Perspectives on the Evaluation of Urban Transportation Systems

Jerry D. Ward,* Department of Civil Engineering, Massachusetts Institute of Technology

The term evaluation is very broad and has different implications for different people: An evaluation can be designed only after it is specified who is evaluating for what purpose against what criteria. The different perspectives of some of the more important actors are discussed here. There is a difference between comprehensive evaluation and functional evaluation. Comprehensive evaluation is concerned with external effects, e.g., pollution, energy consumption, evolution of land-use patterns, and other contributions to the quality of life. Functional evaluation is concerned with system performance parameters, e.g., wait times, coverage, productivity, ridership, and the other variables that characterize the supply characteristics and their appropriateness to the markets served. Two primary conclusions are developed. The first is that intelligent interpretation of an observed set of system descriptors requires a knowledge of the ridership pattern being served. Both potential productivity and service level are sensitive to ridership pattern, and without a knowledge of its nature, other comparisons are uncertain. The second is that the equilibrium between supply and demand is unstable. This implies that an observed trend in ridership will, in the absence of external change, probably continue. Thus, the rate and direction of change in ridership is the single most important evaluative measure, because it portends the future of the system and can be an early warning signal of a need for change.

NATURE OF THE PROBLEM

Evaluation of transportation systems is not a simple problem. First, the persons who are the evaluators are a diverse set of actors who have different perspectives, entertain different objectives, and therefore evaluate the transportation system (or the subelements thereof) by using different, often unarticulated, and changing criteria.

Second, the system being evaluated is often complex and exhibits strong characteristics of scale; i.e., as ridership patterns change, so do costs and levels of service, inducing future change. These interactions between supply and demand and their implications are generally not well understood, but they are crucially important.

Figure 1 illustrates the general framework of the problem. Although there are wide variations in both the degree of interest and the degree of potential impact of the multiple evaluators, it is clear that no single perspective can be taken in arriving at a total evaluation. And the various perspectives change with time and circumstances. A walk to an unsheltered bus stop on a beautiful spring day may be evaluated by riders as perfectly acceptable; it is evaluated in a very different way during a winter rain. A heretofore acceptable cost for installing shelters may be called into question by city officials the day after the sewage treatment plant breaks down.

In the analysts' parlance, we have a multiplicity of different objective functions that cannot be represented explicitly and that change over time.

The analysis and discussion of this is made somewhat easier by differentiating between two not independent, but essentially different, kinds of evaluation. From the definitions given by Hoffman and Goldsmith (1), comprehensive evaluation is that concerned with the impacts the transportation system is supposed to have, and functional evaluation is that concerned with the relations among ridership, service, and costs. These problems and the processes of their evaluation are sufficiently different that they will be discussed separately.

COMPREHENSIVE EVALUATION

The criteria for evaluation should be derivable from the objective(s) that the transportation system in question is intended to serve—e.g., to make the city a nicer place to live, to revitalize the economy of the city center, to reduce the air pollution to a given level, to induce a land-use pattern that is more energy efficient. The problems are that (a) it is often difficult to translate such broad objectives into measurable attainment criteria, (b) the achievement of these objectives almost always concommitantly depends on other conditions that are not functions of transportation performance, and (c) we seldom understand the causal mechanisms well enough to relate these broad impacts to measurable transportation performance characteristics.

Of these three problems, the third is the most limiting; we do not understand in sufficient depth the interactions between transportation system performance and the performance or behavior of economic and societal systems. Gomez-Ibanez (2) points out, for example, how limited is our current ability to relate transportation characteristics to land-use patterns. The wide differences of opinion about the ultimate worth and impact of the Bay Area Rapid Transit in San Francisco and Metro in Washington, D.C., are testimonial to this point.

Even when we can estimate the different impacts of transportation alternatives, they are seldom in a format that allows us to relate them to quality-of-life type objectives easily. The following paragraphs about Metro from the Washington Post (3) illustrate this.

> If there is no subway line from Gallery Place to Greenbelt, $766 million in construction costs will be saved but 21 300 more automobile trips will be taken every day in 1990.
>
> If the subway line to Vienna is halted at Parkington in Arlington County and remaining service replaced with express bus service, $244 million in construction will be saved but Metro will have to buy 540 new buses that will annually consume 10 million gallons of fuel more than the subway would.
>
> Those kinds of choices are spelled out in a new, 5.5-in thick study that was generated to help area politicians decide whether buses, trolleys, or more freeway lanes might be better than subways for some parts of the region.

Another complicating factor is the time horizon. In general, we are dealing with phenomena that change slowly and have long-term impacts that can be very different from their short-term anticipations. In a sense, this is derived from the problem of causal mechanisms. Confidence in long-run predictions is tantamount to confidence in the model on which they are based. However, although models are being improved as our understanding of the underlying phenomena grows, we still have a way to go. With the present state of the art in dynamic, long-run models, we should at least be cautious. Furthermore, people are impatient, and the politician is acutely aware that the political time cycle is short. Long-run promises carry a heavy discount rate.

This discussion of broader objectives is primarily relevant to subsidized, publicly owned systems. For them, the real evaluation is a public budget decision. In this arena, the costs and benefits are assessed through the collective opinion of the political actors and their constituencies using a decision calculus that is much more complex and less traceable than the private sector discipline of profit.

In this discussion, the key distinction between comprehensive and functional is the difference between benefit (or impact) of ridership and ridership itself. Even for a simple, publicly supported, neighborhood system serving the handicapped or elderly, the question of whether the system is being operated efficiently can be answered through functional evaluation. But whether the benefit of that system is worth the cost involves comprehensive evaluation and typically is decided politically. Passenger travel is easy to measure; the worth of a particular collection of passenger travel is not.

Functional evaluation, to which we now turn, is much more amenable to measurement and analysis, but it too presents some difficult judgmental problems.

FUNCTIONAL EVALUATION

Figure 2 shows the hierarchy of interactions we are discussing. Comprehensive evaluation is at the top of the figure, and functional is at the bottom. The linkage variable is ridership: At the top, it is related to its effects on other objectives, and at the bottom, it is related to the supply characteristics of the transportation system.

At this level of evaluation, the perspectives that are of primary importance are limited. There are those of the riders and potential riders as they evaluate the service and price characteristics offered them and compare these to their alternatives and then decide whether or not to use the system. Thus, exogenous actions that affect the alternatives should also be considered in the total picture. For example, increasing the price of downtown parking affects the decision on whether to use transit or, alternatively, to join a car pool.

The second important perspective here is that of the system operator or planner who controls the trade-offs between service and cost characteristics. Ideally, the planner makes these trades with sensitivity to the perspectives of system users and potential users as well as to those of the political bodies that oversee the planner's actions and typically approve budgets.

Thus, the third perspective, looking over the shoulder of the operator, is that of the funder of the system. For a private-sector system, this is the owner (who may also be the operator and planner) and, for a public system, it is the polity. The perspective of the polity includes the social, environmental, and economic impacts of the system (comprehensive criteria), but the primary proxy variable is ridership. A well-patronized system garners more support than a sparsely patronized one.

In the following section, the primary perspective assumed is that of the operator or planner, because a good one will be sensitive to the other perspectives also. The planner's objective is improvement in some sense; the possibilities are considered below.

Figure 1. Evaluation framework.

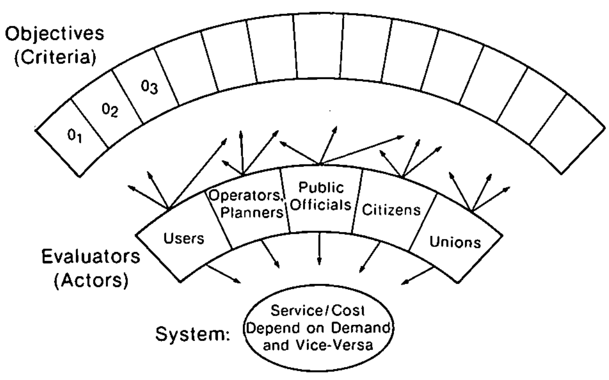
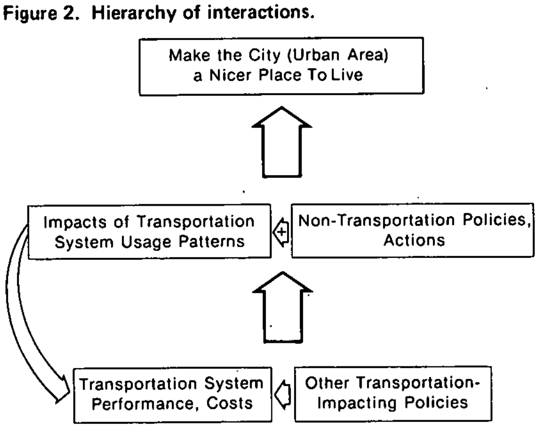


Figure 2. Hierarchy of interactions.



FUNCTIONAL EVALUATION: CONCEPTUAL FRAMEWORK

Most evaluations of transportation systems appear to be largely devoted to descriptions of the system, along with a collection of whatever numbers can be accumulated. There is usually very little (if any) interpretation of the numbers and their significance and almost never any prescription of what changes might be desirable and why. Evaluations tend to be largely descriptive, seldom analytic, and almost never prescriptive. This is understandable; we are dealing with a complex, many-faceted problem.

There are two fundamental kinds of interrelated trade-offs that characterize a transportation system. Evaluating the system really involves assessing how well these trades have been selected. The first are the market-sensitive trades—ridership versus service versus price. The second is the system-efficiency trade—service versus cost.

In developing a framework for the problem of functional evaluation, it is convenient to use the concept of a pattern of desired travel from the point of view of the potential user of public transportation. This is intended to represent the best one could do with public transportation and current technology in attracting ridership, given the competing alternative of the automobile and the extant rules and prices affecting its use. For example, the pattern of desired travel might be the travel pattern that would result if single-ride taxi service were available within 5 min of everyone. Clearly, the price at which it was offered would affect the resultant use; because the purpose here is conceptual, it can be considered free.

Obviously, this hypothetical system would generate a much higher level of ridership than any real system, but it would also involve prohibitively large subsidy. In real life, it would also compound congestion and use energy prodigally. In short, it is not a realistic alternative for a medium-to-large city; it is merely an expository device intended to crudely represent the maximum potential market for public transportation.

The primary technique in all transportation systems for lowering costs, reducing energy consumption, and improving space efficiency is aggregation. Aggregation permits the cargo, whether passengers or freight, to be moved in fewer, higher capacity vehicles that are, ceteris paribus, cheaper and more efficient per unit of capacity than small vehicles. Aggregation is the principle behind car and van pooling, behind shared-ride rather than single-ride taxi service, behind dial-a-ride, and above all behind mass transit. Higher levels of aggregation are the primary reason that full buses are cheaper than full taxis, and full 747s cheaper than full DC-9s.

But there are also costs that must be paid to achieve aggregation. These are usually manifested in some kind of forced distortion in the desired pattern of travel, either spatial or temporal. That is, the users of the system are required to either travel at a different time than they really want, or follow a more circuitous route (that may include their own walking to and from a pick-up point) than they really want, or both. This is the essence of the problem: achieving a reasonable trade-off between the benefits of aggregation and the penalties of service distortion.

Figure 3 illustrates these ideas of temporal and spatial distortion: Figure 3a illustrates a hypothetical pattern of desired travel during three 5-min intervals. But if service is offered only every 15 min, rather than every 5, then all of that travel is aggregated into one set of trips (Figure 3b). Thus, some people either forego the trip or take it at a different time than they prefer. But if they do take the trip, more people travel at the same time and the aggregation is higher.

If the system does not offer origin-to-destination service and therefore requires walking to a pick-up point, then the ideal pattern will be distorted spatially, as illustrated in Figure 3c, in which a bus route that does not go to all origin and destination points was assumed. By choosing the simple route from A to B, the circuity is minimized for the vehicle and for the A and B passengers, but forced on the potential passengers from points C and D. But if the bus route goes through all four points, then the trip for passengers from A and B is more circuitous and therefore takes more time, but the walking for the C and D passengers is eliminated. The operator sees such circuity as extra costs (the price) incurred to achieve the economies of aggregation. Everything involves trade-offs; one attribute is traded for another; what helps one person hurts another.

Figure 4 summarizes the way the actions to reduce operator costs impose walk-wait-circuity distortions on the ideal movement pattern as perceived by potential riders. If potential riders can readjust their travel times to smooth the peaks and valleys in demand, then drivers and vehicles can be used more steadily and avoid idle times, and the system is better utilized. If riders can wait for less frequent service, then more can ride at one time, and the aggregation is higher. If there are fewer pick-up points, then riders must walk more, which makes their total trip more circuitous for them but, for the operator, increases aggregation and reduces vehicle-path circuity.

Figure 5 shows two more examples of how in real systems wait time is traded for walking, or spatial distortion is traded for time distortion. The system shown on the top left has a higher frequency of service and less circuity for the vehicle so that the rider has less in-vehicle time, but the price is more walking to the pick-up point. In the system shown on the top right, the rider trades less walking for longer waits and more in-vehicle time. The market decides which is preferred.

The system shown at the bottom left of Figure 5 offers shuttle service between two collector points rather than individual door-to-door service. If the actual origins and destinations are closely clustered about the chosen collector points, the riders may prefer the shuttle service to door-to-door service because it provides (at the same cost) more frequent service and shorter ride times. But if actual origins and destinations are more randomly located, then door-to-door service may be preferred; circuity and wait time will be traded for walking. An example of this type of trade occurred in the early days of the Batavia, New York, system. Because the ridership patterns observed in providing origin-to-destination dial-a-ride service suggested that using one vehicle as a point-to-point shuttle would be preferred to continuing with dial-a-ride only, mixed service was instituted.

The more general problem of evaluating whether flexible-route service; fixed-route, fixed-schedule service; or some combination is preferred in a given market situation is more complex than simply calculating relative productivities. Some people apparently just will not ride without door-to-door service; others would prefer to walk to a pick-up point rather than make the implied commitment of calling a vehicle to come to their door. The managers of the Teltrans System in Ann Arbor, Michigan, encountered this problem when they were contemplating the substitution of fixed-route service for dial-a-ride service in a neighborhood where ridership was outgrowing the existing service. None of these trades are simple.

Thus, the quest for an absolute scale or set of standards with which to evaluate urban transportation systems is not likely to succeed unless the schema reflect the fact that one attribute must be traded for another. Furthermore, these trades are both market and site specific. A system operator or planner must simultaneously make two kinds of judgments. The first is a market judgment in which he or she must estimate how ridership will re-

Figure 3. Pattern of desired transportation.

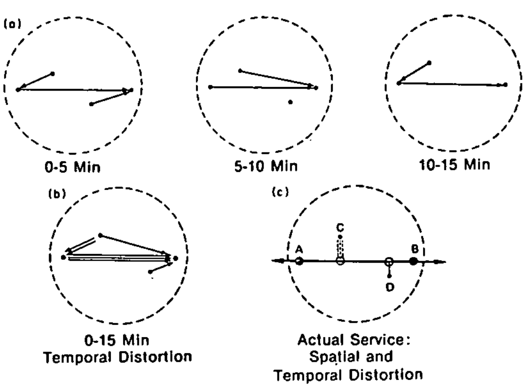


Figure 4. Cost reduction distorts desired pattern.

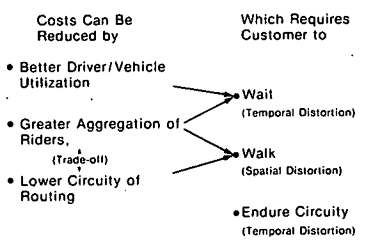
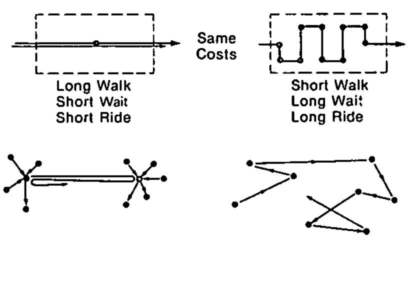


Figure 5. Trading wait time for walking.



spond to various combinations of service (walk-wait-circuity) and price; i.e., what combinations of temporal and spatial distortion of the ideal travel pattern are likely to be the most tolerable to potential riders at a given price. The second is an operating-efficiency judgment in which he or she must estimate how costs will vary as a function of the service (walk-wait-circuity-coverage) offered. Typically, these trades must be decided in the context of a constrained maximum capital-and-operating-subsidy budget. It is not an easy set of decisions. Furthermore, there are numerous pragmatic considerations. Complex operational schemes are confusing to both system operators and system customers, so that there is a premium on simplicity. Changes are slow and usually difficult to bring about, so that frequent experimentation is not usually possible.

The criterion for making these trades is usually some ambiguous version of maximizing the ridership of the target market per unit of subsidy. The qualification "of the target market" is important. For example, a service aimed at helping the elderly that becomes so popular with teenagers that no elderly person wants to ride has missed its objective.

The first of these trades, the ridership versus service versus price relation, is illustrated in Figure 6a. The evaluative question here is whether the spatial and temporal distortions imposed by actual service on the pattern of desired travel are reasonably appropriate to the service propensities of the target market(s); i.e., would offer a different type and pattern of service to attract a higher ridership? This is clearly difficult to answer because we can only deduce desired travel patterns and service propensities from questioning and observed behavior. Both can be very misleading; people often do not know what they want until they experience it, and short-run behavior is often not the same as long-run behavior.

This market-match question entails two different ground rules. The first governs whether more ridership can be attracted by offering a different combination of service and price under the restriction that total deficits do not increase. The second relaxes the constraint on costs. The addition of these dimensions to the original market-match question is illustrated in Figure 6b.

Obviously intertwined in these market questions are the issues of efficiency in choice of modes and utilization of the factors of production. The principal clues to a need for improvement are low load factors, low utilization of equipment and personnel, or a combination of these. Some are unavoidable; transportation, like almost all other forms of human endeavor, has its peaks and valleys of activity. The problem is to minimize their impacts; strategies to do so will be discussed below.

This discussion of a conceptual framework for functional evaluation has not introduced any unfamiliar thoughts; its usefulness depends largely on whether one's personal way of thinking about such problems agrees with the perspective chosen here for exposition. There is, however, one more element of the conceptual problem that is important to the way the market-match and efficiency questions are viewed. This is an aspect of the problem that may be more important than the attention that has been accorded it: the nature of the supply versus demand equilibrium and its implications for system evaluation.

Figure 6. Market-match framework.

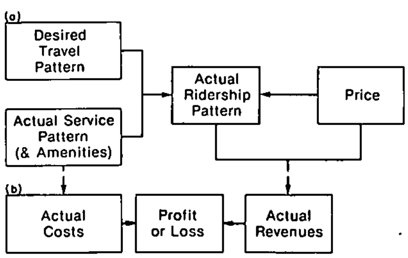
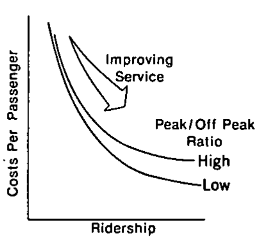


Figure 7. Transportation system costs.



CONCEPTUAL FRAMEWORK: DYNAMIC SUPPLY VERSUS DEMAND INTERACTIONS

The basic determinant of the productivity a transportation system can achieve depends on the pattern of movement being served. A high productivity is possible only if a high level of aggregation of riders is possible. High aggregation is possible only if the number of origins and destinations is relatively limited in comparison with the density of ridership. High utilization is possible only if there is no markedly varying temporal pattern that produces high peak-off-peak ratios. These points are summarized below.

1. Greater aggregation of flows permits larger vehicles and higher service frequency or both, and
2. Flat temporal patterns permit better vehicle and driver use and simpler management.

Figure 7 shows how costs of operation vary with level of ridership. The curves are intended to be illustrative but are based on an actual calculation for a medium-sized city (4, 5). Increased ridership represents higher levels of aggregation, thus lower costs and better service. In general, service improves with ridership for three reasons: (a) once the ridership exceeds that associated with some minimum service threshold, frequency of service can be increased; (b) higher ridership increases the opportunity for more direct routing; and (c) at higher ridership, more infrastructure amenities such as shelters and information systems can be afforded.

The important point is that when ridership is growing and peak-off-peak ratios are improving (or both), then service and productivity should also be improving. This, in turn, may induce still greater ridership. (There is a caveat; if labor or other costs are growing, then productivity improvements can alleviate the growth in total costs, but may not completely offset it.)

Conversely, if the ridership is decreasing, the portents are for further worsening as costs increase and service decreases. Decreasing ridership and increasing peak-off-peak ratios are the early warning signals of deteriorating systems.

An important part of an evaluation, then, is to observe and analyze the direction of change of ridership patterns, so that if the system is deteriorating, corrective actions can be taken and the disequilibrium between demand and supply characteristics reshaped. This should be done on a market-by-market and individual route or neighborhood basis: Aggregate averages hide too much.

To summarize, transportation systems typically operate where marginal costs are falling, and the quality-of-service changes that come with increased service frequency and more direct routing are shifting the de-

Figure 8. Leading indicator: direction of change of ridership patterns.

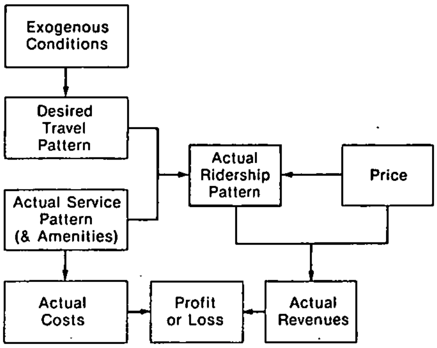
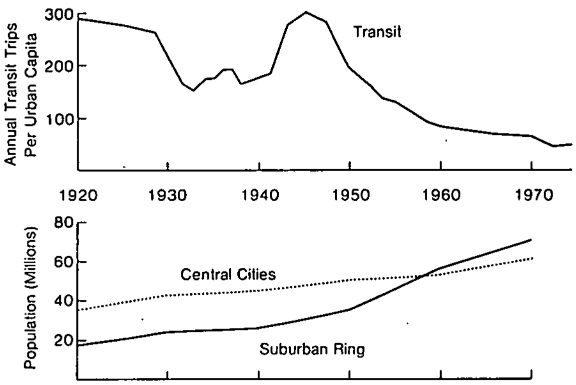


Figure 9. Urban passenger travel.



mand curve. Thus, the equilibrium between demand and supply is unstable. On the one hand, the lower limit is system failure, and on the other, a steady-state is reached as the marginal costs flatten and the market is saturated; no further shifts in the demand curve appear. (External subsidy also shifts operating points.)

STRATEGY ALTERNATIVES FOR FUNCTIONAL IMPROVEMENT

This section completes the general picture and illustrates how specific kinds of actions fit into it.

The most important factors affecting transit—the exogenous factors—which are not shown in the logic structure (Figure 6b), are shown in Figure 8. The evidence is very strong that what has happened to public transportation has not been caused by anything that has happened to transit itself but is the response to one primary exogenous factor: the automobile, first as it has offered competing transportation and second as it has contributed to changing urban spatial form.

If we view the system in the framework of Figure 8, there are three blocks over which the transportation authority potentially has some control. It can influence the patterns of desired travel—the size and nature of the potential market—through changes in the exogenous conditions; it can change prices; and it can change the pattern and amenities of the actual service offered.

There are three classes of exogenous actions that can affect the desired pattern of travel by public mode. The first is the set of actions that can alter the relative attractiveness of the automobile (e.g., changes in parking arrangements and changing prices). These are potentially very powerful because a small shift in the use of automobiles is a large percentage shift in the use of alternative modes. But changes that are very restrictive of automobile use are usually not feasible politically. It does appear, however, that our resistance to such restrictions is slowly decreasing.

The second class are actions (e.g., flextime and other work-hour-staggering schemes) that can change peaking patterns. Car pooling and van pooling also change peak demands on public transit.

The third are actions that could influence spatial organization. Whether this is even possible is controversial, and in any event, it is clearly not a short-term measure.

These classes of actions, with the exception of the borderline case of car pooling, do not involve changes to the public transportation system itself. They are changes that are exogenous to it, aimed at affecting the desired pattern of travel. Changes to the system itself do not change the desired pattern—the basic potential market—but do change the actual pattern because they affect the choices offered to potential users.

The many approaches to changing services and prices have been widely discussed in other places [e.g., by Keyani and Putnam (6) and by Ward (7)]. It is beyond the scope of this paper to further analyze or even enumerate them, except to note that there is still a tendency to evaluate alternatives singly rather than in combination, although the interactions between specific techniques make the latter the preferred approach. But we are still impeded by inadequate analytical tools for synthesis and evaluation of integrated, multielement systems.

If the general decline in public transit has largely been due to a declining market, a shrinking pattern of desired travel, and the service-cost deterioration that followed and an observed decline on a route or in a system portends further decline, what, then, can be done when a deteriorating situation is observed? If it seems reasonably clear that the problem is a declining basic market due to exogenous factors and not simply an ineptly supplied transportation service (and the option of modifying exogenous factors is not possible), then there appear to be only two options. The first is to increase the subsidy, gambling that a greatly improved service will expand market capture and improve the service-cost ratio (and therefore partially absorb the need for continued high subsidy), i.e., move to the right on Figure 7.

The second option is to reexamine the goals of the system, sharpen the definition of the market it is directed toward, and then structure the service to that specific market. For example, if general purpose service is being offered, but the real concern is one small disadvantaged group, then subsidized taxi might be more appropriate. Rethinking goals may become more and more important in the years ahead.

DEMOGRAPHIC INFLUENCES

The magnitude of the ascendency of the automobile has been described elsewhere (8). Between 1950 and 1970, the automobile cut passenger travel by public transportation approximately in half, but created 30 times that amount of new travel.

Although the automobile alone did not cause suburbanization, it has been a major factor in the decreasing densities of both downtowns and suburbs. However, the efficiency of transportation depends on aggregation, and the ability to aggregate depends on the density of potential riders. As densities have decreased, modes that depend on aggregation have increasingly suffered.

Conventional transit seems to require about 17 dwelling units/ha (7/acre) to permit the levels of aggregation for adequate service. This is a fairly high density, corresponding to about 4600 persons/km^2 (12 000 persons/mile^2). Overall average densities in American cities are less than half of that.

The impact of declining urban densities on transit is illustrated in Figure 9. The number of automobiles in the United States began to increase after World War I, reached a plateau of about 0.2 automobile/person in 1930, remained constant during World War II, and is now approximately 0.5 automobile/person. And, as the suburbs grew after World War II, transit use steadily declined. This does not prove causality, but it is not hard to believe that the changing spatial form is a contributing factor.

Thus, the rapid increase in transit unit costs that has been observed is consistent with the idea that, on the average, we are reversing the trend shown by the curve in Figure 7. Again, causality is only inferred, because the unit costs of labor and other inputs have also been increasing. It would take a more careful and disaggregate analysis to establish the relative importances of the various influences. But the general observations developed above about the behavior of transit systems are not inconsistent with the aggregate trends observed.

The evidence is persuasive that the size and density of the central business district, the residential density, and the general spatial arrangement of the city strongly affect transit use. Thus, the evaluation of transit in any of its forms must include these factors explicitly. In particular, comparisons between systems can be very misleading unless differences in demography are specifically considered.

SERVICE STANDARDS

A word on the role of service standards is appropriate. Service standards are primarily a device for operational control, because they provide a standard against which to compare actual operations. They are also a device for communication, because they describe the system. But they do not in any way evaluate whether the system is properly matched to the market conditions, i.e., whether the optimum trade-offs among walk, wait, circuity, price, and subsidy have been selected.

An operator might evaluate his or her own performance as excellent because all of the service standards selected were being achieved, while, at the same time, a potential user was avoiding the system because the standards selected were badly matched to the market.

CONCLUSIONS

The key conclusions are summarized below.

1. Overall cost-benefit and impact issues are largely in the political arena.
2. The key determinant of potential productivity for a given level of service is ridership pattern.
3. Supply-demand equilibria are unstable: The direction of change of ridership pattern is the most important evaluative measure.
4. The options for improvement include the following: (a) exogenously modify latent patterns (work-hour staggering, spatial development patterns, automobile management), (b) new service-characteristic trades to broaden the market, and (c) better service through higher costs.

Although it would be desirable to evaluate our public (and private) transportation systems as to how they contribute to larger objectives such as improving the quality of life, we are unable to do so except very generally because of our relative ignorance of the causal mechanisms. Perhaps one of the first prerequisites of rational evaluation is that one should understand reasonably well the relation between the observable properties of what one is evaluating and the objectives against which it is being evaluated.

The inability to trace causal mechanisms combined with the fact that different persons have different and multiple objectives means that the overall assessment of whether a service is worth its costs is a political decision.

We do have a much better understanding of how the potential service and cost characteristics of alternative transportation modes vary with patterns of ridership. As a first level of functional evaluation, this allows us to assess whether we have appropriately matched the mode and service level with the ridership pattern to be served. It is a reminder that in comparing systems, unless the differences in the types of ridership patterns served are taken into consideration, the comparison can be misleading.

Perhaps the most important conclusion is that the way the ridership pattern is changing is the real clue to the future of the system; if the change is unfavorable, then action to change the trend is required.

Let us think in terms of shaping the ridership to improve the use of public modes. The options are three—exogenous changes, different service characteristics, and better service purchased at the risk of larger subsidy. The functional evaluation then becomes the evaluation of a total strategy for improvement, a prescriptive evaluation rather than simply a judgment of current performance.

The central criterion is ridership response. If ridership can grow appreciably and the largest percentage growth is in the off-peak period, then the promise of reducing the deficit per rider is real. But it seems probable that this cannot be done with incentives alone. The automobile is very attractive, and disincentives or constraints on its use will probably be required, along with staggered work hours and car and van pooling to relieve the peak-hour pressures. However, for automobile restrictions to be acceptable to the public, the alternative offered must be very attractive.

The only way to learn what is possible is by experimentation. Even if pursued vigorously and expeditiously, the experimental process could take a decade, given the time required to bring integrated transportation systems into being and operate them long enough for automobile-use habits to change. To assume otherwise does not seem realistic. Functionally coordinated systems will introduce operational and institutional problems with which we have little or no experience; there is still much trial and error. The experiments now in progress should be continued.

Such experimentation should be managed so that there is continuous interplay with theoretical analyses and simulation modeling as well as with other experiments. The experimental results improve the quality and reality of the theoretical models, and the development of the models improves the ability to extrapolate the experimental results to different situations. The development and consolidation of generically useful tools and knowledge and their broad dissemination are clearly desirable.

Finally, our ability to evaluate performance and prescribe improvements depends on our ability to correctly interpret what we observe.

REFERENCES

1. J. Hoffman and M. E. Goldsmith. A Comprehensive Evaluation of Transportation Alternatives for a Large Urban Area. Proc., 15th Annual Meeting, Transportation Research Forum, Vol. 15, No. 1, 1974.

2. J. A. Gomez-Ibanez. Transportation Policy and Urban Land-Use Control. Department of City and Regional Planning, Harvard Univ., Cambridge, MA, Discussion Paper D75-10.
3. D. Feaver. Metro Choices Detailed. Washington Post, Oct. 18, 1977.
4. J. H. Batchelder and others. Operational Implications of a Major Modal Diversion to Transit. Office of Research and Development Policy, U.S. Department of Transportation, Rept. DOT-TST-76-72, April 1976.
5. J. W. Billheimer and others. Macroanalysis of the Implications of Major Shifts in Integrated Regional Networks. Office of Research and Development Policy, U.S. Department of Transportation, Rept. DOT-TST-76-65, April 1976.
6. B. Keyani and E. Putnam. Transportation Systems Management: State of the Art. Urban Mass Transportation Administration and Federal Highway Administration, U.S. Department of Transportation, Feb. 1977.
7. J. D. Ward. An Approach to Regionwide Urban Transportation. Office of Research and Development Policy, U.S. Department of Transportation, Rept. DOT-TST-75-108, July 1975.
8. Urban Densities for Public Transportation. Regional Plan Association, New York; Tri-State Regional Planning Commission, New York, May 1976.

**Mr. Ward was with the Office of the Secretary, U.S. Department of Transportation when this paper was written.*

Dial-a-Bus Implementation: A Living Example

Joseph S. Silien, Rochester-Genesee Regional Transportation Authority, New York

The process undertaken to transfer a demand-responsive dial-a-bus service from federally funded, experimental-demonstration status to permanent-operation status is described. The particular project described suffered from a local financial condition that was insufficient to support dial-a-bus services, technical difficulties, and problems in perceived attitudes toward the project. The strategy developed included an independent assessment of the program, consultant recommendations modified by local staff, expansion of the service to two additional areas, areawide service to the handicapped and elderly, attempts to negotiate the labor clearances necessary to permit competitive solicitation of service vendors, and approval of an extended demonstration by the Urban Mass Transportation Administration.

Implementation is the most difficult phase of any research, development, or demonstration project—perhaps that is why the term was, until recently, only rarely mentioned. To be sure, any particular project included implementation for the duration of the particular test or as long as the agreed-on combination of federal and local funding lasted. But discussions of permanent implementation are conspicuously absent from the voluminous texts of government contracts. Even when unwritten commitments to implement have been made, events and budgets have often altered even the best of intentions. Recently, acknowledging the problem, the Urban Mass Transportation Administration (UMTA) renamed its Research and Development Program the Technology Development and Deployment Program, in an attempt to establish the essential link between the development of a new concept and its delivery to the public as a permanent part of an improved transportation system.

An examination of the dial-a-bus project in the Rochester, New York, area illustrates the difficulties inherent in permanently implementing a project based on a new technology.

The project had its beginnings in October 1971, when three small buses were placed in demand-responsive service in Batavia. The Batavia Bus Service, or B-line, had been acquired by the Rochester-Genesee Regional Transportation Authority (R-GRTA) in June of that year. A very limited fixed-route service was operated until the transition to dial-a-bus. The dial-a-bus mode was adopted in an attempt to expand transit service to as many residents as possible by using the then-new and innovative technology of demand-responsive service.

R-GRTA had submitted an application to the Urban Mass Transportation Administration (UMTA) for demonstration funds to operate a dial-a-bus in the city of Rochester. But UMTA had determined, under Congressional pressure, that only one dial-a-bus project would be federally supported until such time as success was ascertained. The location of this project was Haddonfield, New Jersey, which was selected after the New Jersey State Department of Transportation committed itself to using its commuter operating authority to assist the project, as well as for reasons of size, density, and traffic potential.

In Rochester, a demand-responsive service was eventually established in the suburban community of Greece by using local resources. The service was named PERT (for PERsonal Transit) and offered door-to-door transportation within the service area and feeder service to fixed-route buses.

In 1974, R-GRTA commissioned a short-range improvement program study. A principal finding of the study report (1) was that PERT services should be expanded to include all suburban areas contiguous to the city of Rochester.

The report projected an annual operating cost, including rationalization of the fixed-route services, of $1.2 million for the entire suburban service. The substitution of PERT for the majority of the urban fixed routes on Sundays and at night was further projected to produce a saving in operating costs. The entire plan was based on the assumption that all of the operating assistance funds available under section 5 of the Urban Mass Transportation Act of 1964 could be drawn down by R-GRTA. Because of insufficient local share, this proved not to be the case.

In part because of the initiative Rochester had taken in implementing the dial-a-bus services in Batavia and Greece and developing an areawide dial-a-bus plan and in part because UMTA was seeking an expanded site for testing computer dispatching software, in March 1975, Rochester was awarded a demonstration grant to develop an integrated adaptable metropolitan transit service program. The Massachusetts Institute of Technology (MIT)