FITNESS, PERFORMANCE AND HEALTH JOURNAL
2024; 3(2): 48-64
Published online 11 13 2024 (https://todoxpublisher.com/index.php/fphj/index)
doi: https://doi.org/10.53797/fphj.v3i2.5.2024
e-ISSN 2821-3785

Develop and validate the event quality Structure scale to evaluate the event quality model

Xuexin Liu [1,2], Jaffry bin Zakaria [1*], Omar Firdaus bin Mohd Said [1], Wei Wang [2], Muhammad Nur Aizuddin bin Abdul Rahman [1]

[1] Faculty of Sport Science and Coaching, Sultan Idris Education University, 35900 Tanjong Malim, Perak, Malaysia
[2] Department of Physical Education, Guizhou Education University, China

***Corresponding Author**, E-mail: jaffry@fsskj.upsi.edu.my

Abstract: This study aims to develop and test a comprehensive set of indicators to assess the event quality in the context of sports tourism. Using the three-round Delphi method, participants evaluated an initial list of event quality model (EQM) indicators based on their experience. Through exploratory factor analysis (EFA) and confirmatory factor analysis (CFA) tests, multivariate normality, multicollinearity, construct reliability (Cronbach Alpha), convergent validity (e.g., principal component analysis), discriminant validity (e.g., Fornell-Larcker criterion) and construct validity were evaluated. Among the 32 candidate indicators, 20 were retained without multicollinearity problems, with strong internal consistency (0.963) and convergence validity (AVE > 0.5, CR > 0.7, SRW > 0.7). Discriminant validity (AVE > R2) and structural validity (P< 0.00, RMSEA = 0.035, CFI = 0.986, $\chi 2$/df =1.373). The results show that the newly developed scale of event quality model (EQM) meets the requirements of relevant indicators and can be applied to practical research.

Keywords: Develop, Validate, Scale, Event Quality Model

1. Introduction

Over the past few decades, the number of events worldwide has steadily increased due to increasing globalization and awareness of the role of events in cultural exchange, commerce and tourism (Menezes et al., 2020), especially as the event management industry continues to grow in importance worldwide. High-quality events attract more participants, create a meaningful experience, enhance the reputation of the host's location, and make quality assessment an essential part of event planning and management (Simasathiansophon, 2021). From local community festivals to international conferences, events are not only social and cultural gathering points, but also important economic drivers, affecting tourism, local economies and community participation. Traditionally, event quality has been assessed by a variety of metrics, including attendance, financial success, and basic attendee satisfaction surveys. While these metrics provide some insight into event performance, they fail to capture the full range of what constitutes a successful and high-quality event (Vahdati et al., 2021). As Reise et al. (2013) stated in their pioneering work on building development emphasizes, the development of reliable and effective measures is crucial. This measure includes evaluating the build of multidimensional. In the context of event quality, this means going beyond simple metrics and developing a model that integrates various elements that together contribute to the overall quality of the event.

Quality has always been a decisive factor in the success of an event, yet what constitutes "quality" is often subjective and multifaceted. Some researchers highlighted that event quality is a multidimensional construct, involving various aspects such as the physical environment, staff attitude, and participant experience (Simasathiansophon, 2021). Not only that, the event quality can include venue facilities, staff service, and entertainment quality, as well as overall ambiance (Kim & Tucker, 2016). The multidimensional nature of event quality also includes factors such as infrastructure, facilities, and the overall organization of the event, which influences not only the attendee experience but also the economic impact on the destination (Andersson et al., 2017). Although there is more and more literature on event quality, the empirical research on event quality lags behind. However, there are still significant gaps in how to effectively measure, evaluate, and improve the quality of events in different contexts (Getz, 2008; Ko et al., 2023). This study addresses the complexity and multi-dimensionality of event quality by introducing the Event Quality Model (EQM), a comprehensive framework designed to systematically evaluate and improve event quality. The development of the EQM follows a rigorous methodology, including an extensive review of relevant literature, interviews with industry experts, and the use of the Delphi method to refine the list of quality indicators. The Delphi method is particularly well suited to this type of research because it allows for the systematic collection and synthesis of expert opinion, ultimately leading to a consensus on the most important dimensions of event quality (Asmelash & Kumar, 2019). Thus, the EQM developed in this study integrates these insights by integrating multiple dimensions of event quality, including game quality, outcome quality, interaction quality, and physical environment quality. This gives a more comprehensive understanding of what constitutes a successful event, providing a powerful measurement tool that covers all dimensions of event quality.

2. Materials & Methods

Participants

Convenient sampling method and snowball sampling method were used in this study. Some scholars suggest that the sample size of CFA ranges from a minimum of 100 to more than 1000, depending on the number of variables and levels of communality (Mundfrom et al., 2005). In addition, it is generally recommended that the CFA sample size be 5 to 10 times the number of parameters to be estimated (Jackson, 2001). Therefore, considering the recovery rate and efficiency of the questionnaire and further ensuring sufficient sample size, this study finally selected random tourists (n=2*175=350) from the sports tourism destinations in two cities as the sample size.

Procedure

Indicator development procedure

A combination of qualitative (expert judgment) and quantitative (survey) methods were used to develop and validate EQM indicators. In order to collect data, based on an extensive literature review and under the guidance of the expert group, we selected and developed the most appropriate scale for questionnaire design, so as to measure the relevant structure of the event quality model. Specifically, we used the previously developed event quality model to evaluate event quality (Jin et al., 2013; Jeong & Kim, 2022), because their model has been widely utilized in the field of sports tourism e.g. (Quirante-Mañas et al., 2023; Jeong & Kim, 2022). For example, Jeong et al. (2019) and Jeong & Kim (2019) slightly modified and developed their study. This includes, 4 sub-structures such as game quality which consists of 4 items, interaction quality which consists of 3 items, outcome quality which consists of 3 items, and physical environment quality which consists of 3 items. In addition, based on the suggestions of experts, combined with the exchanges between sports tourism practitioners, tourists and relevant scholars, we developed the scale of event quality model and added 19 items. In the end, there were 32 potential indicators that were self-developed and collected from literature materials (**Table 1**). In order to reduce the initial list of indicators to a manageable number, the Delphi method was used to evaluate the indicators to fit the purpose.

Table1. Event Quality Questionnaire Items

Sub-Construct	Items code	Measurement Items	Sources
Game quality	GQ1	It was exciting to watch skillful players.	Jeong & Kim (2020)
	GQ2	Skill performance of players was excellent.	Jeong & Kim (2020)
	GQ3	Information about this event was easy to obtain	Jeong & Kim (2020)
	GQ4	Up-to-date information was available on events/teams	Jin, Lee, & Lee (2013)
	GQ5	The team offers excellent game that I want.	Developed
	GQ6	The operating hours and game schedule are convenient.	Developed
	GQ7	I think the game in Guizhou is a very organized sports tourism project.	Developed
	GQ8	I think the game in Guizhou make people curious and want to go to a sports tourism destination.	Developed
Interaction quality	IQ1	The demeanor of the staff was pleasant.	Jeong & Kim (2020)
	IQ2	I enjoyed being with the other spectators.	Jeong & Kim (2020)
	IQ3	Spectators followed the regulations	Jeong & Kim (2020)
	IQ4	I was very impressed with the staff of the race organization in Guizhou.	Developed
	IQ5	I think the event in Guizhou can help me to build friendship.	Developed
	IQ6	I think the event in Guizhou can help me to interact with others.	Developed
	IQ7	I think the event in Guizhou can help me to meet new people with similar interests.	Developed
	IQ8	My interaction with other fans is favorable.	Developed
Outcome quality	OQ1	I view the outcome of this event favorably.	Jeong & Kim (2020)
	OQ2	I enjoyed the social interaction at this event.	Jeong & Kim (2020)
	OQ3	I spent quality time with my friend/family.	Jeong & Kim (2020)
	OQ4	My social interaction in the game event is very positive.	Developed
	OQ5	I think the equipment available for sale in Guizhou sports destinations is of good quality.	Developed
	OQ6	I think shopping in Guizhou tourist destinations is cheaper than other places.	Developed
	OQ7	I think Guizhou tourism program offers a wealth of outdoor activities.	Developed
	OQ8	I think Guizhou sports tourism destination can guarantee the safety of people and property.	Developed
Physical environment quality	PEQ1	The facility was clean and well maintained.	Jeong & Kim (2020)
	PEQ2	I am impressed with the facility design.	Jeong & Kim (2020)
	PEQ3	I believe the facility is safe.	Jeong & Kim (2020)
	PEQ4	I like the ambience of the facility.	Developed
	PEQ5	I like the event venue.	Developed
	PEQ6	I like the well-conserved environment (e.g., nature quality, local culture, ecological environment).	Developed
	PEQ7	The design of the facility is excellent.	Developed
	PEQ8	The signage or scoreboard of the facility is excellent.	Developed

Note, GQ=Game quality; OQ=Outcome quality; IQ=Interaction quality; PEQ=Physical environment quality.

Item Determination for the Delphi Method

The Delphi method is an area where potential researcher bias can be particularly high, and panelists must select expert qualifications, be able to communicate about the topic, and be willing to participate in research that researchers may not be in the best position to evaluate, especially in advance (Avella, 2016). Sterling et al. (2023b) suggested the Delphi method to conduct research, which states eight experts responded during the testing phase. He believes that a project needs six experts to pass the threshold. In addition, Wang et al. (2021) invited three experts to carry out verification when developing and verifying market demand factors related to sports tourism. Therefore, we employ eight experts as panel members. The members of the expert group are mainly based on three aspects. Firstly, researchers who have been engaged in sports tourism research for a long time are very familiar with the relevant situation of sports tourism in Guizhou. Second, government managers have been engaged in sports tourism management for a long time, participating in the formulation of sports tourism activities related policies; Third, managers who have been engaged in sports tourism management for a long time have sports tourism companies that specialize in organizing tourists to participate in sports tourism activities of various sports events.

Meanwhile, Wang et al. (2021), Fetscherin & Stephano (2016), and Rejón-Guardia et al. (2020), in the stage of descriptive statistics, considered it acceptable that the average score of all items was greater than 3.0 (that is, the midpoint of a similar 5-point scale). However, Kim et al. (2015) and Asmelash & Kumar (2019) considers that items with an average score of more than 3.5 points per item are acceptable. To enhance the value of this study, we adopted an average of more than 3.5 per item as acceptable.

Draft Development of Instrument Item

In order to reduce the initial list of indicators to a manageable number without compromising coverage of key issues, the author selected three rounds of the improved Delphi method with reference to the research methodology of Donohoe & Needham (2009), and modified the technique according to the applicability of the implementation of this study. According to the objective situation of index screening and expert suggestions of questionnaire retrieval, the items of each index are adjusted. The first round of Delphi method results found that GQ6, IQ4, OQ7 and PEQ8 had an average score of less than 3.5, which were not considered indicators. In the second round of Delphi method, the retained indicators were reassigned to eight experts to reduce the possible degree of subjectivity (Choi & Sirakaya, 2005). After analyzing the valid questionnaires received, 25 event quality indicators were retained and some of them were modified. Among them, indicators such as GQ8, OQ4, and PEQ7 that were lower than 3.5 were eliminated. In the third round of the Delphi method, the indicators of PEQ5 and GQ7 were removed. Finally, after three rounds, the indicators of event quality were reduced to 23. Before removal, we showed the items to eight expert judges to ensure that they would not result in any loss of content validity (indicator reliability), and they concluded that the items could be removed.

Purification of the indicator development

A pilot study is a small feasibility study designed to test aspects of an approach to planning a larger, more rigorous or confirmatory investigation (Lowe, 2019). Donald et al. (2014) suggest a pilot study with 25-100 participants. Alternatively, Xue & Zhang (2020) conducted a preliminary study with 30 selected tourists in the pilot to determine the facial validity and reliability of these measures. However, according to McCoach et al. (2013), techniques that can be used to measure the validity and reliability of instruments include exploratory factor analysis (EFA) and the reliability analysis. In order to conduct an EFA on a research instrument, several conditions need to be met (Kyriazos, 2018). Firstly, the factor analysis can only be conducted on items with interval or ordinal scales. Secondly, 50 observations are recommended as a minimum size for EFA, but also a sample size of 100 or more is preferred. Finally, the data should be normally distributed. In addition, the ratio of sample to variable, which is a ratio of at least 5:1 is recommended. For example, 5 observations for each variable are analyzed. 10:1 is also acceptable (Kyriazos, 2018).

In summary, based on the exploration of sample size in the literature material, 150 people will be involved in this phase to better conduct the pilot study. The sample size of the pilot study is selected by proportional sampling technique, but the actual study is not involved. The five-point Likert scale was mainly used for the test. Respondents were asked to check the relevance, format and wording of the items and the demographic information of tourists in the questionnaire by answering the scale. Behavioral intent and loyalty to sports tourism destinations. Based on the feedback received in the pilot study, additional modifications and improvements were made to improve its content validity. In addition, the researchers also attempted to control the irrelevance of respondents who fit the characteristics of the following conditions, as described below.

1) Sports tourism teachers in colleges and universities in Guizhou;

2) Students majoring in sports tourism in Guizhou universities or volunteers participating in sports tourism events.

3) Sports tourists who travel to sports tourism destinations in Guizhou.

Ultimately, the translation of the questionnaire in the forward and reverse directions were done by four different bilingual translators. The questionnaire in this study includes the structural scale of the EQM and 6 demographic items. Among the eight demographic indicators, information on gender, age, the educational background, income, occupation and types of sports tourism activities were included. Not only that, the questionnaire includes English and Chinese versions. In this study, a 5-point Likert scale was used to complete the evaluation of potential indicators (the scores of "very important", "important", "normal important", "not important" and "very unimportant" were 5, 4, 3, 2 and 1, respectively), and the correlation was determined according to the scores.

Data collection and Statistical analysis

The data analysis is carried out in two stages. In the first phase, exploratory factor analysis (EFA) is used to obtain dimensions of event quality. In the second stage, confirmatory factor analysis (CFA) is performed for EQM.

3. Results and Discussion

First, the Pearson correlation coefficient is used to examine the correlation between the four sub-dimensions of event quality. This is because correlation matrices are essential for understanding the strength and direction of linear relationships between variables. They are essential for further analysis such as regression models (Daoud, 2017). It can be seen from

Table 2 that the correlation among all variables reaches 0.01 level, indicating that GQ, IQ, OQ and PEQ are strongly correlated. It shows that GQ, IQ, OQ and PEQ are reliable predictors in sports tourism research and can be used in regression models to effectively predict tourism trends and results.

Table 2. Event quality correlation of the four sub-dimensions

	GQ	IQ	OQ	PEQ
GQ	1			
IQ	.584**	1		
OQ	.562**	.571**	1	
PEQ	.597**	.582**	.572**	1

Note, ** Correlation is significant at the 0.01 level (2-tailed). GQ=game quality; IQ= interaction quality; OQ=outcome quality; PEQ=physical environment quality.

Secondly, collinearity statistics are critical for identifying multicollinearity problems, which can distort regression coefficients and undermine the validity of the model (Daoud, 2017). Variance inflation factor (VIF) values greater than 10 or tolerance values less than 0.1 usually indicate problems with multicollinearity (Daoud, 2017). Table 3 shows collinearity statistics for the various predictors associated with sports tourism, expressed by their tolerance and VIF values. The results showed that the VIF values of all predictors in the model are acceptable and all were below the threshold of 10, indicating that there was no serious multicollinearity problem among the predictors. This means that the predictors are independent enough from each other for the regression model to be robust and reliable (Sovia, 2021). Specifically, the VIF values of the predictors GQ1 to PEQ6 range from 1.204 to 3.395, and the tolerance ranges from 0.295 to 0.831, further supporting the absence of multicollinearity (Yuan, 2020).

Thirdly, the normality of each item is examined to evaluate the appropriate extraction method for factor analysis. Table 3 shows the measurement scale of each variable from 1 to 5. The mean values ranged from 1.9 to 3.58, indicating that the average responses to these questions were generally above the midpoint. This shows that the responses to most questions leaned toward positive or higher scores. The normality test using skewness and kurtosis was performed using the z test, and their values in this study did not exceed the recommended ±1.96 (Mishra et al., 2019). The results show that the normal distribution of values is flat with few extreme values (Mishra et al., 2019; ORCAN, 2020).

Table 3. Event quality result of linear analysis, Skewness and Kurtosis peaks

Items	Tolerance	VIF	Mean	Skewness	Kurtosis
GQ1	0.381	2.623	3.04	0.123	-1.031
GQ2	0.388	2.579	2.97	0.112	-0.911
GQ3	0.372	2.686	3.1	0.052	-0.818
GQ4	0.441	2.269	3.25	-0.373	-1.113
GQ5	0.469	2.133	3.01	0.064	-0.91
IQ1	0.329	3.037	3.13	0.015	-1.038
IQ2	0.295	3.395	3.15	-0.091	-0.987
IQ3	0.414	2.415	3.17	-0.271	-0.639
IQ5	0.449	2.225	3.24	0.057	-1.172
IQ6	0.359	2.783	3.11	-0.105	-0.872
IQ7	0.437	2.286	3.13	-0.116	-0.926
IQ8	0.437	2.291	3.39	-0.449	-0.806

continued

OQ1	0.39	2.562	3.13	0.105	-1.146
OQ2	0.462	2.162	3.17	-0.011	-1.142
OQ3	0.396	2.526	3.34	-0.135	-1.096
OQ5	0.408	2.449	3.2	0.015	-1.116
OQ6	0.382	2.62	3.09	-0.046	-1.068
OQ8	0.831	1.204	3.58	-0.611	-0.155
PEQ1	0.376	2.656	3.29	-0.132	-0.945
PEQ2	0.319	3.139	3.08	-0.084	-1.117
PEQ3	0.417	2.397	3.49	-0.256	-0.949
PEQ4	0.381	2.628	3.25	-0.157	-0.994
PEQ6	0.827	1.209	1.9	0.671	-0.448

Note, GQ=game quality; IQ= interaction quality; OQ=outcome quality; PEQ=physical environment quality.

Fourthly, the structural validity process is completed through principal component analysis (PCA) with varimax rotation and unrestricted number of factors for factor analysis. The results showed in **Table 4** that the KMO value is 0.881, which is larger than 0.6 and close to 1.0 (Schreiber, 2021), indicating that the sample is very suitable for factor analysis. The Bartlett test here shows that the Chi-square approximation is 1556.012 and the significance level is 0.000 ($p < 0.05$), confirming that the data is suitable for factor analysis because the variables are correlated (Lloret et al., 2017). Overall, it can be seen that the event quality substructures exceed the recommended values of 0.6 and Barttlet's Test of Spehericity to achieve statistical significance (Shrestha, 2021), supporting the factorial property of the correlation matrix.

Table 4. KMO and Bartlett's Test (first)

Kaiser-Meyer-Olkin Measure of Sampling Adequacy.		0.881
Bartlett's Test of Sphericity	Approx. Chi-Square	1556.011
	df	253
	Sig.	.000

Fifth, principal component analysis showed that the six components with eigenvalues greater than 1 explained 38.645%, 7.882%, 6.710%, 5.972%, 5.012% and 4.920% of the variance, respectively. Among them, IQ8, GQ4, and OQ8 values were above 0.9. According to Dai et al. (2017), abnormally high factor loads may be a sign of poorly regulated factor models. It is suggested to remove these indexes to avoid distortion of factor analysis results. In addition, in factor analysis, GQ3 has a commonality of less than 0.5 and is therefore excluded from the analysis (Shrestha, 2021;Jain & Raj, 2013). However, some researchers argue that while higher communities are generally more desirable, communities below 0.5 should not be automatically discarded, especially in exploratory settings where the goal is to identify underlying factors rather than confirm pre-established models (Auerswald & Moshagen, 2019; Hefetz & Liberman, 2017). Therefore, IQ8, GQ4, OQ8 were removed and GQ3 was retained.

After deleting IQ8, GQ4, and OQ8, PCA analysis was performed again. The results in Table 5 shows that the KMO value is 0.900, higher than the threshold value of 0.6, and the Bartlett speed test is significant. The KMO value increased by 0.019 (KMO=0.881) from the first time. Principal component analysis shows that the difference of the four components with eigenvalues greater than 1 is explained by 44.427%, 52.267%, 59.079% and 64.816%, respectively.

Table 5. KMO and Bartlett's Test (second)

Kaiser-Meyer-Olkin Measure of Sampling Adequacy.		.900
Bartlett's Test of Sphericity	Approx. Chi-Square	1433.985
	df	190
	Sig.	.000

According to previous studies (Jeong & Kim, 2020; Jin et al., 2013; Y. Ko et al., 2011) divided event quality into four sub-dimensions. They are GQ, IQ, OQ and PEQ. According to Lovik et al.(2018), in factor analysis, it is believed that the combination stage of different structural contents can rely on prior classification .Marsh et al. (2020) believes that different structures can be modeled within a single framework in EFA to promote the combination of different structural contents based on previous classifications. Therefore, the event quality is classified according to the definition of the four dimensions. Finally, PCA was performed on 20 sub-items of event quality, namely GQ (including 4 items), IQ (including 6 items), OQ (including 5 items) and PEQ (including 5 items).

Assessing reliability and validity

After examining EFA reliability for all structures, Table 6 showed strong value of factor loading above 0.70 (Shrestha, 2021). The representational event quality constructs consist of four constructs, such as GQ, IQ, OQ, and PEQ. The reliability of EQM was evaluated using the Cronbach Alpha method.

Firstly, the Cronbach's Alpha of the EQM has a value of 0.932, indicating strong internal consistency. Secondly, preliminary results show that every attribute in EQM has a good Cronbach Alpha (α) value above 0.7 (Hair et al., 2019). CR refers to the reliability and internal consistency of the underlying structure. The value of CR was higher than 0.5 (Hair et al., 2019). AVE values greater than 0.5 (Hair et al., 2019) represent the average percentage of variation explained by a measurement item of a potential structure. **Table 7** shows that the AVE values of all constructs are higher than the squared correlation with other constructs, which further supports the point (Fornell & Larcker, 1981). Therefore, the research results show that the developed instrument has good and high reliability.

Table 6. Reliability and validity test results of EQM

Dimension	**Cronbach's alpha**	**Composite reliability**	**Average variance extracted (AVE)**
GQ	0.874	0.875	0.726
IQ	0.917	0.918	0.706
OQ	0.904	0.904	0.722
PEQ	0.95	0.95	0.833

Table 7. The Results of Fornell-Larcker Criterion

Dimension	**GQ**	**IQ**	**OQ**	**PEQ**
GQ	0.852			
IQ	0.677	0.84		
OQ	0.6	0.715	0.85	
PEQ	0.369	0.478	0.445	0.913

Confirmatory Factory Analysis

In the context of sports tourism, CFA helps researchers understand and confirm the factors that influence tourists' behaviors, and loyalty related to sports events and destinations.

(1) First-Order for Event Quality Model

Figure 1 shows the four sub-factors of ETQ and the first-order CFA of their underlying terms. Generally accepted thresholds for interpreting the strength of correlation coefficients are as follows: small (r = 0.10 to 0.29), medium (r = 0.30 to 0.49), and large (r = 0.50 to 1.0) (Schober et al., 2018). Our research shows that the correlation between the substructures of EQ ranges from 0.4-0.782, between medium and large, which is completely acceptable. Based on the results of the measurement model, the model fit assessment results indicate a $\chi2$=226.447, df=164, $\chi2$/df =1.381, CFI=0.986, and RMSEA=0.036. These indices suggest that the model has a good fit, as the chi-square/df ratio is well within acceptable limits, the CFI is very high, indicating excellent fit, and the RMSEA is below the threshold of 0.06, further confirming a good model fit (Asmelash&Kumar, 2019).

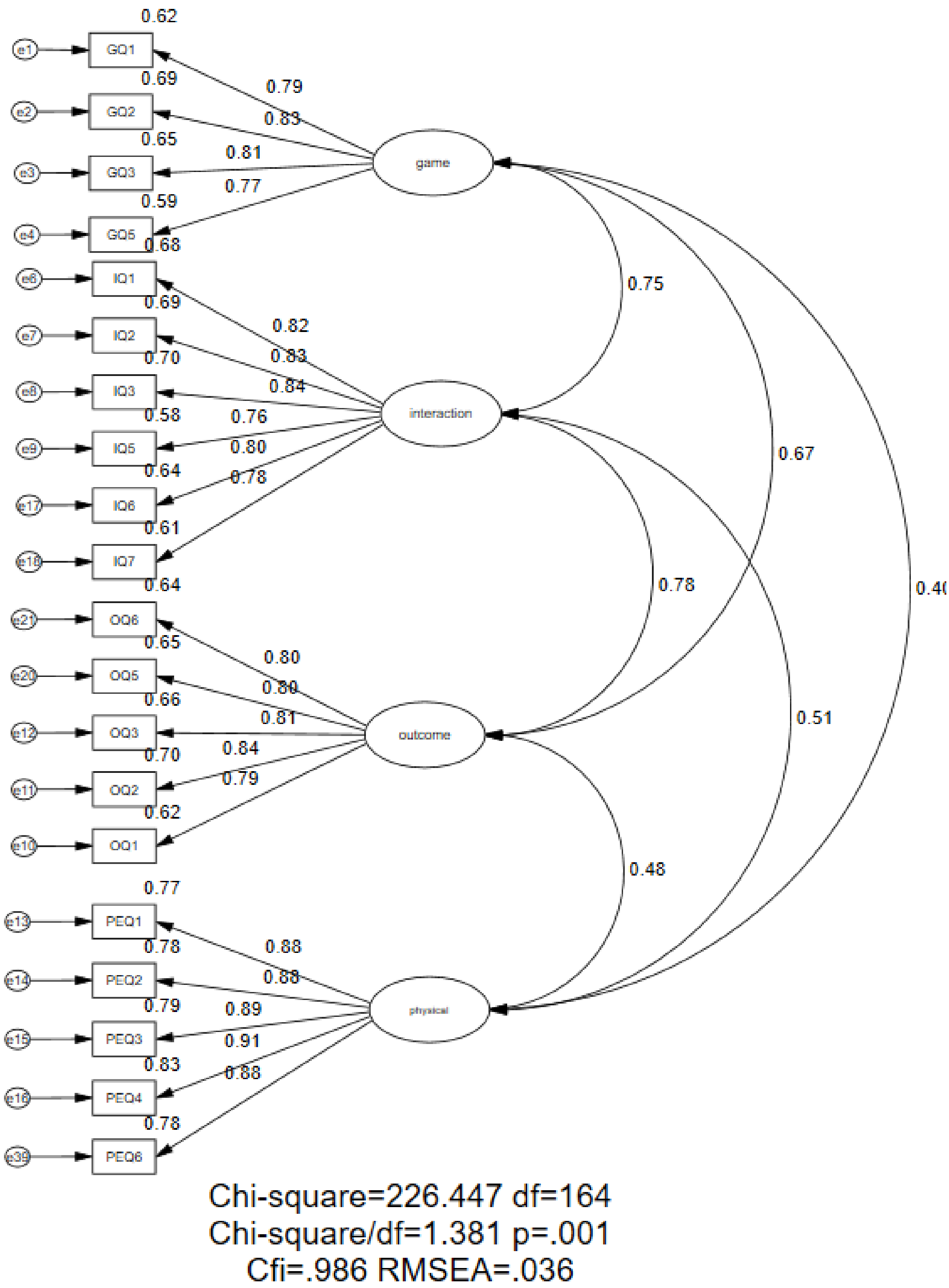

Figure 1. First-order CFA Model of event quality

(2) Second-Order for Event Quality Model

The second-order CFA model was used to analyze the event quality, and the results were significant. **Table 8** shows that the factor loading of each sub-construct on the higher order construct ranges from 0.541 to 0.936. All paths show high significance at p values < 0.000, indicating a robust statistical relationship between substructures and higher-order structures (Jeong et al., 2019b).

Table 8. Regression Weight for Each Construct and its Path

Sub-Constructs	Path	Second-order Constructs	Estimate loading	P
GQ	<---	EQ	0.796	***
IQ	<---	EQ	0.936	***
OQ	<---	EQ	0.84	***
PEQ	<---	EQ	0.541	***

Note, ***indicate highly significance at p-value<0.000GQ=game quality; IQ= interaction quality; OQ=outcome quality; PEQ=physical environment quality.

In **Figure 2**, the second-order for event quality model fit assessment results indicate χ^2=227.919, df=166, χ^2/df =1.373, CFI=0.986, and RMSEA=0.035. These indices suggest that the model has a good fit, as the chi-square/df ratio is well within acceptable limits. The CFI is very high, indicating excellent fit, and the RMSEA is below the threshold of 0.06, further confirming a good model fit.

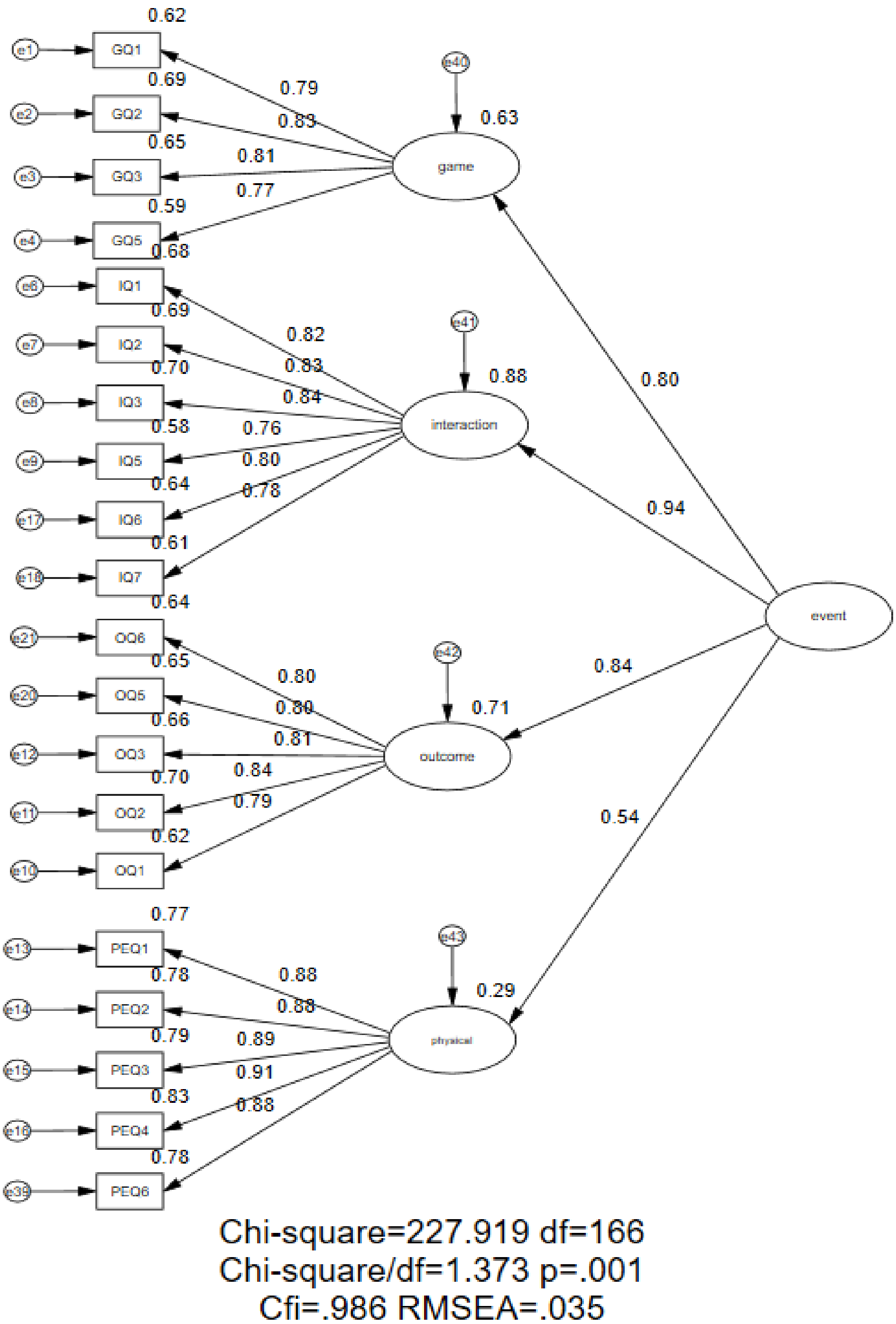

Figure 2. **Second-order CFA Model of Event Quality**

4. Conclusions

This paper aims to further develop and validate a comprehensive event quality model (EQM) on the basis of previous studies, which can be used as a tool to evaluate and improve event quality in different contexts. EQM was developed to utilize a broad multi-dimensional framework that incorporates input from a wide range of stakeholders from event participants, organizers and experts, thus ensuring the robustness of the model (Churchill, 1979). The importance of well-constructed and empirically validated models for assessing event quality is increasingly recognized, but these models are limited in both research and practice (Getz, 2008; Ko et al., 2023). With this in mind, this paper has compiled a comprehensive list of event quality indicators through an extensive literature review and interviews with key informants in the event management industry. The initial list was refined using the three-round Delphi method, which helped reduce the number of event quality indicators from the initial 32 to 23. In the purification phase, the number of indicators was further reduced from 23 to 20, and the final validation phase was concluded with 20 event quality indicators retained.

This study makes some notable contributions to the development and validation of the Event Quality Model (EQM). Firstly, it involves the participation of various stakeholders, including event participants, organizers, industry experts and local authorities, ensuring a holistic perspective on the quality of the event. The involvement of multiple stakeholders, including university faculty, event managers, and visitors, enhances the applicability and validity of the model in different contexts (Lee et al., 2011). Secondly, unlike many previous studies, this study utilizes a multi-dimensional framework to develop EQM. This approach helps ensure that the model takes into account various elements of event quality, including game quality, outcome quality, interaction quality, and physical environment quality. This provides a comprehensive instrument for assessing the overall quality of the event. By including these different factors, EQM provides event organisers and managers with insights into the attendee-experience, enabling them to make informed choices in event planning and management, ultimately ensuring that the events they organise are of high quality and meet the expectations of all stakeholder.

Conflicts of interest: The authors have no conflict of interest

Acknowledgment: The authors would like to thank Guizhou Education University, China and Sultan Idris Education University (UPSI), Malaysia, for their support and collaboration in facilitating this research. Their contributions and resources were invaluable in the successful completion of this study.

References

Andersson, T. D., Armbrecht, J., & Lundberg, E. (2017). Linking event quality to economic impact. Journal of Vacation Marketing, 23(2), 114–132. https://doi.org/10.1177/1356766715615913

Asmelash, A. G., & Kumar, S. (2019). Assessing progress of tourism sustainability: Developing and validating sustainability indicators. Tourism Management, 71, 67–83. https://doi.org/10.1016/j.tourman.2018.09.020

Auerswald, M., & Moshagen, M. (2019). How to determine the number of factors to retain in exploratory factor analysis: A comparison of extraction methods under realistic conditions. Psychological Methods, 24(4), 468–491. https://doi.org/10.1037/met0000200

Bollen, K. A. (2020). When Good Loadings Go Bad: Robustness in Factor Analysis. Structural Equation Modeling: A Multidisciplinary Journal, 27(4), 515–524. https://doi.org/10.1080/10705511.2019.1691005

Churchill, G. A. (1979). A Paradigm for Developing Better Measures of Marketing Constructs. Journal of Marketing Research, 16(1), 64–73. https://doi.org/10.1177/002224377901600110

Dai, C., Lu, K., & Xiu, D. (2017). Knowing Factors or Factor Loadings, or Neither? Evaluating Estimators of Large Covariance Matrices with Noisy and Asynchronous Data. SSRN Electronic Journal. https://doi.org/10.2139/ssrn.2920693

Daoud, J. I. (2017). Multicollinearity and Regression Analysis. Journal of Physics: Conference Series, 949, 012009. https://doi.org/10.1088/1742-6596/949/1/012009

Donald, F., Kilpatrick, K., Reid, K., Carter, N., Martin-Misener, R., Bryant-Lukosius, D., Harbman, P., Kaasalainen, S., Marshall, D. A., Charbonneau-Smith, R., Donald, E. E., Lloyd, M., Wickson-Griffiths, A., Yost, J., Baxter, P., Sangster-Gormley, E., Hubley, P., Laflamme, C., Campbell–Yeo, M., … DiCenso, A. (2014). A Systematic Review of the Cost-Effectiveness of Nurse Practitioners and Clinical Nurse Specialists: What Is the Quality of the Evidence? Nursing Research and Practice, 2014, 1–28. https://doi.org/10.1155/2014/896587

Donohoe, H. M., & Needham, R. D. (2009). Moving best practice forward: Delphi characteristics, advantages, potential problems, and solutions. International Journal of Tourism Research, 11(5), 415–437. https://doi.org/10.1002/jtr.709

Fetscherin, M., & Stephano, R.-M. (2016). The medical tourism index: Scale development and validation. Tourism Management, 52, 539–556. https://doi.org/10.1016/j.tourman.2015.08.010

Fornell, C., & Larcker, D. F. (1981). Evaluating Structural Equation Models with Unobservable Variables and Measurement Error. Journal of Marketing Research, 18(1), 39–50. https://doi.org/10.1177/002224378101800104

Getz, D. (2008). Event tourism: Definition, evolution, and research. Tourism Management, 29(3), 403–428. https://doi.org/10.1016/j.tourman.2007.07.017

Hair, J. F., Risher, J. J., Sarstedt, M., & Ringle, C. M. (2019). When to use and how to report the results of PLS-SEM. European Business Review, 31(1), 2–24. https://doi.org/10.1108/EBR-11-2018-0203

Hefetz, A., & Liberman, G. (2017). The factor analysis procedure for exploration: a short guide with examples / El análisis factorial exploratorio: una guía breve con ejemplos. Cultura y Educación, 29(3), 526–562. https://doi.org/10.1080/11356405.2017.1365425

Jackson, D. L. (2001). Sample Size and Number of Parameter Estimates in Maximum Likelihood Confirmatory Factor Analysis: A Monte Carlo Investigation. Structural Equation Modeling: A Multidisciplinary Journal, 8(2), 205–223. https://doi.org/10.1207/S15328007SEM0802_3

Jain, V., & Raj, T. (2013). Evaluating the variables affecting flexibility in FMS by exploratory and confirmatory factor analysis. Global Journal of Flexible Systems Management, 14(4), 181–193. https://doi.org/10.1007/s40171-013-0042-9

Jeong, Y., Kim, S.-K., & Yu, J.-G. (2019a). Determinants of Behavioral Intentions in the Context of Sport Tourism with the Aim of Sustaining Sporting Destinations. Sustainability, 11(11), 3073. https://doi.org/10.3390/su11113073

Jeong, Y., Kim, S.-K., & Yu, J.-G. (2019b). Sustaining Sporting Destinations through Improving Tourists' Mental and Physical Health in the Tourism Environment: The Case of Korea. International Journal of Environmental Research and Public Health, 17(1), 122. https://doi.org/10.3390/ijerph17010122

Jeong, Y., & Kim, S. (2020). A study of event quality, destination image, perceived value, tourist satisfaction, and destination loyalty among sport tourists. ASIA PACIFIC JOURNAL OF MARKETING AND LOGISTICS, 32(4), 940–960. https://doi.org/10.1108/APJML-02-2019-0101

Jeong, Y., & Kim, S. (2022). The role of a mega-sporting event in attracting domestic tourists: the case of Seoul. Asia Pacific Journal of Marketing and Logistics, 34(2), 306–321. https://doi.org/10.1108/APJML-10-2020-0754

Jin, N. (Paul), Lee, H., & Lee, S. (2013). Event Quality, Perceived Value, Destination Image, and Behavioral Intention of Sports Events: The Case of the IAAF World Championship, Daegu, 2011. Asia Pacific Journal of Tourism Research, 18(8), 849–864. https://doi.org/10.1080/10941665.2012.711336

Kim, K., & Tucker, E. D. (2016). Assessing and segmenting entertainment quality variables and satisfaction of live event attendees: A cluster analysis examination. Journal of Convention & Event Tourism, 17(2), 112–128. https://doi.org/10.1080/15470148.2015.1101035

Kim, W., Jun, H. M., Walker, M., & Drane, D. (2015). Evaluating the perceived social impacts of hosting large-scale sport tourism events: Scale development and validation. Tourism Management, 48, 21–32. https://doi.org/10.1016/j.tourman.2014.10.015

Ko, Y. J., Kwon, H. H., Kim, T., Park, C., & Song, K. (2023). Assessment of Event Quality in Major Spectator Sports: Single-Item Measures. Journal of Global Sport Management, 8(4), 1008–1024. https://doi.org/10.1080/24704067.2021.2001353

Ko, Y., Zhang, J., Cattani, K., & Pastore, D. (2011). Assessment of event quality in major spectator sports. Managing Service Quality: An International Journal, 21(3), 304–322. https://doi.org/10.1108/09604521111127983

Kyriazos, T. A. (2018). Applied Psychometrics: Sample Size and Sample Power Considerations in Factor Analysis (EFA, CFA) and SEM in General. Psychology, 09(08), 2207–2230. https://doi.org/10.4236/psych.2018.98126

Lee, J.-S., Lee, C.-K., & Choi, Y. (2011). Examining the Role of Emotional and Functional Values in Festival Evaluation. Journal of Travel Research, 50(6), 685–696. https://doi.org/10.1177/0047287510385465

Lloret, S., Ferreres, A., Hernández, A., & Tomás, I. (2017). El análisis factorial exploratorio de los ítems: análisis guiado según los datos empíricos y el software. Anales de Psicología, 33(2), 417. https://doi.org/10.6018/analesps.33.2.270211

Lovik, A., Nassiri, V., Verbeke, G., & Molenberghs, G. (2018). Combining Factors from Different Factor Analyses Based on Factor Congruence (pp. 211–219). https://doi.org/10.1007/978-3-319-77249-3_18

Lowe, N. K. (2019). What Is a Pilot Study? Journal of Obstetric, Gynecologic & Neonatal Nursing, 48(2), 117–118. https://doi.org/10.1016/j.jogn.2019.01.005

Marsh, H. W., Guo, J., Dicke, T., Parker, P. D., & Craven, R. G. (2020). Confirmatory Factor Analysis (CFA), Exploratory Structural Equation Modeling (ESEM), and Set-ESEM: Optimal Balance Between Goodness of Fit and Parsimony. Multivariate Behavioral Research, 55(1), 102–119. https://doi.org/10.1080/00273171.2019.1602503

McCoach, D. B., Gable, R. K., & Madura, J. P. (2013). Instrument development in the affective domain. New York, NY: Springer.

Menezes, A., Lima, R. M., Aquere, A. L., & Amorim, M. (2020). An Adaptation of SERVQUAL for Events Evaluation: An Environmental Sustainability Addon. Sustainability, 12(18), 7408. https://doi.org/10.3390/su12187408

Mishra, P., Pandey, C., Singh, U., Gupta, A., Sahu, C., & Keshri, A. (2019). Descriptive statistics and normality tests for statistical data. Annals of Cardiac Anaesthesia, 22(1), 67. https://doi.org/10.4103/aca.ACA_157_18

Mundfrom, D. J., Shaw, D. G., & Ke, T. L. (2005). Minimum Sample Size Recommendations for Conducting Factor Analyses. International Journal of Testing, 5(2), 159–168. https://doi.org/10.1207/s15327574ijt0502_4

ORCAN, F. (2020). Parametric or Non-parametric: Skewness to Test Normality for Mean Comparison. International Journal of Assessment Tools in Education, 7(2), 255–265. https://doi.org/10.21449/ijate.656077

Park, J.-A., Sung, J.-M., Son, J.-M., Na, K., & Kim, S.-K. (2019). Athletes' brand equity, spectator satisfaction, and behavioral intentions. ASIA PACIFIC JOURNAL OF MARKETING AND LOGISTICS, 31(2), 541–558. https://doi.org/10.1108/APJML-05-2018-0176

Quirante-Mañas, M., Fernández-Martínez, A., Nuviala, A., & Cabello-Manrique, D. (2023). Event Quality: The Intention to Take Part in a Popular Race Again. Apunts. Educacion Fisica y Deportes, 151, 70–78. https://doi.org/10.5672/apunts.2014-0983.es.(2023/1).151.07

R Avella, J. (2016). Delphi Panels: Research Design, Procedures, Advantages, and Challenges. International Journal of Doctoral Studies, 11, 305–321. https://doi.org/10.28945/3561

Reise, S. P., Bonifay, W. E., & Haviland, M. G. (2013). Scoring and Modeling Psychological Measures in the Presence of Multidimensionality. Journal of Personality Assessment, 95(2), 129–140. https://doi.org/10.1080/00223891.2012.725437

Rejón-Guardia, F., Alemany-Hormaeche, M., & García-Sastre, M. A. (2020). Ibiza dances to the rhythm of pedals: The motivations of mountain biking tourists competing in sporting events. Tourism Management Perspectives, 36, 100750. https://doi.org/10.1016/j.tmp.2020.100750

Schober, P., Boer, C., & Schwarte, L. A. (2018). Correlation Coefficients: Appropriate Use and Interpretation. Anesthesia & Analgesia, 126(5), 1763–1768. https://doi.org/10.1213/ANE.0000000000002864

Schreiber, J. B. (2021). Issues and recommendations for exploratory factor analysis and principal component analysis. Research in Social and Administrative Pharmacy, 17(5), 1004–1011. https://doi.org/10.1016/j.sapharm.2020.07.027

Shrestha, N. (2021). Factor Analysis as a Tool for Survey Analysis. American Journal of Applied Mathematics and Statistics, 9(1), 4–11. https://doi.org/10.12691/ajams-9-1-2

Simasathiansophon, N. (2021). Running event quality: conceptualization and measurement scale. E3S Web of Conferences, 244, 11017. https://doi.org/10.1051/e3sconf/202124411017

Sovia, R. (2021). Prediction Tourist Visits With Multiple Linear Regressions in Artificial Neural Networks. Turkish Journal of Computer and Mathematics Education (TURCOMAT), 12(3), 1492–1501. https://doi.org/10.17762/turcomat.v12i3.949

Sterling, S., Plonsky, L., Larsson, T., Kytö, M., & Yaw, K. (2023). Introducing and illustrating the Delphi method for applied linguistics research. Research Methods in Applied Linguistics, 2(1), 100040. https://doi.org/10.1016/j.rmal.2022.100040

Vahdati, S., Fathalla, S., Lange, C., Behrend, A., Say, A., Say, Z., & Auer, S. (2021). A comprehensive quality assessment framework for scientific events. Scientometrics, 126(1), 641–682. https://doi.org/10.1007/s11192-020-03758-1

Wang, F. J., Sum, K. W. R., Wong, C. O. J., & Cheng, C. F. (2021). The development and validation of an instrument for measuring market demand factors associated with sports training tourism for sports teams. Asia Pacific Journal of Tourism Research, 26(1), 30–43. https://doi.org/10.1080/10941665.2020.1805477

Xue, L., & Zhang, Y. (2020). The effect of distance on tourist behavior: A study based on social media data. Annals of Tourism Research, 82, 102916. https://doi.org/10.1016/j.annals.2020.102916

Yuan, F.-C. (2020). Intelligent forecasting of inbound tourist arrivals by social networking analysis. Physica A: Statistical Mechanics and Its Applications, 558, 124944. https://doi.org/10.1016/j.physa.2020.124944