Transit Service Evaluation: Preliminary Identification of Variables Characterizing Level of Service

William G. Allen, Jr., Sverdrup and Parcel and Associates, Inc., Silver Spring, Maryland
Frank DiCesare, Transportation Center, Rensselaer Polytechnic Institute, Troy, New York

This paper is an introduction to transit service evaluation and its application to medium-sized bus transit systems. The concept of transit evaluation through the measurement of level of service is discussed in terms of usefulness, past work, theory, and the presentation of a set of characteristic attributes. The need for performance evaluation, since transit is a public service that does not operate under the profit incentive, is presented. Its usefulness for management, governmental policy formulation, and determination of subsidy levels is discussed. The state of the art and practice, including the Pennsylvania Department of Transportation system, is reviewed. A methodology of transportation system evaluation developed by the Rand Corporation is summarized for its potential application to transit service. A preliminary list of service attributes, with the method of measurement identified, is given. It is concluded that transit service can be quantified and evaluated but that considerable effort is necessary to achieve a comprehensive and equitable system.

For the past three decades and until recently, the transit industry has suffered from the spiral of a decrease in patronage leading to increased fares and decreased service until a further decrease in patronage and so on. This nationwide experience and the realization that transit cannot generally pay for itself out of the fare box have driven transit substantially out of the private business sector and into the public service domain. As of 1973, 185 public transit agencies (18 percent of all transit operators) accounted for 91 percent of all transit passengers and 88 percent of all gross revenues (1). After years in which survival rather than progress was their main objective, transit operators are becoming agencies intended to serve the transportation needs and interests of the public.

Public ownership is a result of government recognition of transit as a necessary public service. In this regard, there has been much recent state and federal legislation allocating funds for capital improvements and operating assistance to local public transit properties. However, even as it has solved some of the problems of transit, government subsidy has created others, especially those of incentive and management control. But if transit is not to be operated solely as a profit-making enterprise, then under what general principles should it be run? This question takes on added importance because most transit subsidy funds now originate at the federal and state level, while transit operations are inherently a local concern. To what degree can upper levels of government be expected to pay for transit services that are administered largely outside their control? Or conversely, given a level of federal and state financial participation, how much control should local operators expect to have? The possibility that any government subsidy could diminish local motivation toward efficiency (unless operations are controlled to some extent by guidelines and standards) is of immediate concern. Such guidelines, applied judiciously, could be an integral part of a program to increase transit efficiency and productivity while, at the same time, they help to safeguard the interest of the public, who now have a more or less permanent stake in the provision of mass transit services.

This study, which was sponsored by the New York State Legislature, was motivated by several observations.

1. There is no comprehensive system for transit service evaluation or data collection in New York State.
2. This leads the state to allocate funds for transit operating assistance and capital improvements with little apparent control afterward. Although transit finances are under general scrutiny, the efficiency of transit operations is not.
3. State legislative and administrative officials who have recently assumed responsibility for transit are not yet fully conversant with the subjects of level of service, efficiency, and other factors involved in transit performance.
4. The complex institutional arrangements required for effective operation of transit in the public interest are still in the formative stages. The principal levels and branches of government that interact are the state legislature, the state department of transportation, and the metropolitan transportation authorities.
5. In some cases there is need for local authorities to keep better track of their own operations.

Publication of this paper sponsored by Committee on Public Transportation Planning and Development.
The objective of this paper is to provide some initial information concerning transit service evaluation by

1. Discussing the usefulness of service evaluation to various branches and levels of government,
2. Reviewing current trends in transit evaluation,
3. Discussing transportation evaluation theory as it pertains to transit level of service, and
4. Identifying and describing a set of operating characteristics to be considered in the measurement of transit performance.

It is worthwhile to briefly discuss some areas outside the scope of this work. First, it is assumed that service evaluation is necessary and desirable. Second, the cost and methodology for implementation of an evaluation system are not presented. Third, the project was conceived with the idea of applying it to conventional public bus transit systems of the medium-sized metropolitan areas of upstate New York—the Capital District (Albany-Schenectady-Troy), Central New York (Syracuse), the Genesee Region (Rochester), and the Niagara Frontier (Buffalo).

**BACKGROUND**

**Previous Work**

Several concepts of transit level of service, varying with their context, have been used over the years. Service evaluation has rarely been the sole topic of study but is usually part of a larger issue. Level of service has been variously defined as: speed (2, 3, 4), transit travel time (5), headway (6), operating ratio (7), and service envelopes (limits of economic viability and passenger capacity) (8). There have been few attempts to study or implement systems of service evaluation since the now-defunct National Committee on Urban Transportation of the Public Administration Service (PAS) published two manuals on measuring transit service in 1958 (9, 10). These manuals have been the standard references in transit evaluation since their publication. However, they were meant primarily for use by operators to monitor their own operations, and this has been the extent of much of transit service evaluation thus far. As valuable as it is, it lacks a total community view of transit performance; i.e., transit service has traditionally been viewed as a concern strictly of its providers and users. This paper attempts to broaden the involvement to include all sectors and members of the community, who are ultimately responsible for the success or failure of transit.

The major sources of information on service evaluation besides the PAS manuals are the Pennsylvania Department of Transportation (PennDOT) report, Operating Guidelines and Standards for the Mass Transportation Assistance Program (11), and a supplementary paper by Vuchic, Tennyson, and Underwood on the application of the guidelines (12). The program of guidelines and standards for transit operation described in these two reports is an organized and comprehensive method of subsidy allocation based on the monitoring, evaluation, and improvement of local transit systems. This system, which is based in part on the PAS Manual, is the only statewide transit service evaluation program known to be in operation at the time this paper was written.

**The Operator’s View**

One of the major problems that efforts to evaluate transit encounter is the uncertainty of transit operators who are wary of external appraisal of the level of service they are providing. As part of this research, officials of all four major upstate New York regional transportation authorities were interviewed; the results seem to confirm this assessment. There was a general feeling of cooperation toward being evaluated, since the officials believed that an evaluation would cast a favorable light on their systems, but they expressed concern over the possibility of increased state involvement in local transit affairs and about the cost and administration of an evaluation program.

**PURPOSES OF SERVICE EVALUATION**

**Policy Formulation**

Measuring transit service could generate a great deal of information that would be helpful in formulating policy decisions at all levels of government. Planners could use this measuring system to assess the existing transit situation in relation to what is set forth in the regional short- and long-range mass transportation plans and to resolve controversies about increasing or decreasing transit services. However, in order to know with relative confidence that the level of service is going up or down (or remaining constant), a system that will assist in defining and measuring transit level of service is needed.

**Subsidy**

Another area in which performance measurement can play a significant role is that of the subsidy. (This paper is not concerned with a detailed review of subsidy mechanisms and theory but only with the relationship between financial assistance and service evaluation.) There are three categories of operating subsidy: incentive, sustenance, and innovation. In this context, incentive implies a financial reward for a high level of service. It is well recognized that incentives must be service-oriented to a large degree (12) and apparent that a good system of service criteria and standards can direct the way toward increased efficiency and elimination of wasteful practices.

While it is desirable to reward good service, some provision must be made for the operator who is under severe financial burdens and requires a subsidy just to keep operations from ceasing, as is often the case when a public agency takes over a private operation. This survival subsidy could be used to sustain those systems that need it the most and can show that they will put it to the most productive use in maintaining or possibly improving a minimal level of service. If good faith and conscientious effort are put forth by a public transit agency (as reflected in the level of service), then it is entitled to its fair share of subsidy funds.

The third category is that of innovation. Agencies that want to try new ideas in transit in a responsible manner should have that opportunity. Innovation in transit has been sadly lacking over the past few decades although inventive concepts are necessary to prevent stagnation. Innovation should therefore be included when reviewing transit service.

**Public Information and Involvement**

When discussing level of service and evaluation, one should not forget the large group that is ultimately affected: the public, including those who use transit and those who do not. Under New York State’s sunshine law, public transit agencies are obliged to make their operating records available to anyone. If there were a method for evaluation that the layman could understand, perhaps the public would become more involved in regional transit affairs.
Another motivating force for evaluation is that of the ever-increasing federal role in public transportation. The National Mass Transportation Assistance Act explicitly mentions transit efficiency. The Urban Mass Transportation Administration (UMTA) has also published its interpretation of this act and discussed the matter of improving efficiency and level of service (14). The former administrator of UMTA, Frank C. Herringer, has said (15):

> Greater emphasis is expected in the general area of performance and productivity measurement. We need to know more about the components of efficiency at each level of transit operations. Better information and evaluative tools in these areas will provide transit managers with an increased facility for isolating problem areas and for developing solutions.

This does not indicate that UMTA will become involved in the day-to-day operating decisions of transit authorities. No one (including those at UMTA) sees this as desirable. It does seem, though, that the federal government will soon require reassurances that its (the public's) money is being used wisely at the local level as a prerequisite to distributing funds.

**Transit Agency Management Information**

Finally, evaluation can be seen as desirable in providing information for the transit agencies' own use. Many operators already have some kind of internal evaluation procedures and there are indications that this practice is growing. Managers have a continuous need to know what is happening in their operations; it is this feedback of information that enables sound improvements to be made and efficiency to be increased. This kind of systematically collected operating information is the basis for level-of-service measurements. Again quoting Herringer (15):

> Part of the process of development on a national basis is the use of explicit criteria to guide decision making. While these are not applied in a totally restrictive way, they do contribute to an understanding of goals and progress toward meeting objectives. In effect, they substitute for the profit motive in private industry.

A comprehensive system of uniform data reporting and record keeping would help to establish valid criteria and to facilitate system information gathering.

**EVALUATION THEORY AND ISSUES**

**Hierarchy of Evaluation Methods**

It will be helpful to review a few aspects of the theoretical concepts of evaluation, discuss the hierarchy of methodologies, and then examine evaluation in terms of goals. Various descriptions of the theory of evaluation models have been proposed in many different fields of research. Methods of evaluation can be described on the basis of their complexity, technical input, and completeness. In the context of transportation system evaluation, we have the following (16):

- **Method 0**—an intuitive judgment of the system's attributes by one or more qualified persons,
- **Method 1**—a checklist of all system attributes that are considered significant by all persons involved,
- **Method 2**—the checklist of attributes plus the corresponding performance measures (performance measures are physically measurable characteristics that determine system performance with respect to each attribute and should be based on their appropriateness to the relevant policy or goal structure),
- **Method 3**—a system for setting limits on the variation of attribute values, retaining those values that are acceptable, and eliminating those that are clearly undesirable or infeasible,
- **Method 4**—a listing of the attributes in order of their importance; a system of priorities, and
- **Method 5**—the complete-worth procedure of finding independent worth assessments of the different attribute values, determining a set of weights showing the relative importance of the attributes, and then computing the total worth as a linearly weighted sum of the worths over the attributes.

Each of these methods is obviously a more complex and refined procedure than the one that preceded it. Method 0 was the most commonly used approach in the past; this paper concentrates on methods 1 and 2. PennDOT has attempted to apply methods 3 through 5 (11, 12) but further study is needed. As noted in method 2, it is important to relate performance measurement to a goal structure since evaluation of any kind requires a thorough understanding of what is meant by goals, objectives, standards, and criteria.

**Evaluation Goals and Objectives**

Goals and objectives are generally recognized as equivalent ideas and are defined as the end toward which the action is directed (9). They are necessarily abstract concepts but should always be expressed as explicitly as possible. "To improve public transportation" is so vague as to be meaningless (17). The cooperation of state and local agencies is essential to achieve purposeful definitions of transit goals and objectives.

Criteria are more specific than goals and objectives. A criterion represents a condition or state of the system. Criteria should be clear, realistic, inclusive, and not subject to a wide range of interpretation (18). Even more specific than a criterion is the idea of standard. A standard is a defined level of performance in relation to some goal or value, a set point along the way to the achievement of an objective. Standards must be extremely specific and therefore may not generally be subject to a statewide application.

**Service Standards and Criteria**

Service standards must of course relate as closely as possible to service criteria and to transit goals and objectives. "They should represent public policy objectives in regard to maximum service goals, and not just be related to the economics of transit operations" (19). The judicious use of accepted service standards is the only way to evaluate a single transit system on an absolute basis. The key word here is accepted; since evaluation has not been seen as necessary in past years, industry-wide transit service standards have not been formally adopted. This arises from the previously identified problem that standards must be localized to a high degree since the definition of good bus service can vary widely in different places.

**Quantifying Transit Evaluation**

Adherence to service standards is being used by PennDOT as a basis for evaluating transit systems in Pennsylvania to give a numerical rating for each system. However, even if the idea is accepted, that evaluation of the level of service must be quantified to a substantial degree, the usefulness of a numerical transit grade may still be...
questioned. Although the purpose of evaluation is to objectively assess the performance of a transit system, the use of strictly numerical results must be approached with caution and foresight. Transit operators are understandably concerned about level-of-service measurements, since a poor rating could cause the (possibly) unjustified or misdirected wrath of elected officials and the public. On the other hand, good ratings might be equally deceptive. An alternative to the strictly numerical approach is the use of community value-factor profiles or the goals-achievement matrix described by Wegmann and Carter (20).

One argument in favor of using a single number to rate a transit system is that it facilitates subsidy calculations. Obviously, if level-of-service measurements are to be used to determine operating subsidy payments, then such measurements must be specific, accurate, and readily converted into dollars and cents. A prime example of this is the PennDOT evaluation system, in which transit operators receive an operating subsidy directly proportional to the number of points they score on the rating scale. In addition to the base subsidy for present level of service, there is a bonus or penalty that depends on the change in level of service from the preceding year.

Subsidy and Service Evaluation

It is reasonable to expect that a subsidy mechanism based on all facets of level of service would result in a more equitable and efficient allocation of funds than does the current method of determining transit operating assistance in New York. The formula now in use involves a subsidy based on vehicle-kilometers operated and passengers carried (and, in some cases, on population of the service area). This formula works reasonably well as far as it goes but does not take into account the total level of service provided to the community and is not a particularly effective measure of total performance. In their attempts to increase vehicle-kilometers or even passengers, operators could change their service in a manner that would run counter to the transit needs and objectives of the community.

Another factor in the subsidy issue is the role of politics in determining the allocation of funds. Even though legislatures make the decisions on subsidy policy, politics should not intrude into the daily operations of transit agencies. If subsidies are based mainly on need and incentive as determined by a technical evaluation process, then the benefits of an improved distribution of funds will accrue to all.

If an issue as important as transit operating assistance is to be inextricably related to the evaluation of performance and service standards, these standards and their measurement must be as independent as possible. The guidelines should be agreed upon by the legislature, the transportation agencies, and the citizens' groups, however difficult this may be. Transit standards will be less arbitrary if expertise, coordination, and cooperation are employed in their development, although there may still be valid differences of opinion. The key is to effectively combine the operator's experience, the state's research and planning capabilities, the legislature's policy-making process, and the local citizens' needs and desires. (This does not neglect the role of local government, which also participates in transit funding and acts as a representative of the populace.)

SERVICE CHARACTERISTICS

Identification of Transit Service Characteristics

In the discussion of the theory of evaluation models, method 2 was said to involve the enumeration of attributes and performance measures that contribute to transit level of service. Attributes and performance measures (collectively referred to as characteristics or variables) must be selected by cooperative processes. They must also be chosen so that data collection and manipulation are facilitated. Of utmost importance, however, is that they best reflect the mass transportation objectives of the community. Several of the characteristics that will be presented later in this chapter have been documented in past studies.

Obviously, the identification of such characteristics

Figure 1. Transit service characteristics: quantity of service.
**Figure 2. Transit service characteristics: quality of service.**

### USER RELATED

**Speed:**

- Speed = \( \text{kilometers per hour} \), scheduled or actual, over the whole system  
- Speed ratio = \( \frac{\text{average transit speed}}{\text{average auto speed}} \) for selected routes

**Reliability:**

- On-time performance = percentage of buses one minute early to four minutes late

**Comfort:**

- Interior noise levels = average dBA inside the vehicle
- Loading factor = \( \frac{\text{maximum number of passengers (total seats available}}{\text{c}} \) for each route at the maximum load input
- Floor area = peak hour floor area per passenger
- Percentage of fleet with air conditioning = \( \frac{\text{number of air conditioned buses}}{\text{total number of buses}} \)
- Vehicle jerk = \( \text{kilometers per hour/second} \)

**Convenience:**

- Route directness = \( \frac{\text{number of transferring passengers}}{\text{total number of passengers}} \)
- Hours of service = daily hours of operation
- Stop spacing = average distance between bus stops in meters
- Loading zone quality = subjective measure of bus stop adequacy (e.g., shelters, benches, illumination, bus stop marking, curb length)
- Step height = average curb-to-vehicle step height in centimeters
- Information services = subjective measure of communication between transit agency and public (e.g., schedule coherence, route identification, telephone information, marketing)

**Safety and Security:**

- Accidents and crimes = \( \frac{\text{transit accidents}}{\text{million vehicle-kilometers}} \) and \( \frac{\text{transit crimes}}{\text{million vehicle-kilometers}} \)

**Special Services and Innovations:**

- Subjective measures of the implementation of innovative concepts to benefit the transit related community (e.g., preferential treatment for buses on roads, special service to handicapped, etc.)

### NON-USER RELATED

**System Efficiency:**

- Equipment utilization = \( \frac{\text{scheduled vehicle-hours}}{\text{available vehicle-hours}} \)
- Peaking factor = \( \frac{\text{number of peak hour buses scheduled}}{\text{number of base hour buses scheduled}} \)
- Energy efficiency = \( \frac{\text{passenger-kilometers}}{\text{liter of fuel}} \)

**Pollution:**

- Air Pollution = brake horsepower-hour
- Noise Pollution = average dBA outside the vehicle, usually measured at a distance of 7.8 or 15.6 meters (25 or 50 ft)
- Productivity = pay hours = hours that the bus driver is on duty
- Productivity = pay hours = hours for which the driver is paid

**Demand:**

- Modal split (for work trips) = \( \frac{\text{work trips by transit}}{\text{total number of work trips}} \)
- Service usage = \( \frac{\text{number of annual passengers}}{\text{population}} \)
- Passenger density = \( \frac{\text{number of annual passengers}}{\text{square kilometer}} \)
- Density of usage = \( \frac{\text{passenger-kilometers}}{\text{route-kilometers}} \)
- Vehicle occupancy = \( \frac{\text{average number of passengers per vehicle}}{\text{hour}} \)
- Seat turnover = \( \frac{\text{number of annual passengers}}{\text{annual seat-kilometers}} \)
- Seat kilometers = average number of seats per vehicle \( \times \) vehicle-kilometers

**Desire coincidence = subjective measure of how well bus routes coincide with travel desire lines (i.e., origin-destination patterns)**

**Note:** 1 hour = 0.96 gal.

*See Figure 1.
in this paper is only the first step toward establishing a complete evaluation system, but it is an important procedure. The steps that come later—the setting of priorities and service standards and the ultimate evaluation of existing performance with respect to those standards—depend heavily on the use of thoughtfully considered characteristics of service.

For the purposes of this paper, level of transit service is divided into three major components: quantity, quality, and cost/revenue. These categories should not be regarded too strictly, since several characteristics could be listed in more than one category. Attempts are made to follow existing conventions whenever possible.) Quantity describes how much transit service exists—in other words, the supply. Cost/revenue is considered because it deals with the abstract question of how good the service is. The distinction between quantity and quality is important because more (or less) transit does not necessarily imply better (or worse) transit. Cost/revenue is considered because it deals with economic factors that, although they are dependent on quantity and quality, need to be evaluated separately.

**Listing and Explanation of Characteristics**

The service characteristics recommended in this paper as important to the measurement of transit service are shown in Figures 1, 2, and 3. These service characteristics were selected on the basis of their contribution toward the evaluation of transit performance; this does not purport to be a complete listing of every characteristic that could conceivably influence service levels. On the other hand, some of the variables included could be considered partially or totally irrelevant in certain situations and the many interrelationships among the various categories and variables may give different perspectives on the same attribute.

The categorization or disaggregation of these characteristics takes two forms. First, in setting up a system to measure service, the service viewpoints of passengers, operators (management and labor), all levels of government, and the non-transit-patronizing public must be considered. For simplicity, this report divides these groups into users (passengers) and nonusers (everyone else). (This classification sacrifices some accuracy, since the operator is usually considered separately from those who use transit and those who pay for it, but here the operator's interest will be represented in both categories.) Second, the variables must be disaggregated by areas of the metropolitan region. That is, in monitoring service variables, reference must be made to a specific part of the city. The extent of this geographical breakdown depends on the particular circumstances of the evaluation but should at least distinguish between transit service in the central city and in suburban areas.

**Factors Outside Control of Transit Agency**

Some of these characteristics may be partially or totally out of the control of the regional operating authority or planning commission. Examples of this are stipulations of labor contracts, traffic control and traffic regulation enforcement, service boundaries defined by political divisions, and differences in urban environment. One of the problems of comparing transit operations is that the evaluation must consider differences in urban form and land use, population and employment distribution, topography, and climate. An attempt was made to structure the variables so as to minimize these effects.

**Data Collection and Costs**

Data collection is another major problem facing an evaluation effort, since it is mandatory that all transit operations being monitored use the same methods and express the results in the same format. The concern of the agencies involved is always directed toward the costs and responsibilities of such a system: who is to pay for it and who is to administer it. The state government, with the advice and guidance of all directly affected parties, is the proper agency to carry out the evaluation. The question of evaluation costs should be related to the amount of subsidy: Evaluation costs would probably be small compared with the operating subsidies now being used in many cities. Furthermore, evaluation costs could probably be borne initially by UMTA and possibly later by state governments and regional transportation authorities since UMTA has begun to develop a data collection system of Financial Accounting and Reporting Elements (FARE) that will be implemented in 1977, when its use will be made a mandatory precondition for the granting of UMTA section 5 funds.

**CONCLUSIONS AND RECOMMENDATIONS**

**Conclusions**

1. Aside from the PAS manuals (9, 10) published in 1958 and the recent extension of that work by PennDOT (11, 12), there has been little work done on comprehensive evaluation of transit level of service.

2. Since transit is a public service, not operating under the profit incentive, there exists a need for per-
formance evaluation. This need exists for management at the operating level, for policy formulation at several levels of government, and possibly for determination of subsidy levels. (The issue of basing subsidy in part on performance rather than solely on a demographically based formula is not discussed here.)

3. Transit service can be quantified and evaluated in terms of a finite set of operating characteristics. This assertion is based on the existence of at least one prominent evaluation methodology (16) applicable to public transit service and the present identification and presentation of a preliminary list of transit service attributes.

4. It will require a considerable commitment and effort of government, the transit industry, and the research community to devise and implement a comprehensive transit service evaluation system.

Recommendations

1. A framework for performance evaluation should be developed. This should include establishing goals and relating those goals to the measures of level of service.

2. The techniques and economics of data collection should be studied.

REFERENCES