An Innovative Leadership Effectiveness Measure: Applied Analytic Indicators of High-Consequence Industry Performance

Brent D. Bowen  
*Purdue University, bdbowen@purdue.edu*

Erin E. Bowen  
*Purdue University, West Lafayette, eebowen@purdue.edu*

Dean E. Headley  
*Wichita State University, dean.headley@wichita.edu*

Hatice Küçükönal  
*School of Aviation, Özyeğin University*

Clay Wildt  
*Purdue University, cwildt@purdue.edu*

Follow this and additional works at: [http://docs.lib.purdue.edu/atpubs](http://docs.lib.purdue.edu/atpubs)  
Part of the Management Sciences and Quantitative Methods Commons, and the Organizational Behavior and Theory Commons

[http://docs.lib.purdue.edu/atpubs/5](http://docs.lib.purdue.edu/atpubs/5)

This document has been made available through Purdue e-Pubs, a service of the Purdue University Libraries. Please contact epubs@purdue.edu for additional information.
Abstract

Leadership effectiveness in high-consequence industries has more than a bottom-line fiscal impact; it is linked to critical issues of human safety. Performance, productivity and overall quality of service have to be managed with focus on improvement in systemic safety while simultaneously maintaining a viable and profitable organization. This premise is specifically foremost in the leadership of airline organizations. The Airline Quality Rating has become a recognized and lauded indicator of airline performance in the United States. A valid case is presented herein to confirm that the Airline Quality Rating’s applied analytic methodology effectively provides a tool for assessing organizational leadership. These results provide a benchmark for global adoption in the world airline industry.

© 2012 Published by Elsevier Ltd. Selection and/or peer-review under responsibility of 2nd International Conference on Leadership, Technology and Innovation Management

Keywords: Performance indicators, Quality, Technology, Airline Quality Rating, Airline Passenger Survey, Policy Research Construct

1. Introduction

Performance of organizational leadership and performance of the entire organization have an undeniable dynamic linkage. While this concept may appear self-evident, the mechanisms of the leadership–organizational performance relationship are actually complex, multi-level and multi-dimensional. This is particularly true in the case of the airline industry, in which the performance of corporate leaders occurs within a multi-layered web of national and international government regulations, safety considerations, training requirements, the inter-connection of multiple related yet separately operated systems (e.g., flight, air traffic control, maintenance, security, etc.) and the operation of extremely complex high-technology equipment. To extract the role of the organizational leader with regard to the performance of the organization in such a high-consequence, highly structured setting poses unique challenges both to leaders and to organizational researchers.

High consequence industries have more than the fiscal measures of leadership effectiveness at risk; they face a profound responsibility to maintain safety at the forefront (Bowen, Sabin, & Patankar, 2011). The outlook for airline performance fiscally is mixed, and while the increases in overall systemic safety are improving, the challenge in such settings is primarily one of avoiding complacency in the face of such safety confidence (Bowen et al., 2011). In addition, the impact of new mega-carriers such as United/Continental and Southwest/AirTran on those carriers’ effectiveness, as well as on the industry as a whole, remain to be seen. An examination of past AQR reports reveal that
improved airline performance is not a given. “Look back to the Delta/Northwest and U.S. Air/America West mergers, and you will find that these mergers brought performance problems for the new carriers” (Headley, 2011, p. 1). Being recognized as a provider of quality services produces opportunities for a business to flourish. Airlines are no exception. As a highly competitive business, being known as a provider of good customer outcomes (i.e. quality) offers a company a better chance to succeed. Historical Federal Aviation Administration (FAA) business forecast data affirm that globalization, economic circumstances, and a trend toward aviation alliances have forced U.S. airlines to boost service quality in order to increase competitiveness and survive (Bowen, Headley, & Lu, 2003).

2. Literature Review And Hypotheses

2.1. Airline Industry Performance Measures

In 1999, in the wake of customer dissatisfaction and popular pressure, efforts by the United States Department of Transportation (DOT) and the FAA to protect the rights of airline passengers and the idea of mandating better customer service came to a head. As a result, 14 air carriers voluntarily pledged to improve customer treatment (Bowen & Lu, 2004). This commitment to improve passenger service included: honestly notifying customers of a flight delay, better addressing passengers’ needs due to a delay, allowing passengers a refund on airfare for a changed flight schedule, and, in general, being more responsive to passengers’ complaints (ATA, 2011). To realize this commitment, large airlines launched the Customer First Plan. Although slightly different from airline to airline, the aspects of the Customer First Plan do address the basics of better customer service. As of June 27, 2008, the DOT website provides links to nine major airlines’ Customer Service Plans. More recently, the Air Transport Association’s (ATA’s) “Customer First 12-Point Customer Service Commitment” plan commits member airlines to an expanded list of customer service practices. The only consistently available source for tracking performance on customer outcomes is the Air Travel Consumer Report (DOT, 2011). This monthly report offers great detail, but is not widely read by the flying public.

2.2. Airline Quality Defined

Since 1991, the Airline Quality Rating (AQR) reports have been compiled and made widely available through the popular press and via the Internet. The AQR offers measurement of the airline industry performance and is viewed as insightful and influential in the aviation world (Goodman, 1992; Mann, 2000). The initial objective of the AQR was to develop a method for accurately measuring airline performance quality. Preliminary research indicated that many air carriers were seeking to better control their service quality via quantitative methods (Bowen, Headley, & Luedtke, 1992). Results of the annual AQR have given the airline industry a more coherent approach to meeting this goal of better customer service. Recent upstart JetBlue has exhibited behaviour that serves to validate the AQR reports; in two separate publications that outline the business strategy of the airline, numerous references are given to both the respect of the report within the airline industry as well as its use as a benchmark for performance; during regular business meetings, the DOT metrics and associated ratings are discussed as the benchmark for the company’s operating efficiency and management practices (Peterson, 2004; Wynbrandt, 2004). The AQR reports are widely utilized by the airlines in order to promote service quality and attract potential passengers (Spencer, 1999). The perceived leader in customer service, Southwest Airlines, went so far as to create an internal award for scoring highest in three DOT metrics. Dubbed the “Triple Crown,” the campaign turned relatively abstract DOT statistics into a coveted and sought-after accomplishment (Freiberg & Freiberg, 1996).

The Airline Quality Rating is a weighted average of multiple elements determined to be important to consumers when judging the quality of airline services. Elements considered for inclusion in the rating scale were screened to meet two basic criteria; 1) an element must be obtainable from published data sources for each airline; and 2) an element must have relevance to consumer concerns regarding airline quality. (Headley & Bowen, 1994). Data for the elements used in calculating the ratings represent performance aspects (on time arrival, mishandled baggage, involuntary denied boardings, and 12 customer complaint areas) of airlines that are important to consumers (Bowen, Bowen, & Headley, 2011). Factors included in the rating scale were taken from an initial list of over 80 potential factors (Bowen, Headley, Kane, & Lutte, 1999). All of the elements are reported in the Air Travel Consumer Report maintained by the U.S. Department of Transportation.
2.3. Performance-Based Leadership

Creating an accurate linkage model between the performance of a leader and the success (or lack thereof) of an organization continues to be one of the ongoing challenges among leadership scholars (Yukl, 2002; Northouse, 2010). Despite decades of scholarly research on leadership in organizations going back to the 1930s (House & Aditya, 1997), describing and properly weighting the various factors and contextual variables that may interplay with the leader and the organization’s performance remain somewhat elusive. Developing the leader performance-organizational performance linkage proves even more challenging, however, when the organization under investigation is one operating in a highly regulated, highly structured environment (such as aviation). In these settings, the regulatory limitations, strict training guidelines, pricing limits, and even innovation restrictions create profound barriers for the leader that may restrict his/her ability to act to most rapidly maximize successful organizational performance. In addition, documenting successful performance in these types of settings may be a challenge as well (Bowen, et al., 2011).

Leadership as based on organizational performance has routinely been accomplished through measurements that result in overall quality of actual delivery of goods and services (e.g., Headley & Bowen, 2004), which can be measured both by objective performance data (e.g., via statistics collected by the U.S. Department of Transportation or other comparable entities globally) as well as subjective evaluations by consumers and employees (Herzberg, 1966). The importance of service quality should not be diminished; because of its enhanced visibility in the airline industry, passengers (customers) often use service quality as the sole metric by which the entire quality of an airline is judged (Babbar & Koufteros, 2008). For the U.S. based airline industry, performance data are available in multiple categories over time. These areas include on-time performance, denied boardings, baggage handling, and customer complaints (Bowen & Headley, 2002), all areas of vital importance to the travelling public and of absolute necessity to the managerial leadership of a commercial airline organization.

2.4 Research Questions

Documentation in a performance measurement study (Bowen & Hansen, 1999) found a corresponding relationship between airline organization performance as measured by the Airline Quality Rating and consumer satisfaction as it relates to intended management/leadership direction. This finding, validated by airline CEO Gordon Bethune (1998) best asserted that leaders can identify key performance indicators and make decisions to directly impact those indicators, to both the subjective and objective success of the organization; that it is possible to move an airline from a performance laggard to a performance leader. In replication of this finding, the researchers propose the following questions:

1. Is the Airline Quality Rating valid as a measurement of airline performance;
2. Are the consumers of airline services in agreement that performance is a reflection of leadership performance;
3. Can an objectively calculated AQR score serve as an indicator of leadership effectiveness?

3. Methodology

3.1. Application of the AQR as an Indicator of Performance

The Airline Quality Rating criteria and the weighted average methodology allow a focused comparison of domestic airline performance. Unlike other consumer opinion approaches that have relied on consumer surveys and subjective opinion, the AQR continues to use a mathematical formula that considers multiple weighted objective criteria to arrive at a single, fully comparable rating for airline industry performance (Bowen, Headley, & Kane, 1999). The Airline Quality Rating provides both consumers and industry observers a means for monitoring comparative quality for each airline on a timely basis, using objective, performance based data. Over the years, the Airline Quality Rating has often been cited as an industry standard for comparing airline performance. Currently the AQR stands as the only regularly published rating available for airline performance (AQR reports from 1990 to present are permanently archived and available on-line at www.airlinequalityrating.com, formerly www.aqr.aero). With the continued global trend in airline operations alliances, the need becomes even stronger for the Airline Quality Rating to be used as a standardized method for comparing the quality of airline performance for international operations as well (Headley & Bowen, 1997).
Table 1: Airline Quality Rating Criteria, Weights and Impact

<table>
<thead>
<tr>
<th>Criteria</th>
<th>Weight</th>
<th>Impact</th>
</tr>
</thead>
<tbody>
<tr>
<td>OT (On-Time Performance)</td>
<td>8.63</td>
<td>+</td>
</tr>
<tr>
<td>DB (Denied Boardings)</td>
<td>8.03</td>
<td>—</td>
</tr>
<tr>
<td>MB (Mishandled Baggage)</td>
<td>7.92</td>
<td>—</td>
</tr>
<tr>
<td>CC (Customer Complaints)</td>
<td>7.17</td>
<td>—</td>
</tr>
</tbody>
</table>

NB: Data for all criteria is drawn from the U.S. Department of Transportation's monthly Air Travel Consumer Report (DOT, 2011); data are scaled by passenger volume for equivalency.

3.2. Weighting Performance Indicators

Weights were originally established by surveying 65 airline industry experts regarding their opinion as to what consumers would rate as important (on a Likert scale of 0 [lowest] to 10 [highest]) in judging airline quality. These experts were objectively selected as key informants and included airport executives, airline managers, members of aviation interest groups and organizations, academia, customers, aviation authority figures and aircraft manufacturers (Bowen, Headley, & Lu, 2002). Each weight and element was assigned a plus or minus sign to reflect the nature of impact for that criterion on a consumer's perception of quality. Weights and positive/negative signs are independent of each other. Weights reflect importance of the criteria in consumer decision-making, while signs reflect the direction of impact that the criteria should have on the consumer's rating of airline quality. When all criteria, weights and impacts are combined for an airline over the year, a single interval scaled value is obtained. This value is comparable across airlines and across time periods (Headley & Bowen, 1997).

Figure 1: AQR Formula

The formula for calculating the AQR score is:

\[
AQR = \frac{(+8.63 \times OT) + (-8.03 \times DB) + (-7.92 \times MB) + (-7.17 \times CC)}{8.63 + 8.03 + 7.92 + 7.17}
\]

3.2.1 Reliability and Validity

In the spring of 2002, a nationwide survey of frequent flyers was conducted to facilitate a revisiting of the weighting for the AQR elements. Analysis of the sample of 766 opinions showed no significant change in the relative weights for the AQR elements; in conjunction with this information as well as to maintain comparability across the years, the weights have thus been held constant. This permits researchers to evaluate longitudinal trends both within a particular airline, category of airlines (e.g., low-cost carriers, regional airlines, or ‘legacy’ carriers – those who remain from the years before deregulation of the U.S. commercial airline industry). The reliability of the rating scale was tested through determination of a Cronbach’s alpha calculation. Cronbach’s alpha measures how closely related a set of items are as a group and serves as a measure of internal consistency (UCLA, 2011). The Airline Quality Rating reliability coefficient measured 0.87. This suggests that the AQR is reliable and that factor determination results would be similar for each of the four areas tracked. Mixed method construct validity is enhanced through examination of the analytic performance data in comparison to present and past surveys of actual consumers. This cross-validation provides not only a basis for validation but a representation that performance data from the airlines are consistent with the expectations of the flying public.

3.3. Measuring leadership performance

In order to understand the longitudinal performance of the overall U.S. airline industry, Figure 2 presents a 20-year trend analysis of overall performance on the Airline Quality Rating (Bowen & Headley, 2012). Because three of
the four variables in the AQR formula are negatively weighted (such that they represent a lowered customer service/quality experience), AQR scores are represented on a negative value scale. A lower (i.e., closer to zero) AQR score for an airline represents better overall performance; and lowering scores for the industry indicate improvement in performance industry-wide.

Figure 2: 20-year Industry Average Performance

Creating a theory that bridges the gap between the Airline Quality Rating and management performance is a task that requires the invention of a new model reflecting the relationship between two subjects. The creation of a formula that would indicate the distance away from the industry leader was selected as the most appropriate model to relate the two concepts. Utilizing a benchmark (in this case, the top airline), researchers can theorize that airlines that score highest in the AQR are also those that are well managed by capable leaders and managers. The formula in Figure 3 below depicts the mathematical model used to analyze scores from the Airline Quality Rating.

Figure 3: Airline Performance Formula

\[ \sqrt{\frac{\sum_{i=1}^{n} (A_{x_i} - A_{y_i})^2}{5}} \]

The construction of our equation utilizes a universal mathematical formula, the Euclidean Norm (Walker, n.d.). By using this concept, we can determine the mean distance, over a five year period, of each airline from the top position. Researchers used the following steps to yield a final score: first, the top position \((A_{y})\) is subtracted from the respective airline’s ranking \((A_{x})\). This value is then squared to yield a positive integer. By summing the aforementioned values for all 5 years, we arrive at the aggregated total distance from the top performer over the period. Dividing this total by the number of years in our data interval (five), we arrive at an average score that depicts the mean distance from the top performer over the interval. Finally, calculating the square root normalizes the data and reverts the values to their appropriate position (we must “undo” the square that occurs in the first steps of this
equation). Again, these scores are indicative of how far each airline is (over a five-year period) from the coveted rank of number one (J. Mott, personal communication, July 23, 2012).

Using this formula, researchers calculated an airline performance score for the airlines who were ranked in the top 5 in the most recently-released Airline Quality Ratings. Using their rankings for the past five years, researchers calculated an overall distance from the desired ranking (1st) as an indicator of global performance across this time period.

Table 2: Five-Year Rankings Trend with Airline Performance Score

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>AirTran</td>
<td>2</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0.632456</td>
</tr>
<tr>
<td>Hawaiian</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>0.77460</td>
</tr>
<tr>
<td>JetBlue</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>2.00000</td>
</tr>
<tr>
<td>Frontier</td>
<td>7</td>
<td>7</td>
<td>9</td>
<td>4</td>
<td>4</td>
<td>5.54977</td>
</tr>
<tr>
<td>Alaska</td>
<td>5</td>
<td>11</td>
<td>4</td>
<td>5</td>
<td>5</td>
<td>5.60357</td>
</tr>
</tbody>
</table>

Using the scores to draw relational connections with an airline’s management and leadership quality is the goal of such formulation. As indicated in Table 2, AirTran moved from a second-place ranking (2008 and 2009) to maintain number one in the AQR rankings across the most recent three years of data. To attain such consistently high ratings, AirTran must have scored highly across all four areas of the AQR rankings. AirTran’s relatively low score (0.632456) indicates that the airline’s performance efficiency has been adequate to maintain a high ranking over the five year period; it is possible to theorize that this efficiency may be equated to effective leadership and management practices. Airlines that have a much higher score, such as Alaska Airlines (5.60357), seem to be more prone to erratic movement within the rating scale; changing or ambiguous management and leadership could lead to such sharp movements within rankings. The most validating action of this effectiveness by the leadership team of AirTran is the fact that it was acquired in the past year by Southwest. Southwest, the longstanding and recognized leader among airlines, has been declining in performance. In recognition of this trend, Southwest is attempting to reverse course by acquiring the highest scoring airline today, AirTran. Through this action Southwest may undoubtedly make a determined effort to return to the top of the rankings as they were in the peak of productivity under CEO Herb Kelleher.

4. Conclusion

Measuring the performance of managers and executives is a vital component of an airline’s survival in the post-deregulation industry. In the face of intense competition, managers must effectively balance acceptable costs with high-quality service. Research comparing the opinions of passengers to that of airline managers has shown that the two parties often have differing views regarding the relative importance of various qualities and responsibilities of the airline and its respective employees. In essence, “consumers (passengers) are generally dissatisfied with the service provided by U.S. airlines…they care more about how they are treated than what they are treated to.” (Gourdin, 1992). The experience that each passenger endures while traveling with an airline can be traced directly to the policies and actions implemented by the company’s management personnel. Expanding on this logic, it would suffice to infer that a measure of managers’ effectiveness within the organization has a sizeable impact on the experience of the flying public. This cyclical interchange should be measured and incrementally improved to ensure that the airline is providing the quality of service demanded by customers.

The Airline Quality Rating (AQR) penetrates the customer satisfaction issue much more effectively than opinion gathering; this weighted average model packages readily available material (provided by the U.S. Department of Transportation) in a format that is much easier to digest by industry leaders and the flying public. In terms of consumer decisions and quality judgments, the AQR is much easier for the average traveler to comprehend than the bulky, and often hard to find, metrics provided by the DOT. By presenting data concerning all major domestic airlines
in a consolidated report, consumers are able to compare quality within the industry; this personal perception of quality deduced from the report (even if the passenger has no relative experience) can become an important competitive point that not only seeds opinions about various airlines, but can also be the deciding factor when travelers decide from which airline to purchase a ticket. In all, the AQR serves as “an assessment tool that gives accurate comparative readings based on performance rather than subjective opinion.” (Headley & Bowen, 1997).

Longitudinal performance can additionally be evaluated using an airline performance formula, which evaluates ordinal overall performance data over a specified time period. In so doing, a longitudinal score can be calculated for an airline which may in turn be compared to known information on an airline’s leadership. This longitudinal comparison to various organizational data such as employee turnover, stock price trends, etc continues. Initial findings indicate, though, that using the AQR and longitudinal comparisons map to airline leadership effectiveness; eventually, it is hoped that they may be used for predictive scenarios as well.

Having a system similar to the AQR take hold on a global scale would be an undeniable asset to industry professionals, frequent fliers, and the traveling public. As the aviation industry continues to grow increasingly global, a common tool that could measure both domestic and international airlines within and between respective countries would allow consumers and aviation professionals to compare services and indirectly, the effectiveness of airline management. Government entities and private corporations would have the ability to compare, side-by-side, the quality of various international airlines; these statistics would inadvertently point to various indirect indicators of quality such as safety, management practices, etc. (Headley & Bowen, 1997). Extending such a metric to international airlines will ensure a global standard for quality in airline management.

Acknowledgements

We would like to express our gratitude to all those who gave us the possibility to complete this report. We also want to acknowledge the invaluable assistance of the Applied Human Factors Research Laboratory and the Air Transport Policy Research Group; without these vital research groups, this paper would not have been possible. We would also like to thank Mr. Clay Wildt, Undergraduate Research Assistant, for his assistance in compiling research material and contributing to the construction of the paper’s final draft, and Professor John Mott for assistance in formula derivation.

References


