

---

# The applicability of SERVQUAL in cross-national measurements of health-care quality

---

*William E. Kilbourne*  
*Jo Ann Duffy*  
*Michael Duffy and*  
*George Giarchi*

---

## The authors

**William E. Kilbourne** is Professor of Marketing, Department of Marketing, Clemson University, Clemson, South Carolina, USA.

**Jo Ann Duffy** is Professor of Marketing, Department of Management and Marketing, Sam Houston University, Huntsville, Texas, USA.

**Michael Duffy** is Professor of Counseling Psychology, Department of Educational Psychology, Texas A&M University, College Station, Texas, USA.

**George Giarchi** is Professor of Health and Social Work, University of Plymouth, Plymouth, UK.

---

## Keywords

Health services, Quality, Linear structure equation modelling, Measurement, Nursing homes

---

## Abstract

This study investigates the applicability of a modified SERVQUAL instrument as a means of measuring residents' perceptions of long-term health-care service quality in the USA and UK. The results confirm a stable, four-factor structure that is similar to previously defined service quality dimensions and is invariant across the countries studied.

---

## Electronic access

The Emerald Research Register for this journal is available at  
[www.emeraldinsight.com/researchregister](http://www.emeraldinsight.com/researchregister)

The current issue and full text archive of this journal is available at  
[www.emeraldinsight.com/0887-6045.htm](http://www.emeraldinsight.com/0887-6045.htm)

An executive summary for managers can be found at the end of this article.

## Introduction

On October 10, 2001 the US and the UK signed an agreement to work together to improve quality in health care. Under the Joint Statement of Intent signed by HHS Secretary Tommy G. Thompson and the UK Minister of Health, the Right Honorable Alan Milburn MP, the two countries will share data and experience related to quality of care – including expanding common criteria for measuring health care quality. Such agreements underscore the need for measures of quality that are valid and reliable across countries. Technical dimensions of quality care, such as morbidity/mortality rates, staffing ratios and medical outcomes, may be easier to compare than the service quality dimensions of health care. Service quality dimensions are difficult to measure because they are determined by the recipients of health care and are measured in terms of patient perceptions of the health care experience. To compare health care service quality in two or more countries requires an instrument that is invariant across settings. With the need for long term health care on the rise in both countries as the numbers of older people increases, it is critical that public policy makers and administrators understand what constitutes quality of care within the context of long-term care facilities (nursing homes). Interest in both the quality measurement issue and long-term care shaped the present study, which examines the applicability of the SERVQUAL scale (Parasuraman *et al.*, 1988) in cross-national studies of long-term health care quality.

## Review of the literature

Parasuraman *et al.* (1988) developed a 22-item scale with five dimensions:

- (1) *tangibles* – physical facilities, equipment and appearance of personnel;
- (2) *empathy* – caring, individualized attention;
- (3) *assurance* – knowledge and courtesy of employees and their ability to convey trust and confidence;
- (4) *reliability* – ability to perform the promised service dependably and accurately; and
- (5) *responsiveness* – willingness to help customers and provide prompt service.



The scale, SERVQUAL, has paired questions for expectations and perceptions. Service quality is operationalized as the difference between the measured expectations and perceptions, commonly referred to as the gap score. Since its inception, the scale has received both widespread use and criticism. The history and criticisms of SERVQUAL have been summarized by Grapentine (1998). A major criticism has been over the use of gap scores to measure service quality (Babakus and Mangold, 1992; Bolton and Drew, 1991; Cronin and Taylor, 1992; Woodruff *et al.*, 1983; Teas, 1994; Lee *et al.*, 2000). While the gap scores have been shown to have better diagnostic capabilities, the perception-only measures of service quality appear to have higher convergent and predictive validity. For that reason, the perception scores serve better the purposes of this study.

Another area of criticism has been the dimensionality of SERVQUAL. Cronin and Taylor (1992) advanced a one-factor measure instead of the five-factor measure proposed by Parasuraman *et al.* (1988). In subsequent studies, the number of dimensions underlying SERVQUAL has varied. After completing a comprehensive review of service quality studies, Asubonteng *et al.* (1996) concluded that differences in the number of dimensions appear to be linked to differences among industries. For example, Kettinger and Lee (1994) identified four dimensions in a study of information systems (IS) quality. The tangible dimension was not evident; this is understandable since in the IS industry tangibles are not visible to the customer.

Numerous studies in the health domain have used SERVQUAL to measure quality in US samples (Headley and Miller, 1993; Lytle and Mokwa, 1992; Carman, 1990; Licata *et al.*, 1995; Cronin and Taylor, 1994; Bowers *et al.*, 1994; Babakus and Mangold, 1992; Dean, 1999; O'Connor *et al.*, 1994). Taylor (1994) represents those who contend that the SERVQUAL perception portion would be a valid method to operationalize service quality if it were not unidimensional. While he and Babakus and Mangold (1992) did find only one dimension, Dean (1999) identified four stable dimensions when he used SERVQUAL to compare service quality dimensions in two different health care settings (medical center and maternal and child health center). The four-factor structure – i.e. assurance, tangibles, empathy and reliability/responsiveness (loaded together) – accounted for

approximately 68 percent of the variance in both settings. Gabbott and Hogg (1995) identified six factors: four resembled the ones identified by Dean (1999). Gabbott and Hogg's (1995) study is important because they were studying health care quality in the UK. In the context of long-term health care, SERVQUAL has been the underpinning of research by Kleinsorge and Koenig (1991) and Duffy *et al.* (2001). However, their focus was not on the psychometric properties of the scale, so they provide little insight into the dimensionality or stability of SERVQUAL.

Few authoritative studies addressing the applicability of SERVQUAL outside the US have been published. Three studies related to the present study are Nel and Pitt (1997), Lam (1997), and Kettinger and Lee (1994). Nel and Pitt's (1997) study of service quality in four South African industries is important because it focused on the psychometric properties of SERVQUAL in a setting outside the US. The overall scale alpha coefficient ranged from 0.70 to 0.78, indicating that SERVQUAL performed well in terms of reliability. Individual dimensions with alpha coefficients below 0.70 in more than one setting were tangibles and assurance. Factor analyses yielding solutions ranging from three to five factors resulted. In none of the settings was there evidence to support Taylor's (1994) contention that SERVQUAL is a unidimensional scale. Lam's (1997) study of Hong Kong hospital patients is important because it used a non-US sample and is health-care based. Lam's (1997) results indicate that SERVQUAL is a consistent and reliable unidimensional scale. Coefficient alpha values ranged from 0.71 to 0.90 for the performance perception subscales.

Research into the reliability and validity of SERVQUAL across cultures has focused on information systems users. Kettinger and *et al.* (1995) surveyed Korean, Hong Kong and Netherlands IS students and compared them to IS users in the US. The same four factors found in the US sample were found in the Dutch sample, but not the Korean or Hong Kong samples. The results show that SERVQUAL is invariant across some national settings when the industry is held constant. The present study pursues this line of inquiry into the invariance of SERVQUAL across national settings by studying respondents from the US and UK in the same health-care setting (i.e. long term care). The purpose of this analysis is only to examine the meaning and structure of the SERVQUAL instrument across two countries, the

US and England. Thus we seek to determine whether the number of factors (configural invariance), the degree to which ratings can be compared across the countries, and if the SERVQUAL instrument is a second-order factor model of quality. To do this we will conduct a series of tests required for this level of cross-national invariance.

### Cross-national invariance

The purpose of this research is to determine the applicability of SERVQUAL in cross-national research in long-term care service quality. The specific aims are to establish the dimensionality of service quality in samples from the US and UK and to investigate whether the second-order factor model of service quality is invariant for the two countries. Establishing the cross-national invariance of a survey instrument requires a series of steps with each step imposing more restrictions on the measurement model. The most powerful approach for testing cross-national invariance is generally considered to be multi-group confirmatory factor analysis (Jöreskog, 1971). The approach used here corresponds both to Byrne (2001) and to Steenkamp and Baumgartner's (1998) model, which is ideal for data collected in two or more countries with an instrument that uses the same set of items to measure the constructs in both countries. With measurement invariance assessed in this way, it can be reasonably concluded that measurement operations provide measures of the same constructs. We thus first determine configural invariance to demonstrate that the instrument yields the same factors in both groups with the same pattern of salient and non-salient factor loadings. Next we assess the metric invariance of the items in the instrument to show that the scale intervals are the same across countries. This demonstrates that observed item differences can be meaningfully compared across countries, and that observed differences can be attributed to differences in the underlying constructs. Finally, we will test the invariance of the second order structure across countries to determine if each of the first order factors typical in SERVQUAL are explained by a higher-order factor we call quality.

It should be noted that in each stage of the analysis of invariances, only partial invariance is required. Thus it is not necessary that all items be invariant in metric invariance. Rather, at least one

item other than the one used to assign the scale of measurement must be invariant (Steenkamp and Baumgartner, 1998).

### Methodology

#### Sample

The respondents for the study were 195 nursing home residents (147 females and 48 males) from ten long-term care facilities in the US and 99 nursing home residents from 15 long-term care facilities in the UK. Both samples were drawn from equivalent intermediate care facilities and therefore had similar levels of disability. There was no significant difference in the age of the respondents from the US and UK. A similar screening process was used in both countries with nursing staff being asked to recommend residents who had the cognitive ability to respond to spoken questions. All the residents identified as the cognitively able residents were invited to participate in the study. Only five of those invited to participate chose not to do so. While there is no "correct" sample size for structural equation modeling, some heuristics have been developed. We have chosen maximum likelihood estimation (the most commonly used method), and for this method Hair *et al.* (1998) suggest that the minimum sample size is 100. We are only one short of this rule of thumb in the UK group and consider that sufficient to proceed with the analysis recognizing that we are on the lower bound of acceptability.

#### Procedure

The SERVQUAL questionnaire was first piloted and refined on group of nursing home administrators in the US. The Appendix shows the final version of the questionnaire, which contained the 22 items reworded to fit the nursing home context. When this version of the questionnaire was piloted with a group of UK nursing home administrators, it appeared to need no additional changes. Therefore, the same version was administered in both the US and UK by interviewers trained by the same principal investigator. Shortly after interviews in the UK began, it became obvious that residents were struggling with item 4 ("Appearance of the physical facilities is appropriate"). This item in the original research by Parasuraman *et al.* (1988) had the lowest factor loading (0.47) in the first stage of the purification of the instrument and 0.51 in the

reanalysis on the tangibles subscale, which has proven to be the least reliable (Parasuraman *et al.*, 1988; Carmen, 1990). In a health care study, O'Connor *et al.* (1994) reported that the tangibles scale had inadequate reliability. Based on this information, the researchers decided the value of trying to force the residents to give a response to item 4 was not worth the discomfort caused to the respondents by pushing them to respond.

Residents were asked to rate their perceptions of nursing home services quality using the 22 SERVQUAL items. These items were rated using a seven point Likert scale. The scale is anchored on one end (1) by "disagree very strongly" and on the other end (7) by "agree very strongly". The first four items measure perceptions about the "tangibles" dimension. The order of the other four dimensions was reliability, responsiveness, assurance and empathy, and the scales were measured with five, four, four, and five items respectively. Because of a large number of missing values for item 4 in the tangibles subscale, it was eliminated from the data because list-wise deletion would have reduced the sample too much, and case-wise deletion would have distorted the covariance matrix, affecting the ability of the model to converge.

### Measurement model

To establish the conditions necessary for developing the structural equation modeling approach used here, the measurement model was examined using the procedures developed by Steenkamp and Baumgartner (1998) and Byrne (1998), and all analyses were run using AMOS 4 software (Arbuckle and Wothke, 1995). This involved testing the configural (number of factors) and metric (scale of measurement) invariance of the measurement model across countries and then testing the invariance of the second-order factor structure with quality (QUAL) as the higher-order factor. To assess invariance in the model, the four measures of model fit used were adjusted goodness of fit index (AGFI) (Jöreskog and Sörbom, 1993; Sharma, 1996), Tucker-Lewis index (TLI), comparative fit index (CFI) (Steenkamp and Baumgartner, 1998), and the root mean square error of approximation (RMSEA) (Byrne, 1998). Acceptable fit for the AGFI is 0.8 or higher, and for TLI and CFI it is approximately 0.9 or higher. For the RMSEA the recommended score is 0.08 or less (Hair *et al.*, 1998). Chi-square statistics were relied on less here because the statistic is sensitive to large degrees of freedom (Byrne, 1998). In

addition, sequential chi-squared difference tests (SCDT) (Anderson and Gerbing, 1988) were used to compare successively constrained models in the assessment.

The recommended first step in invariance testing is to establish the measurement model for the two groups independently (Byrne and Campbell, 1999). Preliminary analysis of the data indicated that several individual variables in the UK sample were not correlated with their individual factors, and one variable was highly correlated with two other factors. These were removed from the analysis. Because two of the items removed were in the assurance factor, only two items were left and the factor was removed from the analysis. Because meaningful cross-country comparisons are facilitated when the same number of constructs is in both comparison groups, the assurance factor was removed from the US sample as well. After the variables were removed, the measurement model fit both sets of data, albeit weakly. Modification indices suggested that the fit could be improved by allowing several error covariance terms. This was carried out for each country, and the results indicated that the reduced four factor SERVQUAL model did fit the data well for both countries. The respective statistics for each group are shown in Table I. The acceptable model containing 14 items (three for tangibles, three for responsiveness, three for reliability, and four for empathy) was then assessed for invariance across the two samples.

The second phase of the procedure involved establishing the configural invariance of the model, the purpose of which is to determine if the four-factor model was appropriate for both groups simultaneously. For the usual model comparisons to apply, it is necessary to show that the model for each country has the same measurement variables and latent constructs. This requires only the same pattern of zero and non-zero factor loadings.

The results of this analysis, presented in Table I, indicated that the model was a four-factor model for both the US and the UK. For the model, AGFI was 0.84, TLI was 0.93, CFI was 0.95 and the RMSEA was 0.053. For configural invariance to exist, it is not necessary for the factor loadings to be equivalent across models; only the same pattern of significances need to exist (Steenkamp and Baumgartner, 1998).

Having established that the four-factor model fitted the data for both the US and the UK, we next determined metric (factor loading) invariance. To establish this, we constrained the factor loadings to



Table I Model testing results

Independent	$\chi^2$	df	$\chi^2/df$	AGFI	TLI	CFI	RMSEA	SCDT	Action
US	117	57	2.05	0.87	0.94	0.96	0.074		
UK	88	56	1.57	0.80	0.90	0.93	0.076		
<b>Measurement</b>									
Configural	205	113	1.81	0.84	0.93	0.95	0.053		
Metric	229	122	1.88	0.84	0.93	0.94	0.055	0.004	Free v22
Partial metric	223	121	1.84	0.84	0.93	0.95	0.054	0.020	
<b>Second order</b>									
Configural	244	127	1.92	0.84	0.92	0.94	0.056		
Metric	260	130	2.00	0.83	0.92	0.93	0.059	0.001	Free Tan
Partial metric	249	129	1.93	0.84	0.92	0.94	0.057	0.080	

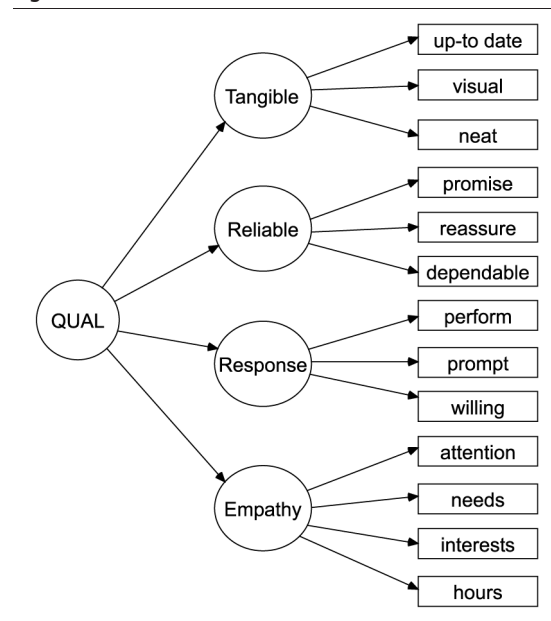
be the same for both groups and retested the model. The results of this analysis indicated that full metric invariance was not supported well. Modification indices suggested that the model would be more appropriate if V22 (hours) were allowed to vary for the two groups. We released this variable and re-evaluated the model. The result was an adequate fit with only slight decreases in each of the test statistics and a SCDT of 0.02. While it is preferable for the SCDT to be above 0.05, a score of 0.02 and acceptable fit on all other criteria was considered acceptable. In addition, *t*-tests for all factor loadings in both groups were significant at less than 0.05. We thus demonstrated that partial metric invariance was achieved for the model, and this was all that was required to proceed with the analysis (Byrne, 1998).

As a final test of the measurement model we examined the alpha reliabilities of the items in each of the constructs. For the US sample, Coefficient alpha for the items in the tangible, reliability, responsiveness, and empathy factors were 0.70, 0.80, 0.75, and 0.87, respectively. For the UK samples, the coefficients were 0.60, 0.76, 0.67 and 0.71. While the US statistics were in the acceptable range, those for the UK were considered slightly low compared to the heuristic of 0.70 as a criterion. Also for the US sample the variation explained was above the normal criterion of 0.50, and those for the UK were somewhat lower than would be desired.

### Second-order confirmatory factor model

The proposed second order factor model is presented in Figure 1. As suggested in the SERVQUAL literature, the latent variables, or dimensions of service quality, should combine to create a single measure of service quality (QUAL). This implies that QUAL is itself a latent variable that underlies the other dimensions of service

Figure 1 Second-order factor model



quality. Thus we now turn to a test of the invariance of the second-order factor model derived from the literature. To do this, we first determined that the second-order factor structure did indeed fit the data for both groups, indicating configural invariance. The results of this analysis yielded adequate test statistics with an AGFI of 0.84, TLI of 0.92, CFI of 0.94, and an RMSEA of 0.056. To test the metric invariance across countries, we constrained the paths from QUAL to the four dimensions of the SERVQUAL instrument in the measurement model to be the same for both the US and the UK.

The results of this analysis indicated that the proposed model fit the data adequately. While the fit statistics were acceptable for the constrained model the SCDT was 0.001, indicating that the constrained model did not fit as well as the full model. The modification indices suggested that

the model fit could be improved by releasing the path from QUAL to tangibles. This was done and the fit reassessed. The result was an acceptable fit on all indices and a SCDT of 0.08, which was acceptable. Again, individual *t*-tests for each of the path coefficients in the final model were all significant at less than 0.05, as required. This indicated that the final model was a second-order factor model with four factors that were partially invariant across the countries studied.

### Implications for research and practice

In this cross-national study of the US and the UK, we have taken the SERVQUAL instrument to task as an invariant measure of service quality in the two countries. The results suggest that it has the potential to serve as a means for comparing perceptions of service quality across countries. We proposed and tested a second-order factor model with tangibles, responsiveness, reliability and empathy as first-order latent constructs and service quality as the second-order latent construct. The results indicate that with only a few minor modifications, the instrument is invariant across the two samples. This means that it does have the potential to be used in the context of long-term health care for these two countries, and that reliable comparisons of construct means can be made between the countries.

There are several caveats involved in this research. The first is that while the SERVQUAL instrument can provide a multinational measure of quality, the variance explained in this study was lower than desirable. While a single study is, of course, inadequate to the task of demonstrating perfect measurement ability, it has been demonstrated here that SERVQUAL has potential as a reliable measurement instrument. The results indicate that SERVQUAL is multidimensional which supports the usefulness of the perception subscale as a robust measure of service quality. Similar studies examining the usefulness of the expectation subscale would also be helpful in future comparisons of service quality expectations variation between countries. Exploring whether the 22 SERVQUAL items capture all aspects of long-term care service quality and why item 4 was a problem for UK nursing home residents are additional research topics stemming from these results. A comparison of the relative importance of the five service quality dimensions to residents in the US and UK would provide a different type of

insight. Reliability has been found to be most important in the US, but elderly residents of nursing homes were not included in those sampled. Is reliability most important to residents or is empathy perhaps the most important dimension in this health care context? Does the perceived most important dimension vary from country to country or from context to context? Further studies comparing different countries and using different contexts are necessary before SERVQUAL's efficacy as a cross-national measure of service quality can be established adequately.

From a long-term care management perspective, this study also confirms the usefulness of SERVQUAL in nursing homes in both the US and UK. Public policy and health and social service departments in both countries have increased surveillance on quality of care in long-term care. However, it has been notoriously difficult to measure a variable as subtle as quality of care or its companion – quality of life. Frequently, and understandably, quality has been estimated through concrete, easily verifiable factors such as food preparation conditions, water temperature and presence of decubitus ulcers. In the US, for example, these “quality indicators” (QIs) are based on data from the Minimum Data Set (MDS), mandatory quarterly assessments on each resident. Residents and staff in nursing homes, however, know that despite a perfect score on QIs, service quality and quality of life can be far from perfect (Duffy and Duffy, 2002). The SERVQUAL instrument clearly captures more subtle quality indicators in a multidimensional way: tangibles, responsiveness, reliability and empathy, as well as the overall (second order) factor of service quality.

By studying residents or family responses to SERVQUAL questions, the nursing home administrator will be able to focus quality improvement efforts where they are most needed. Since resources are scarce in nursing homes, it is important to be able to identify where those resources will do the most good in terms of improving service quality. The information from SERVQUAL also serves as a guide for staff training. For example, if residents identify “lack of courtesy” as an issue, the training session can focus on this issue rather than other topics which may not need to be addressed. Additionally, repeated administration of SERVQUAL at different times can be a valuable tool for CQI (continuous quality improvement) programs as a way of tracking changes in resident perceptions over time. The

SERVQUAL data can function as an evaluation tool for CQI initiatives.

## Limitations

While this study has served to support the psychometric strength of the instrument, it has also supported the utility of SERVQUAL in the context of the nursing home. However, there are some limitations and questions raised as a result of the study. As is frequently the case with SERVQUAL, the best fit was achieved with a short version that eliminated some of the original items. Such shortened versions have been used recently and in the case of Dabholkar *et al.* (2000), for example, a similar subset of items was used for the reliability, responsiveness, and tangibles dimensions. While they suggested in their research in a retail setting that the dimensions of SERVQUAL are best modeled as antecedents to quality rather than quality itself, the bulk of recent research still considers them to be dimensions of quality. Most recently, Bigne *et al.* (2003), Soyoung and Byoung-ho (2002), Swanson and Davis (2003), and Sureshchander *et al.* (2002), have provided evidence that the traditional approach to SERVQUAL is appropriate. While Dabholkar *et al.* (2000) have raised questions about the appropriate use of the items in SERVQUAL, the logic of justification (Hunt, 1991) suggests that the traditional approach used here is justified.

Finally, the purpose of this research was to establish the cross-national validity of the SERVQUAL instrument. To do so, it was necessary to establish the structural and metric invariance of the model to determine if scale results are comparable across countries (Steenkamp and Baumgartner, 1998). This is true for both the first- and second-order factors in the model. Invariance is not imposed on the model, but it must exist *a priori* for meaningful comparisons to be made. The results of the study suggested that the second-order factor was partially invariant, with tangibles left free to vary between countries. This indicates that the parameters estimated are comparable across countries in a meaningful way, providing managers with useful information. It does not suggest that scores on the different dimensions will be identical and thereby managerially irrelevant. The limitation is on comparisons for the tangibles

dimension, which appears to be viewed differently across the countries.

Considering these limitations, administrators will still find the SERVQUAL convenient and reliable to use in the nursing home as a routine measure of service quality between countries. The model appears to be invariant and can be used to make meaningful managerial decisions regarding quality in the nursing home environment.

## Conclusion

With the aging population of Western societies, long-term health care is increasingly becoming an issue to be dealt with in policy discussions. The reported agreement between the US and UK to derive measures of health care quality comparable between the two countries drives the need to find a cross nationally validated measure of service quality. While SERVQUAL has been used and modified with varying degrees of success in different countries as diverse as South Africa and Hong Kong, there have been mixed results when it has been used in cross-national studies (Kettinger *et al.*, 1995). Presently, the cross-national study of SERVQUAL has given scant attention to the health care domain. The main purpose of this paper is to begin the process of assessing the cross-national reliability of the instrument within health care contexts.

## References

- Anderson, J.C. and Gerbing, D.W. (1988), "Structural equation modeling in practice: a review and recommended two-step approach", *Psychological Bulletin*, Vol. 103, pp. 411-23.
- Arbuckle, J.L. and Wothke, W. (1995), *AMOS 4.0 User's Guide*, SmallWaters Corporation, Chicago, IL.
- Asubonteng, P., McCleary, K.J. and Swan, J.E. (1996), "SERVQUAL revisited: critical review of service quality", *Journal of Services Marketing*, Vol. 10 No. 6, pp. 62-72.
- Babakus, E. and Mangold, W.G. (1992), "Adapting the SERVQUAL scale to hospital services: an empirical investigation", *Health Services Research*, Vol. 26 No. 6, pp. 676-86.
- Bigne, E., Molinar, M.A. and Sanchez, J. (2003), "Perceived quality and satisfaction in multiservice organizations: the case of Spanish public services", *Journal of Services Marketing*, Vol. 17 No. 4, pp. 420-42.
- Bolton, R.N. and Drew, J.H. (1991), "A longitudinal analysis of the impact of service changes on customer attitude", *Journal of Marketing*, Vol. 55, January, pp. 1-9.
- Bowers, M.R., Swan, J.E. and Koehler, W.F. (1994), "What attributes deter quality and satisfaction with health care delivery?", *Health Care Management Review*, Vol. 19 No. 4, pp. 49-55.

- Byrne, B.M. (1998), *Structural Equation Modeling with LISREL, PRELIS, and SIMPLIS*, Lawrence Erlbaum Associates, Hillsdale, NJ.
- Byrne, B.M. (2001), *Structural Equation Modeling with AMOS*, Lawrence Erlbaum Associates, Hillsdale, NJ.
- Byrne, B.M. and Campbell, T.L. (1999), "Cross cultural comparisons and the assumption of equivalent measurement and theoretical structure: a look beneath the surface", *Journal of Cross-Cultural Psychology*, Vol. 30 No. 5, pp. 555-74.
- Carman, J.M. (1990), "Consumer perceptions of service quality: an assessment of the SERVQUAL dimensions", *Journal of Retailing*, Vol. 66 No. 1, pp. 33-55.
- Cronin, J.J. and Taylor, S.A. (1992), "Measuring service quality: a reexamination and extension", *Journal of Marketing*, Vol. 56, July, pp. 55-68.
- Cronin, J.J. and Taylor, S.A. (1994), "SERVPERF versus SERVQUAL: reconciling performance-based and perceptions-minus-expectations measurement of service quality", *Journal of Marketing*, Vol. 58, January, pp. 125-31.
- Dabholkar, P.A., Shepard, D.C. and Thorpe, D.I. (2000), "A comprehensive framework for service quality: an investigation of critical conceptual and measurement issues through a longitudinal study", *Journal of Retailing*, Vol. 76 No. 2, pp. 139-73.
- Dean, A.M. (1999), "The applicability of SERVQUAL in different health care environments", *Health Marketing Quarterly*, Vol. 16 No. 3, pp. 1-15.
- Duffy, J.A., Duffy, M. and Kilbourne, W.E. (2001), "A comparative study of resident, family, and administrator expectations for service quality in nursing homes", *Health Care Management Review*, Vol. 26 No. 3, pp. 75-83.
- Duffy, M. and Duffy, J.A. (2002), "The residential care organization in the 21st century", in Hill, R. (Ed.), *Geriatric Residential Care*, Lawrence Erlbaum Associates, Hillsdale, NJ.
- Gabbott, M. and Hogg, G. (1995), "Grounds for discrimination: establishing criteria for evaluating health service", *The Services Industry Journal*, Vol. 15 No. 1, pp. 90-101.
- Grapentine, T. (1998), "The history and future of service quality assessment", *Marketing Research*, Vol. 10 No. 4, pp. 4-21.
- Hair, J.F., Anderson, R.E., Tatham, R.L. and Black, W.C. (1998), *Multivariate Data Analysis*, Prentice-Hall, Englewood Cliffs, NJ.
- Headley, D.E. and Miller, S.J. (1993), "Measuring service quality and its relationship to future consumer behavior", *Journal of Health Care Marketing*, Vol. 13 No. 4, pp. 32-41.
- Hunt, S.D. (1991), *Modern Marketing Theory: Critical Issues in the Philosophy of Marketing Science*, Southwestern Publishing Co., Cincinnati, OH.
- Jöreskog, K.G. (1971), "Simultaneous factor analysis in several populations", *Psychometrika*, Vol. 36, December, pp. 409-26.
- Jöreskog, K. and Sörbom, D. (1993), *LISREL 8: Structural Equation Modeling with the SIMPLIS Command Language*, Lawrence Erlbaum Associates, Hillsdale, NJ.
- Kettinger, W.J. and Lee, C.C. (1994), "Perceived service quality and user satisfaction with the information services functions", *Decision Sciences*, Vol. 25 No. 5/6, pp. 737-66.
- Kettinger, W.J., Lee, C.C. and Lee, S. (1995), "Global measures of information services quality: a cross-national study", *Decision Sciences*, Vol. 26 No. 5, pp. 569-88.
- Kleinsorge, I.K. and Koenig, H.F. (1991), "The silent customers: measuring customer satisfaction in nursing homes", *Journal of Health Care Marketing*, Vol. 11 No. 4, pp. 2-13.
- Lam, S.K. (1997), "SERVQUAL: a tool for measuring patients' opinions of hospital service quality in Hong Kong", *Total Quality Management*, Vol. 8 No. 4, pp. 145-53.
- Lee, H., Lee, Y. and Yoo, D. (2000), "The determinants of perceived service quality and its relationship with satisfaction", *Journal of Services Marketing Research*, Vol. 14 No. 2/3, pp. 217-32.
- Licata, J.W., Mowen, J.C. and Chakraborty, G. (1995), "Diagnosing perceived quality in the medical service channel", *Journal of Health Care Marketing*, Vol. 15 No. 4, pp. 42-9.
- Lytle, R.S. and Mokwa, M.P. (1992), "Evaluating health care quality: the moderating role of outcomes", *Journal of Health Care Marketing*, Vol. 12 No. 2, pp. 4-14.
- Nel, D. and Pitt, L.F. (1997), "The SERVQUAL instrument: reliability, validity in South Africa", *South African Journal of Business Management*, Vol. 28 No. 3, pp. 113-23.
- O'Connor, S.J., Shewchuk, R.M. and Carney, L.W. (1994), "The great gap", *Journal of Health Care Marketing*, Vol. 14 No. 2, pp. 32-9.
- Parasuraman, A., Zeithaml, V.A. and Berry, L.L. (1988), "SERVQUAL: a multiple item scale for measuring consumer perceptions of service quality", *Journal of Retailing*, Vol. 64 No. 1, pp. 12-40.
- Sharma, S. (1996), *Applied Multivariate Techniques*, Wiley, New York, NY.
- Soyoung, K. and Byoungcho, J. (2002), "Validating the retail service quality scale for US and Korean customers of discount stores: an exploratory study", *Journal of Services Marketing*, Vol. 16 No. 3, pp. 223-37.
- Steenkamp, J.-B.E.M. and Baumgartner, H. (1998), "Assessing measurement invariance in cross-national research", *Journal of Consumer Research*, Vol. 25, June, pp. 78-90.
- Sureshchander, G.S., Chandrasekharan, R. and Anantharaman, R.N. (2002), "The relationship between service quality and customer satisfaction – a factor specific approach", *Journal of Services Marketing*, Vol. 16 No. 4, pp. 363-79.
- Swanson, S.R. and Davis, J.C. (2003), "The relationship of differential *loci* with perceived quality and behavioral intentions", *Journal of Services Marketing*, Vol. 17 No. 2, pp. 202-19.
- Taylor, S.A. (1994), "Distinguishing service quality from patient satisfaction in developing health care marketing strategies", *Hospital & Health Services Administration*, Vol. 39 No. 2, pp. 221-36.
- Teas, R.K. (1994), "Expectations as a comparison standard in measuring service quality: an assessment of a reassessment", *Journal of Marketing*, Vol. 58, January, pp. 132-9.
- Woodruff, R.B., Cadotte, E.R. and Jenkins, R.L. (1983), "Modeling consumer satisfaction processes using experience-based norms", *Journal of Marketing Research*, Vol. 20, August, pp. 296-304.



## Appendix

Figure A1 SERVQUAL instrument adapted for use in long-term care (form for residents)

Code \_\_\_\_\_

SERVQUAL Instrument adapted for use in long term care.  
 (Form for Residents)

**DIRECTIONS:** The following set of statements relate to your feelings about your nursing home. For each statement, please show the extent to which you believe your nursing home has the feature described by the statement. Pick one of the seven numbers next to each statement. If you *strongly agree* that your nursing home possesses the feature, circle the number 7. If you *strongly disagree* that your nursing home possesses the feature, circle 1. If your feelings are not strong, circle one of the numbers in the middle. There is no right or wrong answers—all we are interested in is a number that best shows your perceptions about your nursing home services.

Disagree      Agree

1 2 3 4 5 6 7	1. Your nursing home has up-to-date equipment.
1 2 3 4 5 6 7	2. Its physical facilities are visually appealing.
1 2 3 4 5 6 7	3. The staff are well dressed and appears neat.
1 2 3 4 5 6 7	4. The appearance of the physical facilities of nursing homes is in keeping with the type of services provided.
1 2 3 4 5 6 7	5. When your nursing home promises to do something by a certain time, it does so.
1 2 3 4 5 6 7	6. When residents have problems, nursing home is sympathetic and understanding.
1 2 3 4 5 6 7	7. Your nursing home is dependable.
1 2 3 4 5 6 7	8. Your nursing home provides their services at the time they promise to do so.
1 2 3 4 5 6 7	9. It keeps records accurately.
1 2 3 4 5 6 7	10. Your nursing home tells residents exactly when services will be performed.
1 2 3 4 5 6 7	11. Residents receive prompt service from the nursing home staff.
1 2 3 4 5 6 7	12. The staff are always willing to help residents.
1 2 3 4 5 6 7	13. The staff are not too busy to respond to residents' requests promptly.
1 2 3 4 5 6 7	14. Residents can trust the nursing home staff.
1 2 3 4 5 6 7	15. Residents feel safe in their dealings with nursing home staff.
1 2 3 4 5 6 7	16. The staff at your nursing home are polite.
1 2 3 4 5 6 7	17. The employees get adequate support from the nursing home administration to do their jobs well.
1 2 3 4 5 6 7	18. The staff at your nursing homes are sensitive to individual differences among residents.
1 2 3 4 5 6 7	19. The staff at your nursing home gives residents personal attention.
1 2 3 4 5 6 7	20. The staff knows what the needs of the residents are.
1 2 3 4 5 6 7	21. The staff has the residents' interests at heart.
1 2 3 4 5 6 7	22. The staff offers services at times convenient to the residents.

### Executive summary and implications for managers and executives

*This summary has been provided to allow managers and executives a rapid appreciation of the content of the article. Those with a particular interest in the topic covered may then read the article in toto to take advantage of the more comprehensive description of the*

*research undertaken and its results to get the full benefit of the material present.*

#### Service quality in nursing homes

Under an agreement signed by the USA and Britain to work together to improve quality in health care, the two countries will share information and experience on the topic and

expand common criteria for measuring health care quality. Such agreements underscore the need for measures of quality that are valid and reliable across countries. However, the technical dimensions of quality care – such as mortality rates, staffing ratios and medical outcomes – may be easier to compare than the service quality dimensions of health care. The latter are difficult to measure because they are determined by the recipients of health care and measured in terms of patient perceptions of the health care experience. An instrument that does not vary across different settings is needed to compare health care service quality in two or more countries. With the need for long-term health care rising in both the US and Britain as the number of old people increases, policy makers and administrators must be able to understand exactly what constitutes quality care in nursing homes.

#### The SERVQUAL scale

SERVQUAL is a 22-item scale with five dimensions:

- (1) *tangibles* – physical facilities, equipment and appearance of personnel;
- (2) *empathy* – caring and individualized attention;
- (3) *assurance* – knowledge and courtesy of employees and their ability to convey trust and confidence;
- (4) *reliability* – ability to perform the promised service dependably and accurately; and
- (5) *responsiveness* – willingness to help customers and provide prompt service.

The scale has paired questions for expectations and perceptions. Service quality is operationalized as the difference between the measured expectations and perceptions. This is known as the gap score.

Few authoritative studies addressing the applicability of SERVQUAL outside the US have been published, and most research into the reliability and validity of SERVQUAL across

cultures has focused on users of information systems. Kilbourne *et al.* investigate the applicability of a modified SERVQUAL instrument as a means of measuring residents' perceptions of long-term health care service quality in the USA and Britain.

#### The authors' modifications to SERVQUAL

While gap scores have been shown to have better diagnostic capabilities, the perception-only measures of service quality appear to have higher convergent and predictive validity. The perception-only scores were therefore judged to serve the purposes of the Kilbourne *et al.* study better.

Moreover, some researchers have found that better results are sometimes obtained by having fewer than the normal five dimensions underlying SERVQUAL. After initial testing, Kilbourne *et al.* settled upon a four-dimension scale, which excludes "assurance".

#### The research results

The authors' research reveals that the modified, four-dimension SERVQUAL instrument, when applied to residents' perceptions of long-term health care service quality, does not vary across the USA and Britain. SERVQUAL clearly captures quite subtle quality indicators in a multidimensional way – tangibles, responsiveness, reliability and empathy, as well as the overall factor of service quality. Administrators will find SERVQUAL convenient and reliable to use in the nursing home as a routine measure of service quality between countries.

*(A précis of the article "The applicability of SERVQUAL in cross-national measurements of health care quality". Supplied by Marketing Consultants for Emerald.)*