ASSESSING THE UTILITY OF AN INTERACTIVE GRAPHIC COMPUTING SYSTEM: A TRANSPORTATION SYSTEMS DESIGN PROBLEM

Jerry B. Schneider and Dennis Porter,
Departments of Urban Planning and Civil Engineering, University of Washington

An experiment designed to assess the amount of improvement in the quality of a design that can be obtained by using an interactive graphic computing system was undertaken and is interpreted. The problem was design of a bus rapid transit system by five teams of students at the University of Washington. Each team used a man-computer system called UTRANS to search for a design for a BRT system that would satisfy 11 performance measures. None of the teams found a wholly satisfactory design, but the average improvement in performance (design quality) for all teams was close to 50 percent (relative to their initial designs). Further experiments of this type are needed to assess the utility of interactive graphic design tools in the transportation systems field.

•THE POTENTIAL UTILITY of interactive graphic computing systems as designing tools is still an issue of considerable debate (7). Recent hardware and software developments have made it possible to tackle a number of complex design problems with only modest investments in equipment. Still, there is far too little evidence that substantial payoffs (i.e., better designs) can be obtained by implementing an interactive graphic system.

How much improvement in the quality of a design can be expected from the use of an interactive graphic system? Our major hypothesis is that design improvements (using the initial or "first-cut" design as a base) on the order of 25 to 50 percent are likely and improvements between 50 and 100 percent are possible, given a well-trained systems operator and a well-developed interactive graphic computing system. It is probable that improvements of 50 to 100 percent need to be demonstrated before investments in interactive graphic design tools will become attractive to potential users. Improvement is defined, for our purposes, as a measure of increased performance over the initial or "first-cut" design taken as a base. It is recognized that the initial design is not an ideal base from which to measure improvement, and the limitations of using this procedure are discussed.

The substance of this paper is a report on some results of an experiment designed to assess how much improvement in design quality can be expected when the designer is aided by an interactive graphic computing system. The problem to be solved was design of a complex bus rapid transit (BRT) system. Five teams of students were asked to design a BRT system for a portion of the City of Seattle, Washington. Eleven objectives were specified for this problem, and the objective of each team was to find a design that would satisfy all 11 objectives. Each team was to use an interactive graphic computing system, known as the urban transit analysis system (UTRANS), to formulate and evaluate several alternative BRT system designs. The object of the experiment was to determine how much improvement could be obtained when man and computer work together, moving from an initial (unsatisfactory) design to one satisfying as many of the 11 design criteria as possible.

Publication of this paper sponsored by Task Force on Interactive Graphics.

The research design for this experiment was relatively simple and straightforward. The students were grouped into five teams of four or five persons each. One person from each team was given a short period of intensive training on how to operate UTRANS. All students were given some basic instruction on the theoretical and practical characteristics of the UTRAN system. All teams were assigned the same problem: Design a BRT system that will satisfy or exceed 11 objectives that relate to system performance. The definitions of these objectives were discussed in detail. Each team was to work independently, and competition between teams was encouraged. Each team was to have the same amount of computer time (in 2-hour blocks). The teams were formed with a random selection process except that each team had at least one person who had had some previous experience with computers.

In this setting, each team developed a first-cut design on paper and then began the interactive design process. This consisted essentially of looking for ways to improve the current design, making the changes, and then evaluating the modified design. Each team was expected to evolve some type of design strategy that would, more often than not, help the team find a series of successively better designs. This iterative process was to continue until the available computer time had been fully used, and the best design obtained (not necessarily the last design obtained) was to be presented to the class for discussion. Judgments on how to modify the design at each stage of the iteration were to be made by team members by using whatever decision-making procedure suited them.

Four specific evaluation objectives for the experiment were determined:

1. How similar or different are the five best designs, in both visual and quantitative terms?
2. How successful was each team in satisfying the 11 design objectives? How similar or different were the five teams in this respect?
3. How much improvement in performance over the initial design was achieved by each team and for the group as a whole?
4. What were the characteristics of the design strategies evolved by the successful teams? What were the reasons for unsuccessful efforts on the part of any team?

These questions will be discussed later in the paper.

DEFINITION OF THE BRT DESIGN PROBLEM

Network and Demand Data

The problem selected was the design of a peak-period BRT service from many residential origins to a single destination (i.e., a many-to-one service). The northern part of Seattle was chosen as the setting for the problem, and the corridor used is shown in Figure 1. The major destination is downtown Seattle, which is about 100 miles long and 5 miles wide and had a population of 200,000 in 1970. The network included all of the principal residential streets. It is bisected by Interstate 5, which has six major interchanges in the corridor. The density of people who live and work in the corridor is represented by a hexagon at each node in Figure 1. The size of the hexagon is proportional to the number of people who live near the node and who commute daily to downtown Seattle. These data, the network coded with travel times and the demand set, are the two basic elements for the problem.

Behavioral Assumptions

The heart of the UTRAN system is a modal-split model called the n-dimensional logit model (5). This model forecasts how any particular design may be expected to be used. It splits all travel to downtown Seattle among the three modes included in UTRANS: drive, park-n-ride, and walk-n-ride. All 11 performance measures are derived from the results of this modal-split forecast; thus it is most important for the planner (or designer) to understand how it works. Approximately 15 hours of instruction was devoted to this topic, and all of the participants obtained a good understanding of the mechanics of the model. The team's ability to fully understand how the model was formulated and how it worked was probably a major factor in its ability to formulate a successful design strategy.

Performance Measures Used for Evaluation of Alternate Designs

All evaluations of alternative designs were based on the 11 performance objectives. These performance measures, their definitions, and their desired values are given in Table 1. These measures relate to the travel cost, comfort, profitability, and the sociopolitical effects of a BRT system. Other important impact categories were not included in the interests of keeping the complexity of the problem within reason.

These performance measures are not an ideal set for use in this type of experiment inasmuch as some of them are interrelated. Measures that are more independent and, in some cases, less abstract would have been preferred. However, these measures were available, and generation of others would have involved additional programming. This was not possible because of resource limitations. These difficulties did serve a useful educational function in that the students learned about the problems of working with interdependent and abstract measures. While troublesome, these problems were not judged to be serious enough to invalidate the basic experiment.

Estimation of Weights for Performance Measures

When the performance measures had been defined and discussed, each student was asked to complete a partial paired comparisons matrix, with constant sums, to provide his own estimate of the relative importance of each of the performance measures. These results, when analyzed statistically, showed that there was no significant difference among the group average values of these weights. Even though these weights were about equal, some of the teams gave somewhat more thought to improving those measures that ranked highest in this analysis.

Estimation of Minimum-Maximum and Ideal Standards for Performance Measures

As a further step in the definition of the problem, each participant was asked to estimate what he thought the minimum, maximum, and ideal levels of each performance measure should be. These preferences were then analyzed, and a plot of the results for each performance measure for the group was made.

The results of this exercise revealed that some significant differences existed in the interpretation of the definitions of the performance measures. After these differences were identified and discussed by the class, the exercise was repeated, and satisfactory consensus on appropriate values was obtained. The average values for the group are given in Table 2. These values served as a perspective on the range of reasonable variation for the performance measures and were used to establish a common scale for comparison of the best designs of each of the five teams.

Constraints Imposed by the Instructor

A further definition of the problem was made by imposing a set of constraints on the design problem. Some of these constraints were due to hardware or software limitations. Others were included to represent typical environmental constraints that always appear in problems of this type. These constraints were as follows:

1. No more than 60 bus stops and 20 bus lines,
2. No more than 20 park-n-ride lots and no lot larger than 1,000 spaces, and
3. A downtown all-day parking fee of $1.50 and an average walk time downtown of 5 minutes.

This problem definition process together with the presentation of the UTRANS modal-split model required about 15 hours of instructional time and was considered essential to the conduct of the experiment.

CHARACTERISTICS OF THE PARTICIPANTS AND THE PROBLEM-SOLVING ENVIRONMENT

Characteristics of the Participants

Twenty-three students, from both undergraduate and graduate levels, took part in

Figure 1. Seattle study corridor.

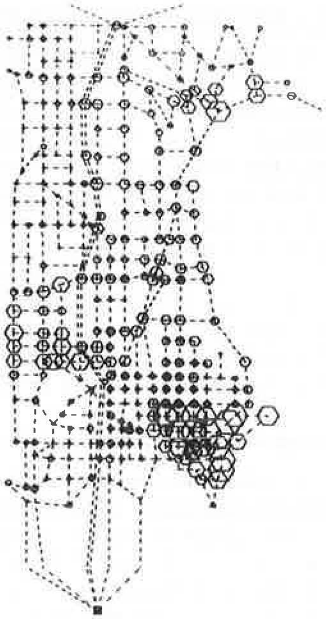


Table 1. Definitions and desired values of performance measures.

Performance Measure	Definition	Specified Satisfactory Level
Profit-loss of bus operations, dollars	Bus fare-box revenues less operating costs	0
Profit-loss of park-n-ride lots, dollars	Park-n-ride lot revenues less operating costs	-250
Walk-n-ride patronage share	Percentage of total demand using walk-n-ride mode	20
Park-n-ride patronage share	Percentage of total demand using park-n-ride mode	30
Bus load ratio, seated patrons	Systemwide efficiency factor of bus operation; measure of overall use of bus carrying capacity	0.80
Percentage of standing room used	Systemwide measure of crowdedness on bus	30
Park-n-ride lot load factor	Systemwide efficiency factor of park-n-ride lot operation; measure of overall use of park-n-ride lot parking capacity	0.80
Walk-n-ride accessibility index	Constructed accessibility measure for walk-n-ride mode	70.0
Park-n-ride accessibility index	Constructed accessibility measure for park-n-ride mode	120.0
Percentage of people within 5-min walk	Proximity measure of closeness of total demand to any bus stop	35
Percentage of people within 5-min drive of bus stop	Proximity measure of closeness of total demand to any park-n-ride lot	70

Table 2. Minimum-maximum and ideal levels of performance measures.

	Group Average Value	
Performance Measure	Minimum-Maximum	Ideal
Profit-loss of bus operation, dollars	-260.0	+72.7
Profit-loss of park-n-ride lots, dollars	-151.0	33.5
Walk-n-ride patronage share, percent	17.0	38.5
Park-n-ride patronage share, percent	22.0	40.7
Bus load ratio	0.63	0.94
Percentage of standing room used	35.0	2.00
Park-n-ride lot usage, percent	0.63	95.0
Walk-n-ride access index	58.4	96.6
Park-n-ride access index	77.0	121.1
Percentage of people within 5-min walk of bus stop	30.0	70.0
Percentage of people within 5-min drive of parking lot	37.7	70.5

this experiment. Of this number, only four had had any previous experience with man-computer systems, and only one had worked with the UTRAN system. They came from a variety of disciplines as given below:

Discipline	Participants
Urban planning	9
Civil engineering	6
Architecture	4
Psychology	2
Anthropology	1
Geography	1
Total	23
Level	
Undergraduate	8
Graduate	15
Total	23

None had had any experience with BRT system design. The most often expressed motivation to participate was the desire to obtain some "hands-on" experience with a man-computer system. Unfortunately, because of the large size of the group and the limited hardware available, only one person from each team was allowed to operate the computer. However, other members of the teams were able to observe the operation of the UTRAN system closely and to participate in the problem-solving (design) process (both on and off line).

UTRANS Operating Characteristics

The origins, evolution, and characteristics of the UTRAN system have been well-documented by Rapp (2, 3, 4), Gehner (1), and Schneider (8) and will not be discussed extensively here. Briefly, UTRANS is operated as shown in Figure 2. It has been structured to assist a planner in generating and evaluating alternative BRT system designs for service in urban activity centers. It is limited to cases where there are many origins and one destination. Two modes of operation are possible. The first is the manual mode where the planner makes a series of design decisions. The second is a partially automated mode, which relieves some of the decision-making burden. In the manual mode the planner is presented with a display of the street network, the demand pattern (i.e., the location of the people who desire to travel to the major activity center), and a display of land values superimposed on the street network. Then the planner lays out a first-cut transit system design by making a series of design decisions and entering them either on the graphic display or on the keyboard of the graphics terminal in the sequence shown in Figure 2.

Once the design decisions have been made, the first-cut transit system design is ready for computer evaluation. Each of these decisions is recorded by the computer and is input into a modal-split model. This model is designed to estimate the proportion of trip-makers that will use each of three modes of getting from their homes to their destination. The modal-split model constitutes the heart of this man-computer system inasmuch as the evaluation of alternative designs is derived wholly from its prediction of the expected performance of each design. This prediction procedure is based on the following assumptions:

1. Each trip-maker selects from among walk-n-ride, park-n-ride, or drive modes.
2. The trip-maker's choice depends on the relative difficulty (or impedance) that he perceives with each mode.
3. The total impedance of a mode is the sum of the impedances associated with the several elements of trip by that mode.
4. Each element of a trip is multiplied by a constant that is proportional to an esti-

Figure 2. Operation of UTRAN system.

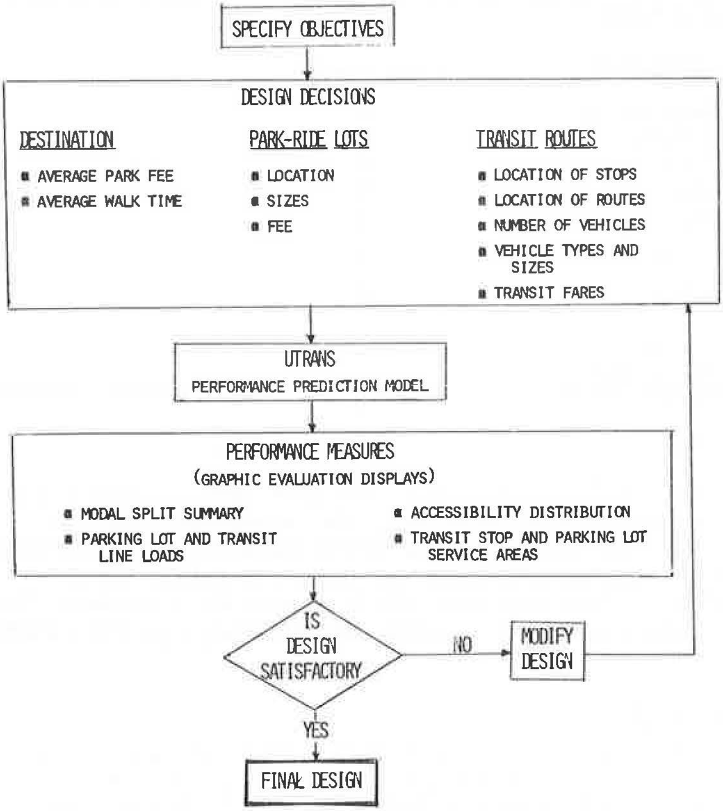
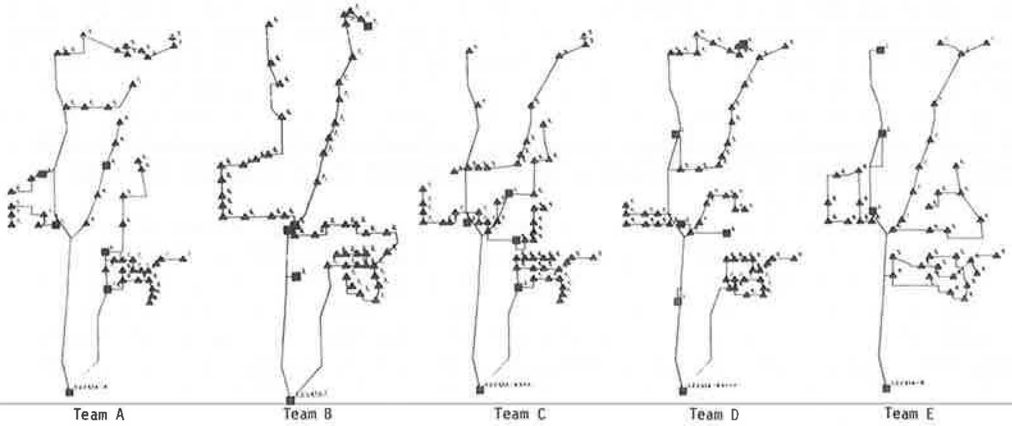


Figure 3. Final designs of each team.



Team A Team B Team C Team D Team E

mate of the relative disutility associated with that type of activity. The constants are called impedance coefficients.

5. The smaller the total impedance of each mode is, the more likely it is that a trip-maker will select it.
6. The share of the available patronage attracted by each of the three modes is inversely proportional (in a negative exponential fashion) to the mode's overall impedance.

The output of the modal-split model is as follows:

1. The percentages of all trip-makers using the three modes,
2. For each transit stop and park-n-ride lot, the volume of patrons who walk or drive (if the stop has a parking lot) to it,
3. For each transit line and parking lot, the costs and revenues of operation (including capital costs), and
4. The total system cost and revenue on a daily (24-hour) basis (the difference between these two figures is the overall daily profit or loss of the system).

This information is presented to the planner in tabular form on the scope face of the graphic terminal. Several additional evaluation displays are available to the planner, all of which are derived from the output of the modal-split model.

In the first cycle the planner structures a first-cut design, and the computer evaluates it and presents him with a variety of displays which he examines. The task then is to develop ideas that should improve the first-cut design by adding park-n-ride lots, changing parking fees, increasing the number of buses on a route, or modifying the original design in some other way. This revised or second design is evaluated by the computer and a new set of evaluations and a listing of performance parameters are displayed. This procedure is repeated until the planner achieves a satisfactory design or is restricted by time constraints. This is the process used by each of the five teams in the conduct of this experiment.

RESULTS OF THE EXPERIMENT

Implementation of the experiment was hindered by hardware malfunctions that interrupted it after the off-line instruction had been completed and the teams were ready to begin work on the UTRAN system. This delay probably reduced the effectiveness of the teams' ability to use the instruction and probably introduced a conservative bias into the results.

The results of the experiment have been grouped into four categories:

1. A description of the physical attributes of the five final designs,
2. An analysis of the performance levels of the five final designs,
3. Measures of the amount of improvement achieved by each of the teams, and
4. A general discussion of the problem-solving strategies used by the teams.

Physical Attributes of the Five Final Designs

One of the questions of interest in this research was whether five teams, working independently, would come up with similar design solutions. Comparisons among complex transit system designs can be made in at least two ways. One query is, Do they look alike? Another is, Are they composed of similar quantities of the various elements (i.e., buses, park-n-ride lots, and so on) that make up a design solution? As an answer to the first question, Figure 3 shows a comparison of the five final designs.

These five designs have a similar appearance. Most involve relatively complex networks, long bus lines, and close bus stop spacing. Park-n-ride lots tend to be located more in the lower half of the network, closer to the downtown destination. The design of team E is a minor exception to some of these qualitative observations. Most teams obviously attempted to provide the greatest service to the nodes of highest demand (Fig. 1). Each team, with the exception of team E; used all the major arterials in the corridor for one or more bus lines. It is possible that the similar appearance of the designs was structured by the dominance (accessibility-wise) of the single Interstate highway and the concentration of demand in the lower right part of the network.

Table 3 gives data relative to the second question. The first three items in Table 3 are quantitative in nature, whereas the other five are more qualitative. The physical elements and operating policies of the five designs are quite similar. This is probably not too surprising because all teams used fairly conservative approaches to the design problem and did little experimentation with unconventional designs. Most teams said that, if they had had more operating time on the computer, they would have tried more unconventional ideas in the design process.

Performance of Final Designs

Figure 4 shows the results of the final designs in terms of the 11 performance measures. Each performance measure is shown as a percentage of the difference between the minimum-maximum standard and the satisfactory level of each measure as given at the beginning of the experiment. As can be seen, none of the teams was able to satisfy all 11 objectives. Teams A and C satisfied eight of the objectives, two teams satisfied seven, and one team satisfied only six of the objectives. Figure 4 shows that all teams were able to oversatisfy some objectives, whereas others proved to be particularly troublesome to nearly all the groups. Data given in Table 4 show how often each of the 11 performance objectives was satisfied. As can be seen in Table 4, all teams were able to satisfy two objectives.

In terms of the 11 performance measures, the final designs were quite dissimilar. Table 5 gives the simple correlation coefficients for all possible pairs of designs. As can be seen, only three of these 10 coefficients are greater than 0.7 while six are less than 0.5. Thus, although the designs are quite similar in physical terms, they are very different in performance terms. The reasons for these variations in performance are many and varied and will be discussed in more detail later.

Measures of Design Improvement Obtained

As discussed previously, the major objective of the experiment was to determine how much improvement in the initial design could be achieved by each of the five teams. The measure used to address this question was as follows:

$$I \times 100 = \sum_{j} \frac{PM_{jf} - PM_{ji}}{PM_{ji}}$$

where

I = percentage of improvement obtained over initial design,
PM_{jf} = value of performance measure j in the final design f, and
PM_{ji} = value of performance measure j in the initial design i.

The average value of I for each team is given below.

Team	I (percent)
A	38
B	100
C	63
D	1
E	28

The average for all teams for all performance measures was 46 percent. The largest overall improvement was obtained by team B (100 percent), followed by team C with 63 percent. Teams A and E made substantial improvements (38 and 28 percent respectively), whereas team D made virtually no progress over its initial solution. The average improvement for all five teams and all 11 performance measures was 46 percent. This figure needs to be interpreted carefully.

It can be argued that the initial solution of a group of novice designers is not a reasonable base for measuring improvement because it was probably either very conserva-

Table 3. Team comparisons of transit system attributes.

	Team				
Attribute	A	B	C	D	E
Number of bus lines	8	7	11	12	8
Number of bus stops	52	54	59	58	42
Number of park-n-ride lots	5	4	4	5	3
Varying bus headways	Yes	Yes	Yes	Yes	Yes
Varying bus types	Yes	Yes	Yes	No	Yes
Varying bus fares	Yes	Yes	Yes	Yes	No
Varying park-n-ride lot sizes	Yes	No	Yes	Yes	Yes
Varying park-n-ride lot fees	Yes	No	Yes	No	Yes

Figure 4. Comparative performance of final designs.

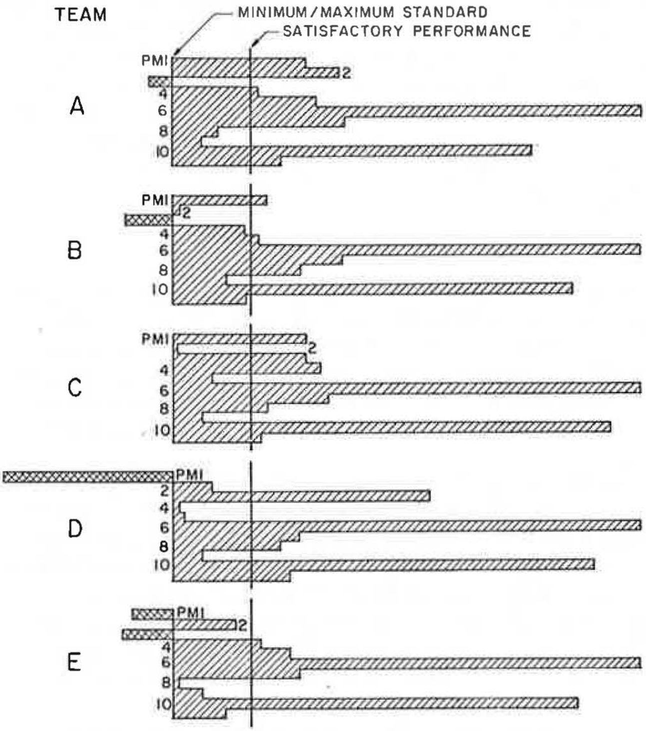


Table 4. Number of times performance objectives satisfied by five design teams.

Performance Measure	Number of Times Satisfied
Percentage of standing room used	5 of 5
Park-n-ride lot load factor	5 of 5
Profit-loss of bus operations	3 of 5
Park-n-ride patronage share	3 of 5
Bus load ratio, seated patrons	3 of 5
Walk-n-ride access index	3 of 5
Percentage of patrons within 5-min walk	3 of 5
Profit-loss of park-n-ride lots	1 of 5
Walk-n-ride patronage share	1 of 5
Park-n-ride access index	0 of 5

Table 5. Correlation coefficients for final designs.

	Team				
Team	A	B	C	D	E
A	1.00	0.03	0.34	-0.68	-0.47
B		1.00	0.86	0.20	0.80
C			1.00	-0.30	0.41
D				1.00	0.74
E					1.00

tive or a wild guess. Although there is some truth in this position, all of the teams worked out their initial design on paper and gave it considerable thought. Still, if we assume that the inexperience of the participants resulted in poor initial designs, we should probably discount the 46 percent overall improvement figure by 10 to 20 percent.

On the other hand, several of the teams were confident that they could have improved their final design substantially if they had had more time to work on it. That two teams were able to achieve improvement proportions of 100 and 68 percent proves that this may be possible. Therefore, it seems likely that, given more time and more experienced designers, improvements in the 60 to 80 percent range would not be difficult to achieve. If this is true, then a general expectation of a 50 percent improvement in design performance seems to be justified in the context of the experiment being reported here. This amount of improvement is large enough to warrant further investments of time and resources in experiments of this type. If similar results are obtained, it will be possible to build a substantial case for the use of interactive graphic design tools in transportation systems design studies.

One further qualification of these results is needed. Each team spent approximately 8 hours working with the UTRAN computer system. Moreover, each team probably spent another 6 to 8 hours planning and discussing the design effort outside the computer room. In addition, some 15 hours of classroom time was used to prepare the participants for the design task. Because it will eventually be necessary to relate improvement obtained with time expended, future studies of this topic should be designed to incorporate this factor. It has not been possible to do so effectively in this study because so much time was spent in learning.

Problem-Solving Strategies

Each team devised a strategy that best matched its perception of the problem. No rigorous analysis of these strategies is attempted here, for they proved to be quite diverse and difficult to compare. Each team generated and evaluated seven or eight design solutions. All but one of the teams were able to improve their designs in successive iterations so that their best designs were their final effort. The most unsuccessful team (team D) generated eight designs but found that the second design was better than the final (eighth) design.

The two most successful teams (as measured by performance improvement achieved) both used an incremental approach to the design problem. They started with a fairly simple design and made those additions to it that they felt were most needed as a result of their evaluation of the previous design. On the other hand, the least successful team began with a very complex design and apparently became quite confused in the process of trying to improve the performance measures. This frustration with the problem led to a fairly negative attitude toward the experiment on the part of some members of this team, and the result (no improvement) was partly a function of this difficulty.

Generally speaking, the results were quite satisfactory when viewed in terms of the human element in the system. Our results tend to confirm those of Sackman (6), who found that access to interactive graphic systems tends to expand or accentuate individual differences. This simply means that a person who has some innate or learned capacity for design work will tend to be aided proportionally more by the computer system than will someone who has less of an innate or learned design ability. In other words, the availability of good interactive graphic design tools may be expected to increase the gap between good and poor designers while the whole spectrum of design quality will be shifted upward.

Much more research needs to be done on the strengths and limitations of the human mind in design situations like the one in this paper. Some work has been done along these lines (6, 9, 10), but much more work remains to be accomplished before we can expect to properly design man-computer systems and to adequately train people to make effective use of them.

CONCLUSIONS

1. The use of the UTRAN interactive graphic system enabled a group of novice designers to make design improvements estimated at close to 50 percent over the initial or "first-cut" solutions. This figure was greater than expectations and substantial enough to warrant further investigations of this type.
2. Four of five design teams were able to make steady progress toward improved performance from their initial design through seven to eight trials to their best design.
3. The designs of teams that began with a simple design and modified it incrementally performed better than other cases. However, the strategies employed by the various teams were quite diverse and are difficult to generalize. A more structured approach to the analysis of design strategies is needed.
4. The designs produced by the five teams were quite similar in appearance and highly similar in the physical elements of which they were composed, but design performance was different.
5. Difficulties with the definition of performance measures, interdependency among performance measures, computer malfunctions, and problem complexity all hindered the ability of the design teams to operate effectively.
6. Replications of the results obtained here are needed before a clear-cut case can be formulated to support the development and use of interactive graphic systems for aiding the transportation systems designer. These results should be regarded as preliminary and indicative of the need for more investigations of this type.

REFERENCES

1. Gehner, C. D., and Clark, J. W. The Urban-Region Transit Analysis System (UTRANS), Volume I: Documentation, Vol. II: User's Manual. Urban Transportation Program, University of Washington, Seattle, Res. Rept. 72-2, 1973.
2. Rapp, M. H. Man-Machine Interactive Transit System Planning. Socio-Economic Planning Sciences, Vol. 6, 1972, pp. 95-123.
3. Rapp, M. H. The Interactive Graphic Transit Simulator: A Tool for Planning Node-Oriented Transit Systems. Urban Transportation Program, Univ. of Washington, Res. Rept. 7, 1971.
4. Rapp, M. H. Planning Demand-Adaptive Urban Public Transportation Systems: The Man-Computer Interactive Graphic Approach. Urban Transportation Program, Univ. of Washington, Res. Rept. 71-4, 1972.
5. Rassam, P. R., Ellis, R. H., and Bennett, J. C. The n-Dimensional Logit Model: Development and Application. Highway Research Record 369, 1971, pp. 135-147.
6. Sackman, H., and Citrenbaum, R. L. Online Planning: Towards Creative Problem-Solving. Prentice-Hall, 1972.
7. Sackman, H. Man-Computer Problem-Solving: Experimental Evaluation of Time-Sharing and Batch Processing. Auerbach Publishers, 1970.
8. Schneider, J. B. Man-Computer Synergistics: An Aid to the Design of Urban Systems. Urban Transportation Program, Univ. of Washington, 16-mm movie, 1972.
9. Schneider, J. B. Solving Urban Location Problems: Human Intuition Versus the Computer. Jour. American Institute of Planners, Vol. 37, No. 2, March 1971, pp. 95-99.
10. Schneider, J. B., Symons, J. G., Jr., and Goldman, M. Planning Transportation Terminal Systems in Urban Regions: A Man-Machine Interactive Problem-Solving Approach. Transportation Research, Vol. 6, pp. 257-273.