Abstract

Purpose – This paper seeks to apply the SERVQUAL model to identify critical performance measures in the airline industry, exploring differences between Eastern and Western expectations of airline service quality and delivery.

Design/methodology/approach – Data from 263 effective questionnaire responses were collected from two locations – Taipei (Taiwan) and London (UK) – to compare differences between the well-documented gap-5 (between perceived and expected levels of service quality) values of respondents from these places of origin.

Findings – The paper generally finds that: there is a statistically significant difference between the perceived and expected levels of service quality in the airline industry; these are affected by such demographic factors as education, occupation and income levels (but not all that were examined); the SERVQUAL model’s dimensions represent appropriately the airline industry; and the gap-5 sizes of these quality dimensions have a significant impact on customer satisfaction and service value; but there does not seem to be a statistically significant difference between the gap-5s of respondents from the two locations.

Research limitations/implications – The paper limited the research to data from two locations, and makes a bold assumption that the two locations make adequate representations of views from the East and West.

Practical implications – Gap-5 and general SERVQUAL analyses seem to apply well to the airline industry. Further, management effort need not be different for the delivery of service quality between Eastern and Western passengers/customers. The findings are generalizable to other sectors for which service quality is an important public sector concern (e.g. household utilities).

Originality/value – A generic framework is presented for how service quality dimensions, and issues of gap-5, relate to overall service quality, customer satisfaction, and service value, in the passenger airline industry.

Keywords Customer services quality, SERVQUAL, Airlines, Performance management, Customer satisfaction

Paper type Research paper

Introduction

The airline industry has had a long and prominent history of growth and success despite the recent notable events that have shaken up passenger confidence – these include the wars of Afghanistan and Iraq, the outbreak of the SARS epidemic in 2003, and the recent surging oil prices. Budget airlines, in particular, have also made a significant impact on the way passengers choose to fly – these are not only local flights, such as those offered by well-known airlines like Easyjet and Ryanair in Europe, or Southwest Airlines in the USA, but also include cross-continental carriers like Oasis (which up until recently) offered flights from UK or Canada to Hong Kong.
Such flights have been known for their ability to offer low prices to passenger, but at the detriment of service quality. The escalating rises in fuel prices and renewed focus on international liberalization have also placed a squeeze on how airline companies manage. It is not just the short haul low-cost carriers which have been “cutting-corners”: 2008’s “credit crunch", which sent Oasis into liquidation, has impacted the way even some of the largest passenger carriers like American Airlines deal with customers (by implementing baggage handling charges), has placed a question-mark on the viability of various business models and the way they deliver value to customers. The quality-price trade-off has been well documented in a number of service sector literatures, especially in British network utilities (Chau, 2002; Chau and Witcher, 2005a; Waddams Price et al., 2008); there is also a groundswell of research on the general importance of service quality (Svensson, 2006). It is therefore interesting to understand the issue of managing service quality in the airline industry, especially as there has been such a prominent increase in the number of worldwide airline passengers (Kasper et al., 2006; Bor, 2003; Doganis, 2001).

The airline industry plays an important role in the service sector itself, as well as contributes significantly to other industries through the ability to transport passengers to their required locations all over the globe (Rhoades and Waguespack, 2008). The industry has significantly taken off since the 1960s with an average annual growth of about 12 per cent, and despite the recent slow down due to a maturing industry and general crises, world passenger traffic has averaged an annual increase of about 5 per cent (Hanlon, 1999). It is suggested that probably the three fundamental factors that affect passenger demand in the airline industry are incomes, fares and service levels (Hanlon, 1999). The core focus of the present research is solely on the service quality aspect (although we accept that cost and price have an influence on service quality).

There are numerous factors that might be fuelling passenger growth, but there is no doubt that service quality plays a critical role in the success of the airline industry (Bamford and Xystouri, 2005; Rhoades and Waguespack, 2005; 2008). There is much evidence to suggest that profitability is linked to service delivery (e.g. Bates et al., 2003) and the general importance of quality in service organizations (Gustafsson and Nilsson, 2003; Di Mascio, 2007). The focus of this paper is on understanding service quality and customer satisfaction, and the ability of airline companies to satisfy customers’ perceptions of what these may be. The services sector literature, which stands firmly in its own right (Edvardsson et al., 2005), has recently shown much interest in the area of passenger expectations and airline quality (e.g. Fodness and Murray, 2007; Chen and Chang, 2005; Gilbert and Wong, 2003). Hence, our research provides a firsthand examination of airline passengers’ expectations of quality standards and suggests areas for improving these, which is proving important (e.g. Armistead, 1989; Rowley, 1998). In the language of the service quality literature, we examine if gap-5 (the gap between customer expectation and perceived performance, as explained later) is the “troubled water” over which a bridge is necessitated, or just another “long and winding road” for future SERVQUAL research (for a review of its value, see Carrillat et al., 2007; Ladhari, 2008).

The originality in the research is that we observe views on service standards from passengers at airports in two countries only, Taiwan and the United Kingdom, to represent views from the East and West. Its value is in how focusing on these two locations (which gives a sufficient comparison of cross-cultural views) can shed light
on understanding if marketing effort, or the management of service quality in the passenger airline industry, on narrowing gap-5 for a range of service requirements, needs to be any different or country-specific. Hence, the research is positioned within the literature base of managing service quality from the marketing management perspective of airline companies. The applicability of gap-5 analysis in this industry, within the conceptual framework presented in this research, may also have useful implications for the international (marketing) management of industries sharing similar characteristics.

**Service quality literature and hypotheses formulation**

*Service quality defined*

Service quality concerns the benefits on the customer side of the company-customer exchange. Carefully integrating issues of quality with productivity, company improvement programs are aimed generally at improving the long-term profitability of the firm (Lovelock and Wirtz, 2004). Hence, company personnel require a common understanding of what exactly the service is, and the meaning of quality (and “service quality”), to address such critical success issues as the measurement of service quality, the identification of causes of service quality shortfalls, and the design and implementation of corrective actions. A range of definitions of “service” surfaced between the 1960s and 1980s, and it was not until the mid-1990s which focused specifically on defining “service quality”. Definitions of service quality range through the disciplines of total quality management (Crosby, 1979) to marketing (Zeithaml et al., 1990).

As most developed economies are now service, rather than product-oriented, industries, service quality holds a prominent position in the marketing/management literature. Service quality is usually regarded as the customer’s impression of the relative inferiority/superiority of a service provider and its services (Bitner and Hubert, 1994; Tsoukas and Rand, 2006) to its competing alternative, and is often considered similar to the customer’s overall attitude towards the company (Parasuraman et al., 1988). It has therefore been popular to conceptualize and establish measures for service quality and explain its relation to the overall performance of companies and organizations; strong implications for management normally extend from these (Chau and Ngai, n.d.).

Based on this emerging trend, the present study is interested in the service gaps between customers’ expected service and perceived service in passenger airline services, and in ascertaining which demographic factors impact on the overall performance expectation of service quality in airlines. There is already a wealth of established extant literature on the variability of perceived and expected levels of service quality (Chang and Yeh, 2002; Chow and Luk, 2005), so it is interesting to explore how this can be extended to the airline context. Similarly, from neighboring studies of service quality (e.g. technology adaptation) Alfans and Sargeant (2000) find that it is possible to relate perceived service benefits to such demographic characteristics as gender, age and income levels. Therefore, two broad propositions are hypothesized:

*H1.* There is a statistically significant difference between expected service and perceived service in airline industry.
$H2$. (a) Gender; (b) age; (c) educational background; (d) occupation; and (e) disposable income – each has an influence on the overall performance expectation of service quality in the airline industry.

Models for measuring service quality
Measuring and evaluating service are not considered as add-ons to the overall delivery of service quality (Smith, 2003; Sigala and Christou, 2006; Voon, 2006; Witell and Lofgren, 2007), and there is already a well-established wealth of extant literature on measuring service quality, particularly in relation to customer expectations (e.g. Zeithaml et al., 1990). The meaning of service quality may vary in different contexts (for a review, see Ghobadian et al., 1994), but we refer to the view generally understood in marketing management. Probably the most prominent two are the disconfirmation model of service quality and the SERVQUAL model.

The disconfirmation model of service quality. The application of quality to the service context was driven in part by the early studies of Carlsmith and Aronson (1963) and Olshavsky and Miller (1972), and more recently Oliver (1997). The disconfirmation model assumes that when the consumer experiences something, s/he will understand it by his/her perception of what has happened. Hence, s/he deals with an assessment of the actual service experience and the reaction to it, to make a link between quality and satisfaction (Kasper et al., 2006). According to this model, if actual experience is better than expectation, there is satisfaction; if it was worse than expected, then there is dissatisfaction.

There are two sides to the model. First, the expectation, individual sources, and environmental sources, together explain their impact upon the level of tolerance in the disconfirmation model. Although the expectations can be positive, in that one looks forward to an event, but equally they can be bad; in which case, one does not look forward to the event. In either case, when one arrives at the service event, or a service transaction, one does so with some predisposition towards it. The second part of the disconfirmation model is making an assessment of what is actually received or experienced. The problem with service performance from the customer’s point of view is that it all depends upon the individual’s perception (and expectation), and individuals react differently in accordance with their personal needs, personality, and perception of events. These differences account for different tolerances to transactional considerations like length of queues, time delay, or reaction to other customers. There is also the likelihood that the customer will compare the level of service quality to that received in other similar public services (Grönroos, 2007).

Simply put, a customer derives quality assessment based on a comparison of expectations relative to performance:

- if expectations (E) are exceeded by perceived performance (P), i.e. $E < P$, then there is a position of customer delight;
- if expectations are not met by perceived performance ($E > P$) then there is customer dissatisfaction; and
- if expectations are met by perception ($E = P$), then the customer will assess the quality of the service to be satisfactory or as predicted (Grönroos, 2007; Bruhn and Georgi, 2006).
The disconfirmation idea has been subject to a series of refinements, and most significant is the identification that services are not one big amorphous event but instead comprises different components that interact to determine the overall level of their quality. The key dimensions of perceived quality are technical quality and functional quality. Technical quality refers to a dimension that describes what the customer gets as the outcome of their interaction with the organization (for example, an airline provides transportation to a destination), and functional quality refers to a dimension that describes the process by which the technical quality is delivered to the customer (for airlines, this may include airport facilities and in-flight comfort).

The SERVQUAL model of service quality. Zeithaml et al.’s (1990) SERVQUAL model is a common model for measuring service quality in different areas (for example: Chow and Luk, 2005; Kvist and Klefsjö, 2006; Tsoukas and Rand, 2006; Kang, 2006). The model (see also Parasuraman et al., 1985), which has been popular in use (eg. Robledo, 2001), makes a distinction between the customers and the organization which consists of managers and front office and back office employees working in various units and levels within the whole service organization. The model assumes that differences between the service desired by the customer and the service finally delivered by the service provider may be caused by the systemic disconfirmations, which comprise the following gaps:


The above four gaps concern the causes of poor service quality in the way the organization is managed. The user quality however is measured by gap-5 as the differences between the expectations and perceptions of customers. Gap-5 depends on the size and direction of the four disconfirmations associated with the delivery of service quality on the marketer’s side. If gaps 1 to 4 are reduced, then service quality can be improved (Kasper et al., 2006).

A consistent issue in managing service quality is the problem of identifying what comprises a service in order to determine the dimensions of the service customers use to assess quality (Lovelock and Wirtz, 2004). SERVQUAL began with developing a more concrete understanding of these service quality dimensions. Five such dimensions were eventually settled on (see for example, Kurtz and Clow, 1998; Johnston and Clark, 2005; Bruhn and Georgi, 2006):

1. Reliability (such as ability to perform the service dependably and accurately) include for example consistency in meeting service promises such as keeping schedules or appointment times, completing tasks on time, ensuring that outcomes are met.
2. Assurance includes competence, courtesy, credibility and security. This dimension would also include staff training in the use of tools and knowledge of their service processes, customer interaction, and the perception that the service is competent.
(3) **Tangibles** (such as appearance of physical facilities, equipment, and personnel) are the elements of the service environment impact upon perceived service quality for instance cleanliness of premises, staff appearance and the appropriateness of things like computers or phones.

(4) **Empathy** includes access, communication and understanding. This combined dimension is really about the communication style of the service organization through its service personnel, its communications including leaflets, instructions, signage and people management.

(5) **Responsiveness** (such as willingness to help customers) refers to the ability of the service to respond to individual customer requirements, such us specifying delivery times, altering aspects of the delivery process, and ensuring that customer remain involved.

The full SERVQUAL model contains a large part of the disconfirmation approach as discussed. However, that process is influenced by the four externalities: word-of-mouth, personal needs, past experience and external communications, which impact upon the formation of expectations. For each service dimension and the total service, a service quality (Q) judgment can be computed as perception (P) less expectation (E), i.e. $Q = P - E$ (where Q is represented by gap-5). These components can be calculated by the same five dimensions – reliability, assurance, tangibles, empathy, and responsiveness. Zeithaml *et al.* (1990) suggest 22 questions (items) that relate specifically to these five dimensions, on which we base the design of our study.

Indeed, there is already a wealth of existing literature on the likely factors that influence expected service quality: airline choice (Etherington and Var, 1984), customer satisfaction (Alotaibi, 1992), customer loyalty (Ostrowski *et al.*, 1993), class options (Makens and Marquardt, 1977), and total transportation offering (Morash and Ozment, 1994). However, in this paper we test specifically the four other externality factors from Zeithaml *et al.*’s (1990) SERVQUAL model, which per se have a stronger marketing focus, to explore further the applicability of gap-5 analysis to the airline context. We therefore hypothesize that:

**H3.** (a) Word-of-mouth; (b) personal needs; (c) past experience; and (d) external communications – is/are an important/contributing factor to customers’ expected service (and therefore choice) of an airline company.

**Customer satisfaction**

We hold to the view that customer satisfaction is the “... consumer’s fulfillment response ... a judgment that a product or service feature, or the product or service itself, provided (or is providing) a pleasurable level of consumption-related fulfillment, including levels of under- or over-fulfillment” (Oliver, 1997, pp. 13-14). Customer satisfaction (or dissatisfaction) is considered as an episodic measure, in that it relates to the most recent service experience, whereas service quality is a global evaluation of all past service experiences. Service quality evaluation is also recognized as an antecedent of customer satisfaction (Kurtz and Clow, 1998; Fournier and Mick, 1999). We also accept that price may be related to the overall customer understanding of service value (Chen *et al.*, 1994) – but we do not explore fully the impact of price because we limit our...
investigation to the gap-5 elements only – as well as accept there is impact from potentially other things, such as communication (Heinonen and Strandvik, 2005).

In simple terms, satisfaction is the result of customers’ assessment of a service based on a comparison of their perceptions of service delivery with their prior expectations. If customers’ perception of the service, the experience, and outcomes match their expectations then they should be satisfied. If their perceptions of the service exceed their expectations then they will be more than satisfied, even delighted; these, broadly speaking, are gap-5s. If their perceptions of the service do not meet their expectations then they may be dissatisfied, even disgusted or outraged (Andreassen, 2001; Schneider and Bowen, 1999). The value of a particular service provision is likely to be affected by the level of quality delivered and degree of customer satisfaction received. We hold to these terminologies in the conduct of this research. Hence, by taking the five dimensions of service quality as possible impacts, two sets of hypotheses are offered:

\[ H4. \] The gap-5s of (a) Reliability; (b) assurance; (c) tangibles; (d) empathy; and (e) responsiveness – have a significant impact on the overall customer satisfaction in the airline industry.

\[ H5. \] Overall service quality, and overall customer satisfaction, each has a significant impact on service value in the airline industry.

Cultural differences in perceptions and expectations
Lastly, we also hypothesize that: \( H6 \) there are cultural differences in the gap-5 problematics of the airline industry, and these are most noticeable between Eastern and Western cultures. This is because, from a basic international management and marketing perspective, the literature accepts there are differences in the way marketing effort should be employed, consistent with the service quality literature (Armstrong et al., 1997). More recently, the specific issues of service quality dimensions of multi-country and cross-cultural comparisons (Malhotra et al., 2005), between East (Korea) and West (USA) (Bang et al., 2005), in Hong Kong (Gilbert and Wong, 2003), and specifically in Taiwan (Cheng et al., 2008) have emerged. The extant literature has tended to advocate some national variation, but more is yet to be conducted outside of the transatlantic routes (Sultan and Simpson, 2000), where the present research is of contributory value.

The contribution of the present research is its focus on the airline services industry and the cross-cultural comparison between Taiwan (representing East) and the United Kingdom (representing West). We believe these two countries are a plausible representation of East and West because Taiwan (otherwise known as the Republic of China) has long benefited from its neighboring mainland China (the largest proportion of Eastern representation), and the United Kingdom has become ever closer to her European Union counterparts as she warmly opens her arms to new Union membership. We argue, without political prejudice, that these two countries provide a sound basis for comparison of customer perception and expectations of passenger airline service quality.

Methodology
The methodology is premised on examining our hypotheses, as shown in Figure 1.
Sample and data collection
A face-to-face administered questionnaire was used to collect data at major international airports of two countries’ capital cities – Taiwan Taoyuan International Airport (TPE – Taipei, Taiwan) and London Heathrow International Airport (LHR – London, UK). Respondents were approached at the check-in areas of the airport, especially during long queues to international locations. This is because it was felt that respondents are more likely to respond when queuing idly than if being engaged in some other activity. To avoid persistent interruption of, and annoyance to, passengers, a number of queues were targeted providing a range of passengers flying to different locations. The quota of at least 200 was targeted to ensure a statistically significant number of 50 (per airport) would be obtained (assuming a low response rate of only 50 per cent). We targeted about five airline queues per airport only, and these did not provide a wide cross-section of normal or budget airlines, so no analysis can be followed on passenger behaviors.

About 300 persons were approached, and 286 filled-in questionnaire responses were obtained, of which 263 were effective responses (all questions complete) over a two-week period of visiting the airports (four visits per airport, totaling eight visits), making an effective response rate of approximately 87.6 per cent. The sample is also a convenience sample because the researchers were based within close vicinity of the airports, which is an acceptable research strategy to employ (Zikmund, 1994). We obtained views from respondents who were aged above 18 years only for the reasons of not requiring additional ethical approval and the likelihood that the views would be mature and well thought through. We added the criterion of passengers who have experience of traveling by passenger airplane (because service quality is an episodic consideration, following the above literature). Further, to avoid complication regarding the extent to which the respondent data collected at the specific locations actually represent that country’s viewpoint, we collected data from outward passengers only who were ordinarily resident of the country.

The profiles of the respondents are shown in Table I. Roughly the same proportion of male and female responses was obtained (146 males, 117 females). The respondents were mainly aged 18-30 ($n = 160$: 61 per cent) followed by those aged 30-40 ($n = 70$: 27 per cent). They mostly had an undergraduate degree background ($n = 197$: 75 per cent) or a
sixth-form/high-school background \((n = 41; 16\text{ per cent})\); only two respondents left school at secondary level \((1\text{ per cent})\) and just one person held a doctoral qualification \((<1\text{ per cent})\). Strangely, the respondents were mostly associated with the marketing sector \((n = 109; 41\text{ per cent})\) or were involved in administration \((n = 41; 16\text{ per cent})\). About half the respondents earned the UK average income level \(£15,000-£30,000\) per annum) falling within the common tax band \((n = 129; 49\text{ per cent})\). Lastly, the number/proportion of questionnaire responses collected from the UK location was 161 \((61\text{ per cent})\) and the location in Taiwan was 102 \((39\text{ per cent})\). No other sample characteristic could be represented meaningfully due to insufficient responses.

Given the extent of restrictions to our data collection, we accept that our sample is not a good representation of the general airline passenger population: the reasonably young respondents on low/medium level incomes may be the consequence of our data collection taking place during the summer when there are generally more middle-aged leisure travelers than the usual proportion that is likely to include more business travelers on higher incomes. Our sample may also be the outcome of this younger respondent group being more willing to take part in the research (see Chau and Ngai, n.d.).

<table>
<thead>
<tr>
<th>Respondent groupings</th>
<th>Responses</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gender</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>146</td>
<td>56</td>
</tr>
<tr>
<td>Female</td>
<td>117</td>
<td>44</td>
</tr>
<tr>
<td>Age</td>
<td></td>
<td></td>
</tr>
<tr>
<td>18-30</td>
<td>160</td>
<td>61</td>
</tr>
<tr>
<td>30-40</td>
<td>70</td>
<td>27</td>
</tr>
<tr>
<td>40-50</td>
<td>27</td>
<td>10</td>
</tr>
<tr>
<td>50-60</td>
<td>6</td>
<td>2</td>
</tr>
<tr>
<td>&gt; 60</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Educational background</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(highest attainment level)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Secondary school</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>6th form/high school</td>
<td>41</td>
<td>16</td>
</tr>
<tr>
<td>Undergraduate</td>
<td>197</td>
<td>75</td>
</tr>
<tr>
<td>Postgraduate</td>
<td>22</td>
<td>8</td>
</tr>
<tr>
<td>Doctoral</td>
<td>1</td>
<td>&lt;1</td>
</tr>
<tr>
<td>Student</td>
<td>58</td>
<td>22</td>
</tr>
<tr>
<td>Education</td>
<td>11</td>
<td>4</td>
</tr>
<tr>
<td>Marketing</td>
<td>109</td>
<td>41</td>
</tr>
<tr>
<td>Accounting/finance</td>
<td>16</td>
<td>6</td>
</tr>
<tr>
<td>Retail</td>
<td>13</td>
<td>5</td>
</tr>
<tr>
<td>Purchasing</td>
<td>4</td>
<td>2</td>
</tr>
<tr>
<td>Research</td>
<td>10</td>
<td>4</td>
</tr>
<tr>
<td>Administration</td>
<td>41</td>
<td>16</td>
</tr>
<tr>
<td>Income (per annum, Pound Sterling, or stated in New Taiwanese Dollar equivalent in the Chinese version)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt; £15K</td>
<td>86</td>
<td>33</td>
</tr>
<tr>
<td>£15K-£30K</td>
<td>129</td>
<td>49</td>
</tr>
<tr>
<td>£30K-£45K</td>
<td>29</td>
<td>11</td>
</tr>
<tr>
<td>£45K-£60K</td>
<td>16</td>
<td>6</td>
</tr>
<tr>
<td>&gt; £60K</td>
<td>3</td>
<td>1</td>
</tr>
</tbody>
</table>

Table I. The profiles of the respondents and sample collection

<table>
<thead>
<tr>
<th>Location of data collection</th>
<th>Responses</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>UK (LHR)</td>
<td>161</td>
<td>61</td>
</tr>
<tr>
<td>Taiwan (TPE)</td>
<td>102</td>
<td>39</td>
</tr>
</tbody>
</table>
**Measures used in the questionnaire**

We used a standard blueprint of a flight that visualizes the phases of service delivery by an airline for a customer from the moment of arrival to immediately after flight landing (Bruhn and Georgi, 2006) to help formulate the data collection instrument for the research. Hence, the questionnaire comprised four parts – general questions, expectation factors, service quality, and overall satisfaction (see Appendix 1). The general questions were asked with the purpose of understanding the sample of respondents obtained and to later observe if such categories of respondents had any relation to overall customer satisfaction expectations. The options for each of the questions were based predominantly on British tiers of expectation (e.g. salary levels). The questionnaire was originally designed in English, and a (traditional, vis-à-vis simplified) Chinese version of it was also made available, after considerable attention during the translation process to ensure consistency of meaning; additional attention was also applied in ensuring that the categories related well to both the Taiwanese and Eastern contexts – so for example, “sixth-form college” was changed to “high school”, and the salary levels were not a direct currency conversion from sterling (GBP) into new Taiwanese dollars (NTD), but instead were readjusted to reflect upper and lower level tiers of Taiwanese income earners; and monthly, rather than annual, figures were used as this is more common in Taiwan. A pilot study of 15 respondents from the UK and 15 from Taiwan (30 in total) was conducted with friends and colleagues of the researchers to ensure that the questionnaire made good rhetorical and practical sense before it was deemed sound for public use. Some amendments were made after detailed consultation with the piloted respondents. Such amendments related mainly to the preamble of each section and concerned minor clarification issues; for example, it was unclear how to count the number of times a passenger had flown (to determine the level of experience) so a sentence request to count round trips as once only was subsequently inserted.

The response options of the remaining categories followed Zikmund’s (1994) claim of the simplicity to apply Likert-style bi-polar rating scales. The decision was therefore to apply five response options with associated ratings:

1. disagree/very unsatisfied;
2. tend to disagree/unsatisfied;
3. not sure/neutral;
4. tend to agree/satisfied; and
5. agree/very satisfied.

The choice of five scales was made to ensure simplicity, which proved beneficial for the circumstances under which the respondents were obtained, which itself was likely to have improved the response rate. These scales were used from part 2 of the questionnaire onwards, and the purpose of which was to measure customers’ expectations of the airline industry. Parasuraman et al.’s (1988) SERVQUAL model indicated that four factors affect customers’ expectation: word-of-mouth, personal needs, past experience, and external communications. These were therefore used to examine the main factors that influence passengers’ choice of airlines; the specific questions used for this were consistent with similar existing research on airlines (e.g. Chang and Yeh, 2002).

The same five-point Likert-like rating scale was used in part 3, but for each of the five measures and 23 items, the respondent was asked to rate the importance of each of
the service quality dimension as well as to rate how well the airline had performed, based on the respondent’s most recent complete experience (not the one s/he was about to experience). The measures were: reliability (six items), assurance (four items), tangibles (five items), empathy (four items), and responsiveness (four items), all based on the principal components of the same flight blueprint (Bruhn and Georgi, 2006). Immediate comparisons were therefore possible for each item, as a difference between customer expectation (the importance) and perceived service (performance), recognizable in the SERVQUAL literature as gap-5.

The fourth part asked the respondent to rate overall satisfaction in the airline industry. Only three questions were asked – all of which summarized the previous parts, asking if the respondents felt a high quality service was delivered, if they were on the whole satisfied, and if the service was good value-for-money. Although the questionnaire was fully confidential, as at no point was the respondent’s name asked, additional space was left at the bottom of the questionnaire for respondents who wished to be contacted later for further enquiries about the research; only 11 interested respondents left contact details, and such details were kept on a separate data file.

Findings and data analysis
Ensuring reliability
In addition to conducting the pilot study to ensure that the respondents would understand the questions, standard scale reliability tests (Cronbach alpha coefficients) were performed for the measures used in the questionnaire. This was to ensure that each scale reflected consistently the construct it was measuring. The overall Cronbach coefficients, as Table II shows, is lowest for part 1’s expectation factors (α = 0.668), which is less than the critical threshold of 0.7 as suggested by Nunnally (1978). As 0.668 is not too far off from the critical value, and the only “Cronbach alpha if deleted” for the external communications item (α = 0.684) is only a little larger than the overall Cronbach alpha coefficient, and as we had only four items for this measure, we chose not to remove this item. The overall Cronbach alpha coefficients for the other measures is above 0.7. Similarly, each of the item “corrected item total correlation coefficients” is above the critical value of 0.3. The high reliability values of this questionnaire may be because of the prior pilot test and because we used measures based on well-established prior SERVQUAL research (e.g. Parasuraman et al., 1988).

Service gaps in the airline industry
Our research hypothesis H1 focused on examining the size of the service gaps in the airline industry – that between perception and expectation. Hence, part 2 responses of the questionnaire were calculated for each item’s gap-5 values. For each of the measures, an average of its individual item gap-5 figures was calculated (shown in Table III). From this simple analysis, responsiveness has the largest average gap between the perceived standard and expected performance (0.510), followed by empathy (gap of 0.467), and reliability has the smallest gap-5 value (0.349). The compromise of expectation and perceived performance is least for airline service reliability. In all of the cases, average expectations are higher than average perceived performance.

To examine further whether these gap-5 values for each of the measures are statistically significant, we performed paired-sample t-tests for each of the five pairs. For each of the five pairs (of expected and perceived performances), the mean differences (value of gap-5s) are shown as large enough to be statistically significant: reliability (t = 10.125, df = 262, p < 0.01), assurance (t = 9.008, df = 262, p < 0.01),
tangibles ($t = 7.568, df = 262, p < 0.01$), empathy ($t = 9.908, df = 262, p < 0.01$), and responsiveness ($t = 11.332, df = 262, p < 0.01$). We therefore conclude that $H1$ can be supported that there are significant differences between expected service and perceived performance levels in the airline industry.

The relation of respondent profile to overall performance evaluation
Focusing on hypothesis $H2$, we then examined whether respondent characteristics, such as gender, age, etc, have any influence on the overall performance evaluation of service quality in the airline industry (Table IV). Three specific items, as suggested
<table>
<thead>
<tr>
<th>Pair</th>
<th>Item average</th>
<th>Mean Difference/Gap-5</th>
<th>Std. deviation</th>
<th>Std. error mean</th>
<th>95% Confidence interval of the difference</th>
<th>t</th>
<th>df</th>
<th>p (two-tailed)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Average reliability (importance) – Average reliability (performance)</td>
<td>4.457</td>
<td>0.349</td>
<td>0.559</td>
<td>0.034</td>
<td>10.12</td>
<td>262</td>
<td>0.000</td>
</tr>
<tr>
<td>2</td>
<td>Average assurance (importance) – Average assurance (performance)</td>
<td>4.282</td>
<td>0.373</td>
<td>0.672</td>
<td>0.041</td>
<td>9.00</td>
<td>262</td>
<td>0.000</td>
</tr>
<tr>
<td>3</td>
<td>Average tangibles (importance) – Average tangibles (performance)</td>
<td>4.083</td>
<td>0.405</td>
<td>0.868</td>
<td>0.053</td>
<td>7.56</td>
<td>262</td>
<td>0.000</td>
</tr>
<tr>
<td>4</td>
<td>Average empathy (importance) – Average empathy (performance)</td>
<td>4.195</td>
<td>0.467</td>
<td>0.765</td>
<td>0.047</td>
<td>9.90</td>
<td>262</td>
<td>0.000</td>
</tr>
<tr>
<td>5</td>
<td>Average responsiveness (importance) – Average responsiveness (performance)</td>
<td>4.254</td>
<td>0.510</td>
<td>0.730</td>
<td>0.045</td>
<td>11.33</td>
<td>262</td>
<td>0.000</td>
</tr>
</tbody>
</table>
significant by Hanlon (1999), were considered as part of this measure: overall high service quality, overall satisfaction, and overall service as value-for-money. We therefore took the characteristics as given in part 1 of the questionnaire and considered them against the items of part 4. As such respondent characteristics are categorical and mutually exclusive, and as the possible responses in the questionnaire are also mutually exclusive, we performed Chi-squared tests of association (of course, we made some very vital assumptions in the use of this statistical technique: that the possible responses are different enough for the respondent to decide clearly which to choose, and that the categories are factually different, which is questionable in cases like the level of education, as well as the respondent being certain into which the correct response should fall, so that the variables may be deemed categorical).

The results show a mix of outcomes. Gender does not have any relationship to any of the perceived performances; this is unsurprising, as there is no obvious reason why one gender would judge a service quality dimension any differently from the other. The age variable generally does not seem to have any association with perceived performance, despite some general expectation that age is normally influenced by past experiences; age does for overall service value show a marginal association ($\chi^2 = 20.97$, df = 12, $p = 0.051$), but we suggest that this is too weak to conclude anything significant. The same marginal association is evident in the association of education with overall service quality ($\chi^2 = 26.03$, df = 16, $p = 0.054$) but some association of education with overall service value ($\chi^2 = 47.28$, df = 32, $p = 0.04$), and no association of education with overall customer satisfaction ($\chi^2 = 46.88$, df = 32, $p = 0.063$). With occupation, there seem to be associations with overall service quality ($\chi^2 = 47.28$, df = 32, $p = 0.04$) and with overall customer satisfaction ($\chi^2 = 46.88$, df = 32, $p = 0.044$), although none with overall service value ($\chi^2 = 30.73$, df = 32, $p = 0.531$). Lastly, the results show that income levels have an association with the perception of overall service quality ($\chi^2 = 46.39$, df = 16, $p < 0.01$), overall customer satisfaction ($\chi^2 = 37.33$, df = 16, $p < 0.01$) and overall service value ($\chi^2 = 26.35$, df = 16, $p < 0.05$).

<table>
<thead>
<tr>
<th>Pearson Chi-square, $\chi^2$ value, df, $p$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gender → Overall service quality: 1.558, 4, 0.816</td>
</tr>
<tr>
<td>Gender → Overall customer satisfaction: 6.381, 4, 0.172</td>
</tr>
<tr>
<td>Gender → Overall service value-for-money: 4.467, 4, 0.346</td>
</tr>
<tr>
<td>Age → Overall service quality: 14.091, 12, 0.295</td>
</tr>
<tr>
<td>Age → Overall customer satisfaction: 18.515, 12, 0.101</td>
</tr>
<tr>
<td>Age → Overall service value-for-money: 20.971*, 12, 0.051</td>
</tr>
<tr>
<td>Education → Overall service quality: 26.032*, 16, 0.054</td>
</tr>
<tr>
<td>Education → Overall customer satisfaction: 25.415, 16, 0.063</td>
</tr>
<tr>
<td>Education → Overall service value-for-money: 47.276**, 32, 0.040</td>
</tr>
<tr>
<td>Occupation → Overall service quality: 47.276**, 32, 0.040</td>
</tr>
<tr>
<td>Occupation → Overall customer satisfaction: 46.873*, 32, 0.044</td>
</tr>
<tr>
<td>Occupation → Overall service value-for-money: 30.728, 32, 0.531</td>
</tr>
<tr>
<td>Income → Overall service quality: 46.387**, 16, 0.000</td>
</tr>
<tr>
<td>Income → Overall customer satisfaction: 37.221**, 16, 0.002</td>
</tr>
<tr>
<td>Income → Overall service value-for-money: 26.349*, 16, 0.049</td>
</tr>
</tbody>
</table>

**Notes:** * = $p < 0.05$; ** = $p < 0.01$ (two-tailed)

Table IV. Chi-square tests ($H2$)
Customer expected service in the airline industry

Based primarily on hypotheses H3, we examined what are the main factors that influence passengers’ choice of an airline. For this, we explored the inter-correlations between each of the variables identified as potential causes of customer expected service (Zeithaml et al. 1985; 1990). We further performed a multiple regression predictor analysis on these variables with each service quality dimensions, reliability, assurance, tangibles, empathy, and responsiveness (Bruhn and Georgi, 2006). These are shown in Tables V and VI.

The inter-correlation results show significant positive associations between the expectation variables (Table V); so they are not mutually exclusive, although the correlation coefficients are rather low. The correlation between personal needs and past experience is the strongest \((r = 0.605, p < 0.01)\), followed by that of past experience with word-of-mouth \((r = 0.370, p < 0.01)\). This is probably because humans are sentimental to their past behaviors, and learn from good or bad experiences that in turn shape the nature of their future needs. The smallest inter-correlation is external communications with personal needs \((r = 0.173, p < 0.01)\). This is probably because a personal need takes priority over how a message is communicated about the possibility of such a need.

In examining the impact of the expectation factors on expected service, five regression models were performed in the standard form shown as follows, where REL (reliability = model 1), ASSUR (assurance = model 2), TANG (tangibles = model 3), EMPAT (empathy = model 4), and RESPON (responsiveness = model 5) were regressed against WOM (word-of-mouth), PN (personal needs), PE (past experience), and EC (external communications), with the constant \(\beta_0\) and error term \(\epsilon_i\): \[ REL = \beta_0 + \beta_1 \text{WOM} + \beta_2 \text{PN} + \beta_3 \text{PE} + \beta_4 \text{EC} + \epsilon_i \] (model 1)

For simplicity, we assumed a linear relationship, and applied the forced-entry mode in the regression analysis, as these were already well supported by the extant SERVQUAL literature. This is also supported empirically in all the cases by a high \(F\)-statistic significance \((p < 0.01\) for all) and \(R^2\) coefficient (lowest \(R^2 = 0.311\) for EMPAT; highest \(R^2 = 0.411\) for REL) (see Table VI). We refer in this paper to “standardized beta coefficients” as we trust the well-established literature on service quality, and in this way we have good comparators of variables within each model’s sample.

In model 1, WOM \((\beta = 0.208, p < 0.01)\), PN \((\beta = 0.323, p < 0.01)\) and PE \((\beta = 0.273, p < 0.01)\) show to have a significant impact on REL, but EC \((\beta = -0.001, p = 0.977)\) shows no relationship at all. Similarly, with ASSUR (model 2), WOM \((\beta = 0.179, p < 0.01)\), PN \((\beta = 0.377, p < 0.01)\), and PE \((\beta = 0.156, p = 0.017)\) are significant predictors, and EC is insignificant \((\beta = 0.064, p = 0.234)\). In model 3, TANG has only two significant predictors - PN \((\beta = 0.355, p < 0.01)\) and PE \((\beta = 0.201, p < 0.01)\), and WOM \((\beta = 0.084, p = 0.152)\) and EC \((\beta = 0.071, p = 0.199)\) are insignificant. With EMPAT (model 4), WOM \((\beta = 0.122, p < 0.05)\), PN \((\beta = 0.335, p < 0.01)\) and PE \((\beta = 0.203, p < 0.01)\) are significant, and EC \((\beta = 0.050, p = 0.368)\) is an insignificant predictor. Lastly, with RESPON (model 5), again only WOM \((\beta = 0.140, p < 0.05)\), PN \((\beta = 0.279, p < 0.01)\) and PE \((\beta = 0.253, p < 0.01)\) are significant, and EC is an insignificant predictor \((\beta = 0.056, p = 0.308)\). From this analysis, we can only partly accept H3, and in all cases must reject H3d, and conclude that external communications are not a contributing factor to customers’ expected service in the airline industry.
<table>
<thead>
<tr>
<th></th>
<th>Mean</th>
<th>Std dev</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>10</th>
<th>11</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. WoM</td>
<td>3.71</td>
<td>0.989</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2. PN</td>
<td>4.11</td>
<td>0.935</td>
<td>0.302**</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3. PE</td>
<td>4.17</td>
<td>0.891</td>
<td>0.370**</td>
<td>0.605</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4. EC</td>
<td>3.43</td>
<td>0.896</td>
<td>0.343**</td>
<td>0.173**</td>
<td>0.216**</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5. gREL</td>
<td>0.349</td>
<td>0.558</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6. gASSUR</td>
<td>0.379</td>
<td>0.696</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7. gTAN</td>
<td>0.405</td>
<td>0.868</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8. gEMPAT</td>
<td>0.477</td>
<td>0.781</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9. gRESPON</td>
<td>0.525</td>
<td>0.746</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10. OSV</td>
<td>3.54</td>
<td>1.003</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>11. OSQ</td>
<td>3.77</td>
<td>0.826</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>12. OCS</td>
<td>3.75</td>
<td>0.804</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Notes: * = p < 0.05; ** = p < 0.01 (two-tailed)
Based on hypothesis H4, we then examined the impact of the customer gap-5 values (in terms of the five service quality dimensions) on overall consumer satisfaction. Again, the inter-correlations of these gap-5 values were first explored (as shown in Table V) before applying the regression analysis. All of the inter-correlations are significant, ranging from tangibles and assurance \((r = 0.770, p < 0.01)\) with the strongest correlation, to empathy and reliability \((r = 0.608, p < 0.01)\) with the weakest correlation. In other words, there is a close relationship between the way customers regard a difference between expected and perceived levels of service quality in terms of the five quality dimension measures. It is therefore questionable if these measures are broad enough to cover the full range of airline services if they are so closely related.
The results from the regression analysis are somewhat different in these gap-5 values in predicting overall customer satisfaction (see Table VI). The following basic linear regression model was employed ($R^2 = 0.205, F = 13.237, p < 0.01$), and for the same reasons of being well-established in the literature as explained before we force-fitted the regression in one go:

\[ \text{OCS} = \beta_0 + \beta_1 g_{\text{REL}} + \beta_2 g_{\text{ASSUR}} + \beta_3 g_{\text{TANG}} + \ldots + \beta_5 g_{\text{EMPAT}} + \epsilon_{vi} \]  

where overall customer satisfaction (OCS) is predicted by the gap-5 of reliability ($g_{\text{REL}}$), plus the gap-5 of assurance ($g_{\text{ASSUR}}$), plus the gap-5 of tangibles ($g_{\text{TANG}}$), plus the gap-5 of empathy ($g_{\text{EMPAT}}$), plus the gap-5 of responsiveness ($g_{\text{RESPON}}$), with the standard error term $\epsilon_{vi}$. The results (Table VI) show clearly that $g_{\text{ASSUR}} (\beta = 0.258, p = 0.012)$ and $g_{\text{RESPON}} (\beta = 0.370, p < 0.01)$ are sound predictors of OCS, whereas $g_{\text{REL}} (\beta = 0.83, p = 0.314)$, $g_{\text{TAN}} (0.068, p = 0.488)$ clearly are not. $g_{\text{EMPAT}}$ is marginal, as its critical significance value sits on the borderline of the 95 per cent level ($\beta = 0.174, p = 0.067$). We can therefore only partially accept $H4$.

**Overall customer service evaluation**

The overall customer evaluations at the last part of the questionnaire were also explored for how they interrelated with each other (as shown in the inter-correlation coefficients in Table V, items 10-12). The results show a reasonable association with each other, with the correlation between overall service quality (OSQ) and overall customer satisfaction (OCS) being strongest ($r = 0.722, p < 0.01$). In connection with hypothesis $H5$, we further regressed overall service value (OSV) with OSQ and OCS as predictors (Table VI, model 7) as follows:

\[ \text{OSV} = \beta_0 + \beta_1 \text{OSQ} + \beta_2 \text{OCS} + \epsilon_{vi} \ldots \]  

(model 7)

In this way, we explored whether the well-established view that the term “value-for-money” is really the product of the level of service quality received and on the whole how satisfied the customer is with the whole service experience. Again, we assumed linearity and force-regressed them ($R^2 = 0.528, F = 145.3, p < 0.01$). The results confirm this view as both OSQ ($\beta = 0.310, p < 0.01$) and OCS ($\beta = 0.471, p < 0.01$) are significant, and so we find that $H5$ is supported.

**A bridge between East and West?**

Lastly, hypothesis $H6$ was to observe from the research results and find the most problematic in terms of how well they support the hypothesis. The gap-5 values impacting on overall customer satisfaction ($H4$) was rather problematic. We therefore examined the impact of gap-5 values on overall service quality (instead of overall customer satisfaction, as this study is interested in the quality issues) and then explored whether the results are different for data collected at the UK and Taiwan locations. The values of the gap-5s were first compared using the independent sample $t$-test (see Table VII). Despite some notable differences in the mean values between the 161 UK and 102 Taiwanese respondents, for each of the gap-5 values, no statistically significant difference was found. The Levene’s test for equality of variances confirms the assumption of equal variances of the data in the two contexts for all of the gap-5 values.
Table VII. Independent sample t-test of Gap-5 values for UK and Taiwanese respondents

<table>
<thead>
<tr>
<th>Group statistics</th>
<th>LOC</th>
<th>n</th>
<th>Mean</th>
<th>Std. deviation</th>
<th>Std. error mean</th>
<th>Levene's test for equality of variances</th>
<th>Independent samples test</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>F</td>
<td>Sig.</td>
</tr>
<tr>
<td>gREL</td>
<td>UK</td>
<td>161</td>
<td>0.391</td>
<td>0.572</td>
<td>0.045</td>
<td>2.295</td>
<td>0.131</td>
</tr>
<tr>
<td></td>
<td>Taiwan</td>
<td>102</td>
<td>0.283</td>
<td>0.532</td>
<td>0.053</td>
<td></td>
<td></td>
</tr>
<tr>
<td>gASSUR</td>
<td>UK</td>
<td>161</td>
<td>0.395</td>
<td>0.716</td>
<td>0.056</td>
<td>1.644</td>
<td>0.201</td>
</tr>
<tr>
<td></td>
<td>Taiwan</td>
<td>102</td>
<td>0.355</td>
<td>0.668</td>
<td>0.066</td>
<td></td>
<td></td>
</tr>
<tr>
<td>gTANG</td>
<td>UK</td>
<td>161</td>
<td>0.419</td>
<td>0.922</td>
<td>0.073</td>
<td>2.666</td>
<td>0.104</td>
</tr>
<tr>
<td></td>
<td>Taiwan</td>
<td>102</td>
<td>0.384</td>
<td>0.781</td>
<td>0.077</td>
<td></td>
<td></td>
</tr>
<tr>
<td>gEMPAT</td>
<td>UK</td>
<td>161</td>
<td>0.507</td>
<td>0.786</td>
<td>0.062</td>
<td>0.000</td>
<td>0.998</td>
</tr>
<tr>
<td></td>
<td>Taiwan</td>
<td>102</td>
<td>0.430</td>
<td>0.777</td>
<td>0.077</td>
<td></td>
<td></td>
</tr>
<tr>
<td>gRESPON</td>
<td>UK</td>
<td>161</td>
<td>0.558</td>
<td>0.769</td>
<td>0.061</td>
<td>1.257</td>
<td>0.263</td>
</tr>
<tr>
<td></td>
<td>Taiwan</td>
<td>102</td>
<td>0.474</td>
<td>0.708</td>
<td>0.070</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
values (for \( g_{\text{REL}} \), \( F = 2.295, p = 0.131 \); for \( g_{\text{ASSUR}} \), \( F = 1.644, p = 0.201 \); for \( g_{\text{TANG}} \), \( F = 2.666, p = 0.104 \); for \( g_{\text{EMPAT}} \), \( F = 0.0000, p = 0.998 \); for \( g_{\text{RESPON}} \), \( F = 1.257, p = 0.263 \)), and so we cannot conclude any significant differences between the data collected at the two locations, or any differences in the points of view they represent.

We then examined the impact of the gap-5 values on overall customer service quality for the complete dataset (for gap-5 values of \( g_{\text{REL}} \), \( g_{\text{ASSUR}} \), etc) as well as separately for those respondents collected at the UK location and for the Taiwan location. The linear regression model was therefore in the form:

\[
\text{OSQ}_B = \beta_0 + \beta_1 g_{\text{REL}} + \beta_2 g_{\text{ASSUR}} + \beta_3 g_{\text{TANG}} + \ldots
\]

\[
+ \beta_4 g_{\text{EMPAT}} + \beta_5 g_{\text{RESPON}} + \epsilon_{viii}
\]

and similarly for the UK context (model 9, \( \text{OSQ}_U = f [g_{\text{REL}}, g_{\text{ASSUR}} \ldots \text{etc}] \)) and the Taiwan context (model 10, \( \text{OSQ}_T = f [g_{\text{REL}}, g_{\text{ASSUR}} \ldots \text{etc}] \)) (Table VIII).

In the model \( \text{OSQ}_B \), only \( g_{\text{RESPON}} \) (\( \beta = -0.409, p < 0.01 \)) shows a significant negative impact on overall customer service. This result is the same for model \( \text{OSQ}_U \) (for \( g_{\text{RESPON}} \) (\( \beta = -0.367, p < 0.01 \)). There is a slight difference in the \( \text{OSQ}_T \) model where in addition to \( g_{\text{RESPON}} \) being significant (\( \beta = -0.449, p < 0.01 \)), \( g_{\text{RESPON}} \) (\( \beta = 0.282, p = 0.056 \)) is also positively significant. This means that generally the overall service quality falls as the size of the perceived difference between expected and received quality increases for items relating to responsiveness; but in the Taiwanese context, there is a positive impact on overall service quality as the gap in

<table>
<thead>
<tr>
<th>Model summary</th>
<th>ANOVA</th>
<th>Regression coefficients</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model</td>
<td>( R^2 )</td>
<td>( F )</td>
</tr>
<tr>
<td>8 OSQB</td>
<td>0.193</td>
<td>12.279</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9 OSQUK</td>
<td>0.194</td>
<td>7.461</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10 OSQTW</td>
<td>0.229</td>
<td>5.705</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table VIII. Results of Linear Regression Models 8 to 10

Notes: * = \( p < 0.05 \); ** = \( p < 0.01 \) (two-tailed)
perceived and expected service increases for items relating to reliability. This is an interesting finding, but must be identified with much caution as its level of statistical significance is only marginal ($p = 0.056$), and this finding alone therefore cannot be sound enough to support hypothesis $H6$ that there is a significant difference between the expectations between East and West.

**Discussion and managerial implications**
The results above generally support our prior hypotheses, with perhaps the exceptional finding being that there is no significant difference in the expectations of service level between Taiwanese and British respondents ($H6$); marketing effort therefore need not be different for these two consumer groups (Eastern and Western consumers). The present research cannot afford scope for understanding fully why this is so; the focus is on the characteristics of the passenger airline services, in relation to how this is contributory to the extant literature. Nonetheless, team strategy, company performance and organizational effectiveness are interlinked (for an exposition, see Chau, 2008) we see the need to offer useful implications for those involved directly in the management or marketing of airline-related commodities; we discuss these further in this section.

*The meaning of quality in the airline industry*
We find that our study conforms to the predispositioning literature on how quality is understood in the airline industry (Sultan and Simpson, 2000), and our empirical data seem to have adapted well to the gap-5 interpretation of perceived and expected levels of service quality. Our results, based on $H1$, show that there is a significant difference in the levels of expected (indicating the degree of importance of the) performance and perceived performance (that felt delivered to them) across all the dimensions relating to passenger airline provision. These are particularly evident in services relating to departure and landing, rather than emergency and safety concerns. This may be due to customers not generally expecting the need for such emergency services and therefore form their expectations on those services to which they relate more closely and more frequently. For the marketing manager, it is suggested that the advertisement of the airline should be focused on those items, which are most commonly used as part of an airline journey. These could include targeting ancillary services, such as special car hire offers and specific coach journeys that are only available to those who use a particular airline, or emphasizing the convenience in using the airline’s check-in procedures (this could be in the form of pre-arrival check-in through the internet and simple bag-drop facility).

Our data analysis, based on $H2$, does not suggest strongly that demographic factors relate to customers’ evaluation of service quality – with only income levels and occupation being significant. This is perhaps not an unusual finding, as airline companies have long separated different classes of service on aircraft – such as economy class, first class and business class – to capture the range of customers. A managerial implication for marketers and advertisers is to place different effort on the different segment groups. So for example, in targeting the lower-income groups, it may be more appropriate to advertise a selection of drinks that are most commonly associated with them, such as beers and lagers, and then those drinks such as champagne with the higher-class travelers; the same can be applied with other
commodities, such as the inclusion of tabloid newspapers in the economy class cabins and broadsheet-presses (e.g. Financial Times) in the first-class and business-class cabins. The purpose is to recognize that there are different ways to improve customer satisfaction for different groups of travelers, and overall, this should lead to repeat purchases. Further, any other promotional offers should also recognize class differences; a recent marketing strategy of many airline companies (e.g. British Airways, Virgin Atlantic, and American Airlines) has been to offer a range of credit cards linked to reward points to gain loyalty, each linked to the customer’s affordability and the kinds of rewards most commonly associated with their class interests.

**Appropriateness of dimensions for measuring service quality**

The five dimensions from the extant SERVQUAL literature seem to support well the questions that relate to the range of services in the airline industry. The research also finds generally that our predisposing factors that might influence customer expectations are good predictors of the five service quality dimensions, as shown in our examination of H3. In all of these cases, only the factor of “external communications” is not significant. As an implication for management, this is surprising, given the amount of money airline companies invest in advertising and promotions. It seems that factors that relate more closely to the customers are more effective – such as past experience and word-of-mouth from friends and relatives. This is consistent with the service quality dimension literature which regards matters of “assurance” of greatest importance (Gilbert and Wong, 2003). It is a sensible suggestion to airline companies, that delivering a good service to existing customers is more effective than retaining new ones from advertising. A possible managerial approach may be to keep a good database of customers, and to re-attract them through email offers and newsletters of similar products; the idea is to “assure” them of the continued service quality the airline can offer. The old cliché is probably true – that, it costs five times more to obtain a new customer than to retain an existing one!

**Customer satisfaction, quality and value**

Our suggestion is further strengthened by our examination of H4, which finds that two of the gap-5s of those service dimensions used impact upon overall customer satisfaction; these are the gap-5 of assurance and the gap-5 of responsiveness. In managerial terms, focusing more on the issues that relate to assurance (e.g. luggage delivery) and responsiveness (e.g. a professional service) will have the largest improvement in customers’ levels of satisfaction, leading to repeat purchases. For the marketing manager, the responsibility is not just on managing these areas, but also to publicize such a reputation. Early 2008 saw an internationally publicized incident of poor luggage delivery when the new terminal 5 at London Heathrow Airport, operated by British Airways, caused days of delay to thousands of travelers and hundreds of flights cancelled. Such incidents are likely to widen the gap-5s, and therefore should undoubtedly be avoided.

Lastly, our examination of H5 finds that there is a general link between customer satisfaction, quality and how customers determine overall value in the service. This means that, as satisfaction is a determinant of ultimate profitability (an aim of a company), it is important to ensure that quality and value are maintained in the
services – or at least, to the extent that the customers believe this is the case. As the quality management literature (e.g. Crosby, 1979) confirms that customers are the ultimate determinants of the meaning of quality, it is important for the airline company (perhaps through email questionnaires) to understand what customers believe important, so that the airline company can keep a monitor on what is being delivered, against what customers perceive is being delivered, thus reducing the gap-5s in these areas.

Conclusions and limitations
The paper has explored the general areas of understanding the meaning of quality, establishing dimensions for measuring service quality, and their relation to customer satisfaction, quality and value, in the airline industry. The research finds that the gap-5 understanding of service quality is applicable for the airline context, with a strong tendency of all the aggregate expected levels of service responses turning out greater than those of perceived levels (perceived as delivered based on experience); following the literature, the position of “customer dissatisfaction” is achieved for all of the five service quality dimensions (Kurtz and Clow, 1998) – that is, the level of expected performance (or the importance of that service level) is higher than that of perceived performance. Of course, this is not to say that customers are in general “dissatisfied” in the ordinary conversational sense of the word, but rather that there is sufficient (and statistically significant) scope for improvement in respect of the delivery of service quality in these areas. On balance, these service dimensions can be argued to impact customers’ evaluation of an airline, and therefore the choice of which airline to use. We also find that customer satisfaction and service quality are interlinked, and these relate to service value – that is, customers also make decisions based on their overall judgment of the overall quality of the service and how well this relates to the price paid (or whatever foregone for it). However, we do not find any cultural differences between the views of respondents originating from either the East or the West; a gap-5 analysis does not seem to help provide adequate insights in this respect.

Our research was conducted specifically in the airline industry, but the findings are generalizable to other contexts that share such similar characteristics as tight regulation, open competition, and where service quality is paramount. In the UK context, the most obvious examples are domestic energy suppliers (electricity and gas suppliers), telecommunications/internet providers, and banks/financial institutions (Chau, 2002; Chau and Ngai, n.d.). In each of these service sectors, the customer has the choice of a number of providers and the freedom to switch between them, should s/he not be satisfied with the level of service quality received. They are regulated (as the service is considered important in each of their own right) by the respective regulators, Ofgem (Office of Gas and Electricity Markets), Ofcom (Office of Communications), and the FSA (Financial Services Authority). The present findings offer important and potentially transferable management implications for these sectors, as for example in energy the SERVQUAL literature has yet to be consulted in deciding on new policy standards (see Chau, 2009).

Indeed, our research is of course not free from limitations. While we have taken care to ensure rigor in our methodological design, using measures from seminal (Zeithaml et al., 1990) as well as popular literature (e.g. Robledo, 2001; Bruhn and Georgi, 2006), our research is limited only to the responses of 263 respondents at the two airport
locations. The airline industry is enormous, and our study can present only one view. The collection of our data took place over one summer is probably limited to a specific leisure group; a more representative sample of airline travelers is better sought by data collection taking place over a longer span lasting throughout all four seasons of the calendar year.

Additionally, we may have mildly exaggerated our claim about the potential bridging of any significant divergence (although we find none), or gap-5s, between the respondents originating from Taiwan and the UK, and the extent to which any such findings represent adequately views from East and West. Our samples were conveniently employed; and for reasons of sensitivity and confidentiality, no additional conditions were maintained to ensure that the subjects represented any strong views of the cultures of Taiwan/East or the UK/West. Further, as our research is premised on the convenience data sample, any findings from the present research do not account for any changes over time, from which the employment of longitudinal studies would provide a stronger promise (Chau and Witcher, 2005b), and were not timed to capture any specific economic or political influence of airline management. For now, as no bridge is built over troubled water, we advocate and look forward to a further long and winding road of future SERVQUAL research on managing service quality.

References
Bor, R. (2003), Passenger Behaviour, Ashgate Publishing Company, Burlington, VT.


Further reading


Appendix: Questionnaire

Part I:
Please tick the items below that describe you:

1. Gender □ Male □ Female
2. Age group □ 18-30 □ 30-40 □ 40-50 □ 50-60 □ < 60
3. Educational background □ Secondary School □ 6th Form □ Undergraduate □ Postgraduate □ Doctoral
Part II:
Please indicate how strongly you agree with the following statements: 1, 2, 3, 4, 5

7. Word of mouth has a big influence in my choice of airlines.
8. Personal needs has a big influence in my choice of airlines.
9. Past experience has a big influence in my choice of airlines.
10. External communications has a big influence in my choice of airlines.

(Such as editorial coverage in broadcast or print media)

Part III:
Please indicate how strongly you agree with the following statements (Experience/Importance: 1, 2, 3, 4, 5; Expected Performance: 1, 2, 3, 4, 5):
(please state the airline (only if you wish) _____)
[Reliability]
11. Check in was carried out quickly and accurately.
12. Flight departure and arrival was on time.
13. Flight safety record was excellent.
14. Had no worries about luggage getting lost, especially when needing to make a transfer flight.
15. Few machinery problems.
16. Handling customers’ complaints was with a sincere attitude.

[Assurance]
17. Luggage delivery and luggage claims were perfect and speedy.
18. Service staff are professional when dealing with customers’ enquires.
19. Competent at answering customers’ general queries.
20. The airline has good reputation and image.

[Tangibles]
21. Seating in the passenger cabin was spacious and comfortable.
22. The type of aircraft used is modern.
23. The appearance of service staff was neat and tidy.
24. In-flight entertainment was satisfactory.
25. The facilities in the passenger cabin are modern and clean.

[Empathy]
26. The airline takes consumer views seriously.
27. Service staff actively used the passenger’s native language.
28. Flight times were appropriate.
29. The airline maintains a customer service quality control system.

[Responsiveness]
30. Service staff are responsive in meeting customer demands’.
31. Messages on the flight were broadcasted clearly and effectively.
32. Service staff offered services efficiently.
33. Staff was capable of handling emergency services effectively.

Part IV:
Please indicate how strongly you agree with the following statements: 1, 2, 3, 4, 5
34. Overall, I feel the airline offers high quality services to passengers.
35. Overall, I feel satisfied with the airline’s current provision of services.
36. Overall, the service is good value for money.

About the authors
Vinh Sum Chau is Lecturer in Strategy & Strategic Management at the Norwich Business School, University of East Anglia, and faculty member of the UK ESRC Centre for Competition Policy. He is also Chair of the British Academy of Management’s special interest group and conference track on Performance Management. He has taught and published widely in the areas of performance management, strategy implementation, service quality, and customer satisfaction. Vinh Sum Chau is the corresponding author and can be contacted at: v.chau@uea.ac.uk

Yu-Ying Kao was formerly at the Norwich Business School, University of East Anglia, where she obtained her postgraduate qualification and worked on a substantial part of the dataset on which this paper is premised.

To purchase reprints of this article please e-mail: reprints@emeraldinsight.com
Or visit our web site for further details: www.emeraldinsight.com/reprints