

1. Introduction

With the deepening of globalization, the internationalization of local brands continues to accelerate. Cross-cultural endorsement with the help of international sports stars has become an important marketing strategy for brands to build an international image [1]. Among them, the cooperation between western sports stars and Chinese traditional time-honored brands has caused extensive discussion on social platforms for many times because of its high topic attention, which has become a practical phenomenon worthy of attention in the marketing of domestic brands to the sea. Previous studies have shown that in the social media environment, sports star endorsement can affect

© 2026 The Authors. This is an open access article distributed under the terms of the Creative Commons Attribution License 4.0 (https://creativecommons.org/licenses/by/4.0/).

DOI: 10.54254/2754-1169/2026.BH37330

consumers' brand attitude, brand trust, word-of-mouth communication and purchase intention, but this effect will differ due to different cultural markets [2, 3]. On the social media platform, the audience's expression of identity and discussion of sense of conflict coexist. Public discussion topics show complex and diverse characteristics, and their internal and composition and relevance logic have not been fully analyzed and explained.

Existing studies have proved that the professional image of sports stars has cross-cultural universality, which can effectively reduce the cross-cultural cognitive threshold of brands [4]. The embedding of local cultural elements can alleviate the loss of perception caused by cultural violations, but the existing studies mostly carry out the macro effect evaluation from the perspective of enterprise strategy, and lack the detailed disassembly of the audience's discussion content from the micro perspective [5]. The research points out that moderate cognitive contrast and creative novelty can activate the audience's attention and communication willingness, and become an important driving mechanism of advertising communication [6]. The research points out that cross-cultural co branding has audience feedback with both communication gains and perceived risks, but the semantic distribution and hierarchical linkage logic of cross-cultural endorsement for strong contrast have not been systematically explained [7]. At the research method level, LDA topic modeling has become the mainstream tool for social text topic mining. In recent years, relevant studies have further discussed how to improve the semantic coherence and consumer insight value of topic results, but most studies still directly present the algorithm output results, and less systematic semantic aggregation is carried out in combination with the theoretical framework, resulting in the theoretical adaptability and interpretability of topic results to be improved [8].

This study has important academic significance for expanding the application boundary of cross-cultural marketing theory in the context of low cultural fit, and optimizing the Methodology Path of short text theme analysis. At the same time, it can also provide empirical reference for the cross-cultural endorsement practice of Chinese time-honored brands. Based on the above realistic background and research gap, this study is mainly carried out from the following aspects: Taking the cross-cultural sports endorsement event of Chinese time-honored brand Wanglaoji signing a contract with Western football star Haaland as a case, this paper systematically investigates the core discussion topics formed by the public around the event in the context of Weibo, the semantic connotation and intensity distribution characteristics of each topic, as well as the public opinion structure and internal correlation logic between different topics.

At the theoretical level, this study is supported by the source credibility theory, match-up hypothesis and Integrated Framework of Schema Incongruity. Among them, the source credibility theory defines the stabilizing effect of spokesperson professionalism on consumer attitudes, the match-up hypothesis reveals the core influence mechanism of spokesperson image fit on spokesperson evaluation, and the Integrated Framework of Schema Incongruity provides a theoretical anchor for explaining the activation effect of moderate cognitive contrast on audience attention and communication willingness, which together constitute a theoretical anchor for analyzing the public opinion structure of strong contrast cross-cultural endorsement [9-11].

In terms of research methods, this study collected the Weibo platform user comment data, and used the two-stage LDA analysis method of fine-grained algorithm modeling combined with artificial semantic aggregation. Specifically, in the first stage, LDA topic modeling is used to automatically identify and assign the probability of large-scale short texts, and the optimal number of topics is determined according to the consistency index; In the second stage, under the guidance of the above theoretical framework, conceptual integration and hierarchical induction are carried out on the sub themes output by the algorithm, and the initial themes are aggregated into core theme

DOI: 10.54254/2754-1169/2026.BH37330

dimensions with theoretical significance. The advantages of this method are: algorithm modeling ensures the objectivity and reproducibility of topic mining, while artificial semantic aggregation makes up for the lack of theoretical explanatory power of pure algorithm results. The combination of the two helps to improve the theoretical adaptability and interpretability of topic results, so as to more systematically present the public opinion structure in the context of strong contrast cross-cultural endorsement.

The research objective of this study is to systematically reveal the core discussion themes and semantic connotations of the public around the strong contrast of cross-cultural sports endorsement events in the context of Weibo, depict the intensity distribution and hierarchical characteristics of each theme, clarify the mutual penetration and correlation logic between different themes, and then build an overall interpretation framework of cross-cultural sports endorsement audience public opinion.

This study has both theoretical and practical value. At the theoretical level, it supplements the public's perception evidence of strong contrast cross-cultural sports endorsement from the micro perspective, expands the application boundary of the matching hypothesis in low cultural fit scenarios, and optimizes the theoretical adaptation path of LDA topic analysis in short text scenarios. At the practical level, the research conclusion can provide empirical basis and decision-making reference for Chinese time-honored brands to control the creative scale of cross-cultural endorsement, optimize the content communication strategy and public opinion guidance scheme.

2. Research method

2.1. Data acquisition method

Based on the general understanding of Weibo comments, this study selected the Octoparse to obtain the relevant comment data. This study chose to collect all the comments under 17 related Weibo jointly signed by Wanglaoji and Haaland, set a custom collection process in Octoparse, collect while scrolling on the page, and switch to the next link for cyclic crawling when a single blog post slides to the bottom without new comments. The request interval is set to 10 seconds.

A total of 3252 pieces of original data were collected this time. The captured fields include blogger ID, user name, publishing time, body, likes, forwards, comments, etc. only the comment content is retained as the core analysis text, the rest fields are archived, and the data is exported to CSV format. Excel was used to remove the duplicate comment text, and 2437 valid samples were obtained after removing the duplicate.

2.2. Data preprocessing

Use Python to complete text cleaning, word segmentation, stop-word removal and TF-IDF vectorization. Firstly, regular expressions are used to clean up irrelevant content such as special symbols, expressions, links, topic tags, numbers, etc., and empty text and short meaningless comments are deleted to ensure the validity of the sample. Then, Jieba word segmentation tool is used to process the text, and the effect of word segmentation is optimized through the dictionary. Import custom brands and professional dictionaries in the field of sports, unify semantic expression with the help of thesaurus, load the HIT stop-word list and supplement exclusive stop words with research, and filter function words without actual semantics. After the word segmentation is completed, the effect is checked by manual sampling. Using gensim library to build a word dictionary, TF-IDF method is used to generate text vector, reduce the weight of high-frequency

DOI: 10.54254/2754-1169/2026.BH37330

general vocabulary, extract effective public opinion characteristic vocabulary, and generate standardized corpus for LDA modeling.

2.3. Research method: LDA topic modeling

In this study, LDA method is used to model the topic of comment text. LDA is an unsupervised probabilistic topic model that can automatically identify latent semantic structures in large-scale, unlabeled user-generated text and extract core discussion topics. It is the most widely used topic mining method in the current network public opinion research and user perception analysis.

3. Research results

Based on the preprocessed Weibo comment data, the LDA model is trained in turn within the range of topic number K from 1 to 10, and the topic coherence is used as the core evaluation index to determine the optimal initial number of topics at the fine-grained level. Topic coherence is used to measure the semantic relevance of high-frequency words under the same topic. The higher the value, the more cohesive the internal semantics of the topic, the clearer the boundary, and the stronger the interpretability.

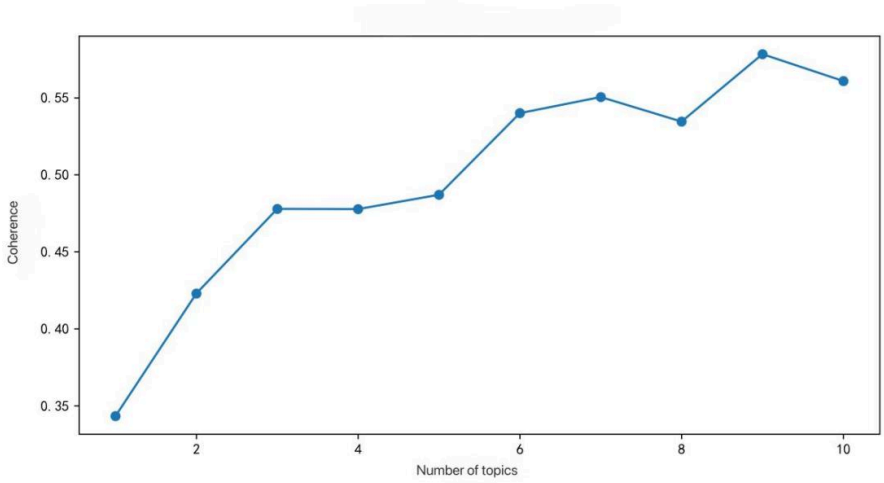

Figure 1. Line chart showing topic coherence against the number of topics (picture credit: original)

As shown in Figure 1, from the perspective of coherence change, the overall value shows a fluctuating upward trend. When the number of topics is between 1 and 3, the coherence is fast, indicating that topic segmentation can significantly improve semantic cohesion. When K=3 to K=4, the coherence remained stable, and the marginal lifting effect was significantly weakened. After K=4, the coherence continued to fluctuate and rise until it reached the peak value of the whole region at k=9. At this time, the lexical semantic relevance within each topic was the highest, the topic discrimination was clear, and the semantic interpretability was the best. When the number of topics increases to k=10, the consistency value drops, indicating that too many topics will lead to excessive semantic splitting, destroy the semantic coherence within a single topic, and reduce the overall quality of the topic.

Based on this, this study finally determines the optimal initial number of topics at the fine-grained level as K=9. Under this value, the topic output from the model not only has the optimal

DOI: 10.54254/2754-1169/2026.BH37330

semantic cohesion, but also avoids the semantic fragmentation caused by excessive splitting, and can completely and accurately capture the subdivision discussion dimension in the comments.

LDA model outputs the membership probability of each comment document to 9 topics. This study uses the maximum-probability assignment principle to classify each comment into the topic with the highest probability, so as to count the document size and proportion of each topic. The results are shown in Table 1.

Table 1. Document distribution across LDA topics

No.	topic	topic-related words
topic 1	brand official announcement attitude evaluation	support, congratulations, World Cup, match performance, success, boost
topic 2	star player athletic performance evaluation	domineering, sure win, awesome, going overseas, co-branding, move, match
topic 3	advertising cultural discrepancy perception	abstract, art style, meme-like, invincible, wonderful, stress-relieving
topic 4	fan emotional expression & identification	congratulations, reduce internal heat, free giveaway, qualify, victory, magic weapon, match-watching
topic 5	product consumption attitude evaluation	herbal tea, co-branding, packaging, foreigners, god-tier, stockpiling, refreshing
topic 6	product advertising creativity evaluation	thunder, excellent, image, texture, adaptation, magic, unbearably tacky
topic 7	advertising entertainment-oriented communication discussion	ad song, catchy, funny, contrast, clash, hardcore, marketing
topic 8	product marketing campaign discussion	cheer on, campaign, hilarious, anticipate, creativity, lucky draw, jersey
topic 9	match-viewing scene marketing evaluation	perks, lucky draw, limited edition, match-watching, advance, schedule, cheer/support



Figure 2. Overall word cloud of the comment corpus (picture credit: original)

From the perspective of distribution characteristics, the volume of the nine fine-grained themes is relatively balanced, and the proportion of single themes is in the range of 9% to 13%. There are no extremely dominant or extremely minority themes, indicating that the fine-grained theme division is

DOI: 10.54254/2754-1169/2026.BH37330

relatively reasonable and can completely and truly capture the subdivision dimension of audience discussion. Figure 2 shows the overall word cloud.

4. Discussion

4.1. The core discussion theme formed by the public around this event in the context of weibo

In the iterative experiment of topic number range K from 1 to 10, combining topic coherence and semantic interpretability, the initial optimal number of fine-grained topics was determined to be 9. The key words are respectively around the nine dimensions of brand official announcement attitude evaluation, star player athletic performance evaluation, advertising cultural discrepancy perception, fan emotional expression & identification, product consumption attitude evaluation, product advertising creativity evaluation, advertising entertainment-oriented communication discussion, product marketing campaign discussion, match-viewing scene marketing evaluation.

Due to the problems of weak semantics, fuzzy topic boundaries and overlapping keywords in Weibo comments, the direct use of nine sub topics is not conducive to systematic analysis. By measuring the semantic overlap of key words, comparing the spatial distance of topics, and combining with the theory of cross-cultural sports marketing, this study carried out artificial semantic aggregation, and finally summarized the nine sub themes into three core discussion categories. The first category is the theme of star player professional image and fan emotional identification; the second category is the theme of cross-cultural advertising creativity and cultural discrepancy perception; and the third category is the theme of brand internationalization strategy and product consumption association. The three types of topics completely covered all the core discussion directions of the Weibo platform public for the endorsement event, and clearly answered all the core discussion topics formed by the public, as shown in Table 2.

Table 2. Topic semantic aggregation and core category classification

core Meta-theme	includes initial topics	characteristic words
star player professional image and fan emotional identification	1, 2, 4	strength, support, handsome, sure win, qualify, anticipation, victory, excellent, reduce internal heat, free giveaway, etc
cross-cultural advertising creativity and cultural discrepancy perception	3, 6	bizarreness, addictive, catchy, awkward, funny, visual style, campaign, interaction, creativity, advertising, etc.
brand internationalization strategy and product consumption association	5, 7, 8, 9	going overseas, globalization, marketing, herbal tea, reduce internal heat, packaging, purchase, perks, limited edition, lucky draw, etc.

4.2. Semantic connotation and intensity distribution characteristics of each core theme

From the perspective of semantic connotation, the three core themes each present a differentiated public opinion expression logic.

First, the theme of star player professional image and fan emotional identification. The theme integrates the relevant contents of Haaland's competitive strength, global influence of stars, and fan traffic transformation in the sub theme. The semantic connotation focuses on the audience's attention carried by top stars and the circle communication effect of fan groups, forming the core expression perspective of fan groups.

DOI: 10.54254/2754-1169/2026.BH37330

Second, cross-cultural advertising creativity and cultural discrepancy perception. The theme integrates the cultural differences between China and foreign countries, the consumption of co-branded products, the question of image adaptation, the limitations of marketing strategies and other related contents. The semantic connotation not only covers the cognition and feedback of ordinary consumers and brand followers on co-branded activities, but also presents the public's discussions and suggestions on the cultural differences between China and Western countries and the tonality matching between brands and stars.

Third, the theme of brand internationalization strategy and product consumption association. The theme integrates cross-border brand communication, traditional culture going to sea, advertising and other related content, and semantic connotation combines cross-cultural output and Chinese and Western cultural barriers, reflecting the public's multiple cognitive logic of Wanglaoji's overseas joint marketing. On the one hand, it pays attention to the role of cross-cultural cooperation in promoting the internationalization of local brands. On the other hand, it also points out that the cultural differences between China and the West may lead to the limited expansion of brand audiences and the distance between overseas audiences' understanding of the core of local culture.

From the perspective of topic intensity and document distribution characteristics, the sample size of each topic has obvious stratification. The theme documents of brand internationalization strategy and product consumption association account for the highest proportion and the strongest theme intensity, which belong to the mainstream discussion direction of the public under this marketing event. The sample size of the topics of star player professional image and fan emotional identification communication takes the second place, forming a stable middle-level discussion perspective. The number of cross-cultural advertising creativity and cultural discrepancy perception documents is the least, which belongs to the critical discussion perspective of the relative edge. At the same time, each comment has a corresponding attribution probability of multiple topics, and the attribution probability of a single text of the mainstream topic is significantly higher than that of the minority controversial topic, reflecting that the public discussion as a whole is dominated by neutral and identity expression, and the voice of questioning is relatively low.

4.3. Comment structure and internal correlation logic between different topics

The overall comment content is not independent of each other, but forms a hierarchical theme structure around brand cross-cultural marketing. Among them, brand cross-cultural marketing constitutes the core issue, and the discussion of star value communication and cultural adaptation constitutes two related branches. The three are interdependent and have a close internal correlation logic.

First of all, the topic of star value is the precondition for the discussion of the other two types of topics. Haaland's global influence and huge fan base are the premise for this cross-border endorsement event to trigger extensive discussion. It is precisely based on the attention of the stars that the public has discussed brand marketing and cultural communication. At the same time, it has also generated various doubts about the image of the stars and the cognitive differences between China and foreign countries, which constitute the starting point of the discussion of the whole public opinion system.

Secondly, there is a complementary relationship between brand value identification and cultural design. The two share the core discussion carrier of cross-cultural cooperation. Some audiences recognized the value of this endorsement from the perspective of brand breaking and cultural output, while others questioned it from the perspective of cultural differences and brand adaptation. The two

DOI: 10.54254/2754-1169/2026.BH37330

together constitute a multi perspective review, forming a comprehensive public evaluation of this transnational endorsement.

The last three themes have the characteristics of mutual penetration and cross correlation. A large number of comments are mixed with multiple themes. While recognizing the value of stars and the advantages of brand marketing, some netizens will also point out the shortcomings of cross-cultural cooperation. It shows that public opinion is not a single and extreme expression, but a compound cognition of both identity and questioning. The three types of themes are intertwined and together constitute a cross-cultural sports endorsement evaluation system.

5. Conclusion

On the whole, the results of this study show that overseas sports stars' endorsement of local brands on the Weibo platform will form a layered, clear and diversified public opinion system. International brand communication is the mainstream discussion, followed by fans-circle evaluation, and cultural adaptation disputes account for a minority. It is further concluded that the strong contrast cross-cultural sports endorsement can form a public opinion pattern with the flow of stars as the bottom support and the coexistence of positive communication and cultural query. The flow of overseas stars brings significant exposure and communication opportunities for brands, but the differences in cultural cognition between China and the West will also give rise to divergent public opinions. This study supplements the empirical evidence of strong contrast cross-cultural sports endorsement from the perspective of micro audience, and expands the application boundary of the matching hypothesis in low cultural fit scenarios. At the same time, it provides data support and theoretical reference for domestic time-honored brands to optimize overseas sports endorsement strategy, advertising creativity and network public opinion management. Future research can expand multi platform comparison and cross case comparison to further test the universality of cross-cultural endorsement public opinion.

References

[1] Li, Y., Liu, B., & Xie, L. (2022). Celebrity endorsement in international destination marketing: Evidence from eye-tracking techniques and laboratory experiments. *Journal of Business Research, 150*, 553–566.
[2] Moreira, F., Pereira, I. V., Santos, J. D., & Pires, P. B. (2023). Sports celebrity endorsement impact on purchase intention. *Revista Galega de Economía, 32*(3).
[3] Crespo, C. F., Ferreira, A. G., Ribeiro, F. M., & Popa, V. (2025). The effect of celebrity endorsement on brand equity: A cross cultural comparison. *Journal of International Consumer Marketing, 37*(1), 1–22.
[4] Lee, Y., & Koo, J. (2015). Athlete endorsement, attitudes, and purchase intention: The interaction effect between athlete endorser-product congruence and endorser credibility. *Journal of Sport Management, 29*(5), 523–538.
[5] Hu, K., Kim, C., & Kang, E. Y. (2025). Cross-cultural analysis of influencer marketing: Meta-analysis and meta-regression of factors and effects. *Journal of Current Issues & Research in Advertising, 47*(2).
[6] van Berlo, Z. M. C., Meijers, M. H. C., Eelen, J., Voorveld, H. A. M., & Eisend, M. (2024). When the medium is the message: A meta-analysis of creative media advertising effects. *Journal of Advertising, 53*(2), 278–295.
[7] Osmanoglu, O., Özsomer, A., & Biliciler, G. (2024). Local–global cobrand positioning and consumer evaluations. *Journal of International Marketing, 32*(4), 81–98.
[8] Büschken, J., Otter, T., & Allenby, G. M. (2025). Accounting for formative and reflective topics in product review data for better consumer insights. *Journal of Marketing Research, 62*(6), 1045–1062.
[9] Ohanian, R. (1990). Construction and validation of a scale to measure celebrity endorsers' perceived expertise, trustworthiness, and attractiveness. *Journal of Advertising, 19*(3), 39–52.
[10] Kamins, M. A. (1990). An investigation into the "match-up" hypothesis in celebrity advertising: When beauty may be only skin deep. *Journal of Advertising, 19*(1), 4–13.
[11] Lee, E. J., & Schumann, D. W. (2004). Explaining the special case of incongruity in advertising: Combining classic theoretical approaches. *Marketing Theory, 4*(1–2), 59–90.