Modelling and evaluating service quality measurement using neural networks

Ravi S. Behara
Department of Information Technology and Operations Management, College of Business, Florida Atlantic University, Boca Raton, Florida, USA
Warren W. Fisher
Department of Management, Marketing and International Business, Stephen F. Austin State University, Nacogdoches, Texas, USA and
Jos G.A.M. Lemmink
Faculty of Economics and Business Administration, Maastricht University, Maastricht, The Netherlands

Keywords *Service quality, Neural networks, Modelling, Customer satisfaction*

Abstract *Effective measurement and analysis of service quality are an essential first step in its improvement. This paper discusses the development of neural network models for this purpose. A valid neural network model for service quality is initially developed. Customer data from a SERVQUAL survey at an auto-dealership network in The Netherlands provide the basis for model development. Different definitions of service quality measurement are modelled using the neural network approach. The perception-minus-expectation model of service quality was found not to be as accurate as the perception-only model in predicting service quality. While this is consistent with the literature, this study also shows that the more intuitively appealing but mathematically less convenient expectation-minus-perception model out-performs all the other service quality measurement models. The study also provides an analytical basis for the importance of expectation in the measurement of service quality. However, the study demonstrates the need for further study before neural network models may be effectively used for sensitivity analyses involving specific dimensions of service quality.*

Introduction

It is well recognized that service quality is multifaceted and that it is ultimately evaluated in the minds of the customers (Sasser *et al.*, 1978; Grönroos, 1982; Lehtinen and Lehtinen, 1982; Parasuraman *et al.*, 1985). Customer evaluation of service quality is an exercise in the assessment of services along an acceptable-not acceptable continuum that involves the analysis of multiple criteria. Attempts to model and understand this have typically involved linear mathematical representations. Such representations are usually considered adequate approximations of what is intuitively known to be a more complex process of human decision making.

However, there have been related developments in other academic disciplines that provide the opportunity to further investigate the modelling of customers' evaluation of service quality. One development of interest is the

International Journal of Operations & Production Management,
Vol. 22 No. 10, 2002, pp. 1162-1185.

DOI 10.1108/01443570210446360

artificial intelligence approach of artificial neural networks. Artificial neural networks, or neural networks for short, have been able to successfully approximate complex non-linear functions and thereby facilitate pattern recognition and pattern classification (Hornik *et al.*, 1989; Hornik, 1991). A neural network is a parallel information-processing structure with a capacity to learn, and is based on the biology of the human brain as a network of neurons (nerve cells). A neural network, whether viewed as a "brain metaphor" of information processing or as a biologically inspired analytical tool, has been successfully applied to several types of problems such as classification, prediction, control, and surface fitting (Hecht-Nielsen, 1990). As such, it may be useful to investigate the appropriateness of this approach to study customer evaluation with respect to service quality.

This paper applies a neural network approach to modelling customer evaluation of service quality. Three distinct issues emerged when considering such an approach. First, the possibility of modelling customer evaluation of service quality adequately using neural networks is to be considered. To address this, the development of an adequate neural network representation of customer evaluation of service quality needs to be investigated. Second, since a neural network can be considered to be a "brain metaphor" of information processing, it may be possible to get some insight into issues related to how service quality is currently being measured and evaluated. And finally, it would be of interest to investigate whether neural network models of service quality can be of use to managers in their attempts to identify and improve targeted aspects of services that have the greatest returns in the overall evaluation of quality by customers. This is especially significant in the current environment of increasing demand for the effectiveness of investment in quality improvement efforts.

This paper addresses the issues identified above. After a brief introduction to neural networks, the development of a neural network model of service quality is presented. A systematic diagram representation of service quality, termed Reverse SERVQUAL, forms the basis for a neural network model. The model data are based on a SERVQUAL customer survey of customers of an auto-dealership network in The Netherlands. The model is initially evaluated and then used in experiments and sensitivity analyses to address the issues raised above. Implications for further research are then discussed. The BrainMaker Professional (1993) neural network simulation software is used in this study.

Neural networks

A neural network, or, more appropriately, an artificial neural network, is a model or methodological tool that is a loose adaptation of the processes by which the brain is thought to operate (McMillen and Henley, 2001). These models consist of a network of artificial neurons or nodes that are representative of neurons or nerve cells in the human brain. Neural networks are designed to mimic human mental activity, such as learning and pattern recognition. These characteristics have made it possible to apply them in diverse fields such as engineering, geology, ophthalmology, statistics, and economics (Cooper, 1999).

A number of successful business applications of neural networks have been discussed in the literature, particularly in financial services. Neural networks have out-performed other mathematical modelling tools in predicting corporate bond ratings and profitability (Surkan and Singleton, 1990). Other applications include credit card fraud (Rochester, 1990), bank failure prediction (Tam and Kiang, 1992), savings and loan associations failures (Salchenberger *et al.*, 1992), mortgage underwriting judgment (Collins *et al.*, 1988), stock price pattern recognition (Kamijo and Tanigawa, 1992), bankruptcy classification (Fletcher and Goss, 1993; Udo, 1993), bond ratings (Singleton, 1990), and predicting percentage change in the *S&P 500* (Fishman *et al.*, 1991). A neural network model for option pricing was also found to be more effective than traditional models in a study of recent volatile markets (Yao *et al.*, 2000).

There is also a growing interest in applying neural network approaches to address issues in other services. A neural network approach is used to forecast the failure rate of buses and rolling stock equipment at an Italian bus-manufacturing company (Bellandi *et al.*, 1998). It may be noted that failure rate forecasting is critical to both suppliers of buses and the buyers who are transportation service providers, with reference to reliability, safety and maintainability. The complex demand-supply relationship in Hong Kong's urban taxi services has also been investigated using neural network models (Xu *et al.*, 1999). Increasing domestic traffic and an aging aircraft fleet have led the Federal Aviation Authority in the USA to initiate new safety research. As a part of this effort, neural network models have been developed to forecast service problems related to specific groups of aircraft structural components (Nordmann and Luxhoj, 2000). However, the application of neural networks to model qualitative and intangible aspects of services is just emerging in the literature. Mozer *et al.* (1999) use neural networks to predict customer churn behavior, the loss of customers who switch from one carrier to another, in the wireless telecommunications industry. Their models include qualitative variables such as customer service and have out-performed the more traditional logit regression models. Neural network models that include qualitative criteria have also been developed in the area of higher education to predict the academic performance of graduate business students (Hoefer and Gould, 2000). After such initial successes in applying neural networks to specific service applications, it may be appropriate to extend its use to address more general and theoretical issues in services.

Neural network structure

An artificial neural network, or neural network for short, is a multi-layer network structure or a directed graph of simple interconnected processing elements called artificial neurons or nodes. The most successful applications to prediction and classification use a feed-forward design (Burke and Ignizio, 1992; BrainMaker Professional, 1993; Lawrence, 1994). Such designs typically have at least three layers of nodes. The first or input layer consists of input nodes that uniquely represent each input or predictor variable. The second or hidden layer consists of hidden nodes that are internal representations, and they facilitate the propagation of feed-forward information from the input layer to the output layer. In general, there could be one or more hidden layers between the input and output layers. The third or output layer consists of output nodes that represent the model's classification decision, with one node for each output (Figure 1).

In the context of this paper, where the neural network is used to model customers' decision processes, the input layer can be considered to represent the physical and psychological cues from the service, the hidden layer plays the role of the cognitive processes that mediate between the cues and the semantic output, and the output layer represents the semantic labels that customers give to the quality of their service experience (McMillen and Henley, 2001).

Information enters the network through neurons of the input layer, and output is generated at the neurons of the output layer. This is referred to as a feed-forward network. There can be one or more hidden layers of neurons between the input and output layers. Information is transmitted through the network via connections between neurons, with a connection weight associated with each connection or arc. These weights represent the network's learning, its embedded knowledge, and must be set to proper values by training the network.

Each neuron in a network is a processing element that performs summation and transfer functions in converting input to output. The summation function evaluates the signed weighted sum of all inputs at a given node. The resulting

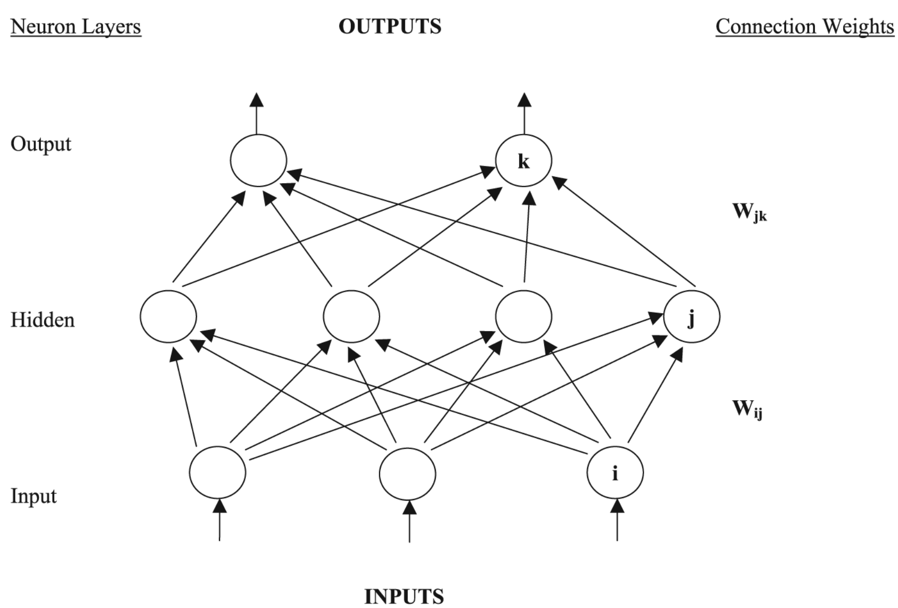


Figure 1. Typical feed-forward three-layer neural network

total input is passed through the transfer or activation function to create output. Network behavior depends substantially on the character of the neuron transfer function. These are typically linear or non-linear squashing functions. For example, the semi-linear sigmoid function:

$$f(x) = (1 + e^{-x}) - 1$$

is a typical transfer function used in neural networks. Other functions such as the linear threshold and step transfer functions are also employed. The choice of the transfer function is based on its ability to emulate the information processing being modeled.

The number of neuron layers, connection weights and transfer function define the architecture of a neural network. As noted above, a network must be trained with data sets for a given problem to establish the correct weights. The data are presented to the network as sets of input and corresponding output data, and such sets are called facts. The most common training procedure is known as back-propagation (Rumelhart *et al.*, 1986a, b). It works by starting with random connection weights, presenting sets of inputs or facts to the network, and letting the network calculate the outputs. The calculated outputs are then compared with known correct values, and a formula, called the generalized delta rule, is used to update network weights. Each pass-through of the facts is called a run, and additional runs are conducted until the calculated outputs are close to the correct values. At this point, the network learns to predict the output pattern. This is termed the supervised training of the network. The back-propagation training algorithm has become the accepted standard process in training the multi-layered, feed-forward neural network (Wilson and Sharda, 1992). It attempts to minimize an error measure, such as the sum of squared error, during the training process. This type of neural network forms the basis for model development in this study and the BrainMaker software is used to implement it.

Modelling service quality

While neural networks were inspired by the structure of the human brain, there appears to be a gap in the literature, when it comes to applying neural networks to model the human decision-making process. This paper attempts to address that gap by developing neural network models of human characteristics of pattern recognition and pattern classification involved in the evaluation of service quality, since it appears that neural networks can better exploit and represent the non-linear relationships that are inherent in such human processes.

The research discussed here is based on the applicability of certain types of neural networks for prediction and classification, that is the ability to predict an output or a classification for a given set of inputs. This usage is sometimes called pattern recognition. These types of networks can often be developed into generalized models that accurately predict outputs or classify input data. Bailey and Thompson (1990) indicate that neural networks also have the ability to effectively model heuristic criteria. These characteristics of neural networks can prove to be useful when modelling an individual customer's decision making. Hence the following proposition:

P1. A valid neural network model of customer evaluation of service quality can be developed.

SERVQUAL is a widely used instrument to measure customer evaluations of service quality (Parasuraman *et al*., 1988, 1991), and involves the measurement of service quality as the difference in customer expectations and perceptions scores. In this instrument, service quality is measured along the five conceptually distinct, yet interrelated dimensions of tangibles, reliability, responsiveness, assurance, and empathy. An overall measure of quality in the form of an average score across all these five dimensions can also be determined.

While SERVQUAL is a leading instrument in measuring service quality, a number of studies have raised concerns about it (Babakus and Boller, 1992; Cronin and Taylor, 1992; Teas, 1993; Peter *et al*., 1993; Buttle, 1996). These critiques can be considered under the categories of conceptual, methodological and analytical issues; and are summarized in the literature (Kettinger and Lee, 1994; Buttle, 1996; Van Dyke *et al*., 1999). Parasuraman *et al*. (1994) provide detailed responses to these critiques, arguing that the conceptual and empirical evidence they present casts doubt on the alleged severity of the concerns of the SERVQUAL approach and on the purported improvements provided by the alternative approaches proposed by the critiquing researchers. Subsequently, these arguments have largely been refuted by the critiquing researchers (Cronin and Taylor, 1994; Teas, 1994). But they all identify some of the unresolved areas in this debate and articulate a set of relevant research questions. Key issues of the debate and the proposed research agenda in the measurement of service quality that are pertinent to this study are discussed below.

The primary issue raised in the literature is the conceptualization of service quality as the difference between perception and expectation (Cronin and Taylor, 1992). But this disconfirmation of expectations conceptualization of service quality is well established in the conceptual work in service quality (Helson, 1964; Sasser *et al*., 1978; Grönroos, 1982; Bolton and Drew, 1991; Parasuraman *et al*., 1994). In addition, Parasuraman *et al*. (1994) reiterate that there is strong theoretical and empirical evidence to support that assessment of performance occurs with reference to some norms or standards. They also recognize an area for additional research as the need to identify the most appropriate way to incorporate expectations into service quality measurement. However, Cronin and Taylor (1994) reiterate their view against the disconfirmation-based SERVQUAL scale of measuring service quality, and continue to be proponents of their perceptions-only approach to service quality measurement. Some empirical investigations involving issues related to expectations and performance gaps have been undertaken (Johnson and

Mathews, 1997; Chenet *et al.*, 2000). This debate over the approach to measure service quality provides an opportunity for further investigation through the use of a neural network model:

P2. A neural network model of customer evaluation of service quality can be used to evaluate the gap method and the perception-only approach to measuring service quality.

P3. Neural network models provide an approach to evaluate other ways of incorporating expectations into service quality measurements.

Further, the use of the perception and expectation gap measure of service also raises related analytical concerns about its low reliability, poor discriminant validity, and spurious correlations (Peter *et al.*, 1993; Cronin and Taylor, 1992; Van Dyke *et al.*, 1999). Parasuraman *et al.* (1994) indicate that, on the issue of discriminant validity, SERVQUAL performs just as well as the perception-only-based SERVPERF (Cronin and Taylor, 1992) on various validity criteria. They also highlight the fact that inter-factor correlations in SERVQUAL have always been recognized as being present (Parasuraman *et al.*, 1988, 1991). In fact, they identify that the nature and causes of these interrelationships between factors representing the service quality dimensions are a useful area for further research. The neural network model developed in this study (*P1*) provides an approach to explicitly model interactions between these factors through the interconnections between the network nodes.

The diagnostic value of different approaches to measuring service quality is also in dispute. Parasuraman *et al.* (1990, 1994) show that service quality measurements that include customer expectations provide more information to managers than those that focus on perceptions only. However, Cronin and Taylor (1992) argue otherwise by stating the popularity of the perceptions-only approach. But Parasuraman *et al.* (1994) consider that the greater variability of SERVQUAL scores across the various dimensions provides for a better method of identifying areas of deficiency in service quality in companies. In addition, they acknowledge the lower predictive power of their gap model compared with the perceptions-only model, and suggest the need for a trade-off between predictive power and diagnostic value of the two methods of service quality measurement. Ultimately it must be recognized that the purpose of measurement is to improve service quality. Managers should be able to understand where limited resources must be applied in order to achieve the greatest improvements for their customers. As such, the two approaches of measurement of gaps and perceptions-only should be evaluated for their ability to provide assistance to managers in this regard. It is suggested here that neural network models can be used for this purpose:

P4. Neural network models of the gaps and perceptions-only service quality measurement approaches can be used to understand the impact of quality improvement efforts on the customer evaluation of service quality.

The debate regarding the gaps and perceptions-only approaches to measuring service quality is still unresolved. There are valid issues and suggestions on either side of this debate. This study contributes to the debate through a non-statistical analytical approach using neural network models of service quality. Such an approach has not been considered before in the literature. The propositions stated above are evaluated through experiments involving the development and use of neural network models. The SERVQUAL instrument is used as a basis for the development of such models. Variations to the neural network models developed are then investigated through experiments.

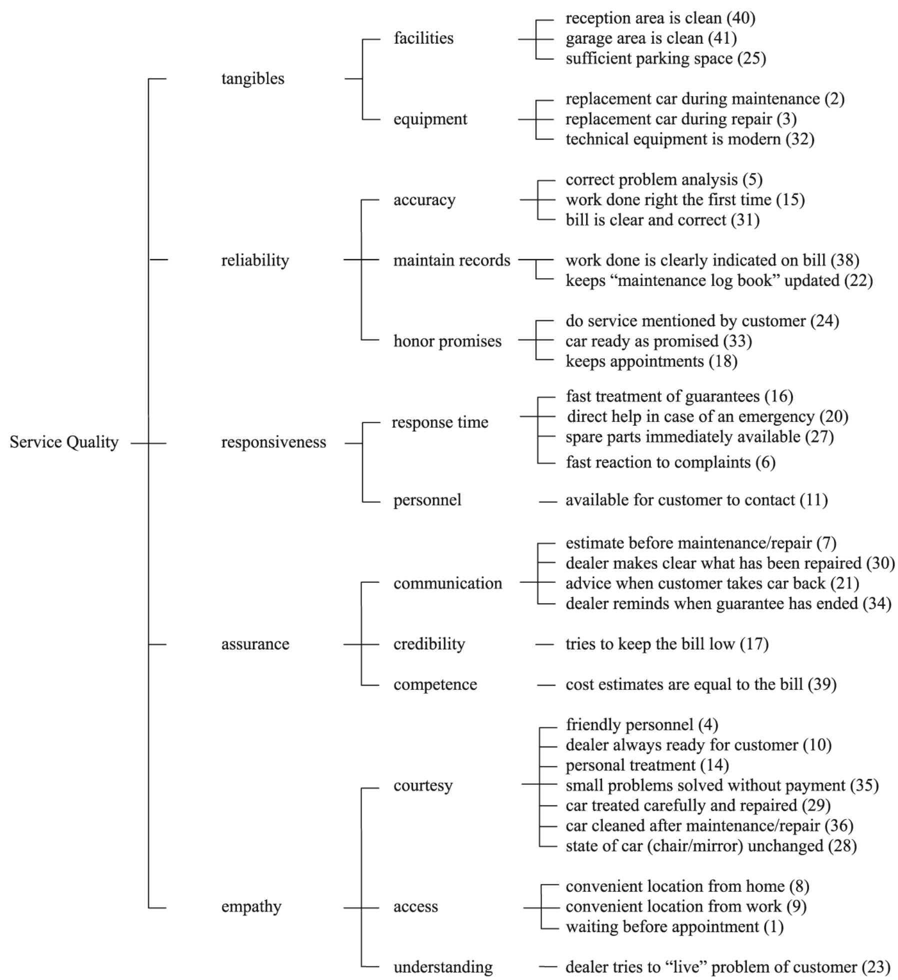
Service quality neural network model

Reverse SERVQUAL model

An initial step in the development of a neural network model for service quality evaluation is to develop a conceptual model that is amenable to neural network modelling. Such a conceptualization is realized through the use of a systematic diagram representation of the SERVQUAL model. The systematic diagram is a technique developed to search for the most appropriate and effective means of accomplishing given objectives (Mizuno, 1988). The technique represents a hierarchy of means that help achieve a given goal. By visually depicting a sequence of intermediate goals and means, the systematic diagram identifies the key items that allow for the achievement of the primary goal. When considering service quality as the primary goal, the quality of service to the customer needs to be translated into identifiable substitute characteristics that must be satisfied to achieve the intended overall quality.

Service quality can be represented as a hierarchy of goals and means in the shape of a systematic diagram such as the reverse SERVQUAL model (Lemmink and Behara, 1992). The five SERVQUAL determinants effectively represent the primary means to achieve the goal of quality service. These primary means may be considered as intermediate goals that are achieved by a secondary set of means. But the five SERVQUAL determinants have been synthesized by combining items. These items represent the secondary means. The secondary means may also be viewed as intermediate goals. These goals are achieved by tertiary means. In the context of SERVQUAL, tertiary means are represented by individual statements for measuring customer expectations and perceptions of service quality. These tertiary means also represent issues that are a target for continuous improvement by managers. The systematic diagram representation of service quality in the auto-services study is shown in Figure 2.

As shown in Figure 2, the primary goal of service quality is broken down into primary, secondary and tertiary means to achieve that primary goal, as the chart is read left to right. At the far right are the tertiary means or characteristics that correspond directly to questions the auto-dealership survey used for this research. This systematic diagram is a hierarchical representation of the means to achieve service quality. The neural network for this research is based on the Reverse SERVQUAL model illustrated in Figure 2.



Note: Variables numbered as in Table I

Figure 2.
Reverse SERVQUAL: systematic diagram of service quality

Survey data

Data were collected from customers who evaluated service quality at an auto-dealership network in The Netherlands using a SERVQUAL-based customer survey. This research uses responses to 73 questions on a seven-point Likert scale. The first 36 of the questions asked about expectations concerning specific service system characteristics (for example, reception area is neat). The question topics are summarized in Table I and are cross-referenced in Figure 2.

The missing numbers represent questions not appropriate for this study. Another 36 questions addressed perceptions of the same characteristics, i.e. how good the customers perceived service. Gaps were calculated as the

No.	Variable descriptions	No.	Variable descriptions
1	Waiting before appointment	22	Keeps "maintenance log book" updated
2	Replacement car during maintenance	23	Dealer tries to "live" problem of customer
3	Replacement car during repair	24	Service mentioned by customer is carried out
4	Friendly personnel	25	Sufficient parking space
5	Correct problem analysis	27	Spare parts immediately available
6	Fast reaction to complaints	28	State of car (chair/mirror) unchanged
7	Estimate before maintenance/repair	29	Car treated carefully
8	Convenient location from home	30	Dealer must make clear what has been repaired
9	Convenient location from work	31	Bill is clear and correct
10	Dealer always ready for customer	32	Technical equipment is modern
11	Available for customer to contact	33	Car ready as promised
14	Personal treatment	34	Dealer reminds when guarantee has ended
15	Work done right the first time	35	Small problems are solved without payment
16	Fast treatment of guarantees	36	Car cleaned after maintenance/repair
17	Tries to keep the bill low	38	Work done is clearly indicated on bill
18	Keeps appointments	39	Cost estimates are equal to the bill
20	Direct help in case of an emergency	40	Reception area is clean
21	Advice when customer takes car back	41	Garage area is clean

Note: Variable numbers and descriptions are the same for corresponding perceptions, expectations and gaps. Missing numbers represent questions not appropriate for this study

Table I. Description of input variables

difference between the perceptions and expectations for each corresponding pair of questions. The expectations, perceptions and gaps were used as inputs to the neural network in several combinations, as described below. A single question regarding overall customer evaluation of the service quality served as the output. There were a total of 161 customer responses to the survey.

Neural network experiments

The survey data were used in several experiments that were conducted to develop the network architecture suitable for representing service quality. Initially, experiments were conducted to identify the network configurations and parameters that produced the best network performance by generating outputs that most closely matched actual customer evaluations of service quality. An appropriate configuration was identified and subsequently used in experiments that addressed the four research propositions posited in the preceding section. The following discussion reviews these experiments, the results and the implications of using neural networks to model and evaluate service quality measurement.

Network parameters

Several parameters affect the speed of training and the quality of the trained network, which is the ability to correctly calculate outputs from the inputs. Learning rate controls the percentage of network error from the current run that is applied to the network by the back-propagation procedure. Larger learning rates tend to converge more quickly but may produce oscillations between relatively poor solutions. In some cases, better results are obtained by beginning with a relatively large learning rate and lowering it during training. For example, the learning rate might begin at 0.9 for the initial untrained network and be reduced toward a lower limit of 0.1, as the calculated outputs approach the desired values. The smoothing or momentum factor controls the percentage of error from past runs that is applied to the network at each update. Larger smoothing factors sometimes converge more rapidly by overcoming local optima. Lawrence (1994) and Jain and Nag (1995) provide detailed formulae illustrating the use of learning and smoothing parameters. Another parameter used in some implementations is the training tolerance (BrainMaker Professional, 1993, pp. 3-16). If used, the network is only updated by those facts whose calculated output differs from the desired output by more than the tolerance, which is usually expressed as a percentage. As with the learning rate, better results are sometimes obtained by starting with a relatively large tolerance and reducing it during training.

It is normally desirable for the network to be able to generalize. This is the ability of the network to calculate correct output(s) for facts that it has not previously seen. This is easily tested by use of a hold-out sample. That is, the network is trained with only part of the data set. It is then tested using facts from the hold-out sample that were not used for training. The network's ability to generalize is dependent on number of both training runs and hidden nodes (Burke and Ignizio, 1992; BrainMaker Professional, 1993; Lawrence, 1994). If too few training runs are conducted, the weights will not have been adjusted to their proper values. If too many runs have been conducted, however, the neural net may simply memorize the facts and be unable to generalize. A similar effect occurs with too many hidden nodes, while too few hidden nodes leave the network with inadequate computational power to learn the fact set. The hidden nodes can be organized into more than one layer, but there is little evidence that this produces better results (Lawrence, 1994).

Successful studies in the literature report using a variety of network parameter settings. Salchenberger *et al.* (1992), who used a neural network to predict savings and loan association failures, used a single hidden layer with 75 percent of the number of nodes in the input layer. The solution was by back-propagation using a variable learning rate beginning at 0.9 and going down during training. Momentum was initially set at 0.6 and was adjusted upward during training, which continued until no further improvement could be obtained. Philipoom *et al.* (1994) used a neural network to determine internally-set due dates in a job shop. The authors used 23 input nodes and a single hidden layer of nine nodes. The network's ability to generalize was tested with a 20 per cent hol-dout sample after every 25 runs, and the optimal network was selected as that yielding the best test results on the hold-out sample, but the learning parameters were not reported. Lenard *et al.* (1995) used a neural network to predict the going-concern uncertainty. The best performance resulted from using a number of hidden nodes that was approximately 75 per cent of the number of input nodes, while the learning rate was set at 1.0 and momentum at 0.9. Training continued until no further improvement could be obtained. Jain and Nag (1995) used a neural network to determine initial public stock offering prices. The authors used 11 input nodes and experimented with numbers of hidden nodes ranging from four to 12 in a single layer. The learning rate and momentum were both started at 0.4 and adjusted downward to 0.05 during training, which continued until no further improvement could be obtained. Ability to generalize was tested with a 50 per cent hold-out sample and was found to be good with all configurations tested.

Experiments for this research were conducted with the BrainMaker Professional (1993) software. Network parameters have a wide variety of options, with the following defaults: one hidden layer with the same number of nodes as inputs; learning rate was set at 1.0; smoothing (momentum) at 0.9; training tolerance at 0.1; and 10 per cent of facts are held out for testing.

Experiments for network design

Facts or data sets of inputs and outputs were presented to neural networks in four different input configurations: gaps (36 inputs), perceptions (36 inputs), perceptions and expectations (72 inputs). Gap experiments were conducted separately with gap = [perception – expectation] and gap = [expectation – perception]. While service quality is usually defined as the difference between perception and expectation, i.e. [perception – expectation], the use of the negative difference of [expectation – perception] was made as an experimental transformation. Each of these input configurations resulted in four distinct network structures. Network parameters were identified for each of these network structures as follows.

Because a variety of parameter settings have been used successfully in the literature, the optimize network option within BrainMaker's Genetic Training Option add-on (Genetic Training, 1993) was used to automate the process of testing many parameter combinations. Tests included combinations similar to BrainMaker defaults and those similar to combinations reported to perform well in the literature. The setting combinations used are shown in Table II.

These combinations were tested in full factorial, producing 72 parameter combinations for networks with 36 inputs, and 90 combinations for the network with 72 inputs. Each combination was tested at 30 different run lengths from 50 to 1,500.

One disadvantage of conducting a large number of parameter experiments is that some of the combinations may perform well on the hold-out sample by accident. In such cases, the network happens to perform well on this particular hold-out sample but would not perform well on others. A double hold-out

Table II.
Network setting combinations used

Network structure component	Settings used
Hidden nodes	One layer with 24, 28, 32, and 36 hidden nodes for the perception and gap experiments (36 inputs); 36, 45, 54, 63, and 72 hidden for the expectation and perception experiments (72 inputs)
Learning rate	Constant rate of 0.1; variable 0.5 to 0.1; variable 0.9 to 0.1
Smoothing (momentum)	0.1, 0.5, and 0.9
Training tolerance	Constant rate of 0.1; variable 0.2 to 0.1
Testing tolerance: 0.2 (20 per cent)	This parameter applies to hold-out sample tests. A given fact is reported as "good", if the calculated output is within 20 per cent of the correct value
Training length	1,500 runs with testing every 50 runs

approach was used to reduce the possibility of this occurrence. Specifically, 10 per cent or 16 randomly selected facts were removed from the original data set of 161 facts. This formed the secondary hold-out data set. Of the remaining 145 facts, BrainMaker removed an additional 10 per cent or 14 for testing during the optimization runs. This formed the primary hold-out data set. The network is trained using the remaining 131 facts. The network configurations are then tested using the 14-fact primary hold-out data to identify the best performers along the criteria of least average error, least root mean square error (RMSE), and largest percent good or correctly predicted outputs. The networks performing best on these criteria were retested on the 16-fact secondary hold-out data set. In this manner, the secondary hold-out tests of the networks were with data not previously used for either training or testing during the optimization runs.

Experiment for P1

This proposition states that a valid neural network model of customer evaluation of service quality can be developed. The first neural network model was developed to investigate this possibility. The input configuration for the network was defined by using SERVQUAL-based gap (perception – expectation) scores. This resulted in 36 input nodes, a hidden layer, and an output layer consisting of one node representing the overall evaluation of service quality. Experiments were conducted to determine network parameters, as discussed above. For each optimization run, BrainMaker reports the network configurations that performed the best using the primary hold-out sample on three separate criteria: the best performers along the criteria of least average error, least root mean square error (RMSE), and largest percent good or correctly predicted outputs. Each of those configurations was then retested using the secondary hold-out sample. The results are presented in Table III.

As shown in Table III, the network configuration with least average errors using the primary hold-out sample also performed the best on all three criteria, when using the secondary hold-out sample. Hence a network with 75 per cent accuracy in predicting customer evaluations with a low (0.6944) average error was developed. A greater accuracy would be preferred, and may be achieved with a larger training data set. However, this experiment clearly shows that a valid neural network model of customer evaluation of service quality can be developed.

Experiment for P2

The second proposition suggests that a neural network model of customer evaluation of service quality can be used to compare the P-E gap method and the perception-only approach to measuring service quality. Since a neural network can be viewed as a simulation of customer evaluation, it provides a new approach to address the debate regarding the appropriateness of each of these methods of measuring service quality. Two distinct network models were compared in this experiment. The network developed for *P1* above provided the model for the P-E gap approach to measuring service quality. A second network was then developed, in which the configuration consisted of 36 input nodes for perception-only input data. Once again, the network configurations were tested with both the primary and secondary hold-out data. The results are shown in Table IV.

Criteria with primary hold-out	Hidden nodes	Training tolerance	Learning rate	Smoothing	Training runs	RMSE[a]	Average error[a]	% good
Least RMSE	28	0.1	0.9 to 0.1	0.5	400	0.9636	0.7928	62.5
Least avg. error	24	0.2	0.1	0.9	450	0.8792	0.6944	75.0[b]
Highest % good	28	0.1	0.9 to 0.1	0.5	400	0.9636	0.7928	62.5

Notes: [a] BrainMaker scales all data and reports scaled errors. The above numbers are the actual errors, which are four times the BrainMaker report results for this research
[b] This configuration, denoted the "gap – average error" network, produced the best overall

Table III. Testing gaps (P-E) network configurations with secondary hold-out sample

Criteria with primary hold-out	Hidden nodes	Training tolerance	Learning rate	Smoothing	Training runs	RMSE[a]	Average error[a]	% good
P-E gap model								
Least RMSE	28	0.1	0.9 to 0.1	0.5	400	0.9636	0.7928	62.5
Least avg. error	24	0.2	0.1	0.9	450	0.8792	0.6944	75.0[b]
Highest % good	28	0.1	0.9 to 0.1	0.5	400	0.9636	0.7928	62.5
Perceptions-only model								
Least RMSE	28	0.2 to 0.1	0.1	0.9	50	0.6899	0.5975	68.75
Least avg. error	36	0.1	0.1	0.5	200	0.7021	0.5623	81.25[b]
Highest % good	36	0.1	0.9 to 0.1	0.5	50	0.7370	0.6388	81.25

Notes: [a] BrainMaker scales all data and reports scaled errors. The above numbers are the actual errors, which are four times the BrainMaker report results for this research
[b] These configurations denote the best performing networks

Table IV. Comparing P-E gap and perception-only network configurations with secondary hold-out sample

The results show that the perception-only model is better at correctly predicting customer evaluation of service quality than the P-E gap model. This is consistent with the reported superior predictive power of perception-only measure of service quality as opposed to the P-E gap measure of service quality (Cronin and Taylor, 1992; Parasuraman *et al.*, 1994). In addition, the best performing P-E model required 450 training runs compared with 200 runs for the best perception-only model, and only 50 runs for a comparable (81.25 per cent good outputs) perceptions-only model (see Table IV). This supports the accepted notion that over-training of networks typically degrades its ability to generalize. Since similar levels of accuracy require this over-training, it may be surmised that the P-E construct for measuring the input to the neural network service quality evaluation model is not as effective as a perceptions-only approach. Hence the above experiments show that the perception-only measure of service quality is a better predictor of customer evaluation of service quality than the P-E gap measure.

Experiment for P3

The third proposition indicates that the neural network modelling method developed in this study provides an approach to evaluate other ways of incorporating expectations into service quality measurements. Expectations have typically been used to determine gap scores. The gap is defined as perceptions minus expectations or (P-E). Defining gap as P-E is mathematically appealing, because the gap is positive when perceptions exceed expectations, thereby indicating a positive customer evaluation. In addition, a negative gap score indicates that the service being evaluated needs to be improved along the dimensions identified as performing below expectations. Parasuraman *et al.* (1994) and Teas (1994) discuss inconclusively other approaches to calculating the gaps under various circumstances. These are variations on the P-E definition of the gap.

This paper suggests a different approach to defining gaps. It may be generally assumed that most customers enter a service situation with some expectations (Johnson and Mathews, 1997). These expectations are formed either by previous experiences of the same or similar service, previous experiences of a different type of service, or simply expectations generated by the customer independently of any specific service experience. Hence customers usually undertake a service experience with some preconceived expectations, and thereafter develop a perception of that experience. So service quality measurement as a disconfirmation of expectations could be measured as expectations minus perceptions or E-P. This is not intuitive, since it is a reverse negative disconfirmation approach to measurement and involved a double negative. A positive E-P score implies that customer expectations were not met. This proposed E-P model of service quality was tested using neural networks.

A second approach to including expectations in service quality measurement is now considered. Customer expectations are generally accepted as a part of the service experience, but their exact role in the overall evaluation of service quality is still debatable. Since the neural network may be considered as a "brain metaphor", it may be allowed to represent the interactions of expectations and perceptions without a relationship between them being predefined. This approach was tested by designing a network that used all the 36 expectations and 36 perceptions as 72 independent inputs. The model developed is referred to as the expectations and perceptions or E&P model.

The performances of the two models to measure service quality proposed above were compared with the performance of the two traditional P-E and the perceptions-only measurement models. The results are shown in Table V.

The E-P model performed the best when tested with the secondary hold-out data. It has the best predictability as given by the highest percentage of correct predictions, and the lowest errors. In addition, the specific network trained to its level of accuracy in a short duration of only 50 runs. This ensures that the specific model is not developed through a memorization of the data, but is more an effective representation of customer evaluation. The expectations and perceptions (E&P) model performed as well as the perceptions-only model with respect to percentage of correct predictions. However, the perceptions-only model predicted with slightly better accuracy, while the E&P model trained in a shorter duration. The traditional P-E gap model was the worst performing model with the least percentage of correct predictions, most training runs, and

Criteria with primary hold-out	Hidden nodes	Training tolerance	Learning rate	Smoothing	Training runs	RMSE[a]	Average error[a]	% good
P-E gap model								
Least RMSE	28	0.1	0.9 to 0.1	0.5	400	0.9636	0.7928	62.5
Least avg. error	24	0.2	0.1	0.9	450	0.8792	0.6944	75.0[b]
Highest % good	28	0.1	0.9 to 0.1	0.5	400	0.9636	0.7928	62.5
Perceptions-only model								
Least RMSE	28	0.2 to 0.1	0.1	0.9	50	0.6899	0.5975	68.75
Least avg. error	36	0.1	0.1	0.5	200	0.7021	0.5623	81.25[b]
Highest % good	36	0.1	0.9 to 0.1	0.5	50	0.7370	0.6388	81.25
E-P gap model								
Least RMSE	24	0.1	0.9 to 0.1	0.9	1,450	1.1205	0.9008	56.25
Least avg. error	24	0.1	0.9 to 0.1	0.9	50	0.6477	0.5554	87.5[b]
Highest % good	24	0.1	0.9 to 0.1	0.9	50	0.6477	0.5554	87.5
Expectations and perceptions E&P model								
Least RMSE	36	0.2 to 0.1	0.1	0.9	100	0.7808	0.6816	75.0
Least avg. error	36	0.2 to 0.1	0.1	0.9	50	0.7363	0.6408	81.25[b]
Highest % good	36	0.2 to 0.1	0.1	0.9	100	0.7808	0.6816	75.0

Notes: Test results on secondary (16-fact) hold-out of those networks producing the best RMSE, average error, and number (#) good on the respective primary (14-fact) hold-out tests during the optimization runs. See text; [a] BrainMaker scales all data and reports scaled errors. The above numbers are the actual errors, which are four times the BrainMaker report results for this research; [b] This configuration, denoted the "gap – average error" network, produced the best overall results on the secondary hold-out. It is identical to the "gap – number (#) good" network

Table V. Comparing new and traditional approaches to measuring service quality (secondary holdout sample used to test models) least accuracy. This experiment shows that there is initial evidence that the more intuitive E-P gap model of service quality that assumes that customers typically enter a service experience with some expectations is better than both the mathematically convenient P-E gap model and the perceptions-only model. Further, the perceptions-only model that is presented in the literature as an alternative to the traditional P-E model is similar to a neural network model that uses perceptions and expectations without any predefined relationship. This supports the view that expectations do and should play a role in the measurement of service quality. It also shows that neural networks can be used to evaluate new approaches to measuring service quality.

Experiment for P4

The fourth proposition in this study suggests that neural network models of service quality measurement can be used to understand the impact of quality improvement efforts on the customer evaluation of service quality. The experiment conducted was essentially a sensitivity analysis using the neural network models developed in this study. Since the E-P gap model performed better than any other model, it was used in this experiment instead of the traditional P-E gap model. The perceptions-only model was also used, as it is the second best performing model, and is presented in the literature as an alternative approach to gap models of service quality. The least average error network was used in each case, as it represented the best performing network (see Table V). The impact of changes in service performance along the various dimensions (network inputs) on customer evaluation of service quality (network output) was simulated using these two models. This was done using the BrainMaker sensitivity analyzer option to analyze inputs with multiple facts (BrainMaker Professional, 1993). Using the secondary hold-out sample, each input variable was systematically varied up and down 10 per cent (+10 per cent) from its original value for each fact, and the resulting change in the output variable was recorded. The results were averaged across the 16 facts of the secondary hold-out, and scaled to fall within the range of 0 to 1 in BrainMaker. To accomplish this, each of the 16 secondary facts was applied one at a time, to the trained network in BrainMaker. For each fact, one input at a time was changed, holding all other inputs at their original value. Then the scaled output was calculated separately with the current input increased by 10 per cent and the current input decreased by 10 per cent. The overall scaled change in output is given by:

$$\text{scaled change in output} = \frac{\text{scaled output for 10\% increase in input} - \text{scaled output for 10\% decrease in input}}{2}.$$

Thus the result obtained was the scaled output change per 10 per cent change in input. The calculation was repeated for every input and for every fact, and then averaged across the facts, yielding a single mean scaled change in output for each input service criterion. These results presented in Tables VI (E-P gap model) and VII (perceptions-only model) give the mean effect or change in the overall evaluation of service quality (network output) for a ±10 per cent change in the corresponding dimension of service performance (network inputs).

The inputs shown in Table VI are E-P gaps. Each gap is increased, as the corresponding inputs are varied from –10 to +10 per cent of their base value. This represents a widening of the gap or difference between expectations and perceptions that should logically result in a reduced overall service quality evaluation by the customer. Such a reduction in service quality is identified by

	Input	Mean effect
Gap 18	Keeps appointments	–0.0329
Gap 4	Friendly personnel	–0.0328
Gap 32	Technical equipment is modern	–0.0293
Gap 24	Service mentioned by customer is carried out	–0.0234
Gap 29	Car treated carefully	–0.0228
Gap 16	Fast treatment of guarantees	–0.0227
Gap 9	Convenient location from work	–0.0192
Gap 10	Dealer always ready for customer	–0.0174
Gap 36	Car cleaned after maintenance/repair	–0.0141
Gap 17	Tries to keep the bill low	–0.0115
Gap 31	Bill is clear and correct	–0.0091
Gap 3	Replacement car during repair	–0.0086
Gap 39	Cost estimates are equal to the bill	–0.0085
Gap 23	Dealer tries to "live" problem of customer	–0.0073
Gap 6	Fast reaction to complaints	–0.0067
Gap 34	Dealer reminds when guarantee has ended	–0.0067
Gap 25	Sufficient parking space	–0.0036
Gap 11	Available for customer to contact	–0.0016
Gap 8	Convenient location from home	–0.0008
Gap 40	Reception area is clean	0.0008
Gap 38	Work done is clearly indicated on bill	0.0032
Gap 30	Dealer must make clear what has been repaired	0.0047
Gap 20	Direct help in case of an emergency	0.0058
Gap 7	Estimate before maintenance/repair	0.0074
Gap 1	Waiting before appointment	0.0085
Gap 2	Replacement car during maintenance	0.0097
Gap 21	Advice when customer takes car back	0.0104
Gap 33	Car ready as promised	0.0106
Gap 15	Work done right the first time	0.0155
Gap 41	Garage area is clean	0.0166
Gap 27	Spare parts immediately available	0.0186
Gap 14	Personal treatment	0.0207
Gap 5	Correct problem analysis	0.0208
Gap 35	Small problems are solved without payment	0.0228
Gap 28	State of car (chair/mirror) unchanged	0.0240
Gap 22	Keeps "maintenance log book" updated	0.0259

Note: Average scaled output change per 10% change in scaled input. Gaps are (expectations – perceptions), so negative effects are the norm

Table VI.
Sensitivity analysis (mean effect of inputs for "gap – average error" network)

	Input	Mean effect
Perception 3	Replacement car during repair	0.0186
Perception 6	Fast reaction to complaints	0.0185
Perception 17	Tries to keep the bill low	0.0169
Perception 4	Friendly personnel	0.0142
Perception 16	Fast treatment of guarantees	0.0132
Perception 34	Dealer reminds when guarantee has ended	0.0121
Perception 18	Keeps appointments	0.0120
Perception 11	Available for customer to contact	0.0118
Perception 40	Reception area is clean	0.0100
Perception 38	Work done is clearly indicated on bill	0.0099
Perception 21	Advice when customer takes car back	0.0094
Perception 9	Convenient location from work	0.0092
Perception 20	Direct help in case of an emergency	0.0086
Perception 14	Personal treatment	0.0032
Perception 32	Technical equipment is modern	0.0032
Perception 39	Cost estimates are equal to the bill	0.0032
Perception 30	Dealer must make clear what has been repaired	0.0024
Perception 28	State of car (chair/mirror) unchanged	0.0023
Perception 24	Service mentioned by customer is carried out	–0.0002
Perception 8	Convenient location from home	–0.0009
Perception 15	Work done right the first time	–0.0018
Perception 33	Car ready as promised	–0.0022
Perception 31	Bill is clear and correct	–0.0023
Perception 10	Dealer always ready for customer	–0.0031
Perception 29	Car treated carefully	–0.0041
Perception 5	Correct problem analysis	–0.0044
Perception 36	Car cleaned after maintenance/repair	–0.0046
Perception 25	Sufficient parking space	–0.0065
Perception 27	Spare parts immediately available	–0.0084
Perception 41	Garage area is clean	–0.0096
Perception 35	Small problems are solved without payment	–0.0120
Perception 7	Estimate before maintenance/repair	–0.0125
Perception 23	Dealer tries to "live" problem of customer	–0.0133
Perception 1	Waiting before appointment	–0.0144
Perception 22	Keeps "maintenance log book" updated	–0.0163
Perception 2	Replacement car during maintenance	–0.0183

Note: Average scaled output change per 10% change in scaled input

Table VII.
Sensitivity analysis (mean effect of inputs for "perception – average error" network)

a corresponding negative network average output value. The output scores in Table VI should therefore be negative. They are also scaled, as stated above, due to the nature of representation in BrainMaker. About half the inputs (when the gaps are increased) create negative service quality changes, as expected. However, positive or increased service quality results were obtained, when the gaps for the other inputs were increased. This anomaly could be due to the noisiness of the survey data. Noisy data exist when customers responding to the survey have similar evaluation on individual service criteria but very different evaluations of the overall service quality. This results in similar input data for the neural network, with very different corresponding outputs. Since neural networks learn through pattern matching of input-output combinations, the noisiness in data can lead to some degree of ambiguity in network performance. As seen in earlier experiments in this study, the networks developed here were robust enough to provide adequate levels (up to 87.5 per cent accuracy (see Table V)) of accuracy in predicting service quality evaluation, when tested against previously unused (secondary) data. However, the same networks now show unexpected output behavior for about half the inputs. This shows that the neural networks developed in this study to model service quality are adequate for predicting overall evaluation of customers but not robust enough for sensitivity analyses, indicating a need to investigate the issue in future research.

Table VII shows the results of sensitivity analysis using the perceptions-only model. The inputs are customer perceptions of the various criteria of service, and each is varied from –10 to +10 percent of their base value. This represents an increase in customer perception that should logically result in an increased overall service quality evaluation by the customer. Such an increase in service quality is identified by a corresponding positive change in the network average output value. These scaled output scores are shown in Table VII. As in the previous sensitivity analysis discussed above, about half the inputs (when the perceptions are increased) create positive service quality changes, as expected. Many of these were the same input items as those in the E-P model. However, negative or decreased service quality results were obtained, when the perceptions for the other inputs were increased. This anomaly is similar to that discussed above and could also be attributed to the noisy data and model robustness for sensitivity analysis. In addition, it is interesting to note that the magnitudes of the scaled outputs were larger in the E-P gap model sensitivity analysis than in the analysis using the perception-only models. That is, the same percentage change in gaps (inputs) produced a larger change in the overall evaluation of service quality (output) than a similar magnitude change in perceptions (inputs). This provides some initial analytical support of the view that the use of expectations provides a richer and more sensitive model of service quality than when only perceptions are used. This is also an interesting issue that may be studied further.

Conclusion

The main contribution of this paper is twofold. It provides a new approach to modelling customer evaluation of service quality through the use of neural networks. It also contributes to the ongoing debate on service quality measurement by providing an analytical framework through which issues may be further investigated.

In this study, customer responses about individual service system characteristics in an auto-dealership network were used as inputs to a neural network model of service system quality using a SERVQUAL-based systematic diagram of service quality. Four significant conclusions can be identified.

First, the model was able to predict overall service quality as viewed by customers with a 75 per cent accuracy using totally new input data that were not previously used for network training or configuration tests. This seems to be relatively accurate considering the noisy nature of the survey data. While one study is not being presented as conclusive evidence, this study shows that it may be possible to use neural networks to adequately model individual human decision-making characteristics, as suggested by the biological roots of neural networks.

Second, the perceptions-only model of customer evaluation of service quality out-performed the P-E gap model in accuracy. This is consistent with the literature, further validating the neural network approach to modelling service quality.

Third, the best results were obtained when an E-P gap model was used. This configuration out-performed models based on perceptions-only, perceptions and expectations, and the traditional P-E model. The superior performance of the E-P gap model may have occurred, because it is intuitively more likely that customers recognize the gap of what was not delivered relative to predefined expectations, making this E-P gap a better representation of the negative disconfirmation concept. But the service quality gap has traditionally been defined as P-E (perception – expectation) in order to get a mathematically appealing definition for service quality that results in a negative number when expectations have not been met. Further research is needed to determine whether this specific result of the study is a consistent phenomenon and, if so, it provides a basis for the re-evaluation of the traditional gap definition of service quality. The relatively poorer performance of the perceptions-only model supports the notion that expectations do and should play a role in the measurement of service quality.

Fourth, sensitivity analyses experiments using the E-P and perceptions-only models showed that, while they are adequate for predicting overall evaluation of customers, they are not robust enough for sensitivity analyses. The sensitivity tests had limited success in that about half the service attributes had effects that were opposite in direction from what was expected. This could possibly be attributed to the noisy nature of the survey data, as suggested in the preceding section, indicating a need for further research to investigate the usefulness of such neural network models to managers in their efforts to achieve effective quality improvements. However, about one-third of the service attributes did stand out as having strong effects on service quality, indicating some promise for managers in their search for targeted improvements that provide the greatest return. These include friendly personnel (input 4), fast treatment of guarantees (input 16), keeping promises concerning appointments (input 18), offering replacement car during repair (input 3), convenient location from work (input 9), and trying to keep the bill low (input 17). Further, the sensitivity of the E-P model was found to be greater than that of the perceptions-only model. The importance of data regarding perceptions for managers is identified in the literature through anecdotal information. This study provides some initial analytical support to the view that the use of expectations provides a richer and more sensitive model of service quality than when only perceptions are used.

Evaluation of service quality is largely based on human perception. The challenge of effectively measuring and modelling this subjective evaluation needs to be addressed in the quest for providing improved services to customers. Overall, this paper demonstrates that neural networks have the potential to be a valuable approach to understanding customer evaluation of service quality and providing a promising approach to data mining in the domain of service quality.

References

Babakus, E. and Boller, G.W. (1992), "An empirical assessment of the SERVQUAL scale", *Journal of Business Research*, Vol. 24, pp. 253-68.

Bailey, D. and Thompson, D. (1990), "Developing neural-network applications", *AI Expert*, September, pp. 34-41.

Bellandi, G., Dulmin, R. and Mininno, V. (1998), "Failure rate neural analysis in the transport sector", *International Journal of Operations & Production Management*, Vol. 18 No. 8, pp. 778-93.

Bolton, R.N. and Drew, J.H. (1991), "A multistage model of customers' assessments of service quality and value," *Journal of Consumer Research*, Vol. 17 No. 4, pp. 375-84.

BrainMaker Professional (1993), *BrainMaker Professional Neural Network Simulation Software User's Guide and Reference Manual*, 4th ed., California Scientific Software Press, Nevada City, CA.

Burke, L.I. and Ignizio, J.P. (1992), "Neural networks and operations research: an overview", *Computers and Operations Research*, Vol. 19 No. 3/4, pp. 179-89.

Buttle, F. (1996), "SERVQUAL: review, critique, research agenda", *European Journal of Marketing*, Vol. 30 No. 1, pp. 8-32.

Chenet, C., Tynan, C. and Money, A., (2000), "The service performance gap: testing the redeveloped causal model", *European Journal of Marketing*, Vol. 34 No. 3/4, pp. 472-95.

Collins, E., Gosh, S. and Scofield, C. (1988), "An application of a multiple neural network learning system to emulate mortgage-underwriting judgments", *Proceedings of the IEEE International Conference on Neural Networks*, Vol. 2, pp. 459-66.

Cooper, J.C.B. (1999), "Artificial neural networks versus multivariate statistics: an application from economics", *Journal of Applied Statistics*, Vol. 26 No. 8, pp. 909-21.

Cronin, J.J. and Taylor, S.A. (1992), "Measuring service quality: a re-examination and extension", *Journal of Marketing*, Vol. 56 No. 3, 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, pp. 125-31.

Fishman, M., Barr, D. and Loick, W. (1991), "Using neural networks in market analysis", *Technical Analysis of Stocks and Commodities*, April, pp. 18-25.

Fletcher, D. and Goss, E. (1993), "Forecasting with neural networks: an application using bankruptcy data", *Information and Management*, Vol. 24 No. 3, pp. 159-67.

Genetic Training (1993), *Genetic Training Option User's Guide and Reference Manual*, 3rd ed., California Scientific Software Press, Nevada City, CA.

Grönroos, C. (1982), *Strategic Management and Marketing in the Service Sector*, Swedish School of Economics and Business Administration, Helsinki.

Hecht-Nielsen, R. (1990), *Neurocomputing*, Addison-Wesley Publishing, Reading, MA.

Helson, H. (1964), *Adaptation Level Theory*, Harper & Row, New York, NY.

Hoefer, P. and Gould, J. (2000), "Assessment of admission criteria for predicting students' academic performance in graduate business programs", *Journal of Education for Business*, Vol. 75 No. 2, p. 229.

Hornik, K. (1991), "Approximation capabilities of multiplayer feed-forward networks", *Neural Networks*, Vol. 4, pp. 251-7.

Hornik, K., Stinchcombe, M. and White, H. (1989), "Multilayer feed-forward networks are universal approximators", *Neural Networks*, Vol. 2, pp. 359-66.

Jain, B.A. and Nag, N.B. (1995), "Artificial neural network models for pricing initial public offerings", *Decision Sciences*, Vol. 26 No. 3, May/June, pp. 283-302.

Johnson, C. and Mathews, B.P. (1997), "The influence of experience on service expectations", *International Journal of Service Industry Management*, Vol. 8 No. 4, pp. 290-305.

Kamijo, K. and Tanigawa, T. (1992), "Stock price pattern recognition: a recurrent neural network approach", *Proceedings of the International Joint Conference on Neural Networks*, San Diego, CA, Vol. 1, pp. 215-21.

Kettinger, W.J. and Lee, C.C. (1994), "Perceived service quality and user satisfaction with the information services function", *Decision Sciences*, Vol. 25 No. 5/6, pp. 737-66.

Lawrence, J. (1994), *Introduction to Neural Networks*, California Scientific Software Press, Nevada City, CA.

Lehtinen, U. and Lehtinen, J.R. (1982), "Service quality: a study of quality dimensions", working paper, Service Management Institute, Helsinki.

Lemmink, J. and Behara, R.S. (1992), "Q-matrix: a multi-dimensional approach to using service quality measurements", in Knust, P. and Lemmink, J. (Eds), *Quality Management in Services*, Van Gorcum, Assen/Maastricht, pp. 79-87.

Lenard, M.J., Alam P. and Madey, G.R. (1995), "The application of neural networks and a qualitative response model to the auditor's going concern uncertainty decision", *Decision Sciences*, Vol. 26 No. 2, pp. 209-27.

McMillen, R. and Henley, T. (2001), "Connectionism isn't just for cognitive science: neural networks as methodological tools", *The Psychological Record*, Vol. 51, pp. 3-18.

Mizuno, S. (1988), *Management for Quality Improvement: The Seven New QC Tools*, Productivity Press, Cambridge, MA.

Mozer, M., Wolniewicz, R., Johnson, E. and Kaushansky, H. (1999), "Churn reduction in the wireless industry", *Proceedings of the Neural Information Processing Systems Conference*, San Diego, CA.

Nordmann, L.H. and Luxhoj, J.T. (2000), "Neural network forecasting of service problems for aircraft structural component groupings", *Journal of Aircraft*, Vol. 37 No. 2, pp. 332-8.

Parasuraman, A., Berry, L.L. and Zeithaml, V.A. (1990), "Guidelines for conducting service quality research", *Marketing Research*, December, pp. 34-44.

Parasuraman, A., Zeithaml, V.A. and Berry, L.L. (1985), "A conceptual model of service quality and its implications for future research", *Journal of Marketing*, Vol. 49, pp. 41-50.

Parasuraman, A., Zeithaml, V.A. and Berry, L.L. (1988), "SERVQUAL: a multiple-item scale for measuring customer perceptions of service quality", *Journal of Retailing*, Vol. 64 No. 1, pp. 12-40.

Parasuraman, A., Zeithaml, V.A. and Berry, L.L. (1991), "Refinement and reassessment of the SERVQUAL scale", *Journal of Retailing*, Vol. 67 No. 4, pp. 420-50.

Parasuraman, A., Zeithaml, V.A. and Berry, L.L. (1994), "Reassessment of expectations as a comparison standard in measuring service quality: implications for further research", *Journal of Marketing*, Vol. 58, pp. 111-24.

Peter, J.P., Churchill, G.A. Jr and Brown, T.J. (1993), "Caution in the use of difference scores in consumer research", *Journal of Consumer Research*, Vol. 19, March, pp. 655-62.

Philipoom, P.R., Rees, L.P and Wiegmann, L. (1994), "Using neural networks to determine internally-set due-date assignments for shop scheduling", *Decision Sciences*, Vol. 25 No. 5/6, pp. 825-51.

Rochester, J. (Ed.) (1990), "New business uses for neurocomputing", *I/S Analyzer*, February, pp. 1-17.

Rumelhart, D.E., Hinton, G.E. and Williams, R.J. (1986a), "Learning representations by back-propagating errors", *Nature*, Vol. 323, pp. 533-6.

Rumelhart, D.E., Hinton, G.E. and Williams, R.J. (1986b), "Learning internal representations by error propagation", in Rumelhart, D.E. and McClelland, J.L. (Eds), *Parallel Distributed Processing: Explorations in the Microstructure of Cognition, Vol. I: Foundations*, The MIT Press, Cambridge, MA, pp. 319-62.

Salchenberger, L.M., Cinar, E.M. and Lash, N.A. (1992), "Neural networks: a new tool for predicting thrift failures", *Decision Sciences*, Vol. 23 No. 4, pp. 899-916.

Sasser, W.E., Olsen, R.P. and Wyckoff, D.D. (1978), "Understanding service operations", *Management of Service Operations*, Allyn & Bacon, Boston, MA.

Singleton, J.C. (1990), "Neural nets for bond rating improved by multiple hidden layers", *Proceedings of the International Joint Conference on Neural Networks*, Vol. 2, San Diego, CA, pp. 151-62.

Surkan, A. and Singleton, J.C. (1990), "Neural networks for bond rating improved by multiple hidden layers", *Proceedings International Joint Conference on Neural Networks*, San Diego, CA.

Tam, K.Y. and Kiang, M.Y. (1992), "Managerial applications of neural networks: the case of bank failure predictions", *Management Science*, Vol. 38 No. 7, pp. 926-47.

Teas, R.K. (1993), "Expectations, performance evaluation and consumers' perceptions of quality", *Journal of Marketing*, Vol. 57, October, pp. 18-34.

Teas, R.K. (1994), "Expectations as a comparison standard in measuring service quality: an assessment of a reassessment", *Journal of Marketing*, Vol. 58, pp. 132-9.

Udo, G. (1993), "Neural network performance on the bankruptcy classification problem", *Computers and Industrial Engineering*, Vol. 25 No. 1-4, pp. 377-80.

Van Dyke, T.P., Prybutok, V.R. and Kappelman, L.A. (1999), "Cautions on the use of the SERVQUAL measure to assess the quality of information systems services", *Decision Sciences*, Vol. 30 No. 3, pp. 877-91.

Wilson, R. and Sharda, R. (1992), "Neural networks", *OR/MS Today*, August, pp. 36-42.

Xu, J., Wong, S.C., Yang, H. and Tong, C. (1999), "Modelling level of urban taxi services using neural network", *Journal of Transportation Engineering*, Vol. 125 No. 3, pp. 216-23.

Yao, J., Li, Y. and Tan, C.L. (2000), "Option price forecasting using neural networks", *Omega*, Vol. 28 No. 4, pp. 455-6.