A first lecture/revision in exploratory and confirmatory factor analyses for tourism students, researchers, and practitioners

Yulin Liu

Queensland University of Technology, Brisbane, Australia

ARTICLE HISTORY Received 24 October 2020; Accepted 24 November 2020

Introduction

Factor analysis (FA) is a prevalent quantitative technique for tourism researchers to measure “unstructured concepts” involving “values, attitudes, motivations, risks, satisfaction, and beliefs” (Turner & Vu, 2014, p. 183). For example, constructs like destination image and travel intention can be measured with instruments like site-inspection checklist of destination characteristics and traveller questionnaire in rating attitudinal and/or behavioural statements. Blind use of the technique, however, causes a concern about methodological rigour as today’s user-friendly statistical software requires analysts to have little understanding of FA theory or methodology (Mikulić & Ryan, 2018; Stewart, 1981). From a research educator’s perspective, this note intends to foster the essential, conceptual knowledge of FA among tourism students, researchers, and practitioners.

Methodology

The author observes that many FA learners focus on numerous technical decisions in conducting FA, guided by often inconsistent recommendations scattered in the literature (Schmitt, 2011). Although conscientious learners of FA should develop both procedural skills and conceptual knowledge, the latter is more transferable but often overlooked (Rittle-Johnson & Alibali, 1999). This note emphasizes the conceptual comparison in measurement model specification and research application between exploratory and confirmatory factor analyses rather than their procedural details. Notwithstanding, this note cites a curation of literature on factor analytic methods for curious readers to acquire further theoretical knowledge and practical skills. This short communication aims to present a “small but mighty” (Wen et al., 2020) “narrative synthesis” (Furunes, 2019) of must-reads as a first lecture (or revision) in FA for novice learners (or seasoned researchers) in the tourism field.

Results

FA is a collection of multivariate statistical procedures for explaining a set of interrelated observable variables (a.k.a. measures, indicators, or items) by a smaller set of latent variables (a.k.a. factors, constructs, or scales). The rationale is that measures being strongly interrelated are likely to manifest the same factor. Factors indicate important qualities in data by providing a dimensional structure (Stewart, 1981). Exploratory factor analysis (EFA) and confirmatory factor analysis (CFA) are two general types of FA in tourism literature. EFA uncovers factor(s) underlying a number of measures in data, whereas CFA tests whether specified relationships of factor(s) and measure fit data.

CONTACT Yulin Liu y68.liu@qut.edu.au Queensland University of Technology, Brisbane 4001, Australia

© 2020 Informa UK Limited, trading as Taylor & Francis Group
Both EFA and CFA are conventionally based on the common factor model. It proposes that observed measures are influenced partially by an underlying common factor(s) and partially by their unique measurement errors. Figure 1(a,b) illustrate the typical factorial structures of EFA and CFA models of five indicators influenced by two constructs and individual measurement errors. The double-arrowed arc between two factors of each model in Figure 1 represents the inter-factor correlation, which exists between most constructs. Measurement errors are usually assumed to be independent. The only difference between Figure 1(a,b) is: each measure is loaded by both factors in the EFA model, while each measure is loaded by only one factor in the CFA model. Indeed, “EFA and CFA are mostly differentiated by including cross-factor loadings or not” (Schmitt, 2011, p. 316).

Besides the common factor model, the component model is also frequently found in the EFA of tourism studies (Turner & Vu, 2014). The main difference between those two models is that the component model assumes no measurement errors, which common factor model accounts for. Principal component analysis (PCA) is the most popular component model–based EFA method. Figure 1- c shows a typical PCA model. Although PCA’s assumption of no measurement errors can be unrealistic, PCA and common factor model-based EFA often produce similar factorial structures, especially when items are strongly correlated and/or the number of item increases (Schmitt, 2011; Stewart, 1981). Furthermore, unlike the common factor model, the component model specifies the direction of influence from measures to constructs. As a data-reduction technique, PCA is more suitable for deriving fewer components as linear combinations of a larger set of measures while retaining most original information (Widaman, 2012).

The direction of influence between factors and measures differentiates reflective measurement from formative measurement (Jarvis et al., 2003). Orthodox FA models specify reflective items that take influence from latent constructs. In contrast, formative items (a.k.a. causal indicators) exert influence onto latent variables. Reflective items of a construct constitute a “sample” of potential items that reflect the construct; formative construct requires a “census” of relevant indicators that are unexchangeable and collectively define the construct (Murphy et al., 2009, p. 732). As the first step of construct operationalization, conceptualizing measurement indicators to be reflective or formative is critical for developing valid and robust tourism theories (Mikulić & Ryan, 2018).

The choice between reflective and formative indicators of latent variables depends on contextual ontological assumptions and can be a complex controversial issue (Bagozzi & Yi, 2012). For example, “will recommend to others” and “like to revisit in the future” clearly reflect attitudes towards travel intention (Murphy et al., 2009); whereas, “accessibility”, “variety of amenities”, and “cultural attractions”, among several others, have been taken as reflective or formative indicators of destination image in published studies with/out contextual justifications (Mikulić & Ryan, 2018). As the use of constructs in tourism research grows, more critical thinking is required to ensure rigorous construct validity for tourism to advance as a field (Murphy et al., 2009). There could be practical or statistical advantages to use formative specification (Schmitt, 2011); nevertheless, the decision must be grounded in the

![Figure 1. Factorial structures of EFA, CFA, and PCA models.](image-url)
theoretical relationship between indicators and latent construct(s). Otherwise, it is impossible to replicate research or verify tourism theories with inconsistent construct operationalizations from study to study (Mikulić & Ryan, 2018).

Methodologically, EFA is an inductive method in the framework of data reduction (Mazzocchi, 2008), while CFA is a deductive method in the framework of structural equation modelling (SEM) (Bagossi & Yi, 2012). An EFA solution reveals the data’s optimal factor loadings according to the analyst’s decisions on optimization criterion for the fitting model (i.e., factor extraction method), inter-factor correlation (i.e., factor rotation methods), and a number of factors to extract. EFA is more suitable for exploring new measures’ factorial structure based on factor loading pattern, particularly when relevant conceptual frameworks are not well established. EFA can also be used to further explore poorly fitting CFA models (Hurley et al., 1997; Schmitt, 2011) because CFA models are often found to be too restrictive, e.g., assuming cross-factor loadings to be zero (Marsh et al., 2009).

Compared to EFA, CFA is newer and provides a wider range of statistical measures of model’s goodness of fit, scale reliability, and construct validity (Bagossi & Yi, 2012; Fornell & Larcker, 1981). CFA also enables advanced types of testing in factorial structure, e.g., flexible constraints of model parameters, higher-order factors (Rindskopf & Rose, 1988), method bias (Podsakoff et al., 2012), measurement invariance between groups or over time (Meredith, 1993; Schmitt & Kuljanin, 2008), and smooth transition to SEM to explain factors (Widaman, 2012). Tourism researchers should synergize EFA and CFA when developing or validating construct measurement (Turner & Vu, 2014).

Practically, FA can be used either as an exploratory or as a confirmatory technique, depending on analysis objective (Schmitt, 2011). EFA can be performed in a confirmatory manner to empirically evaluate a theoretically proposed or previously published measurement model. This practice was commonplace in earlier studies before CFA became increasingly accessible since the 1990s along with the popularity of statistical software. “Historically, hypothesis testing of factor loadings within EFA has been given little consideration due to computational complexity of the estimated standard errors, [which] have recently become more readily available for EFA in programs, such as Mplus” (Schmitt, 2011, p. 311). Meanwhile, researchers often conduct CFA in an exploratory fashion by searching for factor model specification with the assistance of SEM modification indices and model fit indices (Hurley et al., 1997; Marsh et al., 2009; Schmitt, 2011). Nonetheless, this operation must be justified with firm theoretical reasoning and never driven only by statistical merits, in order to minimize overcapitalizing on chance factors in one study (Hurley et al., 1997).

FA can be confused with another family of broadly used multivariate statistical procedures in tourism research, i.e., cluster analysis. Using FA as a clustering procedure is “at best an extreme perversion of the method” (Stewart, 1981, p. 51). A factor is a qualitative dimension of a domain space and does not indicate the distance between entities on that dimension. Scores of factor(s) position entities in a factor space on the dimension(s). Popular statistical packages (e.g., SPSS®, AMOS®) implement several methods for estimating factor scores (Mazzocchi, 2008; Widaman, 2012). Clusters are thus easier to understand because they are defined by relatively adjacent entities mapped in the space structured by the factor(s). “There will almost always be more clusters than factors … [and] it is also possible to obtain meaningful factors in the absence of clusters … [when the entities] are homogeneously distributed in factor space” (Stewart, 1981, p. 52). Notably, FA is also widely employed as a preliminary data-reduction step to cluster analysis using factor scores in (tourism) market segmentation studies (Mazzocchi, 2008; Turner & Vu, 2014). However, such a “factor-cluster segmentation” approach has been criticized for being inferior to directly clustering raw measures “with respect to the identification of true heterogeneity in the data” (Dolnicar & Grün, 2008, p. 70).

Conclusion

This note presents an essential introduction to EFA and CFA in terms of their conceptual similarities and differences in model specification and research application. Readers of interest in factor analytic methods are strongly encouraged to further explore the references for in-depth discussions about
a range of extended topics (e.g., item scaling, sample size, distributional assumptions, model estimation, and evaluation, results in interpretation and presentation), which are not unfolded here.

Both EFA and CFA offer insights in multivariate data’s covariance/correlation structure yet from two complementary perspectives. Both factor analytic methods are instrumental for developing theories in marketing and management fields (Hurley et al., 1997; Stewart, 1981) and for “solving tourism-related problems” (Turner & Vu, 2014, p. 207). However, there are no universally accepted decision rules about how to best use EFA and/or CFA (Hurley et al., 1997). This debate shows no sign to settle soon as “both EFA and CFA remain popular and continue to be expanded and updated despite more than 100 years existence of FA” (Schmitt, 2011, p. 304). Tourism researchers and practitioners should acquire at least a fundamental conceptual understanding of FA theory to avoid misspecification or misinterpretation of construct measurement (Turner & Vu, 2014). This is imperative in the current digital age: computers nowadays allow analysts to effortlessly perform FA on easily available data, but one may turn on a computer, and turn off their brain (Mikulić & Ryan, 2018).

**Disclosure statement**

No potential conflict of interest was reported by the author.

**Notes on contributor**

**Yulin Liu** teaches statistical methods and research design in business and engineering disciplines. His main research interests include travel behaviour, tourist behaviour, tourism marketing, project management, methodological innovation, and teaching statistics and scientific research in higher education.

**ORCID**

Yulin Liu [http://orcid.org/0000-0002-6893-6724](http://orcid.org/0000-0002-6893-6724)

**References**

Bagozzi, R. P., & Yi, Y. (2012). Specification, evaluation, and interpretation of structural equation models. *Journal of Academy of Marketing Science*, 40(1), 8–34. [https://doi.org/10.1007/s11747-011-0278-x](https://doi.org/10.1007/s11747-011-0278-x)

Dolnicar, S., & Grün, B. (2008). Challenging “factor–cluster segmentation”. *Journal of Travel Research*, 47(1), 63–71. [https://doi.org/10.1177/0047287508318910](https://doi.org/10.1177/0047287508318910)

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](https://doi.org/10.1177/002224378101800104)

Furunes, T. (2019). Reflections on systematic reviews: Moving golden standards? *Scandinavian Journal of Hospitality and Tourism*, 19(3), 227–231. [https://doi.org/10.1080/15022250.2019.1584965](https://doi.org/10.1080/15022250.2019.1584965)

Hurley, A. E., Scandura, T. A., Schriesheim, C. A., Brannick, M. T., Seers, A., Vandenberg, R. J., & Williams, L. J. (1997). Exploratory and confirmatory factor analysis: Guidelines, issues, and alternatives. *Journal of Organizational Behavior*, 18(6), 667–683. [https://doi.org/10.1002/(SICI)1099-1379(199711)18:6<667::AID-JOB874>3.0.CO;2-T](https://doi.org/10.1002/(SICI)1099-1379(199711)18:6<667::AID-JOB874>3.0.CO;2-T)

Jarvis, C. B., MacKenzie, S. B., & Podsakoff, P. M. (2003). A critical review of construct indicators and measurement model misspecification in marketing and consumer research. *Journal of Consumer Research*, 30(2), 199–218. [https://doi.org/10.1086/376806](https://doi.org/10.1086/376806)

Marsh, H. W., Muthén, B., Asparouhov, T., Lüdtke, O., Robitzsch, A., Morin, A. J. S., & Trautwein, U. (2009). Exploratory structural equation modelling: Integrating CFA and EFA. *Structural Equation Modeling: A Multidisciplinary Journal*, 16(3), 439–476. [https://doi.org/10.1080/10705510903008220](https://doi.org/10.1080/10705510903008220)

Mazzocchi, M. (2008). Factor analysis and principal component analysis. In M. Mazzocchi (Ed.), *Statistics for marketing and consumer research* (pp. 218–246). SAGE Publications. [http://doi.org/10.4135/9780857024657.n10](http://doi.org/10.4135/9780857024657.n10)

Meredith, W. M. (1993). Measurement invariance, factor analysis, and factorial invariance. *Psychometrika*, 58(4), 525–543. [https://doi.org/10.1007/BF02294825](https://doi.org/10.1007/BF02294825)

Mikulić, J., & Ryan, C. (2018). Reflective versus formative confusion in SEM based tourism research: A critical comment. *Tourism Management*, 68(3), 465–469. [https://doi.org/10.1016/j.tourman.2018.05.002](https://doi.org/10.1016/j.tourman.2018.05.002)
Murphy, J., Olaru, D., & Hofacker, C. F. (2009). Rigor in tourism research: Formative and reflective constructs. Annals of Tourism Research, 36(4), 730–734. https://doi.org/10.1016/j.annals.2009.04.005
Podsakoff, P. M., MacKenzie, S. B., & Podsakoff, N. P. (2012). Sources of method bias in social science research and recommendations on how to control it. Annual Review of Psychology, 63(3), 539–569. https://doi.org/10.1146/annurev-psych-120710-100452
Rindskopf, D., & Rose, T. (1988). Some theory and applications of confirmatory second-order factor analysis. Multivariate Behavioral Research, 23(1), 51–67. https://doi.org/10.1207/s15327906mbr2301_3
Rittle-Johnson, B., & Alibali, M. W. (1999). Conceptual and procedural knowledge of mathematics: Does one lead to another? Journal of Educational Psychology, 91(1), 175–189. https://doi.org/10.1037/0022-0663.91.1.175
Schmitt, N., & Kuljanin, G. (2008). Measurement invariance: Review of practice and implications. Human Resource Management Review, 18(4), 210–222. https://doi.org/10.1016/j.hrmr.2008.03.003
Schmitt, T. A. (2011). Current methodological considerations in exploratory and confirmatory factor analysis. Journal of Psychoeducational Assessment, 29(4), 304–321. https://doi.org/10.1177/0734282911406653
Stewart, D. (1981). The application and misapplication of factor analysis in marketing research. Journal of Marketing Research, 18(1), 51–62. https://doi.org/10.1177/002224378101800105
Turner, L. W., & Vu, C. (2014). Factor analysis. In L. Dwyer, A. Gill, & N. Seetaram (Eds.), Handbook of research methods in tourism: Quantitative and qualitative approaches (pp. 183–211). Edward Elgar Publishing.
Wen, J., Wang, W., & Kozak, M. (2020). Small but mighty: A newfound respect for brief research communications. Annals of Tourism Research, 84(3), 102970. https://doi.org/10.1016/j.annals.2020.102970
Widaman, K. F. (2012). Exploratory factor analysis and confirmatory factor analysis. In H. Cooper, P. M. Camic, D. L. Long, A. T. Panter, D. Rindskopf, & K. J. Sher (Eds.), APA handbook of research methods in psychology: Vol. 3. data analysis and research publication (pp. 361–389). American Psychological Association. https://doi.org/10.1037/13621-018