Open Research Online

The Open University's repository of research publications
and other research outputs

Assessing the impact of response styles on cross-cultural service quality evaluation: a simplified approach to eliminating the problem

Journal Item

oro.open.ac.uk

Assessing the Impact of Response Styles on Cross-Cultural Service Quality Evaluation: A Simplified Approach to Eliminating the Problem

Nina L. Reynolds and Anne M. Smith

Nina L. Reynolds, Professor of Marketing

Address: Bradford University School of Management, Emm Lane, Bradford, BD9 4JL, United Kingdom.

Email: n.l.reynolds@bradford.ac.uk

Tel: +44(0)1274 234393

Fax: +44(0)1274 546866

Anne M. Smith, Reader in Marketing

Address: Centre for Strategy and Marketing, The Open University Business School, Michael Young Building, Milton Keynes, MK7 6AA, United Kingdom.

Email: anne.smith@open.ac.uk

Tel: +44(0)1908 858672

Key words: comparative research, response styles, service quality

Assessing the Impact of Response Styles on Cross-Cultural Service Quality

Evaluation: A Simplified Approach to Eliminating the Problem

ABSTRACT

With the proliferation of comparative research, it is important to recognize some of the inherent limitations of cross-cultural measurement. This paper examines the impact of response styles on substantive conclusions of cross-cultural service quality research. We use relatively simple analysis methods in conditions where more sophisticated approaches are unlikely to be robust. We demonstrate how ANCOVA and partial regression can be used to assess both differences in mean scores and differences in relationships. Our results demonstrate that conclusions drawn from analysis that ignores the potential impact of response styles differ from those drawn when response styles are considered. For researchers our findings imply that attempts to understand and explain cultural differences in service quality expectations, and relationships between perceptions and overall quality assessments, may be impeded by the presence of response styles. A further impact relates to the assessment of 'gaps' or a 'zone of tolerance' in service quality evaluation. For managers our conclusions have implications relating to the use of research findings as a basis for market segmentation, service design, staff training and other resource allocation decisions. In particular we question the use of such research as a basis for comparative service evaluation across cultures.

INTRODUCTION

This paper addresses a subject of considerable concern to cross-cultural service researchers, that is, how to eliminate the impact of response styles from substantive cross-cultural research findings. Managers and researchers often consider comparable scores as indicative of comparable attitudes. However, non-comparability can arise from two sources. The first is associated with the *content* of the question. Here there is 'true' variation reflecting substantive differences of interest such as different levels of perceived service quality across cultures. The second relates to respondents' unconscious reactions to the method of presenting the item and is a form of method bias known as 'response style' (Cronbach 1950). Importantly for our study, response styles have been shown to vary across cultures (Bresnahan et al. 1999; Chen, Shin-ying, and Stevenson 1995; Clarke 2000; Diamantopulos, Reynolds, and Simintiras 2006; Harzing 2006; Johnson et al. 2005; Si and Cullen 1998). In particular we examine a potential solution to assessing the impact of response styles. This involves relatively uncomplicated analysis techniques that can be applied to comparatively small data sets (Podsakoff et al. 2003). More sophisticated techniques often need large samples and may also require that response styles are measured independently (De Jong et al. 2008; Weijters, Schillewaert, and Geuens 2008). Where this is not possible, not accounting for response style contamination would leave a potential rival hypothesis for cross-cultural research findings.

This study focuses on the impact of response styles on cross-cultural service quality research. Here studies are dominated by survey methods using assumed interval scales, similar to SERVQUAL (Parasuraman, Zeithaml and Berry 1994a), for the measurement of expectations

and perceptions. Such scales tend to be the most problematic in terms of encouraging 'response styles' (Smith and Reynolds 2002). Theoretical understanding of service quality evaluation across cultures is impeded if the conclusions of comparative studies reflect response style effects rather than true variations between cultural groups. Similarly, managers may use such findings to make decisions with respect to resource allocation, market segmentation or service development. Considerable resources may be allocated to service quality programs that account for cross-cultural differences, though these differences may simply be the presence of differential response styles across cultures. Consequently, the findings of this study are relevant to both theory and practice.

The next section highlights some of the major findings of previous cross-cultural service quality research. This is followed by an explanation of response styles, and their potential impact on service quality measurement. Then, after describing the research methodology, we illustrate the impact of four response styles on the comparability of cross-cultural service quality data. Four cultural groups are selected to assess how likely an impact is to occur (van de Vijver and Leung 1997). As the response style most consistently considered in the cross-cultural environment is extreme responding (e.g., De Jong et al. 2008), the analysis considers this first. Subsequent analysis considers the three main groups of response style (Diamantopoulos, Reynolds, and Simintiras 2006) and examines how they impact on substantive findings. Finally, we provide suggestions for researchers on how to deal with response style issues in comparative research, and we discuss the theoretical and managerial implications of our findings.

MEASURING SERVICE QUALITY ACROSS CULTURES

Service quality is typically conceptualized as a consumer evaluation involving comparisons of expectations and perceptions of a service. Parasuraman, Zeithaml, and Berry (1994a) describe expectations as bi-level; that is, adequate expectations (the level of service the customer will accept) and desired expectations (the level of service the customer hopes to receive) exist alongside each other. The difference between the two is the zone of tolerance. If perceived service falls below adequate expectations, poor perceived service quality will result; between the two levels of expectations, adequate service quality will result; and above the desired expectation level, superior service quality will result. Service exchanges take place in a social and economic context where norms and values determine the appropriateness of service process features and, in particular, employee behaviors (Winsted 2000). Culture reflects the norms and values of a society. There are, therefore, cross-cultural variations in what is expected, or important, to consumers in service delivery.

Many cross-cultural service quality studies adopt Hofstede's (1980) cultural dimensions of individualism-collectivism and power distance as an approach to explaining differences in the five factor framework of service attributes (reliability, responsiveness, assurance, empathy and tangibles) developed by Parasuraman, Zeithaml, and Berry (1988; 1994a; 1994b). Individualism-collectivism has been described as a fundamental aspect of cultural variation (Bond and Hwang 1986). Individualism emphasizes 'self' and that the individual (and immediate family) is an end in himself, whereas collectivism emphasizes the views, needs and goals of the collective (or group). Consumers from individualistic cultures have higher service expectations generally

(Donthu and Yoo 1998), particularly emphasizing outcome (or reliability) rather than process (Mattila 1999). Consideration of process elements highlights responsiveness (willingness to help and prompt service), and particularly speed of service, as an attribute valued by Western individualist cultures (Malhotra et al. 2005). Eastern, collectivist cultures, however, are more tolerant of failure (Imrie, Durden, and Cadogan 2000), emphasizing personal relationships and the human interaction elements of service such as assurance (knowledge and courtesy of employees and ability to inspire trust and confidence) (Raajpoot 2004) and 'empathy' (individualized attention and convenience) (Kettinger, Lee, and Lee 1995). Although Donthu and Yoo (1998) suggest that individualist consumers have higher expectations of both assurance and empathy, others (Furrer, Liu and Sudharshan 2000) suggest that this is a measurement issue. After assessing the relative importance of attributes these latter authors argue that assurance and empathy are more important to consumers from collectivist cultures. Additionally, disparities with respect to the 'empathy' factor may be attributable to the combination of individual/ personal attention (similar to assurance) and convenience (similar to responsiveness) in the perceptions of consumers. The role of tangibles is less clear and it has been suggested that tangibles, such as the appearance of staff and facilities, are more important to Westerners (Mattila 1999); to consumers in developing countries (Kettinger, Lee, and Lee 1995); or to 'medium individualists' (Furrer, Liu, and Sudharshan 2000). These last authors highlight the relationship between tangibles and 'power-distance' (i.e., the acceptance of inequality of power distribution in a society (Hofstede 1997)), illustrating how this may differ according to the consumer's position relative to the service provider.

Factors other than culture can also influence the effectiveness of marketing practices (Vanderstraeten and Matthyssens 2008). Economic factors, in particular the level of economic development, can be determinants of differences in service evaluation. Low- and middle-income economies are sometimes referred to as developing economies although income does not necessarily reflect development status (The World Bank 2009). Malhotra et al. (1994) highlight that developing countries generally differ from developed countries in terms of financial and technological affluence, the level of sophistication of educational and communication systems, and the nature of commerce/competition. Customers in developing countries have higher tolerance levels, particularly for the content or timing of a service (Raajpoot 2004), and lower quality expectations generally (Malhotra et al. 1994; 2005). Lower expectations may also result from experience of lower levels of service quality (Witowski and Wolfinbarger 2002).

While information about consumers' expectations and how they differ between cultures is of considerable interest and importance to both managers and researchers, it is consumers' perceptions of the service and their overall assessment of performance that provide vital information for decision-making. In an experimental setting, respondents from different cultures will report differing quality perceptions of the same service encounter (Laroche et al. 2004). Few cross-cultural researchers, however, measure consumer perceptions and relate these directly to overall service quality measures. Exceptions include Cunningham, Young, and Moonkyu (2002), who found 'assurance' to be a significant factor in determining service satisfaction for Korean but not US respondents, and Witowski and Wolfinbarger (2002) who found 'responsiveness' to be an important determinant of overall service quality for both US and German respondents.

In summary, previous research indicates that expectations of assurance and empathy are likely to be higher in developing and collectivist cultures, whereas expectations of responsiveness and reliability are likely to be higher in developed and individualist cultures. While there is less evidence with respect to the relationship between perceptions and overall quality evaluations, there is some evidence to support the proposition that assurance will be an important predictor in developing collectivist countries. Conversely, previous studies indicate that responsiveness will play a more important role in developed individualist countries. Where researchers find differences or similarities attributable to cultural groupings (such as those outlined above) inferences and implications are drawn; for example as to how organizations should adapt their services and/or segment their markets. However, cross-cultural differences in response styles may hide real differences or produce spurious differences thus invalidating any conclusions drawn. The potential impact of response styles is discussed in the next section.

RESPONSE STYLES IN CROSS-CULTURAL RESEARCH

Every response given on a research instrument is made up of two elements: the 'true' response that reflects the level of the attribute being investigated and response bias. One type of response bias is response style, a habit (or momentary attitude) that causes the respondent to earn a score that is different from the one that would have been earned if the same question were presented in a different form (Cronbach 1950). Response styles distort measurement and can obscure what is actually happening. They fall into three broad types: (1) the use of *particular* response categories (e.g., extreme and mid-point responding), (2) the *spread* of responses (e.g., index of dispersion), and (3) the respondent's reaction to the item or response category *direction*

(e.g., tendency to rate to the right of centre) (Diamantopoulos, Reynolds, and Simintiras 2006). Each will affect substantive measurement slightly differently. Differences in extreme responding, for instance, can result in a differential increase/ decrease in the magnitude of respondents' scores which would be compounded on multi-item unidirectional scales such as SERVQUAL (Baumgartner and Steenkamp 2001).

Cross-cultural variation occurs with each group of response styles. Collectivist cultures tend to avoid extremes (Chen, Shin-ying, and Stevenson 1995; De Jong et al. 2008), and Eastern cultures are more likely to use the middle response category than Western cultures (Si and Cullen 1998). Differences in response styles are also found within Western cultures (Diamantopoulos, Reynolds, and Simintiras 2006), indicating that response styles cannot be ignored even with research restricted to relatively similar nations. Nevertheless, cross-cultural differences are not always apparent in respondents' use of particular response categories (Cheung and Rensvold 2000), suggesting a need to check for response styles on a study-by-study basis. Cross-cultural differences also exist in responses styles that are associated with the individual's spread of responses (Diamantopoulos, Reynolds, and Simintiras 2006), and respondents' reactions to the direction of the stimulus (acquiescence) also vary across cultures (see Harzing 2006).

Response style differences are of interest because of their implications for cross-cultural research results (Chun, Campbell, and Yoo 1974; Stening and Everett 1984). Response styles influence the foundations of many analysis techniques by increasing (or decreasing) correlations between variables and/or changing the magnitude of a variable's value (Baumgartner and

Steenkamp 2001; Podsakoff et al. 2003; van de Vijver and Leung 1997). When response styles change correlations, they can affect any analysis technique that relies on relationships between variables, including Cronbach's alpha, regression, factor analysis, and structural equation modeling (Greenleaf, Bickart, and Menon 1997; Heide and Grønhaug 1992). By changing the magnitude of the mean value, response styles can influence the results of comparative tests, such as t-tests and analysis of variance (ANOVA) (Cheung and Rensvold 2000; Diamantopoulos, Reynolds, and Simintiras 2006). Indeed, response styles can be a plausible rival hypothesis for cross-cultural research results (Stening and Everett 1984).

RESEARCH DESIGN

We aimed to examine the influence of response styles on substantive findings at two levels, the impact of (i) extreme responding in isolation, and (ii) multiple response styles. Data relating to expectations and perceptions of retail banking services were collected from respondents from four cultural groupings as discussed below.

Sampling and Research Context

To investigate the impact of response styles on cross-cultural service quality research, multiple cultures must be considered. Cultures likely to show differences in response styles, and in service quality evaluations, were required. Previous research had established that Eastern and Western cultures show differences in response styles (Chen, Shin-ying, and Stevenson 1995), and in how service quality was evaluated (Donthu and Yoo 1998; Mattila 1999). At the macro-

sampling level, we selected two paired cultural[i] groups to reflect the differences discussed earlier with respect to service quality evaluation. Hofstede's studies (e.g., 1997) located China (score 20) and Africa (East score 27, West score 20) at the opposite end of an individualism-collectivism continuum from Great Britain (including England and Scotland) (score 89). Similarly, two of our cultural groups, China (score 80) and Africa (East score 64, West score 77), scored highly on power distance. Conversely, Great Britain was low in power distance (score 35). The level of economic development may also influence what is considered as an adequate level of service quality (e.g., Malhotra et al. 1994; 2005; Raajpoot 2004), and some findings have indicated that this also impacts on response styles (Johnson et al. 2005). The World Bank (2009) classified Great Britain as 'developed' and China and Africa (East and West) as 'developing'. If there are cross-cultural differences, variation between similar groups (e.g., English and Scottish) would be a stronger test of the theory than if dissimilarities are found between highly distinct groups (e.g., Chinese and English). In contrast, if no true differences exist, then no differences between highly distinct groups would prove a stronger test of the theory than no differences between similar groups (van de Vijver and Leung 1997).

University students fulfilled the study requirements at the individual level. The sample was drawn from a number of universities within the U.K. Students were used to represent the characteristics of their nations (Netemeyer, Durvasula, and Lichenstein 1991), and because they "represent the upwardly mobile middle and upper classes, which are the target markets chosen by most corporations in foreign countries" (Ueltschy et al. 2004, p. 904). The Chinese and African students in the sample consisted mainly of 'sojourners' (Teske and Nelson 1974), expecting to leave the UK after completing their studies and unlikely to assimilate to the host,

culture even if they temporarily modified their behavior to 'fit in'. The supposition that African and Chinese students had been in the U.K. a relatively short time was supported by the amount of time they had held their bank accounts (African average of 1.3 (std dev = 1.47), Chinese 1.6 (std dev 1.88) years). In contrast, English and Scottish respondents had held their accounts for an average of 6.1 (std dev = 4.64) and 7.6 (std dev = 5.73) years respectively. In addition, cross-cultural comparability calls for homogeneous samples to control for extraneous factors (Reynolds, Simintiras, and Diamantopoulos 2003). In this context, the use of students helped to restrict demographic factors (e.g., age, education) related to response style (Greenleaf 1992), and the service used helped to isolate differences due to multiple service contexts. In addition, as Soueif (1968) found cross-cultural differences in extreme responding among male respondents, but not female respondents, we noted and accounted for gender in the analysis. In total 638 responses were obtained (123 African, 190 Chinese, 164 English, 161 Scottish). Finally, banking services were chosen because they (1) had been the subject of a substantial number of published mono- and cross-cultural service quality studies; (2) were a relevant service to our sample; and, (3) were used on a regular basis involving both face to face and other, for example, electronic encounters.

Research Instrument

We developed and pre-tested a self-completion questionnaire initially with groups and then with individuals from each of the cultural groups. Survey questionnaires were distributed during business and management lectures at a number of UK universities where time (approximately 25 minutes) was allowed for completion. Respondents were asked to consider

their experiences with respect to their own bank and to record their perceptions. They were also asked to record their expectations of banking services.

Measures

For each of the 21 statements included in the SERVQUAL battery respondents were required to assess their expectations in terms of adequate and desired service quality and their perceptions of the service quality received on a nine-point scale anchored by low (1) and high (9) (Parasuraman, Zeithaml, and Berry 1994a). Many authors (see, e.g., Cronin and Taylor 1992; 1994) argue for the adoption of perceptions measures for the prediction of overall service quality evaluations, and we adopted this approach. Table 2 shows the means, standard deviations, and reliability levels for all the sub-scales summated scores. Overall, the reliability of the service quality dimensions was acceptable for cross-cultural research, with all reliabilities above .60 (Craig and Douglas 2000).

Quality: The 'overall service quality' scale includes five-items measured on a seven-point scale (Cronin et al. 1997). Low scores indicated high-quality evaluations. Reliability levels were acceptably high for cross-cultural research from .763 (Chinese) to .933 (English) (Craig and Douglas 2000).

Insert Table 1 about here

Measurement Invariance

None of the three measurement models (adequate expectations, desired expectations, and perceptions of service quality) showed equivalent factor structures, the most basic level of measurement invariance, across the four groups. That is, acceptable model fit statistics were not achieved when the factor structures of the models were constrained to be equal (using AMOS 16.0). According to Steenkamp and Baumgartner (1998, p.78), measurement invariance assesses whether "cross-national differences in scale means might be due to true differences between countries on the underlying construct or due to systematic biases in the way people from different countries respond to certain items." Consequently, as we selected cultures with the expectation of differences in response styles (i.e., cross-cultural differences in systematic biases), it was not surprising that measurement invariance did not occur.

Response Style Measures

We used four measures to operationalize response styles (Table 2). According to Diamantopoulos, Reynolds, and Simintiras (2006) response styles can be grouped into three types: extreme and mid-point responding captured respondents' use of *particular response categories*; the index of dispersion, which measures the evenness of responses, captured the *spread of respondents' responses*; and the tendency to rate to the right of centre captured the impact of the *direction of the response categories*. This study was part of a larger survey that examined cross-cultural evaluations of service quality and related constructs[ii]. Although the

discussion here focuses on expectations, perceptions, and overall service quality evaluations, we calculated the response style measures from all the items on the questionnaire (100+).

Insert Table 2 about here

RESULTS

This section examines the impact of response styles on substantive conclusions of cross-cultural service quality research. We focus on two main issues. First, we assess the impact on cross-cultural differences in expectations (adequate and desired). Second, we examine the impact on relationships between perceptions of the five sub-dimensions and evaluations of overall service quality. Three sets of results are presented; when (1) response styles are ignored, (2) only extreme responding is considered, and (3) multiple response styles are considered.

Impact of Response Styles on Comparisons of Expectations

The cross-cultural comparisons of service quality expectations are summarized in Table 3[iii] under three conditions:

1. Cross-cultural differences in adequate/ desired service quality without accounting for response styles (i.e., Observed scores);

2. Cross-cultural differences after extreme responding has been removed by entering it as a covariate in ANCOVA (i.e., Extreme responding taken into account); and

3. Cross-cultural differences after all four response styles have been removed (i.e., All

response styles taken into account).

Condition 1: Observed Scores

There are significant cross-cultural differences across all five sub-dimensions for both adequate and desired service quality. If substantive conclusions were drawn from these findings, they could include the following:

a)   The Chinese have lower adequate and desired expectation levels than the other groups on all dimensions of service quality except tangibles.

b)   The English and Scottish (i.e., two similar cultures) do not differ in their expectations of service quality.

c)   The Africans generally have the highest expectations of adequate service quality.

These results indicate that Eastern (Chinese) respondents have lower expectations of service quality than Western (English/ Scottish) respondents and that African respondents' adequate service levels are higher than those of the more developed countries. Based on these results researchers might recommend that fewer resources need to be devoted to improving services for Chinese consumers. Nevertheless response biases may contaminate these results. If this has occurred, any recommendations based on these results will also be flawed.

Insert Table 3 about here

Conditions 2 & 3: Extreme Responding, and All Responses Styles, Taken into Account

Table 3 also shows the impact of extreme responding, and of multiple response styles, on cross-cultural differences in service quality expectations. Response styles were taken into account by entering them as covariates in ANCOVA. Referring to the conclusions drawn previously (a-c), response styles modify all but (b). It now becomes clear that expectations, particularly adequate expectations, do not always differ cross-culturally. Of particular interest is that the impact of removing response styles is apparent in two ways: (1) changing the overall significance of cross-cultural differences, and (2) altering the grouping of cultures in the post-hoc tests.

Impact of Response Styles on Significance Levels

Controlling for extreme responding (in ANCOVA) results in fewer significant differences (when $\alpha = .01$), now neither adequate levels of reliability nor desired levels of tangibles differ cross-culturally. Accounting for multiple response styles further clarifies whether cross-cultural differences in adequate or desired service quality exist – now, only half of the results show significant differences. These results show the necessity of considering more than just extreme responding when assessing the impact of response styles. The addition of mid-point responding, index of dispersion, and rating to the right of centre, reveals cross-cultural similarity in adequate expectations of service quality (except for tangibles); a result very different to those found under condition 1.

In summary, when response styles are accounted for, it can still be concluded that there are cross-cultural differences in the desired level of service quality in all dimensions except tangibles. However, except for tangibles, it can no longer be concluded that there are cross-cultural differences in expectations of adequate service quality.

Impact of Response Styles on Grouping of Cultures

Another area where response styles have an impact is on the results of the post-hoc tests for differences between cultures. There are three ways that removing response styles can change the post-hoc tests (Table 3): (1) by revealing clearer cross-cultural differences (e.g., for adequate expectations of tangibles, the Chinese are similar to the English and Scottish with observed scores, but clearly different from the English and Scottish when extreme responding is removed); (2) by changing the grouping of cultures (e.g., for desired service quality on assurance, responsiveness, and empathy the African group moves from a group that contains the English and Scottish to a group with the Chinese); and (3) by blurring the differences between groups (e.g., for adequate levels of assurance, responsiveness, and empathy when the effect of extreme responding is removed differences are no longer found between the Chinese and Scottish groups). In contrast to the third effect, the first and second effects clarify where differences between cultures occur, making the interpretation of the substantive results easier.

Impact of Response Styles on Predicting Overall Quality Evaluations

Sub-dimensions of service quality (e.g., SERVQUAL dimensions) should correlate

highly with overall service quality measures; that is, they should have predictive (or concurrent) validity. A problem of response styles, however, is that they can alter the relationships between measures. Table 4 presents the results of the regression analysis that examines how well the perceived service quality sub-dimensions (together with age and gender) predict overall service quality evaluation within each group.

Condition 1: Observed Scores

The percentage of variance in overall service quality evaluation explained by the *observed* scores of the perceived service quality dimensions ranges from 22% (Chinese) to 44% (Scottish). Neither gender nor age impact on quality. None of the sub-dimensions of service quality are significant predictors of overall quality for the Chinese group. For African respondents the negative coefficients associated with perceived assurance and reliability indicate that an increase in these increases perceptions of overall quality (as low scores indicate high overall quality), whereas an increase in tangibles leads to lower perceived levels of overall quality. Increasing perceptions of responsiveness for English respondents increases overall perceived quality. The Scottish results are more complex. Increasing empathy increases overall quality, however, the Scottish respondents' gender influences the impact of both reliability and empathy. This indicates that gender influences the relationship between the service quality sub-dimensions and overall quality in this group.

Insert Table 4 about here

Condition 2: Extreme Responding

The percentage of variance in overall quality explained when we partial[iv] out extreme responding from the service quality dimensions ranges from 20% (Chinese) to 39% (English and Scottish). Removing extreme responding does not result in any changes in significant coefficients with the Chinese or English groups. However, with the African group, removing extreme responding leaves only two significant coefficients – perceived assurance and tangibles – and for the Scottish, removing extreme responding leaves no main effects – that is, empathy is no longer significant.

Condition 3: Multiple Response Styles

Partialling out all four response styles from the service quality sub-dimensions results in a further reduction in the percentage of variance in overall quality explained. When the coefficients themselves are considered, removing all response styles results in the same significant coefficients as those found when only extreme responding was removed.

CONCLUSIONS AND IMPLICATIONS

The aim of this study was to highlight the problem of response styles in cross-cultural service quality research and to demonstrate a relatively uncomplicated method to remove them. This paper shows to cross-cultural researchers in general, and cross-cultural service quality researchers in particular, the importance of accounting for response styles when analyzing data.

The methods shown are useful because they have fewer restrictions than other techniques for removing method bias (Podsakoff et al. 2003), and require no modification to the research design to make their use feasible. ANOVA and its derivatives (e.g., ANCOVA) can be applied to show the impact of response styles on cross-cultural differences. The impact of response styles on relationships between variables was assessed via partial regressions in this study. In brief, and for this data set, response styles affected cross-cultural differences in expectations as well as impacting on the relationships between perceived service quality sub-dimensions and overall quality evaluation. Note that differences in response styles are not the problem *per se*, rather, it is their impact on substantive conclusions that may be problematic.

We found that conclusions drawn from the initial analysis that did not account for response styles were clearly different from those which did. Researchers argue that expectations are generally lower in developing, collectivist countries than in developed, individualist cultures (Donthu and Yoo 1998; Malhotra et al. 1994). Our findings suggested that, when response styles' impact is removed, this was only true of desired expectations and for adequate expectations of tangibles. In summary, the progression of results for adequate levels of service quality (Table 3) showed that: (1) when no response styles were considered there were significant cross-cultural differences; (2) removing the impact of only extreme responding blurred the differences in the post-hoc tests; and, (3) removing the influence of multiple response styles further reduced the number of significant cross-cultural differences. Removing response styles lent support to the conclusion that developing and collectivist countries have lower 'desired' expectations than consumers in developed and individualist countries, but it also changed the nature of these differences. For example, observed scores indicated that Chinese

consumers had the lowest 'desired' expectations whereas, after removing response styles, it was African respondents that score lowest here. The removal of response styles also clearly grouped together respondents from Africa and China and those from England and Scotland.

With respect to the impact of perceptions of bank service performance on overall quality evaluations, the findings provided support for the role of assurance in developing, collectivist cultures and for responsiveness in developed, individualist cultures. These findings, however, were restricted to African and English respondents. Other relationships lost their significance when response styles were removed; that is, reliability for African respondents and empathy for the Scottish were no longer significant. In view of the conflicting evidence from cross-cultural research studies (particularly with respect to the SERVQUAL dimension of 'empathy'), the impact of response styles on these relationships is of particular interest. Indeed, taking response styles into account may help researchers and managers to clarify areas of cross-cultural similarity/ differences. Nevertheless, care needs to be taken if the researcher/ manager assesses the impact of one response style and uncovers no impact, as a lack of impact from one type of response style cannot be relied upon to predict the same for another.

Implications for Theory

The theoretical implications for cross-cultural service quality researchers extend beyond the results presented here. Since the early 1980s, a significant amount of research has focused on: (1) determining the nature of expectations in terms of dimensions or traits (Parasuraman, Berry, and Zeithaml 1991; Parasuraman, Zeithaml, and Berry 1988; 1994a); (2) defining and

measuring consumers' expectations of services (Parasuraman, Zeithaml, and Berry 1994b); and (3) examining the nature of the gaps between expectations and perceptions, as well as the relationship between them and overall service quality evaluations (Cronin and Taylor 1992; 1994; Parasuraman, Zeithaml, and Berry 1988; 1994a). Cross-cultural service quality research develops this earlier work by making comparative assessments between nations/ cultures of interest. Consequently, the impact of response styles on mean scores, and of mean scores on gap scores, needs to be considered. For example, the impact of extreme responding on observed scores is such that the observed score of a respondent with a true score greater (less) than the mid-point of the scale is increased (decreased) by extreme responding (Baumgartner and Steenkamp 2001). Thus, the impact of extreme responding on gap scores depends on whether the true scores of the two components are on the same side of the mid-point or not. When they are on the same side of the mid-point, extreme responding will have less of an impact on gap scores than when they are on opposite sides of the mid-point. Table 5 illustrates the four possible scenarios. As the table shows, the most likely impact of extreme response style is a reduction in the gap score; however, this is not necessarily the case, and thus those interested in cross-cultural service quality research need to pay particular attention to response styles when using gap scores. Consequently, researchers and practitioners who are interested in assessing relationships between desired and adequate expectations, the 'zone of tolerance' (Parasuraman, Zeithaml, and Berry 1994a), and perceptions in order to determine whether consumers evaluate service quality as poor, adequate or superior, may find that conclusions are impacted by response styles. The use of gap scores in service quality research generally has been questioned (Brown, Churchill and Peter 1993). This study provides further evidence of problems which may arise in the cross-cultural context.

Insert Table 5 about here

Managerial Implications

Managers need to account for response style differences when assessing the level of

service quality required among bank branches serving different cultural groups, or on an

international scale, if they are to ensure that they are making 'accurate' comparisons. The

importance of the service sector to the growth in world economies and to organizations in their

international marketing strategies has been emphasized (Laroche et al. 2004; Malhotra et al.

2005). Comparative information about consumers' expectations and the impact of perceptions on

overall quality evaluation is vital to service management in determining, among other things, the

need for adaptation of the service design, the pattern of resource allocation, and the nature of

training. However, apparent differences (or similarities) in consumers' expectations and

perceptions may be a function of their operationalization and/or measurement. Consequently, it

is vital that the impact of response styles on comparative data is recognized as substantive

conclusions can be changed by response style contamination. For example, managers examining

the initial observed scores from this study would believe that Chinese respondents' expectations

of an adequate level of service are more easily satisfied (Table 3). They would also believe that

the minimum level of service required for African respondents is relatively high. Taking

response styles into account shows that neither is the case.

Additionally, the removal of response styles impacts on the relationship between service quality dimensions and overall service evaluation, again suggesting that a shift in managerial attention may be needed. This could have implications for the development of international market segmentation strategies. As managers can account for response styles through analysis, the cost implications associated with making valid and fair comparisons are minimal. Ideally, this should be done using the more sophisticated methods that are available (e.g., De Jong et al. 2008; Weijters, Schillewaert, and Geuens 2008). However, when these methods cannot be applied, this paper demonstrates a way to determine whether response styles are influencing substantive conclusions. Even if response styles are not accounted for each time the data are analyzed, having some knowledge of their impact would contribute to the interpretation of the results and, thus, to many areas of decision making.

LIMITATIONS AND FURTHER RESEARCH

Although the findings of this study are specific to this data set, they act as a powerful example of the impact of response styles for all cross-cultural researchers and service quality researchers in particular. Specifically, cross-cultural studies that do not fit the requirements of more sophisticated methods can still assess the impact of response styles using the methods applied here (Podsakoff et al. 2003). While the methods used in this study can be applied by other researchers, it would be useful if further research made a direct comparison between these methods and the more sophisticated methods that are now available (e.g., De Jong et al. 2008; Weijters, Schillewaert, and Geuens 2008). The sophisticated methods for assessing response styles could be combined with the techniques used to assess measurement invariance (Mullen

1995; Steenkamp and Baumgartner 1998). This combination of techniques was not possible in this study because of the relatively small sample sizes for the four groups and the number of variables and items in the model. Nevertheless, using these methods alongside the techniques employed in this study would allow for an assessment of their relative utility. In addition, while we have focused specifically on service quality, assessing the impact of response styles on other constructs (e.g., service satisfaction, assessment of switching costs) could provide an avenue for future research.

While we refer to China, and East and West African countries as developing countries throughout, we did not measure respondents' comparative income levels. Such comparisons are difficult due to factors such as differences in exchange rates and cost of living indicators. Yet a related factor here, for further research, is the extent to which differences in respondents' evaluations of service quality are explained by economic as opposed to cultural factors. Additionally, the extent to which the same measures (e.g., SERVQUAL) are relevant to countries at various stages of development, or where preferred usage patterns and channels differ (e.g., face-to-face or electronic encounters), should also be a focus of future research.

A potential contaminating factor in this study is that the survey instrument is not translated into the native language of each respondent group. While this allows the problems associated with translation to be minimized (Green and White 1976), there is some evidence to show that changing language impacts on the responses given by bilingual respondents (Ji, Zang, and Nisbett 2004). This difference may be partially explained by response styles manifesting differently according to language (Harzing 2006; Weijters, Geuens, and de Wulf 2008). Further

research could, however, investigate whether cross-language differences are a result of how languages themselves are processed as suggested by Ji, Zang, and Nisbett (2004), or whether it is the language's impact on response styles that causes the differences in responses.

The sample also has its limitations. Specifically, the use of African and Chinese students studying in the UK allowed for relatively homogeneous samples to be compared across cultural groups, which was in line with the objectives of this research (Reynolds, Simintiras, and Diamantopoulos 2003). However, it could be argued that students away from home are: (1) not typical of students from that cultural group; and (2) assimilated into the culture of their host nation. If our foreign student sample had moved from being typical examples of students from their respective cultures and towards the host culture, then the fact that response styles impact on the results is more, rather than less, impressive as there would have been more similarities between the cultural groups if this movement had occurred (van de Vijver and Leung 1997). This, then, serves to emphasize the importance of taking response styles into account. In addition, while it is possible that these students had assimilated to the host culture, as the vast majority of these students are only in the UK for the duration of their studies – i.e., as sojourners (Teske and Nelson 1974) – they are more likely to have adapted superficial behaviors to their temporary environment than to have adapted their norms, values and attitudes.

Despite these limitations, this study clearly shows how response styles impact on cross-cultural service quality research, and how relatively uncomplicated analysis techniques can be used to determine the impact of response styles. As the potential for response style

contamination on substantive conclusions can be considerable, the additional work required to

identify and eliminate their impact is advisable.

Endnotes

---

[i] Although culture is different from nationality, self-defined nationality is commonly used by researchers to identify respondents into cultural groups. Hofstede (1997) argues that it makes *practical* sense to collect data from different national, rather than cultural, groups. This advice particularly applies when the focus is on issues that separate or unite nations. In addition, the boundaries between cultures are often found to mirror the boundaries between countries (Gudykunst and Kim 1997). Finally, only those UK respondents who self-identified as English or Scottish, rather than British, are included in this study.

[ii] To ensure that the tendency to rate to the right of centre was not influenced by the respondents underlying attitudes, half of the item response categories were presented from low to high scores, and half were presented from high to low.

[iii] Table 3 also displays the power of the statistical tests and their effect sizes; however, because of length considerations, we discuss only significance and post-hoc comparisons of cross-cultural differences in the text.

[iv] The following steps were used to partial out the impact of response styles. Using Assurance to illustrate, when accounting for extreme responding, the first step was to regress Assurance onto extreme responding ($y = \alpha + \beta x + \varepsilon$, where y = Assurance, x = extreme responding, and ε the residual). The resulting residual scores represent each respondents' level of Assurance independent of the influence of extreme responding (Assurance*). Assurance* then replaced Assurance (i.e., the raw scale score) as the independent variable in the regression predicting overall quality. When multiple response styles were considered the residuals were calculated using multiple regression with each response style representing an independent variable in the equation (i.e., $y = \alpha + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_4 x_4 + \varepsilon$, where y = Assurance, x_1 = extreme responding, x_2 = mid-point responding, x_3 = index of dispersion, x_4 = tendency to rate to the right of centre, and ε the residual). The resulting residual scores – Assurance** – represent each respondents' level of Assurance independent of the influence of the response styles considered.

References

Baumgartner, Hans and Jan-Benedict E.M. Steenkamp (2001), "Response styles in marketing research: a cross-national investigation," *Journal of Marketing Research*, 38(2), 143-56.

Bond, M.H. and K.-K. Hwang (1986) *The Social Psychology of Chinese People, Chapter 6 in The Psychology of the Chinese People*, Bond, M.H. (ed) Oxford University Press: Oxford, 213-66.

Bresnahan, M.J., R. Ohashi, W. Liu, R. Nebashi, and C-C. Liao (1999), "A comparison of response styles in Singapore and Taiwan," *Journal of Cross-Cultural Psychology*, 30(3), 342-58.

Brown, T.J., G.A. Churchill, and J.P. Peter (1993). "Improving the measurement of service quality," *Journal of Retailing*, 69(1), 127-39.

Chen, C., L. Shin-ying, and H.W. Stevenson (1995), "Response style and cross-cultural comparison of rating scales among East Asian and North American students," *Psychological Science*, 6(3), 170-5.

Cheung, G.W. and R.G. Rensvold (2000), "Assessing extreme and acquiescence response sets in cross-cultural research using structural equation modeling," *Journal of Cross-Cultural Psychology*, 31(2), 187-212.

Chun, K., J.B. Campbell, and J.H. Yoo (1974), "Extreme response style in cross-cultural research," *Journal of Cross-Cultural Psychology*, 5(5), 464-80.

Clarke III, I. (2000), "Extreme response style in cross-cultural research: an empirical investigation," *Journal of Social Behavior and Personality*, 15(1), 137-52.

Craig, C.S. and S.P. Douglas (2000), *International Marketing Research*, 2nd ed. Chichester: John Wiley & Sons, Ltd.

Cronbach, L.J. (1950), "Further evidence on response sets and test design," *Educational and Psychological Measurement*, 10, 3-31.

Cronin, J.J. and S.A. Taylor (1992), "Measuring service quality: a re-examination and extension," *Journal of Marketing*, 56(July), 55-68.

--------- and --------- (1994), "SERVPERF versus SERVQUAL: Reconciling performance-based and perceptions-minus-expectations measurement of service quality," *Journal of Marketing*, 58, 125-31.

Cronin, J.J., M.K. Brady, R.R. Brand, R. Hightower Jr., and D.J. Shemwell (1997), "A cross-sectional test of the effect and the conceptualization of service value," *The Journal of Services Marketing*, 11(6), 375-91.

Cunningham, L.F., C.E. Young, and L. Moonkyu (2002), "Cross-cultural perspectives of service quality and risk in air transportation," *Journal of Air Transportation*, 7(1), 3-26.

De Jong, G. Martin, Jan-Benedict E.M. Steenkamp, Jean-Paul Fox, and Hans Baumgartner (2008), "Using item response theory to measure extreme response style in marketing research: A global investigation," *Journal of Marketing Research*, 45(1), 104-15.

Diamantopoulos, A., N.L. Reynolds and A. Simintiras (2006), "The impact of response styles on the stability of cross-national comparisons," *Journal of Business Research*, 59(8), 925-35.

Donthu, N. and B. Yoo, (1998), "Cultural influences on service quality expectations," *Journal of Service Research*, 1(2), 178-86.

Furrer, O., B.S.-C. Liu, and D. Sudharshan (2000), "The relationships between culture and service quality perceptions: basis for cross-cultural market segmentation and resource allocation," *Journal of Service Research*, 2(4), 355-71.

Green, R. and P.D. White (1976), "Methodological considerations in cross-national consumer research," *Journal of International Business Studies*, 7(2), 81-7.

Greenleaf, E.A. (1992), "Measuring extreme response style," *Public Opinion Quarterly*, 56(3), 328-51.

---------, B. Bickart, and G. Menon (1997), "Measuring how dispersion attenuates bivariate correlations in rating scale data and correcting these effects," Working paper, New York University Mark-97-2.

Gudykunst, W. and Y.Y. Kim (1997), *Communicating with Strangers: An Approach to Intercultural Communications.* 3rd edition, New York: McGraw Hill.

Harzing, Anne-Wil (2006), "Response styles in cross-national survey research: A 26 country study," *International Journal of Cross-Cultural Management*, 6(2), 243-66.

Heide, M. and K. Grønhaug (1992), "The impact of response styles in surveys: a simulation study," *Journal of the Market Research Society*, 34(3), 215-30.

Hofstede, G. (1980), "Motivation, leadership and organization. Do American theories apply abroad?" *Organisational Dynamics*, 9(1), 42-63.

Hofstede, G. (1997), *Cultures and Organizations: Softwares of the Mind,* New York: McGraw Hill.

Imrie, B.C., G. Durden, and J.W. Cadogan (2000), "Towards a conceptualization of service quality in the global market arena," *Advances in International Marketing*, 10(1), 143-62.

Ji, Li-Jun, Zhiyong Zang and Richard E. Nisbett (2004), "Is it culture or is it language? Examination of language effects in cross-cultural research on categorization," *Journal of Personality and Social Psychology*, 87(1), 57-65.

Johnson, Timothy, Patrick Kulesa, Isr Llc, Young Ik Cho, and Sharon Shavitt (2005), "The relation between culture and response styles: Evidence from 19 countries," *Journal of Cross-Cultural Psychology,* 36(2), 264-77.

Kettinger, W. J., C.C. Lee, and S. Lee (1995), "Global measures of information service quality: a cross-national study," *Decision Sciences*, 26(5), 569-88.

Laroche, M., L.C. Ueltschy, A. Shuzo, and M. Cleveland (2004), "Service quality perceptions and customer satisfaction: evaluating the role of culture", *Journal of International Marketing,* 12(3), 58-85.

Malhotra, N.K., F.M. Ulgado, J. Agarwal, and I.B. Baalbaki (1994), "International services marketing: A comparative evaluation of the dimensions of service quality between developed and developing countries," *International Marketing Review*, 11(2), 5-15.

---------, ---------, ---------, G. Shainesh, and W. Lan (2005), "Dimensions of service quality in developed and developing economies multi-country cross-cultural comparisons," *International Marketing Review*, 22(3), 256-78.

Mattila, A.S. (1999), "The role of culture and purchase motivation in service encounter evaluations," *Journal of Services Marketing*, 13(4/5), 376-89.

Mullen, M.R. (1995), "Diagnosing measurement equivalence in cross-national research," *Journal of International Business Studies*, 26 (3), 573-96.

Netemeyer, R.G., S. Durvasula, and D.R. Lichtenstein (1991), "A cross-national assessment of the reliability and validity of the CETSCALE," *Journal of Marketing Research*, 28(3), 320-27.

Parasuraman, A., L.L. Berry, and V.A. Zeithaml (1991), "Refinement and reassessment of the SERVQUAL scale," *Journal of Retailing*, 67(4), 420-50.

Parasuraman, A., V.A. Zeithaml and Berry, LL (1988), "SERVQUAL: A multiple - item scale for measuring consumer perceptions of service quality," *Journal of Retailing*, 64(1), 14-40.

---------, ---------, and --------- (1994a), "Alternative scales for measuring service quality: a comparative assessment based on psychometric and diagnostic criteria," *Journal of Retailing*, 70(3), 201-30.

---------, ---------, and --------- (1994b), "Reassessment of expectations as a comparison standard in measuring service quality: implications for future research," *Journal of Marketing*, 58(February), 6-17.

Podsakoff, Philip M., Scott B. MacKenzie, Jeong-Yeon Lee, and Nathan P. Podsakoff (2003), "Common method biases in behavioral research: A critical review of the literature and recommended remedies," *Journal of Applied Psychology* 88(5), 879–903.

Raajpoot, Nusser (2004), "Reconceptualizing service encounter quality in a non-western context," *Journal of Service Research*, 7(2), 181-201.

Reynolds, N.L., A.C. Simintiras, and A. Diamantopoulos (2003), "Theoretical justification of sampling choices in international marketing research: key issues and guidelines for researchers," *Journal of International Business Studies*, 34(1), 80-9.

Si, S.X. and J.B. Cullen (1998), "Response categories and potential cultural bias, effects of an explicit middle point in cross-cultural surveys," *The International Journal of Organizational Analysis*, 6(3), 218-30.

Smith, A. and N.L. Reynolds (2002), "Measuring cross-cultural service quality: A framework for assessment," *International Marketing Review*, 19(5), 450-80.

Soueif, M.I. (1968), "Extreme response sets as a measure of intolerance of ambiguity," *British Journal of Psychology*, 49, 329-34.

Steenkamp, Jan-Benedict E.M. and Hans Baumgartner (1998), "Assessing measurement invariance in cross-cultural research," *Journal of Consumer Research*, 25(June), 78-90.

Stening, B.W. and J.E. Everett (1984), "Response styles in a cross-cultural managerial study," *Journal of Social Psychology*, 122, 151-6.

Teske, Raymond H.C. and Bardin H. Nelson (1974), "Acculturation and assimilation: A clarification," *American Anthropologist*, 1(2), 351-67.

Ueltschy, Linda C., Michel Laroche, Robert D. Tamilia, and Peter Yannopoulos (2004), "Cross-cultural invariance of measures of satisfaction and service quality," *Journal of Business Research*, 57(8), 901-12.

Vanderstraeten, Johanna and Paul Matthyssens (2008), "Country classification and the cultural dimension: a review and evaluation," *International Marketing Review*, 25(2), 230-51.

Van de Vijver, Fons and Kwok Leung (1997), *Methods and Data Analysis for Cross-Cultural Research*, London: Sage Publications.

Weijters, Bert, Niels Schillewaert, and Maggie Geuens (2008), "Assessing response styles across modes of data collection," *Journal of the Academy of Marketing Science*. 36(3), 409-22.

---------, Maggie Geuens, and Kristof de Wulf (2008), "Assessing differences in response styles in Europe: language versus nationality" *Marketing Landscapes: A Pause for Thought*, 37th EMAC Conference, Brighton, 27th-30th May 2008.

Winsted, K.F. (2000), "Service behaviors that lead to satisfied customers," *European Journal of Marketing*, 34(3/4), 399-417.

Witkowski T.H. and M.F. Wolfinbarger (2002), "Comparative service quality: German and American ratings across service settings," *Journal of Business Research*, 55(11), 875-81.

World Bank (the) (accessed 31-05-2009)

http://web.worldbank.org/WBSITE/EXTERNAL/DATASTATISTICS/0,,contentMDK:20420458~menuPK:64133156~pagePK:64133150~piPK:64133175~theSitePK:239419,00.html

Table 1 Scale descriptives

Service quality		African			Chinese			English			Scottish		
		Mean	Standard deviation	Reliability	Mean	Standard deviation	Reliability	Mean	Standard deviation	Reliability	Mean	Standard deviation	Reliability
Adequate	**Assurance**	6.94	1.42	.714	6.05	1.29	.781	6.34	1.43	.797	6.62	1.09	.700
	Reliability	6.65	1.61	.743	5.85	1.34	.761	6.49	1.28	.791	6.68	1.15	.735
	Responsiveness	6.82	1.42	.656	5.85	1.34	.790	6.37	1.29	.796	6.61	1.15	.752
	Empathy	6.95	1.39	.798	5.65	1.32	.790	6.13	1.38	.813	6.37	1.18	.779
	Tangibles	7.06	1.34	.703	5.82	1.27	.744	5.48	1.61	.811	5.51	1.48	.758
Desired	**Assurance**	7.66	1.21	.789	7.17	1.24	.817	7.81	1.05	.755	7.96	.85	.639
	Reliability	7.39	1.51	.731	7.07	1.35	.809	8.87	1.07	.761	7.98	.86	.672
	Responsiveness	7.50	1.36	.764	7.05	1.37	.832	7.84	.94	.744	7.90	.96	.690
	Empathy	7.54	1.25	.787	6.91	1.40	.863	7.65	1.11	.805	7.70	.98	.722
	Tangibles	7.59	1.20	.678	6.86	1.36	.827	6.83	1.37	.737	6.80	1.39	.743
Perceived	**Assurance**	6.36	1.78	.673	5.82	1.37	.612	6.19	1.65	.758	6.44	1.50	.793
	Reliability	6.32	1.75	.694	5.55	1.41	.617	6.15	1.75	.853	6.26	1.59	.809
	Responsiveness	6.38	1.78	.732	5.70	1.37	.700	6.09	1.60	.755	6.32	1.51	.789
	Empathy	6.28	1.64	.717	5.43	1.47	.776	5.77	1.69	.805	6.16	1.46	.812
	Tangibles	6.61	1.75	.671	5.76	1.36	.683	5.87	1.84	.775	6.07	1.60	.715
Overall Quality		3.22	1.48	.900	3.61	.89	.763	3.01	1.10	.933	2.98	1.04	.906

Table 2 Response style calculations

Extreme responding	The percentage of responses on the end points of the scale.
Mid-point responding	The percentage of responses in the middle response category.
Index of dispersion	The index of dispersion is calculated using the following formula: $$D = h\left(n^2 - \sum_{i=1}^{h} n_i^2\right)\Big/ n^2(h-1)$$ where h = the number of possible scale values (categories) that can be used when recording responses, n = the total number of possible responses to all items, and n_i = the number of response in the ith scale category.
Rating to the right of centre	The percentage of all responses (excluding the mid-point) that are to the right of the scale's mid-point.

Table 3 Cross-cultural differences in scale scores [a]

		Observed scores				Extreme responding taken into account				All response styles taken into account			
		p-value	*Effects size[b]*	*Power*	*Post-hoc tests[c]*	*p-value*	*Effects size[b]*	*Power*	*Post-hoc tests[c]*	*p-value*	*Effects size[b]*	*Power*	*Post-hoc tests[c]*
Adequate	*Assurance*	.000	.062	1.000	(CE)(ES)(SA)[d]	.005	.021	.869	(CES)(SA)[d]	.589	.003	.185	---
	Reliability	.000	.063	1.000	(C)(EAS)[d]	.033	.014	.698	---	.030	.015	.710	---
	Responsiveness	.000	.073	1.000	(C)(ES)(SA)[d]	.001	.025	.927	(CE)(ES)(SA)[d]	.228	.007	.384	---
	Empathy	.000	.111	1.000	(C)(ES)(A)[d]	.000	.056	1.000	(CE)(ES)(SA)[d]	.012	.018	.806	---
	Tangibles	.000	.135	1.000	(ESC)(A)[d]	.000	.131	1.000	(SE)(C)(A)[d]	.000	.092	1.000	(SE)(C)(A)[d]
Desired	*Assurance*	.000	.096	1.000	(C)(AES)[d]	.000	.076	1.000	(AC)(ES)[d]	.000	.073	1.000	(AC)(ES)[d]
	Reliability	.000	.107	1.000	(CA)(ES)[d]	.000	.108	1.000	(AC)(ES)[d]	.000	.108	1.000	(A)(C)(ES)[d]
	Responsiveness	.000	.095	1.000	(C)(AES)[d]	.000	.084	1.000	(AC)(SE)[d]	.000	.080	1.000	(AC)(SE)[d]
	Empathy	.000	.095	1.000	(C)(AES)[d]	.000	.064	1.000	(CA)(SE)[d]	.000	.059	1.000	(AC)(SE)[d]
	Tangibles	.000	.045	.998	(SEC)(A)[d]	.046	.013	.653	---	.031	.014	.704	---

[a] Main effects of gender and culture were also tested: However, as neither the main effect of gender, nor any interaction effects between gender and culture were significant, only the results for culture are displayed.

[b] Partial eta squared.

[c] Lowest mean scores first.

[d] A = Africa, C = China, E = England, and S = Scotland

Table 4 Using perceived service quality to predict overall quality[a] (direction of significant coefficients indicated)

		Observed scores				Extreme responding taken into account				All response styles taken into account			
		Africa	China	England	Scotland	Africa	China	England	Scotland	Africa	China	England	Scotland
	Gender	.411	.079	.610	-.499	.400	.085	.039	-.542	.369	.085	.022	-.313
	Age	.178	-.021	.059	.087	.181	-.027	.036	.079	.161	-.047	.010	.092
Perceived service	Assurance	-.441**	.183	-.145	-.071	-.455**	.183	-.143	-.074	-.428**	.252	-.133	-.055
	Reliability	-.375*	-.039	-.090	-.028	-.308	-.001	-.077	-.015	-.305	.043	-.105	-.035
	Responsiveness	.134	-.003	-.428*	-.207	.120	-.019	-.432*	-.229	.093	.037	-.405*	-.214
	Empathy	-.074	-.301	-.011	-.329*	-.007	-.294	-.039	-.272	-.002	-.297	-.014	-.253
	Tangibles	.380**	-.287	.118	-.076	.420**	-.283	.126	-.080	.404**	-.256	.088	-.094
Gender interactions[b]	Age	-.147	.003	-.274	.382	-.158	.016	-.045	.619	-.119	.059	-.016	.372
	Assurance	.328	-.395	-.534	-.530	.101	-.131	-.172	-.117	.108	-.164	-.176	-.081
	Reliability	.644	.291	.073	-1.509*	.094	.071	.029	-.412*	.187	.067	.069	-.426*
	Responsiveness	-.377	-.143	.406	.780	-.065	-.033	.160	.182	-.071	.022	.171	.239
	Empathy	.029	-.044	-.109	1.679**	-.038	-.016	-.059	.386*	-.037	.013	-.094	.344*
	Tangibles	-.722	.315	-.211	-.186	-.290	.102	-.058	.002	-.244	.024	-.022	.042
	R²	.346	.215	.422	.436	.341	.203	.389	.388	.298	.136	.339	.327
	Significance	.000	.000	.000	.000	.000	.000	.000	.000	.000	.033	.000	.000

[a] Low scores indicate high overall quality evaluations

[b] Significant results indicate a difference between male and female respondents on the variable concerned.

* $p < .05$, ** $p < .01$

Table 5 Impact of extreme responding on service quality evaluations (1st component minus 2nd component)

	When true score is	First component of gap score (Perceived or desired service)	
		Above mid-point	*Below mid-point*
Second component of gap score (Desired or adequate service)	*Above mid-point*	Potentially no impact on gap scores[a]	Increase in the magnitude, decrease in the value, of the gap scores[b]
	Below mid-point	Increase in the magnitude and value of the gap scores	Potentially no impact on gap scores[a]

[a] When both components have a true score above (below) the mid-point of the scale: When the true score of the first (second) component is already close to, or at, the endpoint of the scale, and extreme responding moves the observed score of the second (first) component toward the endpoint of the scale (i.e., toward the observed score of the first [second] component), this will then result in a decrease in the magnitude of the gap score.

[b] The magnitude of the gap score increases, however, as it has a negative sign, its absolute value decreases.