

Definition and measurement of quality of service for local public transport management†

By WILLIAM T. PULLEN

Transport Operations Research Group,
University of Newcastle upon Tyne, U.K.

2008

Following in the footsteps of the quality revolution in manufacturing industry, service businesses are awakening to the need for quality management processes. In the U.K., local public transport undertakings are facing a future in private ownership and operating in a deregulated market place. The need to maintain market share and increase profitability has focused attention on quality of service. The paper concentrates on the definition and measurement components of quality management processes for local public transport services. Existing methods based on passengers' waiting times, lost mileage, expanded sets of measures and indices are reviewed. It is concluded that a single measure is unlikely to encompass all aspects of quality of service. Similarly, indices which collapse measures of different attributes of quality of service into a single abstract value appear to have limited appeal. Instead, a limited set of quality of service attributes can be defined, based on local operating circumstances and the goals and objectives of local interest groups. Both relevant performance measures, and psychometric measures, can be selected as indicators of the defined set of quality of service attributes. The application of a given set of quality of service indicators would then involve a comparison of measured values across time.

1. Introduction

Until recently, it was widely held that the 'growing quality revolution in manufacturing firms' had largely gone unnoticed in service businesses (Berry 1991). However, there is now a recognition in some sectors of service industry, such as financial services (Goodstadt 1990) and telecommunications (Bolton and Drew 1991), that the management and improvement of quality is a matter of survival in a global market place. Initiatives include staff awareness schemes, quality service action teams, market research and service standards. In the U.K., local public transport undertakings are facing a future in private ownership and operating in a deregulated market place. The need to maintain market share, and increase profitability is awakening managers to the importance of quality of service.

In the development of a total quality model applicable to large to medium service businesses, Berry (1991) identified the following key components in the total quality management process:

- (a) realizing the need for improvement;
- (b) committing to quality;
- (c) structuring for quality;
- (d) determining customers needs;
- (e) setting the culture blueprint;

† This article won Dr Pullen the Transport Reviews Prize for the competition year ended 30 April 1992.

- (f) designing the quality process;
- (g) quality planning;
- (h) quality improvement teams;
- (i) unit level quality;
- (j) training;
- (k) vendor quality;
- (l) awareness and promotion;
- (m) recognition and celebration;
- (n) monitoring and measuring; and
- (o) enhancing the process.

The focus of this paper is on the 'monitoring and measuring' component of the quality management process. This is examined with reference to local public transport services. A review is therefore undertaken of existing methods for definition and measurement of quality of service for local public transport. This is then used as a basis for suggesting a method of definition and measurement which could act as a component in a quality management process for local public transport.

2. Definitions of quality of service for local public transport

The ability to define, and hence measure, quality of service is central to its management and improvement. In setting out the terminology for standards and regulations, the Advisory Council for Applied Research and Development defined quality as 'those features and characteristics of a product or service that bear upon its fitness for purpose' (ACARD 1980). In order to define quality of service for local public transport, therefore, it is first necessary to identify the features and characteristics that bear upon its fitness for purpose.

For local public transport services, sets of attributes, and their associated performance measures, are often considered under broad categories of efficiency and effectiveness (Fielding *et al.* 1978, OECD 1980, Silcock 1981). Efficiency measures deal with the processes that produce the bus services, while effectiveness measures are concerned with how well the services produced meet the objectives set for them. From this classification, measures of quality of service are generally, although not universally, held to be a subset of effectiveness measures. By examining how these measures were developed it may be possible to move towards a definition of quality of service. Table 1 lists the broad categories of efficiency and effectiveness measures employed in three studies.

In the first example in table 1, the study did not address the concept of quality of service (Fielding *et al.* 1981). The second study included service quality as part of effectiveness, but listed accessibility as a separate category (OECD 1980). In the third review of performance measures, quality of service was again included as a category of effectiveness, although accessibility was this time defined as an integral part of quality of service (Silcock 1981). Quality of service was therefore comprised of measures of accessibility, reliability, comfort, convenience and safety.

An alternative classification has been put forward which does not start from the traditional basis of efficiency and effectiveness (Allen and DiCesare 1976). In this study, performance measures were based on categories of quantity of service, quality of service and cost/revenue. Quality of service measures were further subdivided into user and non-user categories. User-related measures were speed, reliability, comfort,

- (i) the operator, responsible for planning and operating services, the success of which determines profitability;
- (ii) central government, responsible for enacting legislation which determines the administrative framework within which public transport operates;
- (iii) local authorities, responsible for planning and procuring socially necessary services;
- (iv) users, who decide whether they will use the services provided and, if so, in what way;
- (v) operator personnel, whose working conditions are directly affected by the management methods and level of service;
- (vi) trade unions, representing the personnel;
- (vii) operators of services other than public transport, whose costs are affected by public transport services that are provided;
- (viii) residents, subjected to various nuisances (safety, visual and acoustic intrusions) or by certain prejudices such as perceived loss of property values;
- (ix) vehicle manufacturers, who must comply with set standards such as exhaust emissions or step heights; and
- (x) the scientific and technical community, who undertake research and development.

In order to move towards a definition of quality of service for local public transport, a general pool of attributes and associated performance measures must first be considered. From these a package of measures can be developed which will vary from place to place. The package will be uniquely determined by a variety of local circumstances as outlined above. Also, it will be dependent on the variety of perspectives encompassed, so that 'transit standards will be less arbitrary if expertise, coordination and cooperation are employed in their development' (Allen and DiCesare 1976).

3. Measuring the quality of service of local public transport

In order to measure the quality of service for local public transport, it is first necessary to have a working definition. As there is no single definition, it is therefore apparent that there will be no single method of measurement. This section will therefore review some studies that have been concerned with measuring the quality of service of local public transport and report on the methods used.

3.1. *Passengers' waiting times*

London Transport were pioneers in introducing a system to monitor regularly the quality of its bus services. The Business Planning Department initiated a Quality of Service Indicator (QSI) monitoring scheme in 1976 (Shields 1978). This was intended to provide management with a broad picture of how services were performing and whether there were any significant trends. This may be regarded as a form of quality monitoring.

The development of QSIs started from the assumption that although quality of service was comprised of several attributes, the main concern of the passenger was the reliability of the service. The QSI monitoring scheme therefore expressed quality of service in terms of passenger waits at bus stops. Direct measurement of passengers' waiting times at bus stops was considered impractical, as several routes often used a single bus stop for boarding and alighting. It would therefore be costly to conduct interviews with all passengers to obtain their exact intentions. Also, passengers can

For a punctual bus arrival, the assumptions gave an expected value of 5 minutes for AWT, which in turn would give an EWT value of zero. If all buses were 1 minute late then an EWT of 1 minute was introduced. On the other hand, if all buses were 1 minute early, then 10% of the passengers were deemed to have missed their bus and were required to wait for the next. These 10% of passengers have the same value as 1 minute wait time under the assumption of uniform arrival of 100% of passengers in the 10 minute period before the arrival of buses. This EWT is then required to be incorporated into the calculation for the next arrival.

The use of these measures established London Transport's QSI monitoring programme. However, while there is evidence to suggest that reliability is a major concern for passengers in London (GLC/LTE 1976), it is not the only characteristic of bus services that affects fitness for purpose. Features such as accessibility, comfort, convenience and safety are also important (Gault 1980). Also, while waiting time is considered an important part of perceived reliability, it may be that other features also contribute to passengers' understanding of reliability. Factors such as in-vehicle delays through congestion or breakdowns may also be important.

The assumptions regarding passenger arrival patterns are crucial to the measurement of waiting time by indirect methods. These assumptions were refined and developed in a study in Newcastle upon Tyne using AWT and EWT to measure reliability as a proxy for the passengers' view of performance (Gault and Doherty 1979). There was evidence to support the theory that for services with headways less than 12 minutes, passengers tended to arrive in a random fashion. However, for longer headway services, evidence suggested that passengers did not in fact arrive uniformly in a 10 minute period before the scheduled arrival of the bus. Rather, the distribution was normal with a peak about 5.5 minutes before the scheduled bus departure.

3.2. *Lost mileage*

Another measure that has been used to set quality of service standards is lost mileage. Lost mileage is defined as the difference between the number of scheduled miles and the number of miles operated. In practice, it is impossible to achieve 100% of scheduled mileage over a long period of time. This is due to a number of unpredictable factors such as vehicle breakdowns, driver sickness and weather conditions, which inevitably curtail the operation of some services. While a certain amount of lost mileage is therefore inevitable, it is obviously desirable that it should be kept to minimum.

On this basis, quality standards have been set by Cheshire County Council which are used to improve the quality of service. These are based on lost mileage, as 'There are a number of criteria to measure quality of service but the most important one is lost mileage. Other standards for timekeeping, quality of vehicle, driver behaviour, display of route number and destination are all secondary to the provision of the service.' (Dobson 1987)

Contracts let by Cheshire County Council specify the timetables, and the contract is for the provision of the full service. Given that 100% of scheduled mileage is unobtainable, a standard is set below which penalties are incurred. To operate this procedure contractors are required to submit regular lost mileage returns which are then backed by on-the-road monitoring and checking of waybills where necessary.

However, lost mileage cannot be regarded as a wholly accurate measure of quality of service. Lost mileage is caused by several factors, and each can have a different impact on service provision. If there is a shortage of vehicles or drivers then control

how close the measures used came to coinciding with the passengers' view of quality of service. Also, the authors made no mention of how the sample of 408 journeys was obtained. No information was given on the age or sex structure of the sample or on the journey purposes of those involved. It is therefore difficult to be certain that the findings of the study would in fact reflect the passengers' perception of quality of service in Budapest.

In San Francisco, USA, the management of the Bay Area Rapid Transit (BART) system have sought to develop methods for monitoring quality of service (DuPlessis 1984). The initial aim was to provide an assessment mechanism to provide objective data from the perspective of the passenger. Direct passenger surveys were discounted as they were deemed to incorporate subjective judgements and varied interpretations. They were also considered to emphasize negative aspects of the services.

The first step in this procedure was to determine the features and characteristics of the system that could be used to evaluate services. This was achieved by reviewing trip patterns and complaints reports. The parameters thus identified were further refined by meetings between Marketing and Field Services Managers. The items included in the final selection were listed as follows:

- (a) station information:
 - (i) agent in or out of station agent booth;
 - (ii) agent in uniform;
 - (iii) supervisor present;
 - (iv) BART police present;
 - (v) brochures available;
 - (vi) equipment operable (fare gates, ticket machines, elevators, escalators, etc.);
 - (vii) cleanliness (station, restroom, and elevator);
 - (viii) announcements heard over PA system;
 - (ix) number of rule violations.
- (b) boarding information:
 - (i) waiting time;
 - (ii) destination signs working;
 - (iii) train exterior cleanliness;
 - (iv) train operator watching doors;
- (c) trip information:
 - (i) trip time;
 - (ii) car interior cleanliness;
 - (iii) car loading;
 - (iv) BART police on car;
 - (v) rule violations committed by patrons; and
 - (vi) announcements heard on car.

The information required was collected by a group of survey staff using statistical work-sampling techniques. Data collection was in a format suitable for direct key punching, and a computer program was written to compile and analyse the data.

The surveyors examined both stations and trains. In stations, a full or abbreviated check was undertaken. The abbreviated check included a determination of station cleanliness based on standards provided by BART on a scale of one to five. Also,

Table 2. Hypothetical level of service on BART system.

Characteristic	LOTS value	Points
Adjusted speed	A	24
Delay	A	10
Space	B	20
Acceleration	C	6
Temperature	A	12
Ventilation	B	4
Noise	A	5
Total	B	81

Source: Botzow (1974).

modes that most reflected the quality of service (Botzow 1974). The LOTS method was proposed in order that existing systems could be managed and improved, and new systems could be built on the basis of their ability to fulfil a desired level of service. The features used to reflect quality of service were based on the passengers' environment and travel speed for line haul systems. The characteristics of travel speed and comfort were defined as:

- (a) speed;
- (b) delay;
- (c) space;
- (d) horizontal acceleration;
- (e) vertical acceleration;
- (f) jerk;
- (g) temperature;
- (h) ventilation; and
- (i) noise.

Appropriate values for these characteristics were then defined in six LOTS categories, from A (excellent) to F (not suitable). Weights were then applied to each characteristic based on its importance. These were expressed in terms of points, so that a service achieving an A LOTS value for each characteristics would score 100, and a service scoring a F LOTS value for each characteristic would have an overall score of 0. A hypothetical analysis of the LOTS values for a trip during the peak hour with all seats filled on the BART system was calculated. In this case seven characteristics were measured, with the two acceleration characteristics above given as one, and the omission of 'Jerk'. This is set out in table 2. The BART system achieves a favourable rating of 81 points, and therefore an overall LOTS value of B. All individual characteristics achieved LOTS values of A or B with the one exception of acceleration which received a LOTS value of C.

The author proposed that a set of uniform national standards of service could be developed along similar lines to the LOTS concept. This was regarded as desirable to facilitate daily management and the uniform improvement of public transport. National Government grants might then be based on local conditions and local plans for attainment of appropriate minimum levels of service. Priorities could include upgrading those routes that exhibit the lowest level of service. Local government subsidies could also be based on the maintenance of a specified level of service. These suggestions move the LOTS concept toward quality management process, through the

As the contribution of attributes sums to 10 and the contribution of aspects also sums to 10, the quality of service index will therefore sum to 100. The index is applied to a special transport service by relating it to a score of satisfaction for particular attributes.

The score of satisfaction is derived from performance measures through a scoring function. The scoring function is set so that the maximum score possible, representing complete satisfaction with an attribute, is 1. The minimum score possible, representing no satisfaction, is 0. The scoring function can be represented as follows:

$$S_{ik} = f_i(m_{i1k}, m_{i2k}, \dots, m_{ink})$$

where S_{ik} is the scoring function for attribute i for provider k ; $f_i(\cdot)$ is the functional form of attribute i ; and m_{i1k}, \dots, m_{ink} are the performance measures for attribute i for provider k .

The final quality of service score for a provider of a special transport service is therefore given by

$$Q_k = \sum_j \sum_i w_j w_{ij} S_{ijk}$$

where Q_k is the quality of service score for provider k ; w_{ij} is the relative contribution of attribute i to aspect j ; w_j is the relative contribution of aspect j to quality of service; and S_{ijk} is the score for attribute i under aspect j for provider k .

The quality of service index was proposed as an evaluative tool. This could be for self-evaluation by management of special transport services, or comparative evaluation for organizations wishing to place contracts or provide subsidy. Another use postulated was the analysis of general policies for the organizational form of providers of special transport services (McKnight and Pagano 1983).

This attempt to construct a quality of service index relies on a more considered approach to definition of aspects and attributes than was apparent in the LOTS concept. With the inclusion of 42 attributes it may be possible to cover a variety of local circumstances and perspectives. However, there is a negative side to such comprehensive coverage. This relates to the use of a ranking task to construct the index. Ranking is considered to be more difficult to perform than rating (Dobson 1974), especially when the list of attributes is long. Thus, of 236 questionnaires returned to the researchers, only 148 (63%) were usable. This could prove problematic in situations with an initial low response rate. Also, the measurement and calculation of performance measures for 42 attributes is likely to require much expertise and specialized equipment.

In addition, there are problems in the construction of an index with weights based on perceived importance. It can be shown that the perceived importance of attributes is inversely related to performance (Stern 1982). That is, if an attribute can be shown to be performing poorly, it is likely to be perceived as important by users of a service. Conversely, if an attribute is performing well, it is less likely to be perceived as important relative to other attributes. This raises concerns in the calculation of an index based on weights of perceived importance and performance scores. These concepts are unlikely to change independently of each other. This may lead to difficulties in the interpretation of values across time.

Finally, there are questions regarding the intrinsic worth of indices. As these examples demonstrate, they are both complicated to construct and apply. The researcher will arrive at a single measure but will be required to retrace steps in order to understand what the measure actually means. There is little to be learned from the

Table 3. A suggested quality of service monitoring scheme for local bus services.

Quality of service attributes	Time period <i>T</i> 1		Time period <i>T</i> 2	
	Performance measure	Psychometric measure	Performance measure	Psychometric measure
Attribute 1				
Attribute 2				
—				
—				
Attribute <i>n</i>				

aged where monitoring of the quality of service of local bus services would take place under a scheme developed from one similar to that in table 3.

The most obvious drawback of such a proposal is that there is little information available on the relationship between performance measures and passengers' perceptions of the attribute. The situation could therefore arise where movement in the physical performance indicator is accompanied by a different movement in passengers' perceptions. It would not be possible to ascertain if this was purely a functional relationship, or if other factors were at play.

Information from the field of psychophysics suggests that there are indeed different subjective magnitudes of response to similar physical intensities in different stimuli (Schiffman 1984). It may be that the relationship between perceived and real changes for individual attributes will differ from each other and that such relationships may not be linear. However, little is known of these relationships, and it is likely that they will vary between localities.

This feature of different subjective magnitudes of responses further strengthens the case against pulling together packages of indicators into quality of service indices. The method outlined in table 3 keeps each individual attribute separate, and separates the performance and perception measures. At the same time, it is possible to identify relative change in each over time. This itself may aid the identification of subjective magnitudes of responses for individual attributes on a local basis. Care would be required to establish the degree to which each of the selected attributes was independent of the others.

Through this method the 'purpose' of services are defined locally and indicators can be developed to measure their 'fitness'. It would then be possible to undertake an assessment of whether resources would be best directed towards improving physical performance of a given attribute or, if more could be gained from concentrating on perceptions, say through an advertising campaign. Such methods could be developed within the framework of quality management processes for local public transport undertakings. Remedial action would therefore be called for if performance fell below specified standards. These methods could also be put to use as a basis for monitoring the quality of service on contract services operated for local authorities. This could be used to help ensure value for money and to help raise the standard of quality of service on offer to passengers.

5. Conclusions

In the introduction to this paper, it was postulated that some of the issues from quality management processes that have evolved in manufacturing industry could have

fixed-route, many-to-many-type bus services under different operating schedules. Quality of service is assessed in relation to OTP (on time performance) and deviations from dispatching headway (DOH). When using the OTP criterion, the model is used to simulate bus travel times between two given points, and to compute the probability distribution and the parameters of the deviations of actual arrival times from scheduled arrival times. The details of the study and the results are described. It is shown that the Bus-Monitor can be utilized to perform several functions that are presently done manually. These and other findings are discussed.

(Author)