

Attribute Importance Research in Travel and Tourism: Are We Following Accepted Guidelines?

Kenneth R. Bartkus

Jon M, Huntsman School of Business, Utah State University

Megan McDonald

Jon M. Huntsman School of Business, Utah State University

Follow this and additional works at: <https://scholarworks.umass.edu/ttra>

Bartkus, Kenneth R. and McDonald, Megan, "Attribute Importance Research in Travel and Tourism: Are We Following Accepted Guidelines?" (2016). *Travel and Tourism Research Association: Advancing Tourism Research Globally*. 45.

<https://scholarworks.umass.edu/ttra/2011/Oral/45>

This is brought to you for free and open access by ScholarWorks@UMass Amherst. It has been accepted for inclusion in Travel and Tourism Research Association: Advancing Tourism Research Globally by an authorized administrator of ScholarWorks@UMass Amherst. For more information, please contact scholarworks@library.umass.edu.

Attribute Importance Research in Travel and Tourism: Are We Following Accepted Guidelines?

Kenneth R. Bartkus
Jon M. Huntsman School of Business
Utah State University

and

Megan McDonald
Jon M. Huntsman School of Business
Utah State University

ABSTRACT

The purpose of this study is to examine the reporting practices and potential challenges in the measurement of attribute importance in travel and tourism research. The results indicate that a majority of studies evaluated in this study utilized Likert scaling which has a tendency to produce end-piling or upward bias. As such, it becomes more difficult to identify meaningful differences among attributes. Recommendations for addressing this potential limitation and for better reporting of results are provided.

Keywords: *attribute importance, Likert scaling, measurement, upward bias*

INTRODUCTION

That research should be properly reported is a *sine qua non* of science. Without it, research results are at best misleading and, at worst, worthless and perhaps even dangerous. Unfortunately, problems sometimes occur even with the best efforts of researchers and the editorial review process. In such instances, the scientific community is charged with identifying inaccuracies and providing clarification. Past examples include the reporting of faulty citations (Wright and Armstrong 2008), the misapplication of statistical methods (Spirer and Jaffe 1984), the use of inappropriate analytic techniques (LaTour and Miniard 1983), and common misconceptions about the reporting of effect sizes (Breugh 2003). In doing so, the scholarly community meets its obligation to improve the practice of science.

The current study follows this tradition by examining two potential problems with regard to the measurement and analysis of attribute importance: (1) the tendency for Likert scaling to produce upward bias and (2) the extent to which researchers make inferences about mean differences without conducting appropriate statistical tests. Should the results indicate that researchers are adhering to generally accepted guidelines, this would be evidence of robust scholarship. Alternatively, should the evidence suggest that researchers are not following guidelines, additional clarification and comment is needed to improve the quality of future research.

LITERATURE

There is little debate regarding the significance of attribute importance in strategic decision making. It is a mainstay of well-established decision models such as importance-performance analysis (Rood and Dziadkowiec 2010). Even the mere measurement of attribute importance provides valuable information by identifying target audience preferences which can then be used by decision makers to develop products and services that best meet the needs and wants of consumers. For example, Kim, Bergman, and Raab (2010) measured how mature diners perceived the importance of attributes at three types of restaurants. The mean scores were then ranked and a subsequent factor analysis was conducted to identify common themes.

Given that attribute importance is an essential component of the decision-making process, appropriate techniques are needed to ensure that the measurement and analysis are appropriately conducted and reported. In this regard, we identified two issues that have the potential to create challenges in the measurement and reporting of attribute importance: (1) the use of Likert scaling and (2) the analysis of mean differences.

The use of Likert scaling is a potential concern because it has a tendency to produce ratings that crowd together at the high end of the scale, resulting in upward bias. (Matilla 1999; Bacon 2003; Cohen 2003; Hood 2003; Lee, Soutar, and Louviere 2008) Lee, Soutar, and Louviere (2008) note this is due to the fact that Likert scales do not require respondents to make trade-offs in their evaluations of importance. Hence, respondents have a greater propensity to rate attributes similarly, which essentially defeats the purpose of measuring attribute importance. Nonetheless, Likert scaling remains popular in this area, perhaps because Likert scales can be efficiently administered and respondents find them familiar and easy to answer. (Chrzan and Golovashkina 2006)

Several recommendations have been provided to correct for the crowding tendency. Abalo, Varela & Mazano (2007), for example, argue for the use of rank ordering in place of Likert scaling because it requires respondents to evaluate attributes jointly rather than independently. The ranks are then adjusted by a transformation procedure that is intended to improve discrimination. Alternatively, Sampson and Showalter (1999) suggest that importance ratings should not be measured independently, but as a function of performance. This procedure, unfortunately, restricts its use to products and services for which the respondent already has experience. Louviere and Woodworth (1983) recommend abandoning Likert scaling altogether in favor of a “best-worst” method that requires respondent’s to select choices among alternative attributes. In this way, ratings on any individual attribute are dependent on consideration of alternative choices, much like the real world. While this technique provides statistically valid discrimination, it is also more difficult to administer.

Even if appropriate procedures are used to control for upward bias, proper analytic procedures are still needed to test for differences among the attribute means. Because studies typically compare more than two means, the use of the standard t-test needs to be adjusted to account for the increased probability of making a Type I error; in essence, to control for a “false discovery” effect. The literature provides numerous recommendations such as Bonferroni. (see Hwang, Lai, Oh, 2010 for a review).

Finally, it is recommended that limitations of the research be included in the conclusion section of the study so that the results can be placed in proper context. (Lee and Greenley 2009) Clearly, if responses exhibit evidence of upward bias, this would constitute a limitation.

This study examines the extent to which researchers are adhering to the recommendations in each of these areas. To provide structure for the analysis, two general guidelines were used:

- G1: Appropriate statistical tests should be conducted when comparing means.
- G2: Evidence of upward bias should be acknowledged as a limitation.

METHODOLOGY

A literature search for empirical research involving attribute importance in travel and tourism was conducted for the years 2000-2010. The search included peer-reviewed articles without regard for the nature of the publication (i.e., it included non-travel and tourism journals).

The search resulted in the identification of 25 articles that were specific to the measurement of attribute importance. Each article was then evaluated for the measurement technique used, whether and how mean differences were reported, and the reporting of relevant limitations. While the sample may or may not represent a full audit of all available literature in this area, an attempt was made to ensure completeness. The sample size appears sufficient for the purposes of this study.

RESULTS AND DISCUSSION

Of the 25 articles examined in this study, all but one utilized Likert scaling to measure attribute importance. Of the 24 that used Likert scaling, 19 (79.2%) presented a ranking of the means and made reference to the most, and sometimes least, important attributes. No evidence was identified to indicate that statistical tests were conducted to support the inferences. Instead, the reporting appears to have been based solely on stated magnitude of the means. For example, one study noted that “Ambiance was rated as the most important attribute, followed by lodging features and customer service.” Another noted that “The least important variables, as rated by the respondents, were nightlife, recreational facilities, tourist attractions, shopping facilities, and climate” Five of the 24 articles did not make reference to mean differences.

Further analysis revealed that the data for all 24 studies contained some evidence of upward bias in the means (i.e., the majority of means were above the mid-point). For the study that utilized a ranking procedure, no procedure was used to transform the data, which is contrary to recommendations by Abalo, Varela & Mazano (2007).

Only three of the 24 studies acknowledged limitations due to upward bias. One mentioned it with no further comment. Another attributed the upward bias to a small sample size (n=76) rather than the Likert procedure. Finally, one study suggested that adding more response categories might have enabled a wider distribution of responses (the study utilized a four-point scale). No study attributed the upward bias to any of the arguments presented in the literature review (or any alternative argument that might be referenced from the literature).

CONCLUSION

This study examined how attribute importance has been measured, analyzed, and reported in travel and tourism research. The results appear to have important managerial and research implications. First, the study cautions against the simple use of Likert scaling without also accounting for potential upward bias. At a minimum, it is recommended that researchers calculate and report levels of skewness in order to determine the extent to which it is a problem. Second, managers and researchers are reminded that inferences of mean differences are meaningful only in the context of statistical significance. For researchers (and the editorial review process), this also implies a need for greater oversight of reporting.

It is also important to note that most of the articles reviewed in this study conducted additional statistical analyses not related to the initial rankings (e.g., factor analysis). Hence, it is not appropriate to generalize the findings of this study to the overall quality of the reviewed research. Additionally, care should be taken not to generalize the results of this study to other disciplines, although we suspect similar findings may be evident in other areas as well. Finally, the purpose of the current study is to improve science in travel and tourism research rather than to personally criticize the research process, per se. Continual improvement in techniques and reporting are a natural part of the evolutionary process in science and our intent was merely to make a contribution toward this effort. We hope that the results will be accepted in this spirit.

REFERENCES

- Abalo, J., Varela, J. & Mazano, V. (2007). Importance values for Importance–Performance Analysis: A formula for spreading out values derived from preference rankings. *Journal of Business Research*, 60(2), 115-121.
- Bacon, D.R. (2003). A comparison of approaches to Importance–Performance Analysis. *International Journal of Marketing Research*, 45(1), 55–71.
- Breaugh, J.A. (2003). Effect size estimation: Factors to consider and mistakes to avoid. *Journal of Management*, 29(1), 79-98.
- Chrzan, K. & Golovashkina, N. (2006). An empirical test of six stated importance measures. *International Journal of Market Research*, 48(6), 717-740.
- Cohen, S. (2003). Maximum difference scaling: improved measures of importance and preference for segmentation. *Proceedings Sawtooth Software Conference Proceedings*, San Diego, CA, 61–74.
- Hood, J. N. (2003). The relationship of leadership style and CEO values to ethical practices in organizations. *Journal of Business Ethics*, 43, 263–273.
- Hwang, Y., Lai, J. & Oh, S. (2010). Evaluations of FWER-controlling methods in multiple hypothesis testing. *Journal of Applied Statistics*, 37(10), 1681-1694.
- Kim, Y., Bergman, C. & Raab, C. (2010). Factors that impact mature customer dining choices in Las Vegas. *Journal of Foodservice Business Research*, 13(3), 178-192.
- Kahle, L. R. (1983). *Social values and social change: Adaptation to life in America*. New York: Praeger.
- LaTour, S.A. & Miniard, P.W. (1983). The misuse of repeated measures of analysis in marketing research. *Journal of Marketing Research*, 20(1), 45-57.

- Lee, J.A., Soutar, G. & Louviere, J. (2008). The Best-Worst Scaling Approach: An alternative to Schwartz's Values Survey. *Journal of Personality Assessment*, 90(4), 335-347.
- Lee, N. & Greenley, G. (2009). What makes a good article? Generating an insightful manuscript. *European Journal of Marketing*, 43(5/6), 577-582.
- Louviere, J. J. & Woodworth, G.G. (1983). Design and analysis of simulated consumer choice or allocation experiments. *Journal of Marketing Research*, 20, 350-367.
- Mattila, A. (1999). Consumers' value judgments. *Cornell Hotel and Restaurant Administration Quarterly*, 40, 40-45.
- Rood, S.A. & Dziadkowiec, J. (2010) Why use importance performance analysis in mystery shopping? A USA-Poland comparative answer. *Journal of Quality Assurance in Hospitality & Tourism*, 11(1), 1-16.
- Sampson, S. & Showalter, S (1999). The performance-importance response function: observations and implications. *Service Industries Journal*, 19(3), 1-25.
- Spirer, H.F. & Jaffe, A.J. (1984). Misuses of statistics: Lessons for statisticians, non-statisticians, students and teachers. *American Journal of Economics & Sociology*, 43(2), 205-216.
- Wright, M. & Armstrong, S.J. (2008). The Ombudsman: Verification of citations: Faulty towers of knowledge? *Interfaces*, 38(2), 125-132.