Monograph:
May, A.D., Bonsall, P.W., Carsten, O.M.J. et al. (1 more author) (1988) The Evaluation of Route Guidance Systems. Working Paper. Institute of Transport Studies, University of Leeds , Leeds, UK.

Working Paper 266

Published paper

May, A.D., Bonsall, P.W., Carsten, O.M.J. van Vuren, T. (1988) The Evaluation of Route Guidance Systems. Institute of Transport Studies, University of Leeds. Working Paper 266

Working Paper 266

May 1988

THE EVALUATION OF ROUTE GUIDANCE SYSTEMS

A D May

P W Bonsall

with additional contributions from

O M J Carsten

T van Vuren

This work was sponsored by the Transport and Road Research Laboratory.

CONTENTS PAGE NO.

CONTENTS PAGE NO.

1 BACKGROUND

1.1 Objectives

1.1.1 We were commissioned by the Transport and Road Research Laboratory to:

"collaborate with the German government and their representatives who are responsible for conducting the LISB trial in Berlin in order to produce an agreed methodology, which is acceptable in both Germany and the UK, for assessing the automatic route guidance systems which will be provided in Berlin and London." The brief suggested a number of aspects to be included, and required detailed proposals, timescales and costs for implementation in London.

1.1.2 The background to the brief lies in decisions to introduce pilot automatic route guidance systems in the two cities. The principles of the systems are similar, and have been described in detail elsewhere (Jeffery, 1987). In brief, they involve:

(i) a central computer which retains information on a specified road network, which is updated using real time information from the equipment users;

(ii) infra red beacons at selected junctions which transmit information to equipped vehicles and receive information from those vehicles;

(iii) in-vehicle equipment which includes a dead-reckoning system for position finding, a device for requesting guidance and specifying the destination, a micro-computer which selects the optimal route, and a display which indicates when a turn is required on the main network, and the compass direction and distance to the final destination;

(iv) transmission from the equipped vehicles of origin, requested destination, links used since passing the last beacon and, for each link, the time of entry and departure and time spent delayed.

It is this travel time information which is used to update the central computer's knowledge of the best routes.

1.1.3 At the time of our commission, in January 1988, plans for the Berlin pilot were well advanced. The system was to come into operation in April 1988, with a three month period in which vehicles transmit travel time information without receiving guidance, followed by nine months in which they receive guidance. Naturally, plans for evaluation were largely fixed, and the emphasis in our study has been on obtaining details of these plans, identifying opportunities for comparison with the pilot in London, and suggesting minor modifications to the Berlin plans.

1.1.4 By contrast, the plans in London are in an early stage of development. A small scale demonstration is being implemented in April 1988, and the private sector is being invited to submit proposals for operating a pilot system which, it is envisaged, would be operational during 1990. Proposals for the London pilot were being developed during the course of our study, and we saw our role as suggesting the types of evaluation procedure which should be specified in the final guidelines issued to potential operators.

1.1.5 The objectives for the pilot as specified in a draft of the guidelines were:

"(1) to provide a large scale demonstration to potential backers

(2) to determine the financial and technical viability of a full system.

In addition .... to determine driver reaction to the various aspects and uses of the system and to provide an appreciation of the problems which should be experienced in operating and maintaining the system."

1.1.6 In discussion with the Department of Transport, it was made clear that the objectives of an evaluation should be to provide sufficient information on the technical, ergonomic and operational performance of different levels of route guidance to enable decisions to be taken on the advisability of extending the pilot into a full scheme, of introducing similar schemes elsewhere, and on the most appropriate specification for such schemes.

1.2 Study Approach

1.2.1 In meeting these objectives, we started by setting out our own ideas of the elements which should ideally be included in a comprehensive evaluation. These were then discussed in detail with those involved both in Germany and the UK to identify omissions and priorities, and to highlight those aspects of the plans for Berlin and London which might hinder the implementation of these proposals.

1.2.2 As part of this process, discussions were held with Siemens, SNV and the Technical University of Berlin in Germany, and with the DTp, TRRL, Plessey, the AA and Westminster City Council in the UK. In addition, views were sought in writing from BASt, BMFT, GEC and the RAC. We acknowledge with thanks the helpful comments received.

1.2.3 Based on these comments we identified a number of specific technical requirements of an evaluation, and a series of technical problems which needed to be overcome. The majority of the work programme focused on these problems and on the timetabling and costing of the resulting proposals. It was appreciated at the outset that it would not be possible in the short time available to resolve all the outstanding technical issues. With the agreement of the client, we have therefore concentrated on the main technical issues, and identified those on which further work will be required.

1.2.4 At the same time, it became clear that there were a series of policy issues which needed to be resolved, particularly arising from the decision that the London pilot should be mounted by the private sector, and the resulting uncertainty over the role of a public sector evaluation. We have spent some time in discussing these issues with those concerned, and making recommendations.

1.3 Outline of the Report

1.3.1 The report is in two parts. The first, which includes Chapters 2 to 5, covers the policy-related issues, and recommends the issues to be included in a technical evaluation. The second, including Chapters 6 to 13, covers our technical proposals. Our recommendations are summarised in Chapter 14.

1.3.2 Chapter 2 outlines our proposals for the elements which should ideally be included in an evaluation, and the reasons for their inclusion. Chapters 3 and 4 then consider in turn the implications for this specification of the plans for the pilots in Berlin and London. Chapter 5 summarises our conclusions on the content of the evaluation methodology for London, and highlights the policy decisions required.

1.3.3 Chapter 6 identifies the possible information sources for each indicator, and chapter 7 considers the specific information requirements for the prediction of the effects of a full scale system. Chapter 8 develops an experimental design which provides the necessary information from the most appropriate sources.

1.3.4 Chapters 9 to 12 then consider in detail each of the main information sources. Chapter 9 deals with the use of test vehicles. Chapter 10 discusses the information required from participants and control groups. Chapter 11 considers information from the in-vehicle units in the equipped vehicles. Chapter 12 outlines the other sources of data which can be used.

1.3.5 Chapter 13 considers priorities among these technical proposals, particularly in the light of the desire to maintain comparability with Berlin. It provides initial cost estimates for our higher priority proposals.

2 EVALUATION REQUIREMENTS

2.1 The Identification of Requirements

2.1.1 As noted in Chapter 1, we have interpreted our brief as being to identify the evaluation requirements necessary to meet the objectives specified for the London pilot, and to recommend those requirements which should be specified in the final version of the guidelines issued to potential operators.

2.1.2 We started by preparing our own list of possible elements of such an evaluation, and discussed these ideas with those concerned, both in Germany and the UK. This chapter presents the resulting list of ideal requirements of an evaluation. Inevitably decisions already taken impose limitations on the achievement of these ideal requirements. These limitations are identified for Berlin in Chapter 3 and for London in Chapter 4.

2.1.3 The objectives specified for the London pilot are quoted in para 1.1.5. The first is "to provide a large scale demonstration to potential backers". This does not of itself make any particular requirements on evaluation, except to the extent that potential backers will need information on the costs of operation, on the performance and reliability of the system, and on the likely size of the market.

2.1.4 The second objective is "to determine the financial and technical viability of a full system". Although not specifically stated, it is understood that the viability should be considered for full scale systems both in London and elsehwere. Financial viability is largely covered above in terms of costs of operation and likely revenue. Technical viability can be viewed from the point of view of the operator, the user and government. It includes both the performance of the individual elements of the system, and the impact of route guidance on transport system performance.

2.1.5 In addition the guidelines refer to determining "driver reaction to the various aspects and uses of the system" and to providing "an appreciation of the problems which should be experienced in operating and maintaining the system". Driver reactions will be fundamental to the size of the market for a full scale system, and the way in which they use the system will help determine its impact on the transport system as a whole. Problems for operation and maintenance will be of concern partly to the operator and to those contributing information, and partly to government.

2.1.6 It is clear that this set of requirements for evaluation is closely interlinked, and includes issues of concern to the operator, to the user, to government and to third parties involved in operation. Operators are likely to be concerned with technical performance, costs of operation, and levels of demand and hence revenue. Levels of demand will depend on user response, which will be affected by ergonomic performance and users' perceptions of the benefits from the system and its safety. These perceptions will depend on, but not necessarily be identical to the true benefits and safety implications. Government will be concerned with the true benefits of the system to individual users, with the safety implications, and with the network effects, whether in terms of changes in congestion or environmental intrusion. Third parties will be concerned with the costs to them of contributing to the system, and also with any benefits which they might obtain from it.

2.1.7 All of the above issues can be identified in the pilot, and be used to assess the effects of the pilot in its own right. It is of much greater importance to operators, government and users, however, to be able to judge the performance of a full scale system in London or elsewhere, and it is essential that the pilot provide enough information to enable reliable judgments to be made. This makes particularly significant demands on the design of the evaluation.

2.2 Specific Evaluation Requirements for the Pilot

2.2.1 The review of objectives in section 2.1 identified eleven issues which ought ideally to be considered in evaluating a pilot scheme. these are considered in more detail in this section.

2.2.2 Technical Performance Each of the elements of the system, as described in para 1.1.2, needs to be tested for long enough, and in sufficient different road and vehicle environments, to provide an assessment of its reliability. In addition, it will be necessary to assess the accuracy of the network data base, and the frequency and accuracy with which travel time data is updated. It will also be necessary to confirm that there are no unforseen side effects (such as interference with other communications) stemming from the operation of the equipment. Such information will be of importance to potential operators and users, but also to government in deciding whether to promote or license the system.

2.2.3 Ergonomic Performance It will be necessary to assess the ease with which the in-vehicle equipment can be used, both for keying in the destination and in following the guidance given. Such tests will be needed with a wide range of drivers, journeys and vehicles. The information will be of importance to operators and to potential users, and the safety implications, which are covered later, will be of concern to government.

2.2.4 True Benefits and Disbenefits of Guidance The benefits of guidance will need to be assessed in terms of savings in time, operating costs, and other parameters of concern to the user. Information will be needed for different types of user and of journey. Such benefits may arise in route finding on and off the network, in destination finding having left the network and, potentially, in using any special features of the system such as journey preview facilities or parking information. Knowledge of these benefits will be of use to the operator for promotional purposes and to government as part of its assessment of the benefits of a route guidance system.

2.2.5 Perceived Benefits and Disbenefits of Route Guidance The user will in practice respond to his or her perception of the benefits outlined in 2.2.4 above. That perception may differ from reality if knowledge of the true conditions on the alternative route is imperfect, and this is particularly likely to happen on unfamiliar journeys and in volatile traffic conditions. These perceptions will be of importance to users, but also to operators and government, because they will influence demand.

2.2.6 Responses of Users In the light of their perceptions of benefits, users may elect not to continue using the system, not to use it for certain types of journey, or not to follow the guidance given throughout the journey. Some may attempt, but fail, to follow the guidance given. Generally, those provided with guidance may increase their level of journey making.These responses are likely to vary by journey and driver type. Some responses will be of direct concern to the operator; all will concern government, particularly if they lead to reduced benefits or to traffic diverting from the preferred network.

2.2.7 Safety Implications The way in which the in-vehicle equipment is used, the way in which drivers manoeuvre in traffic while following guidance, and the routes that they take, can all have an effect on the safety of the road system. Such issues are of concern to operators, government and users.

2.2.8 Driver Stress and Perceived Safety As with benefits, the driver's perceptions of safety may differ from reality, and these perceived safety effects may affect drivers' willingness to purchase and use the equipment. In addition, drivers' experiences of stress and fatigue may be affected. All these factors will be of concern to operators and, potentially, to government.

2.2.9 Network Effects The routes which drivers take as a result of their level of use of the system may have environmental effects as well as safety ones, particularly if drivers are diverted away from environmentally sensitive routes or elect to use 'rat runs' in preference to guided routes. At higher levels of use, the rerouteing of traffic may lead to changes in the level and distribution of congestion or the performance of other network control strategies, either of which could have effects on non-users. These network effects will be of particular concern to government.

2.2.10 Costs to Operators Costs of implementing, operating and maintaining the system will largely be borne by the operator and be of direct concern to him. Government may also be interested in these costs as part of an assessment of cost-effectiveness relative to alternative technologies, or in order to determine the basis for licensing operation or, potentially, if it wished to sponsor a system itself.

2.2.11 Costs to Third Parties In addition to the operator's costs, third parties may experience costs in the establishment or operation of the system. These could include provision of basic information on the network, access to roadside and communication equipment, and information on roadworks and emergencies. In addition to the direct concern to third parties, government will be interested in these costs in terms of their implications for cost-effectiveness.

2.2.12 Benefits to Third Parties Conversely, third parties may receive benefits from the system. The most obvious example is likely to be information on travel patterns and traffic conditions. Such benefits will obviously be of interest to government, but may also be of concern to the operator, if such information can be marketed either to government or to others in the private sector.

2.3 Effects of a Full Scale System

2.3.1 All of the information in section 2.2 can be obtained from a pilot system in which a limited number of vehicles are equipped, and the pilot itself may be sufficient for resolving some issues such as ergonomic performance. In many ways, however, the effects of a full scale system cannot simply be extrapolated from the pilot.

2.3.2 In particular, a full scale system will provide much more information and hence be able to update the identification of best routes more rapidly and reliably. Conversely, it will lead to the potential for large proportions of the traffic stream to be diverted, thus having potentially dramatic effects on the distribution of congestion in both space and time. This may well in turn influence the level of demand for, and use of route guidance. A full scheme may provide additional benefits by providing a link to traffic signal controls. However, the nature of the market and the patterns of use may be significantly different from the pilot.

2.3.3 It seems clear to us that these effects can only adequately be addressed by simulating the performance of a full scale system. Such a simulation will need to be calibrated on the basis of results from the pilot, and collection of the necessary data must therefore form an important part of the pilot.

2.3.4 In addition to determining whether a full scale system should be implemented, it will be necessary to determine the most appropriate design for that system. As well as the level of use (which will be determined largely by market forces), the parameters which will need to be tested can be divided into three categories:

(a) those which can readily be tested in the pilot, including the density of the network, the density of beacons on that network, the use of external information, the way in which optimal routes are identified and, to a limited extent, the frequency with which they are updated;

(b) those which would only become apparent in a full scale system, and would need to be assessed by simulation, including the procedures for dealing with higher proportions of guided vehicles, the interaction with other traffic control systems, and the performance in other locations;

(c) those which could be tested provided that special facilities were incorporated in the pilot, such as route previews, multiple destination finding, fleet location and parking information.

It would seem desirable to vary as many of the elements in categories (a) and (c) as is practicable to provide experience on the cost-effectiveness of, and driver response to, different types of operation.

3 THE LISB TRIAL IN BERLIN

3.1 Background

3.1.1 The Berlin route guidance pilot scheme is entitled LISB (Leit- und Informationssystem Berlin). It is being jointly funded by the Federal Government, the State of Berlin, and an industrial consortium led by Siemens. The project is well advanced; the roadside equipment has already been installed, and the first equipped vehicles are expected to begin transmitting information from June 1988.

3.1.2 The intention is to obtain information from equipped vehicles over a three month period and then start providing equipped vehicles with guidance from September. Guidance would be provided for at least a six month period, during which the project would be being carefully monitored.

3.1.3 Fuller details of the scheme are given elsewhere (von Tomkewitsch, 1987; May, 1987). What follows is a brief description of the attributes of particular concern to evaluation.

3.1.4 The equipped fleet will consist of 280 private cars, 205 business cars, 100 taxis and 40 hire cars. In addition it is hoped to sell perhaps 75 sets of equipment to firms whose vehicles would enhance the data provided to the system. This would provide a total of 700 equipped vehicles. An additional 200 commercial vehicles may be equipped under a related programme called TRANSLISB, which is designed to incorporate route guidance and drop scheduling for commercial vehicles.

3.1.5 The private and business vehicles have been selected following response to an advertisement in the press and a screening questionnaire. Selection is on the basis of vehicle type, driver characteristics, predominant journey orientation and familiarity with the network. Some business vehicles have been selected by approaching firms based in particular areas of the city, so that some concentrations of destinations are obtained. The taxis have been included not to provide them with guidance but because they are intensive users of the network, and hence will provide substantial quantities of travel time data. The hire cars have been equipped partly to test the impact on occasional users and partly to test destination finding in a city where most local drivers know the network. Around 10 test vehicles are also being equipped for research purposes.

3.1.6 There is considerable concern over privacy in West Germany, and in a full system drivers will be able to elect not to transmit information to the central computer. In LISB, this option is being denied in return for involvement in the pilot. Moreover, it is intended to use codes input by the driver to identify which driver is using the vehicle.

3.1.7 The equipped network includes some 23 km of motorways and around 660 km of main roads and major 'rat runs'. There is also a secondary network of some 120 km to which vehicles will not be guided, but along which they can be tracked if necessary. Some 220 junctions have been equipped and beacons have been installed at 10 sites on motorways. The shortest inter-beacon spacing is around 300 m; several in the suburbs exceed 5 km.

3.1.8 Optimal routes will be based on travel time minimisation only, although commercial vehicles, if guided, will be restricted to routes which are physically capable of accommodating them. Travel time profiles will be established for each link during the initial three months of operation without guidance. These will be updated on the passage of equipped vehicles, to provide guidance in real time. However, it is accepted that the minimum headway between equipped vehicles is likely to be 5 minutes, implying that real time updating will not be very precise.

3.1.9 It appears that no special features, such as journey preview and parking information are to be provided. There is, however, some interest in providing a link to the traffic signal control system.

3.1.10 Current plans for the evaluation database include: three questionnaires to equipped drivers; interviews of a sample of drivers to identify reasons for not using guidance; brief questionnaires for hire car users; and a questionnaire to a control group of between 1000 and 2000 unequipped drivers who are members of the German motoring organisation, ADAC. This questionnaire, and the first questionnaire for equipped drivers seeking information on attitudes to technology, were distributed in April 1988. It is expected that the second and third questionnaires to equipped drivers will be distributed in the autumn of 1988 and the spring of 1989, and will seek comments on experience with LISB. Some fuller interviews will be conducted at the end of the pilot. Informal group discussions will also be held and comments sought through a newsletter. There are no plans in the LISB pilot to relate any questions (other than for hire car users) to specific journeys or to require drivers to keep manual logs. However, some such records are planned in the context of the Institute for Transport Studies', SERC sponsored, study.

3.1.11 It was clear from our discussions that evaluation in Berlin will place considerable emphasis on learning by experience, and this may restrict the opportunity to conduct a fully controlled scientific experiment.

3.2 Implications for Evaluation Requirements

3.2.1 This section reviews the plans for Berlin against the requirements outlined in Section 2.2.

3.2.2 Technical Performance The technical performance of the system will be monitored throughout the pilot, using the system log and logs of vehicle movements. No automated records records are being kept of the performance of in-vehicle equipment, but drivers' complaints about malfunctions will be logged. The range of types of vehicle, driver and route is probably sufficient to test the equipment in a wide range of conditions.

3.2.3 Ergonomic Performance Detailed assessment of this is to be obtained from driver questionnaires, psychometric studies and interviews.

3.2.4 True Benefits of Route Guidance The intention is to compare, for a sample of regular journeys, the routes taken before receiving guidance with those under guidance. Times taken on the route no longer used will be compared with those on the guided route to identify benefits and disbenefits. Time on the unguided route will be identified by recording the time taken on the individual links at the appropriate time by other equipped vehicles or, failing this, by using the system estimate of link travel time. It is accepted that this method is somewhat approximate and may need to be checked using test vehicles. It will anyway be inappropriate for irregular journeys, and will overlook seasonal differences in routes taken, if they exist. Guidance is only being given on the basis of time minimisation and benefits to those wishing to optimise against other criteria will be difficult to assess. Destination finding benefits are likely to be small, but they may be assessed by comparing those who know their destination with those who do not, and by mounting a small experiment in destination finding. Careful selection of the equipped sample should permit comparison of benefits by user and journey type.

3.2.5 Perceived Benefits of Guidance Questionnaires to drivers will seek their assessment of the benefits they have received from LISB for different types of journey. No direct comparison of perceptions and reality for individual routes is planned. This is, however, an aspect which ITS intend to study in an SERC funded research project in Berlin. At the end of the project, drivers will be asked whether they would purchase the system, and at what price.

3.2.6 Responses of Users It will be technically possible to monitor all the journeys made by equipped vehicles, whether they request guidance or not and whether they follow it partially or fully. It will also be technically possible to record the identity of the driver concerned. It will not be possible to record the route taken off the guided and secondary networks or, particularly, the extent to which autonomous mode guidance is followed. Questionnaires will seek reasons for not seeking or following advice and the responses will be interpreted in the light of the participants' characteristics. However, drivers will be encouraged to follow all guidance given for the first few months of the experiment.

3.2.7 Safety Implications The Germans appreciate the importance of these and Siemens expressed interest in suggestions that we made for using information directly from the in-vehicle equipment. We did detect, however, a feeling by Siemens that it might not always be desirable to publish data on the safety record of technological devices. Their main concern was over the potential for using information on safety effects of the in-vehicle equipment in litigation. There is a proposal for an observational study of drivers in a test vehicle, although there appear to be some misgivings over the research method proposed.

3.2.8 Driver Stress and Perceived Safety Detailed questions on these issues will be included in the driver questionnaires.

3.2.9 Network Effects It will be possible to monitor the extent to which drivers use the guidance network, and they can also be followed on the secondary network. More detailed information on roads used off the networks for which vector chains have been built will not be available. The level of congestion and the number of equipped vehicles are such that no interactive effects of rerouteing on other traffic can be expected. Some aggregate measures of network usage are to be retained as part of the evaluation.

3.2.10 Costs to Operators An estimate will be provided to the State of Berlin on the costs of operation.

3.2.11 Costs to Third Parties In general terms, the work involved in adjusting the network as a result of roadworks, and by the Police in providing information on accidents and incidents will be monitored.

3.2.12 Benefits to Third Parties The State of Berlin is interested in the use of data for traffic signal control and an experiment is planned in the later stages of the pilot. Otherwise there are no plans for measuring such benefits.

3.3 Effects of a Full Scale System

3.3.1 It is the intention that the Berlin pilot will be a forerunner of a full scale system throughout West Germany, involving both interurban and urban roads. It is probable that a full scale system will be implemented if the benefits of the pilot exceed its costs. There are no immediate plans to estimate the performance of a full scale system, but one sub-contractor is estimating the effects of different levels of use.

3.3.2 In the full scale system, drivers will be free not to transmit information if they wish. The effects of this on the amount and representativeness of the information provided is an interesting issue for study, but is not being addressed in the pilot. However, it will be possible to check whether pilot drivers elect to shield their transmitters.

3.3.3 There are no plans to vary any of the operational parameters either over time or space or between drivers during the pilot. It is considered difficult enough to conduct all the analysis necessary in the nine month period without varying the parameters.

3.3.4 Since no special features are to be provided in the LISB pilot, their evaluation does not arise.

3.4 Conclusions

3.4.1 There appears to be general agreement on the content of the evaluation programme for the LISB experiment in Berlin, perhaps because the experiment is being monitored collaboratively by government and industry. The only concern expressed during our discussions was on the merits of a public analysis of the safety implications of the in-vehicle equipment, given the potential legal implications.

3.4.2 The main constraints on comparability with London arise from the nature of the Berlin network and the advanced state of the evaluation plans. The Berlin network is somewhat less congested, and will therefore provide rather different opportunities for real-time variations in the guidance given. More importantly, most users do most of their driving in Berlin and are therefore unlikely to benefit greatly from destination-finding advice.

3.4.3 Most decisions have already been taken on the evaluation programme for Berlin. In particular, the timetable is set, with a three month period without guidance followed by six months with guidance; the pilot sample has already been selected, and initial questionnaires administered; and it has been decided not to vary the nature of the time-minimising guidance being given. To the extent that different procedures are designed for London, comparisons will be more difficult.

4 PLANS FOR THE AUTOGUIDE PILOT IN LONDON

4.1 Background

4.1.1 Plans for the London pilot are much less fully developed. Apart from the decision to use broadly similar technology to that in Berlin, the one definite decision is that it should be mounted by the private sector, with government only concerned in the selection and licensing of the operator and in deciding whether, in the light of the pilot, to proceed to a full scale system. This distinction between Berlin and London is highly significant, and raises several important policy implications for the design of an evaluation procedure and for comparison with Berlin. We return to these in Section 4.4.

4.1.2 The currently envisaged timetable for London is that the demonstration project, with five beacons and 12 equipped vehicles, will be launched during April 1988. Draft guidelines for intending pilot operators will be issued at the same time for consultation, with final guidelines being issued in July. It is intended that the successful operator should have been selected by the autumn, and that pilot implementation should start in January 1989. Opinions vary on the time that will be required for implementation, but it is generally accepted that it will not be possible to begin operation until mid 1990. The duration of the pilot is also uncertain. Some have suggested a period as short as three months, and it is clear that it will be in commercial operators' interest to keep the period as short as possible. The choice of duration for the pilot should be determined to a large extent by the requirements for evaluation. We return to this issue later.

4.1.3 Fuller details will be given in the draft guidelines (DTp, 1988). What follows is a brief description of the system as planned, gleaned from discussions with DTp and potential operators.

4.1.4 Current plans envisage the equipping of some 1000 vehicles, of which 400 would be financed by the pilot and available for evaluation, and the others would be privately financed. The guidelines envisage an emphasis on commercial vehicles and 'working vehicles', but the final choice should be determined primarily by evaluation requirements.

4.1.5 Some considerable concern has been expressed over privacy issues, and there is reluctance to give any suggestion to users that they can be tracked. The draft guidelines envisage two classes of vehicle; Class A vehicles will not be able to be tracked while Class B will (primarily for fleet location and scheduling purposes). All vehicles would have Private Access Numbers (PANs) to protect privacy and restrict access. No consideration appears yet to have been given to the value of identifying vehicles' (and, indeed drivers') routes as is planned in Berlin. We have developed proposals in Chapter 11 which may, subject to technical feasibility, overcome this problem.

4.1.6 The current intention is that the network will consist of all A and B class roads within and immediately outside the M25. The draft guidelines invite potential operators to specify the basis by which they will select the network and consult with highway authorities on that selection. There is considerable concern among London Boroughs about the implications of including minor residential roads. There appears to be no intention at present to include any secondary network, as in Berlin, for tracking vehicles which leave the network.

4.1.7 The draft guidelines envisage dome 300 equipped junctions, but with beacon sites concentrated on three orbital routes, which would probably be the M25, the North and South Circulars, and the Inner Ring Road, and more intensively within a 'wedge' between the M40/A40 and the M3/A316.

4.1.8 Guidance criteria are expected to include time minimisation, cost minimisation, and an impedance factor applied on lower quality roads. It appears likely that drivers will be able to select the optimising criterion that they require, but that the impedance factors will be imposed as a matter of public policy.

4.1.9 Several special facilities for drivers have been considered, such as route previewing, multiple destination finding and information on parking availability. It will be for operators to decide which to offer, but boroughs are very interested in the potential for providing parking information.

4.2 Specific Evaluation Requirements

4.2.1 It is clear that the plans for the London pilot are sufficiently uncertain to retain considerable flexibility in the satisfaction of the evaluation requirements outlined in Section 2.2.

4.2.2 Technical Performance It should be perfectly feasible to record this both in a system log and by monitoring in-vehicle equipment.

4.2.3 Ergonomic Performance It has been suggested that this will be adequately monitored in the demonstration project. We have doubts whether this will provide a sufficient range of drivers, or enable comparisons between experienced and inexperienced users. It should be possible to test these in the pilot, provided that pilot drivers are committed to provide the necessary information.

4.2.4 True Benefits and Disbenefits of Guidance Any restrictions on the identification of drivers' routes and of the driver involved will limit the ability of the pilot to estimate these benefits. We return to these issues in Part II of this report.

4.2.5 Perceived Benefits and Disbenefits of Guidance Appropriately designed questionnaires should be able to identify the perception of benefits in broad terms, but it should ideally be possible to relate these assessments to individually identified journeys and sets of journeys. Again, any restriction in the interests of privacy will make this more difficult.

4.2.6 Responses of Users These require the identification of journeys where guidance is not requested, and where guidance is not completely followed. While the technology should permit this, it will only be feasible in aggregate terms unless the restrictions resulting from concern over privacy can be overcome.

4.2.7 Safety Implications It should be possible to determine the number of guided vehicles taking different routes on the network, and it may also be feasible to identify specific manoeuvres which involve sudden braking. However, if there is no secondary network, it may not be possible to track the use of those roads off the network which may be inherently less safe. Moreover, concern to protect privacy may limit the ability to relate specific events to individual vehicles.

4.2.8 Driver Stress and Perceived Safety It should be feasible to cover this in appropriately designed questionnaires. The Draft Guidelines refer specifically to the need to assess effects on driver stress and fatigue.

4.2.9 Network Effects As noted above, it will be possible to monitor the extent to which guided vehicles use different links on the network, but not, apparently, those links off the network which are likely to be more environmentally sensitive. The Draft Guidelines introduce the concept that there should be no net increase in traffic levels off the strategic network in a local authority area. This may be difficult to demonstrate. Even though levels of congestion are much higher than in Berlin, it seems clear that the number of equipped vehicles will be far too small to induce any redistribution of congestion.

4.2.10 Costs There is no technical reason why these should not be recorded.

4.2.11 Costs to Third Parties Again, there is no reason why these should not be recorded. Indeed, our one contact with local authorities indicated that they, and probably the police, will be required to charge directly for any advice and assistance provided.

4.2.12 Benefits to Third Parties Our contacts to date suggect that little thought has been given to these, although the draft guidelines refer to the provision of flow, census and journey time data. It seems probable that only the last of these will be provided to a level of detail which could be of value; but information on location of congestion and conflicts may also be worth obtaining. It should be possible to present examples of the data and seek reactions as to its value.

4.3 Effects of a Full Scale System

4.3.1 It is clearly the intention that the London pilot should be used as a basis for deciding whether to extend to a full scale scheme in London, and whether to extend provision to other parts of the country. The government will need to make such decisions for licensing purposes, but it will be for commercial operators to decide whether such extensions are likely to be cost-effective. The process for estimating the effects of a full scale system is complex, and we have given considerable thought to it. Our ideas are set out in Chapter 7.

4.3.2 One particularly valuable way in which the relative cost-effectiveness of different extents and types of system can be assessed is to vary operating parameters during the pilot. The Draft Guidelines refer to the possibility of simulating more intensive use within the 'wedge'. We outline a number of possibilities in Chapter 8.

4.3.3 It is not yet clear which additional features will be offered to drivers although, as noted in Para 4.1.9, there is some interest in offering parking information. It should be possible to estimate the effects of this particular feature, although it may well be that such a facility in a full scale system would need to be more sensitive to the dangers of feedback.

4.4 The Implications of Commercial Operation

4.4.1 As noted in Para 4.1.1, the decision that the route guidance pilot should be operated commercially in London raises several important policy issues. It is estimated that it will cost the successful proposer between #5m and #10m to implement the pilot. He will inevitably wish to reduce the risk on that capital to a minimum, and to obtain a return on the investment as rapidly as possible. This affects the design of an evaluation package in three important ways.

4.4.2 The first implication is that the operator will wish to be as certain as possible that there will be a direct transition from pilot to full scale system. This implies a need to understand the basis on which government intends to take a decision, and to be reasonably certain in advance that the system will satisfy the criteria for success. In our discussions with potential operators, the case was made for the Department of Transport specifying in advance the criteria against which the pilot would be judged, and the thresholds for acceptability. While there clearly must be some room for taking into consideration issues which are not apparent at the start of the pilot, we have some sympathy with the view expressed. We recommend that the Department should give consideration to the criteria to be used as part of the overall evaluation, and should specify them in the final guidelines. Our proposals in Part II should be sufficient to meet any such requirements.

4.4.3 The second issue is that of cost of the evaluation. To the extent that the evaluation imposes costs on the operator, he will understandably wish to keep these to a minimum. We consider that the costs of the evaluation package should in any case be carefully justified, and we have given some consideration to evaluation priorities and costs in Part II of this report.

4.4.4 The third consideration is that of the time taken before a decision is taken to implement a full scale scheme. Clearly the operator will wish to keep this to a minimum. There is a real danger that pressure to limit the duration of the pilot will make it impossible to benefit adequately from the experience which the pilot provides. We have given careful consideration to the appropriate timetable for the pilot in Chapter 8.

4.4.5 In addition to these financial interests, the operator will be concerned to protect information which is commercially of value to competitors. Such competition could arise from others wishing to operate a similar system, but is probably more serious from those offering competing systems, whose market position could be expected to benefit from a publicly conducted critical examination of the system.

4.4.6 In our discussions with potential operators, this issue was particularly strongly stressed. Indeed, it was suggested to us that in a commercial operation the information which government required could be limited to ensuring that a system which was commercially viable did not impose any unreasonable disbenefits on users or others. Our own view is that government does have other, more positive, interests as well, in achieving optimum benefits for society, and in providing advice on the relative role, and relative cost-effectiveness, of route guidance as compared with other solutions to urban traffic problems.

4.4.7 Our discussions led us to the view that the information requirements outlined in Section 2.2 could be categorised in three ways:

(a) of public concern, and able to be made freely available;

(b) of public concern, but requiring protection of commercial interests;

(c) of no obvious public concern, and commercially confidential.

4.4.8 The potential operators to whom we spoke suggested an allocation as follows:

(a) safety implications; network effects; benefits to third parties;

(b) technical performance; real benefits and disbenefits; user reactions; perceived safety; costs to third parties;

(c) ergonomic performance; perceived benefits and disbenefits; costs to the operator.

4.4.9 Conversely, discussions with the local authority representative suggested that it would be wrong for any of this information to be withheld from the public, and discussions in the Department of Transport indicated that at the very least the decision should be made by the Secretary of State in discussion with the operator.

4.4.10 We consider that the three way categorisation is a useful one, but that there is a public interest in all of those items which the potential operators wished to place in category (c). In particular, ergonomic performance is closely allied to safety issues, and both it and perceived benefits are likely to influence demand and usage, and hence the benefits obtained by society. Even costs may be relevant if at some stage, perhaps because route guidance proved not to be commercially viable, government wished to sponsor a scheme itself.

4.4.11 The final issue raised under this heading was the question of who should be responsible for the evaluation. It appears to us essential that the evaluation should be conducted in the public sector, probably under the auspices of the Transport and Road Research Laboratory, to avoid any suggestion that it was not being conducted wholly objectively.

5 POLICY CONCLUSIONS

5.1 The objectives of an evaluation should be to provide sufficient information on the technical, ergonomic and operational performance of different levels of route guidance to enable decisions to be taken on the advisability of extending the pilot into a full scheme, of introducing similar schemes elsewhere, and on the most appropriate specification for such schemes (Para. 1.1.6).

5.2 To be comprehensive, an evaluation of a pilot route guidance system requires information on

(i) the technical performance of the system;

(ii) the ergonomic performance of the in-vehicle equipment;

(iii) the true benefits and disbenefits of the guidance given to different categories of user and for different types and elements, of journeys;

(iv) the perceived benefits and disbenefits of such guidance;

(v) the responses of users, in terms of the extent to which guidance is sought and followed;

(vi) the implications for safety of the use of the equipment, and the resulting manoeuvres and routes;

(vii) driver stress and perceptions of safety;

(viii) the impacts, if any, on network performance and travel conditions for non-users and environmental conditions;

(ix) the costs of setting up and operating the system;

(x) the costs to third parties of providing information for the system;

(xi) the benefits to the public sector of data provided by the system. (Section 2.2)

5.3 At the same time it will be necessary to use the pilot scheme to predict the effects of a full scale route guidance system. A full scale system is likely to perform very differently from the pilot, both because more information will be available, and because changes in route will be on a sufficient scale to affect network performance. These effects cannot be monitored directly from the pilot, but will need to be estimated from simulation, which can be calibrated on pilot data. (Section 2.3).

5.4 The pilot should also be used to test the effects of modifying those design parameters which can be varied, and to simulate the effects of varying others. Where operators elect to provide additional facilities, the evaluation should assess the effects of these as well. (Section 2.3).

5.5 There appears to be general agreement on the content of the evaluation programme for the LISB experiment in Berlin, perhaps because the experiment is being monitored collaboratively by government and industry. It has been decided to cover all the issues listed in 5.2 above although some of them will not be intensively treated. The only concern expressed during our discussions was on the merits of a public analysis of the safety implications of the in-vehicle equipment, given the potential legal implications. (Section 3.2).

5.6 The main constraints on comparability with London arise from the nature of the Berlin network and the advanced state of the evaluation plans. The Berlin network is somewhat less congested, and will therefore provide rather different opportunities for real-time variations in the guidance given. More importantly, most users do most of their driving in Berlin and are therefore unlikely to benefit greatly from destination-finding advice. (Section 3.4).

5.7 Most decisions have already been taken on the evaluation programme for Berlin. In particular, the timetable is set, with a three month period without guidance followed by six months with guidance; the pilot sample has already been selected, and initial questionnaires administered; and it has been decided not to vary the nature of the time-minimising guidance being given. To the extent that different procedures are designed for London, comparability will be more difficult. (Section 3.4).

5.8 The main constraint on evaluation proposals for London is the commercial nature of the pilot. Not surprisingly, would-be commercial operators would wish much of the information obtained to be treated as commercially confidential, and would resist the idea of the public sector acquiring certain items of information. Discussions with potential operators led to the identification of three possible categories of information:-

(a) of public concern and able to be made freely available;

(b) of public concern, but requiring protection of commercial interests;

(c) of no obvious public concern, and commercially confidential. (Para. 4.4.7).

5.9 In the light of our discussions, an appropriate categorisation of the items listed in para. 5.2 above would appear to be as follows:-

(a): (vi), (viii), (xi).

(b): (i), (ii), (iii), (iv), (v), (vii), (ix), (x).

(c): None.

However, we accept that the final decision must be for the Secretary of State in consultation with the operator. It should be noted, in this context, that operators may well wish to treat items (ii), (iv) and (vi) as being in category (c), since they are likely directly to affect market size. It will be necessary to convince them of the importance of these items in obtaining an assessment of the likely effects of a full scale system. (Paras. 4.4.8 - 10).

5.10 Commercial considerations may well also argue for a short-lived, low cost evaluation programme. We have not received any specific proposals from potential operators on timescale or cost, but it is clear that any requirements on them will need to be carefully justified. We have borne these considerations in mind in Part II. (Paras. 4.4.3 - 4).

5.11 One other policy consideration which has arisen in our discussions is the question of privacy, and the implications for market size of any suggestion that the system could track individual vehicles. In Berlin this is being dealt with by making it a condition of participation in the pilot that vehicles can be tracked, but by enabling the user in any full scale system to elect not to provide information. Potential U.K. operators are not convinced that this will be sufficient to remove the fear of being tracked. We are satisfied that this is a legitimate concern, but that other means can (as outlined in Chapter 11) be found to obtain information on vehicles' routes. (Para. 4.1.5).

5.12 It will be important for the evaluation to be conducted in the public sector and for every opportunity to be taken, within the limitations imposed by commercial confidentiality, for lessons to be learned for the design and operation of route guidance schemes elsewhere. (Para. 4.4.11).

5.13 Careful thought will need to be given to the ways in which the results of the evaluation are used in reaching a decision on the implementation of a full scale scheme in London (Section 4.3).

6 INFORMATION REQUIREMENTS AND SOURCES

6.1 Categories of Information Required

6.1.1 In Chapter 5 we concluded that the London Autoguide pilot evaluation package ought to encompass the following elements, not all of which would necessarily be evaluated in the public domain:

(i) the technical performance of the system;

(ii) the ergonomic performance of the in-vehicle equipment;

(iii) the true benefits and disbenefits of the guidance given to different categories of user and for different types and elements, of journeys;

(iv) the perceived benefits and disbenefits of such guidance;

(v) the responses of users, in terms of the extent to which guidance is sought and followed;

(vi) the implications for safety of the use of the equipment, and the resulting manoeuvres and routes;

(vii) driver stress and perceptions of safety;

(viii) the impacts, if any, on network performance and travel conditions for non-users and environmental conditions;

(ix) the costs of setting up and operating the system;

(x) the costs to third parties of providing information for the system;

(xi) the benefits to the public sector of data provided by the system.

6.2 Information sources to be used

We suggest that the information required could be supplied from the sources shown as columns in the matrix below:

	A	B	C	D	E	F	G	H	I
Technical Performance	X	.	X	.	X	.	X	X	.
Ergonomic performance	.	.	.	.	X	.	.	.	.
True Benefits	.	X	X	X	.	X	.	.	.
Perceived benefits	.	.	.	.	X	.	.	.	.
Users' responses	.	X	.	X	X	X	.	.	.
Safety implications	.	X	.	X	X	X	.	X	.
Perceived safety and stress	.	.	.	.	X	.	.	.	.
Network effects	X	X	.	X	X	.	.	.	.
Costs to operators	.	.	.	.	.	.	X	.	.
Costs to third parties	.	.	.	.	.	.	.	X	.
Benefits to third parties	X	X	.	.	.	.	.	X	.
Performance of full system	X	X	X	X	X	X	.	X	X

Key: A = System logs — F = Control groups
B = IVU data — G = Operators
C = Test vehicles — H = Third parties
D = Driver logs — I = Equipped, non-guided fleet
E = Driver comments — X = Information source

The following sections discuss, for each category of information required, how we propose that it be met.

6.3 The Technical Performance of the System and its Component Parts

6.3.1 This ought to be monitored primarily via the system log itself. This could be achieved by maintaining a record of uptime/downtime of the central computers, the communications links and the beacons. The record ought also to contain details of the frequency with which the advice is updated and the frequency which which exogenous information is input as well as any changes in the algorithms.

6.3.2 Although the system ought to be able to detect certain types of malfunction there will be others that it is unable to detect. These would include beacon transmission faults, difficulties experienced in receiving guidance at certain locations (e.g. wide roads with a high proportion of buses), most IVU faults and inaccuracies in tracking vehicles. We suggest that these could be detected by users and, more systematically, by roving test vehicles. Information from these sources should be filed centrally.

6.3.3 Users and test vehicles would also be able to detect database faults (as manifest by advice to make a manoeuvre at an inappropriate place) and, although these should have been cleared up before the start of the pilot, it would be useful to have any such instances filed centrally.

6.3.4 Travel time data transmitted by IVU's to the system and travel time predictions made by the system should be logged in such a way that the accuracy of the system's predictions can be monitored.

6.3.5 Any other information about the system's technical performance noticed by the operators or third parties (e.g. problems experienced in entering exogenous data) should also be filed as a matter of course.

6.4 The Ergonomic Performance of the IVU's

6.4.1 This will have been considered prior to the pilot but, for reasons outlined in Chapter Four, we believe that it ought also to be monitored during the pilot. The source of information would be user comments collected via questionnaires administered ideally at two points during the pilot: after a brief period of familiarisation (say a month after commencement of guidance) and after a substantial period of experience (say at the end of the pilot). Of these we regard the second as being the more valuable.

6.4.2 If it is thought desirable that the initial impressions of new users be logged then we would propose that, as in Berlin, questionnaires or interviews might be conducted amongst users of equipped hire cars. We have not pursued this line of thought.

6.5 Objective Benefits and Disbenefits of Guidance

6.5.1 This is problematic because even though it is possible to record the route taken and conditions experienced while following advice, it is very difficult to be sure what route would have been followed, and what network conditions would have been experienced, if advice had not been followed. We propose that two approaches to this problem be adopted.

6.5.2 Firstly, for regularly made journeys, it may be possible to establish, during a monitoring period prior to guidance being given, which route or routes were normally taken without guidance. It may then be possible to infer conditions experienced on these routes after the driver accepts advice to use an alternative route. This inference could be based on the experience of test vehicles following the routes in question or, less precisely, by observing aggregate changes in network conditions on the basis of data from equipped and unequipped (control sample) vehicles. Some of the data could come from the IVUs while some might more effectively be obtained from driver logs.

6.5.3 Secondly, it will be possible for a sample of journeys, to arrange for an equipped test vehicle, which is following advice, to be paired with another vehicle whose driver is not following advice. Both vehicles would start out from the same place at the same time and their different experience would be logged for later comparison. It might sometimes be appropriate for the vehicles to work in groups of three or more in order to reduce the impact of driver-specific effects and to allow for different levels of guidance following. Some of the required data could be obtained from the IVUs while some might more effectively be obtained from driver logs.

6.6 The Perceived Benefits and Disbenefits of Guidance

6.6.1 This could be obtained via questionnaires issued to participants, some of it relating to specific journeys actually made under guidance and some of it based on a distillation of experience gained.

6.7 User Responses

6.7.1 We suggest that an analysis of IVU data and driver logs be conducted to determine in what circumstances a driver seeks advice, in what circumstances he attempts to follow it and in what circumstances he succeeds in following it. It will be necessary to supplement this objective data with drivers' comments on why they did not seek advice and why they did or did not follow it. Without these comments it is possible that the 'objective' data might be meaningless or incorrectly interpreted.

6.7.2 It is obviously possible that the very fact of having one's car equipped with Autoguide might cause one to increase the number, length and timing of one's trips. This effect could only be detected if the pilot is long enough for 'novelty' effects to wear off. It would require monitoring of trip making behaviour via the vehicle or driver log before and during the experiment. Also, in order to correct for seasonal effects, a similar log would have to be kept by drivers in a control group. We consider that a simple reading of the odometer on a regular, preferably daily, basis would be the simplest way of achieving this.

6.8 Safety Implications

6.8.1 These are of obvious concern to all parties. We suggest that data be collected from three sources. Firstly, we suggest that drivers in the sample of equipped vehicles and in the control group be asked to log their involvement in road accidents (including damage only accidents) during the pilot and during a run up period to it. This simple expedient is not expected to be able to detect minor changes in involvement in accidents but is included because of its political appeal and because, if the changes were very significant, if would be a very powerful indicator. It would be sensible to monitor seasonal trends in accidents during the period via published accident statistics as well as via the control group.

6.8.2 Secondly, we suggest that it would be worthwhile to use IVU data to note changes in the frequency with which drivers undertake dangerous manoeuvres while under full guidance, in autonomous mode and while not under guidance. The problem here would be to determine which manoeuvres are to be classified as dangerous but obvious examples would include making banned turns, U-turns and proceeding the wrong way up one-way streets. The same data source might also be used to log changes in guided vehicles' usage of roads with differing safety records, for example to detect a shift away from the use of urban radials with shop frontage and any shift towards use of motorways.

6.8.3 Thirdly, we suggest that it would be worth using information logged by the IVUs to determine the frequencies with which vehicles are involved in sudden decelerations while under full guidance, while in autonomous mode and and while not under guidance. The idea behind this proposal is that a change in the incidence of sudden braking might be indicative of a change in the number of near miss traffic conflicts. We should emphasise that the technical feasibility of this suggestion is not yet established, (although Herr von Tomkewitsch of Siemens has indicated that it should not prove too difficult to achieve the necessary modification to the IVU software) and that the link between conflicts and accidents is not irrefutably established. ITS is separately preparing a proposal for work to establish the practicality and validity of the procedure. The value of the information might be further enhanced if interviews were conducted with drivers being asked to comment on the 'deceleration log' of their very recent journeys.

6.9 Driver Stress and the Perceived Safety of Using Autoguide

6.9.1 Changes in levels of driver stress or fatigue may occur when route guidance is provided by Autoguide. Although these could be monitored 'objectively' by trained psychologists accompanying the drivers or even by physiological monitoring we do not think such techniques are appropriate in the Autoguide pilot. We do, however, suggest that appropriately designed questionnaires could be used to gather the necessary information from the drivers themselves.

6.9.2 We suggest that it would be worth collating users' comments on the perceived safety of using Autoguide (compared to the alternative, be it attempting to follow signposts or to read a map balanced on the steering wheel) partly because they are valuable in their own right and partly because they could be invaluable if none of the other 'objective' methods of detecting safety effects were to produce statistically significant results.

6.10 Changes in Network Usage

6.10.1 Network usage should be monitored using the IVU data supplemented by information from driver logs. It will be particularly important to record usage of environmentally sensitive links which are not included in the secondary network, still less the guidance network. In the case of IVU data this creates some technical issues as to how tracking might be done which are addressed in Chapter 11. Note that the IVU data might be accessed directly or, perhaps more efficiently, via the system log.

6.11 The Costs Incurred by the System Operators

6.11.1 This information would need to be available in a fairly disaggregate form (see Section 12.3 for details) and the full co-operation of the operators in supplying objective data would obviously be required.

6.12 The Costs Imposed on Third Parties

6.12.1 In some cases a commercial market for the information in question may already exist in which case these costs would have been transferred to the operators, but in other cases (e.g. provision of information about network characteristics or exogenous incidents) a market may not yet exist and the third parties would have to be asked what they would charge if there were one.

6.13 Benefits to Third Parties

6.13.1 The value of any data produced as a byproduct of the scheme will be difficult to estimate, since during the pilot it is likely to be of limited volume and representativeness and so no commercial market will yet exist. Nonetheless we believe that it would be worth providing examples of outputs (e.g. journey time data, incident detection, flow monitoring and the location of congestion or conflicts) to potential purchasers of the information before seeking their valuation of it.

6.13.2 The information would be retrieved from the system log using appropriate software (the costs, prediction and operation of which would be an example of one of the costs incurred by the scheme operator). Depending on the distribution of participant vehicles' travel in the network it may be appropriate to arrange for equipped test vehicles to use particular parts of the network in order to enhance the quality of the planning information that the system can provide.

6.14 The Performance of a Full Scale System

6.14.1 This would, we believe, have to be estimated via a simulation model rather than directly from the results of the pilot scheme. However, data for calibration of such a simulation model could be gained from the pilot. Further details of what we recommend are given in the following chapter.

6.15 Further Details on Sources

6.15.1 Chapter 7 discusses the way in which a simulation model could help to predict the effects of a full scale scheme and indicates the information that would be required from the pilot to calibrate such a model. Subsequent chapters consider requirements for experimental design (Chapter 8), the use of test vehicles (Chapter 9), information from participants and control groups (Chapter 10), data from IVUs (Chapter 11) and other data sources (Chapter 12).

7 PREDICTION OF THE PERFORMANCE OF FULL SCALE SCHEMES

7.1 Introduction

7.1.1 It is not a simple matter to extrapolate the findings of the pilot scheme into likely impacts of a full scale scheme in London or elsewhere. There are four main reasons for this:

- a scheme with a large number of participants should be capable of producing more accurate and more rapidly updated journey time estimates and thus better guidance, provided that the data handling problems can be overcome;
- a scheme, whose participants formed a significant minority, or even a majority of road users would potentially suffer from feedback effects, since a route recommended on the grounds of its lack of congestion might, by virtue of being recommended, become busier than an alternative route;
- a scheme, whose participants formed a significant minority, or even a majority, of road users, might be able to provide information of great value to traffic planning and, via a link with the UTC system, real-time control;
- the type of people, and the type of journeys for which they seek advice, might be different in a fully commercial scheme.

7.1.2 It is clear that the impact and interaction of these effects might be extremely complex and could not be estimated without some form of simulation model. In this chapter we indicate what might be required of such a model and what it, in turn, would require from the pilot scheme.

7.2 The Outputs Required of a Simulation Model

7.2.1 The model would be required to produce a number of indicators. We identify the following as being particularly significant.

7.2.2 Firstly, it would need to estimate the distribution (including variability) of travel times and costs experienced by four different groups of drivers:

- those driving equipped vehicles and following advice completely;
- those driving equipped vehicles but following advice only in part;
- those driving equipped vehicles but not following advice at all;
- those driving unequipped vehicles (since they may differ from the above).

7.2.3 Secondly, it would need to indicate the distribution, in time and space, of traffic flow and congestion on the network indicating, in particular, the different impacts on different types of link, e.g.:

- major arterials;
- city centre roads;
- minor roads on the guidance network;
- minor roads not in the guidance network;
- other environmentally sensitive roads.

7.2.4 Thirdly, the extent and quality of planning data produced as a byproduct of the guidance system would need to be assessed.

7.2.5 Other issues, such as safety, could be extrapolated directly from experience gained from the pilot and do not therefore have to be included in the simulation model except insofar as they are related to changes in the pattern of use of the network.

7.3 The Basic Structure of the Simulation Model

7.3.1 The model would ideally be able to represent the performance of schemes based in a range of types of city with different types of network, different levels of congestion and patterns of demand. For this reason we propose that it be based on a hypothetical network with a hypothetical origin-destination matrix network, whose characteristics could be modified to represent different cities with differing network and demand characteristics. If this requirement to estimate impacts in a variety of cities did not exist it might be desirable to use a London network and origin-destination matrix but this could prove very difficult to achieve. For practical reasons it would probably be necessary to restrict coverage to a segment of the total city and, perhaps, to only a part of the day.

7.3.2 The model would ideally be able to represent the performance of schemes having different specifications (e.g. different beacon densities, different guidance network densities and different travel time estimation procedures) and we therefore suggest that it be able to accept variation in these.

7.3.3 The full range of input parameters might be: the street network; the origin destination matrix; the market penetration; the guidance system parameters; the route choice behaviour of non-users and the behavioural response of equipped drivers.

7.3.4 The street network of the (hypothetical) city would have to include almost all links and not just those known to the guidance system. If a 'real' network is required this would imply a significant effort in data collection and network building to the required level of detail.

7.3.5 The origin destination matrix of the (hypothetical) city would have to include all trips and not just those by equipped vehicles and would have to indicate the trip start times. If a 'real' matrix is required it might be possible to base it on existing data (e.g. the LTS matrix for London) but a considerable amount of work would be required to get it into the correct form and some information (e.g. trip start times) could prove particularly difficult to find. Variation in the matrix would allow the effects of different levels of underlying congestion in the (hypothetical) city to be tested.

7.3.6 The assumed market penetration of the IVUs under different guidance system design and performance scenarios would need to be disaggregated by vehicle type and driver characteristics.

7.3.7 The guidance system design would be specified via critical system parameters including:

- method and rate of travel time matrix updating;
- density of beacons;
- density of the guidance network and the criteria for inclusion or exclusion of specific links;
- the density of any secondary system used for vehicle tracking;
- the nature of any link with the UTC system;
- the nature of guidance criteria (e.g. user optimal or system optimal, time or cost minimisation);
- the nature of any mechanism designed to reduce feedback effects (e.g. proportional routeing);
- the use made of any external information;
- the availability of special facilities such as parking information.

7.3.8 Assumptions on the route choice behaviour of non-users would ideally reflect the different behaviour of different types of people making different types of journey.

7.3.9 Assumptions on the extent to which, under a range of guidance system design scenarios, users can and do follow advice, including an indication of how this varies with personal characteristics, vehicle type, journey purpose and type of advice available (e.g. full guidance or autonomous mode only). An estimate would also have to be made of the effect on driver behaviour of a perception on their part that feedback was occurring.

7.4 Sources of Data for the Simulation Model

7.4.1 With reference to the parameters listed in 7.3.3 above we suggest that the necessary information could be derived as follows.

7.4.2 The street networks and the origin destination matrices could be loosely based on those of cities for which route guidance is being, or might be, considered.

7.4.3 The assumed market penetration of the IVU's could be based on the results of: the market research already carried out on behalf of DTp (contract TRRL 539); the LISB trial and attitudinal work conducted amongst participants in the London pilot. The London pilot would ideally allow consideration of feedback from guidance system parameters to market penetration.

7.4.4 The effects of a wide range of guidance system parameters could be simulated fairly readily, but there would be little point in attempting to represent scenarios for which there was no evidence of the behavioural response. The range of scenarios to be tested would therefore be effectively constrained by the experience gained during the London pilot. However, the opportunity should be taken in the Pilot to vary at least the density of beacons, the density of the network, the optimising criteria and the rate of updating of guidance. Further details of the issues involved are provided in Section 8.2 of this report.

7.4.5 Assumptions on the route choice behaviour of non-users could be based on state-of-the-art assignment techniques augmented, perhaps, by anything learned during the London trial about route choice while not under guidance. This would produce a reasonable (but not perfect) basis for modelling the behaviour of non-users.

7.4.6 Assumptions on the extent to which users can and do follow advice would be determined from appropriately disaggregate observation of the behaviour of users during the London pilot. Further details of the requirements that this places on the pilot as a source of data with which to calibrate the simulation model are given in Section 7.6.

7.5 <u>Technical Aspects of the Simulation Model</u>

7.5.1 There is clearly considerable work to do in building a simulation model such as that outlined above. We suggest that such work should be put in hand without delay in order that it is ready to accept information from the pilot scheme and produce the required results in what will be a very restricted time scale between the end of the pilot and decisions on licences.

7.5.2 The model will clearly have to be purpose-built but should obviously incorporate mechanisms employed in the guidance algorithms used in the pilot along with features from advanced assignment models such as CONTRAM and SATURN. We are aware of existing work at Southampton University on behalf of Plessey on a route guidance simulation model incorporating some of the features we have outlined above. With appropriate modifications, some of which would require considerable extra work, the Southampton/Plessey model might form an input to the required simulation. We should also draw attention to a joint proposal by the University of Leeds and Southampton to conduct work in this area under an SERC rolling programme grant.

7.5.3 We do not see it as our role here to present detailed recommendations on the structure of the simulation model but we do make the following observations.

7.5.4 For the sake of realism the simulation model ought, if possible, to contain a representation of delays at junctions and during specific turning movements rather than simply via link-based speed-flow relationships. This is particularly important if an investigation of the effect of a linkage between the guidance and UTC systems is to be carried out.

7.5.5 The simulation model ought to represent variability and uncertainty in journey times since this will influence the perceived quality of advice from the guidance system. Work currently underway at Leeds University is attempting to model this phenomenon for part of the London network and might form an input to the simulation.

7.5.6 Since the route guidance system is itself dynamic it will be necessary to build this feature into the simulation. This implies that drivers will not be assumed to complete the journey along the route to which they have been initially assigned but that the model will be based on time slices (perhaps of variable length) at the end of which the current position of each driver (or group of drivers) will be assessed and recorded prior to a new time slice during which they would continue their journey in the light of the latest information and changed circumstances. This dynamic model structure should be capable of representing feedback effects (whereby the advice affects the road conditions which in turn affects the advice).

7.5.7 The simulation model will have to be able to represent the effect of the type of advice given on the drivers' ability or desire to follow it. Thus a sub-routine might determine that advice to follow a road in a counter-intuitive direction or to ignore an environmentally sensitive link, might be rejected by a given proportion of drivers. It might similarly suggest that a given proportion of drivers will fail to follow advice at junctions of a given type. These proportions would obviously have had to be established during the pilot.

7.6 Data Required from the Pilot for the Simulation Model

7.6.1 In order that a variety of scheme specifications can be tested in conjunction with a variety of assumptions about market penetration, the data from the pilot will need to be disaggregated according to the characteristics of the participant, the type of advice being received, the type of journey being undertaken and the part of the network in which it is being made.

7.6.2 The main sources of behavioural information will be the IVU data and the driver log. These will indicate to what extent, and in what circumstances, a driver seeks and follows advice from the system. Additional information from questionnaires completed by the drivers will also be a useful input to the prediction of market penetration.

7.6.3 The performance of full scale schemes will require some information on the extent to which participants react differently to different qualities of advice and in particular whether advice based on a travel time database which is receiving a high level of input from equipped vehicles provokes a different response from that based on less complete data. If the pilot is to provide evidence of this effect there is obviously a need to ensure (perhaps via the use of test vehicles or a fleet of vehicles equipped only to transmit travel time data) that high levels of travel time data input are achieved for at least part of the time in part of the network.

7.6.4 Estimates of behavioural response and of market penetration will require account to be taken of the extent to which participants in the pilot differ from the population at large. This implies a contribution from the population control group to the preparation of data for the simulation model.

7.6.5 Objective assessments of the net impact on safety (derived again from IVU data and driver records of involvement in accidents, with the latter interpreted in the light of the information from the sample control group and, perhaps, underlying accident trends) are not strictly necessary for the simulation but might be able to be incorporated onto it.

8 EXPERIMENTAL DESIGN

8.1 Objectives

8.1.1 We consider that a prime objective of the evaluation should be to provide sufficient information on the technical, ergonomic and operational performance of different levels of route guidance to enable decisions to be taken on the advisability of extending the pilot into a full scheme, of introducing similar schemes elsewhere, and on the most appropriate specification for such schemes.

8.1.2 This implies that the evaluation should have two main components:

a) an assessment of the pilot scheme itself and
b) an assessment, using information derived from the pilot scheme, of the probable performance of full scale schemes which might differ from the pilot not only in scale but also in operational details and network characteristics.

8.1.3 Effective assessment of the pilot and of possible full scale schemes will require the pilot scheme to have provided information at a sufficiently disaggregate level to distinguish the effects of

- variation in system parameters (such as beacon density and travel time matrix updating procedures)
- variations in street network conditions (e.g. density of streets, level of congestion) and
- variations in the user profiles including types of journey being made and the frequency with which destination finding is required.

This in turn implies that the pilot scheme and evaluation package should be designed with these dimensions in mind.

8.1.4 For practical reasons it is appropriate to consider the variation in system parameters and street network conditions together under the heading of 'system specification'. We will now consider what variants are necessary under this heading before turning to a consideration of required variation in user profiles.

8.2 Variation in system specification

8.2.1 We wish here to represent a range of street network conditions and of system design parameters. There are three ways in which we could seek to incorporate the necessary range into pilot schemes:

- Temporally whereby the system is varied such that at one time it operates in one manner and at other times it operates in another manner.

- Spatially whereby the system would operate differently in different areas (some degree of spatial variation is, of course, inescapable given the heterogeneity of the London road network).

- By dividing the users into separate groups whereby each group would be treated differently - effectively experiencing a different system.

8.2.2 Other things being equal we would prefer not to introduce temporal variation unless the period of the experiment is sufficient to allow users to become familiar with each variant of the system and thus allow us to observe behaviour uncomplicated by learning curves. However, even if sufficient time is available, there is the risk that a user's experience of one system will colour his perception, and affect his usage of, a subsequent system. There is also the risk of a variant of affirmation bias whereby a user exaggerates the difference between two systems in order 'to be helpful'. This effect could be tested for by having a subset of users who were told that the system had been changed when in fact it had not.

8.2.3 The problem with spatial variation is that a mobile user will experience different conditions in different areas and it may be difficult to determine which conditions are having the greatest impact on his perception and subsequent behaviour.

8.2.4 Division of the users into separate groups to receive different treatment is theoretically attractive but it has unfortunate implications for sample size requirements.

8.2.5 We will now consider each of the variations which we think ought to be included in the pilot and comment on how they might best be accomplished.

8.2.6 **Type of road network.** Here the distinction is between those links in the city centre, those elsewhere in the inner areas and those outside. Payoffs from guidance might be different in the different areas. For London the city centre might be defined as the area inside the ring of roads linking the main line termini, the inner area being between this inner ring and the North and South Circular Roads and the rest as being beyond that. This distinction is obviously spatial. To effect it, it will be necessary to distinguish, in the system log and in participant responses, between links in the three areas.

8.2.7 **Density of guidance network.** Here we wish to distinguish between a network that provides guidance along all links and one which is skeletal and covers only major links. Obviously a very skeletal network would be cheaper, but we need to know how unsatisfactory it would be in operation. It is of course already envisaged that the density of the guidance network will vary spatially (being denser in the West London corridor) and we see this as a useful element in the pilot design. We would, however, urge that the guidance network be designed to allow for a part of central London to be served by a dense network and part by a sparse network. This is necessary in order that the different effects of road network density and guidance network density can be assessed.

8.2.8 **Frequency of beacons.** A lower density of beacons will result in more frequent dropping into autonomous mode and there may be differing reactions to autonomous mode as opposed to guidance mode. If so, this dimension could be assessed by comparing reaction to the two modes, and modeling for different densities of beacons. The proposed pilot system already provides spatial variation in beacon density but, in order that the different effects of guidance network density and beacon frequency can be identified, it is important that the beacons are not always most frequent where the guidance network is at its densest. If this is not possible it would be necessary to detect the beacon density effect by having some beacons 'invisible' to some users.

8.2.9 **Inclusion or exclusion of 'ratruns' from the guidance network.** There will inevitably be a desire on the part of the system designers to include some links on their network that the local authority might prefer to exclude on environmental grounds. At the margin it may be difficult to determine whether a given link ought, or ought not to be included. To test the hypothesis that an equipped driver's behaviour and perception might be influenced by whether or not such ratruns are included, we propose that such links be available to a subset of participants and not to the rest. We do, of course, understand that there is no need to experiment with the inclusion of <u>very</u> sensitive routes because they would never be included, but we do see a need to test the effect of including or excluding the marginal cases. A temporal variation in the inclusion of such links would suffer from the problems already mentioned, while a spatial variation (allowing ratruns in some parts of the network, but not in others) would almost certainly result in some bizarre recommended routes when the "ratrun border" was crossed.

8.2.10 **Procedures for dealing with temporary diversions.** The quality of guidance, as perceived by the users, could be significantly affected by the nature of the procedures for dealing with temporary diversions made necessary by major road works or incidents. The distinction can be drawn between a system that simply finds best routes on the guidance network avoiding the affected links and one that allows for extra local links to be added to the guidance network for the duration of the emergency. This distinction would be easy to test on a case by case basis (with some diversions being made on temporary links and others not), but the results would be difficult to interpret because of the importance of other local factors. We therefore conclude that this is another dimension that should be varied on a person-by-person basis, with some users able to use temporary local links and others not.

8.2.11 In addition to these sources of variation, there is another dimension to be considered: the sophistication of the software and algorithms used by the system to estimate link times and to recommend routes. Different levels of sophistication can clearly be expected to result in different qualities of advice and are likely to produce different user reactions. We are aware of the form of the procedures used in LISB and envisaged for London but we think that it would be shortsighted to assume that there might not be pressures either to implement simpler procedures (if, for example, it was apparent that a large proportion of the benefits could be achieved at a fraction of the expense) or more complex procedures (if, for example, those proposed do not prove adequate). This being the case we suggest that the pilot ought to allow for tests of procedures with different levels of sophistication.

8.2.12 We suggest that it would be useful to include the following (in increasing order of sophistication):

(i) use of historic data with different journey times stored for each link for (say) each half hour period during the day with separate profiles for weekends. This system would need to be capable of incorporating exogenous information on scheduled distortions to the network such as roadworks and public events;

(ii) as (i) but with the ability to accept real-time exogenous information e.g. on accidents which affect network capacity;

(iii) as (ii) but with the ability in real time to modify estimated link times on the basis of real time information from automatic traffic counters and/or occupancy detectors;

(iv) as (ii) or (iii) but incorporating real time link travel time information from equipped vehicles, with such information coming in at a rate no higher than one vehicle per 5 minute period (a rate unlikely to be exceeded very often during the Pilot as currently envisaged);

(v) as (iv) but with information coming in at a rate of more than one vehicle per 5 minute period (which might be achieved in a full scale scheme and which could be 'simulated' on some links during the pilot by judicious use of test vehicles (see Chapter 9).

8.2.13 With the exception of the distinction between (iv) and (v) it is almost certainly impractical to suggest that these different modes of operation could occur simultaneously in different parts of the network. If the number of participants could be increased above 400 we would recommend that variants (i)-(v) be provided for different subgroups of users. We understand that the specification of Autoguide allows for different classes of user to receive different types of advice. It should therefore be possible, albeit perhaps at the cost of increased processing power in the central computers, to run several different algorithms simultaneously.

8.2.14 It may not, however, be technically or economically feasible to run all five algorithms simultaneously. Notwithstanding the caveats in 8.2.2, the alternative method of ensuring that all systems are tested would be to run them sequentially. We estimate that this could be achieved in a 36 week programme: 4 weeks for familiarisation, 8 weeks for the first method, and 8 for each of three subsequent ones.

8.3 Dimensions of the Design of the Participant Sample

8.3.1 In selecting participants for the evaluation, it is necessary to ensure that all the most important dimensions that will be used in the evaluation are represented. Thus, for example, if it is considered that men and women may react differently to guidance, then both men and women need to be included in the evaluation. Furthermore, each gender needs to be included in sufficient numbers that further splits can be made along other dimensions of interest, e.g. age.

8.3.2 In practice it will almost certainly be necessary to select the participants by some kind of convenience sampling, with their true selection probabilities from the population of drivers in London unknown. But it is still convenient to use the shorthand of "sample" to denote the participants selected.

8.3.3 In fixing the dimensions of the sample there are two conflicting considerations; firstly, to identify <u>all</u> the important factors (independent variables) related to the sample that will affect the evaluation criteria (dependent variables); and secondly, to be as economical as possible in selecting the factors, so that crucial sample size is not wasted on factors or interactions that are relatively unimportant.

8.3.4 There is an element here of cart-before-the-horse in that one has to have hindsight before the event, and hence optimal design is unlikely to be achieved in practice. However, there is a large background of knowledge about driver behaviour, which ought to ensure both that crucial factors are not excluded and that superfluous factors are not included.

8.3.5 It is now appropriate to examine the dimensions of the sample in more detail. All of these factors will be related in some way to driver behaviour, since we are attempting to evaluate how different kinds of drivers react to Autoguide. But we can make a distinction between those kinds of factors that are inherent in London drivers in general, and those that are more related to the guidance system being installed. The former category includes such factors as age, sex, and driving patterns, while the latter includes the participants' use of the fully beaconed West London sector as compared with that in the partially beaconed area. For convenience, we may term the former "inherent" and the latter "experiential."

8.3.6 The inherent dimensions which most obviously influence driver behaviour are age and sex. Both of these are known to have a strong relationship to driving behaviour, and both are related to acceptance of new technology. There is considerable evidence that, given current acculturation, older people and women are more reluctant to use "high-tech" equipment, and will probably take longer to adjust to the new equipment. We suggest that, in addition to the two sexes, three age bands could be adopted with boundaries at ages 30 and 50 (chosen for their rough correspondence with family life cycle break points as well as their psycho-social significance).

8.3.7 Another dimension is the urgency with which drivers are trying to reach their destinations. There is a continuum between, at one extreme, aggressive drivers who for whatever reason (journey purpose, value of time, personality) are single-minded about trying to reach their destinations in the shortest possible time, and at the other extreme those drivers who are more willing to sacrifice journey time for some other personal or community benefit. An individual driver's position on this spectrum is, of course, difficult to determine and it may instead be necessary to use journey purpose and socio-economic group (as revealed by occupation) as proxies. It might be worth identifying two journey purposes (work or commute versus the rest) and two socio-economic levels (A and B versus the rest).

8.3.8 A driver's perception and use of Autoguide is likely to be affected by the extent to which, on regular journeys, he tends always to use the same route or to experiment with a variety. We suggest that this continuum can be represented by three gradations.

8.3.9 The benefits of Autoguide as experienced by the user are likely to depend quite significantly on whether he tends to use ratruns or to stay on the main roads. We therefore suggest that it would be useful to distinguish between these two types of behaviour.

8.3.10 The type of vehicle driven may also be considered to be inherent to the behaviour of the vehicle/driver combination: HGVs and LGVs will be engaged in different types of driving from car drivers, while, among car drivers, professional drivers will have different route patterns from private motorists.

8.3.11 We have thus identified: two sexes, three age groups, two journey purposes, two levels of SEG, three degrees of experimentation with alternative routes, two levels of use of ratruns and four vehicle/driver combinations. Were we to seek to identify separately each combination of these factors this would imply 576 (=2x3x2x2x3x2x4) separate groups each requiring to be separately evaluated. Such a procedure would clearly be cumbersome, expensive, and superfluous. It is cumbersome because of the very large number of groups that need to be managed in the design and analysis. It is expensive for the same reasons, and because it demands a very large sample. And it is superfluous because not all the potential subgroups are of interest: for example, there is little to be gained by splitting HGV drivers on their sex, journey purpose, or socio-economic group.

We recommend instead that the participants simply be selected so as to provide sufficient people in each class on each of the seven dimensions outlined above and to ensure that each class is made up of a representative range of people (in terms of the conflicting dimensions).

8.3.12 For each of the seven dimensions we suggest the following sub-categories and conflicting dimensions as a starting point for design.

sex -	(male) - age, journey purpose, SEG, diversions, ratruns, vehicle type; (female) - age, journey purpose, SEG, diversions, ratruns, vehicle type;
age -	(under 31) - sex, journey purpose, SEG, diversions, ratruns, vehicle type; (31-50) - sex, journey purpose, SEG, diversions, ratruns, vehicle type; (over 50) - journey purpose, SEG, diversions, ratruns, vehicle type;
journey purpose -	(work/commute) - sex, age, SEG, diversions ratruns; (other) - sex, age, SEG, diversions, ratruns;
SEG -	(A/B) - sex, age, journey purpose, diversions, ratruns; (other) - sex, age, journey purpose, diversions, ratruns, vehicle type;
diversions -	(always uses same route) - sex, age, SEG, journey purpose, ratruns, vehicle type; (sometimes uses alternative routes) - sex, age, SEG, journey purpose, ratruns, vehicle type; (uses a wide variety of routes) - sex, age, SEG, journey purpose, vehicle type;
ratruns -	(prefers to stick to main roads) - sex, age, SEG, journey purpose, diversion, vehicle type; (often uses ratruns) - sex, age, SEG, journey purpose, diversion, vehicle type;
vehicle type -	(HGV) - age, diversions, ratruns; (LGV) - age, diversions, ratruns; (private car) - sex, age, journey purpose, SEG, diversions, ratruns; (car used on business) - age, diversions, ratruns.

8.3.13 Turning now to the 'experiential' dimensions, we suggest that usefulness of Autoguide to a given driver will depend in part on the extent to which he has to travel in unfamilar parts of the network and to locate destinations previously unknown to him. It will also depend on the amount of driving he does in different parts of London where the street network and/or the provision of Autoguide differ.

8.3.14 Any given driver will obviously experience all of these conditions to a greater or lesser extent. But it will be important, when selecting the participants, to ensure that, as far as possible, there is likely to be, for each class of driver, a range of experience of different types of driving in twelve different parts of London. The three different types of driving being:

- trips to familar destinations in familar parts of the network;
- trips to unfamilar destinations in fairly familiar parts of the network;
- trips to unfamiliar destinations in unfamiliar parts of the network;

and the twelve parts of London being the product of: three types of street network (central London, other areas inside the North and South circulars and, the rest of London); two guidance network densities (dense and sparse); and two beacon frequencies (frequent and infrequent).

8.3.15 Over and above the requirement to ensure that the participants display a range of characteristics it will be important to ensure that they are chosen so as to maximise the probability that sufficiently high usage (and thus relatively frequent updates of the travel time matrices) are achieved in appropriate parts of the network (see 8.2.12 et seq.). In practice this will probably mean that participants should be selected partly on the basis of whether or not they are likely to make use of roads in 'target' corridors.

8.3.16 In chapter 10 we outline our ideas for the procedures that might be adopted for the selection of participants. In summary, we suggest that, as in Berlin, advertisements in the press and broadcast media are likely to produce a pool of applicants wishing to participate in the evaluation in exchange for having their vehicles equipped.

8.4 The Need for "Control" in the Experimental Design

8.4.1 The potential arises in the evaluation for bias due to the non-typicality of the participants. For example, an above average involvement in road accidents associated with drivers of equipped vehicles might reflect their above average mileage in urban traffic, or the aggressive driving style of those who are attracted to new technology, rather than being attributable to Autoguide. Such difficulties arise because, however careful the selection from among the applicants (see 8.3.15), the participants in the trial are likely to be younger, more technologically orientated and to cover greater mileages than the population at large. Although it can be argued that they would be similar in many ways to the first wave of purchasers of Autoguide, they might not be typical of those who would be potential users of a more widely available system.

8.4.2 Problems can also arise because of the trend or seasonal effects. For example an apparent increase during the pilot, in mileage driven or the number of unfamiliar destinations sought by equipped drivers might reflect seasonal effects such as a tendency to travel further afield during the summer months.

8.4.3 One statistical method for controlling for seasonal and trend effects is to use a time series approach, such as intervention analysis. However, this requires a long study period, (at least twelve months) both before and after the "intervention" (here the introduction of route guidance). This would be both impractical and costly.

8.4.4 An alternative solution is to use an unguided control group matched to the participants and to observe this control group both while the participants are and are not under guidance. We believe that the inclusion of such a control in the study design is essential. We term it the 'sample' control group.

8.4.5 A further issue arises: the sample will not be drawn at random from all London drivers. It will instead be drawn from those attracted by the process for selecting participants. The need will arise to assess how the sample differs in makeup, attitudes, and driving patterns from the general population of London drivers. This will provide important information for the modelling process, and in particular for the assessment of the likely impact of widespread use of guidance. We suggest that a second control group, drawn at random from all London drivers, be used here. We would term it the 'population' control group. The full range of information required from the participants and the sample control group would not be required of the population control group. A set of survey questions on attitudes, driving patterns, and current route-finding procedures would suffice. The same questions would also have to be asked of the participants and the sample control groups.

8.5 Adequacy of Sample Size and Duration of Experiment

8.5.1 Ideally, all the potential uses of the participant and sample control groups would be taken into consideration before deciding on the appropriate sample sizes and on the length of the evaluation period. However, we accept that it is extremely unlikely that the number of participant vehicles will exceed 400 and that this number ought to be taken as given. We concentrate instead on estimating, for the most important evaluation criteria, whether sample sizes are adequate for the kind of disaggregation outlined in section 8.3 and what the implications are for the duration of the evaluation. In this latter context we start by assuming that each experiment should last for about eight weeks to ensure that a sufficient range of conditions is covered. In practice, even eight weeks may be insufficient since major network disturbances can last for at least as long as this.

8.5.2 Precision Required The five criteria to be used in this investigation of sample sizes are:

1. The ability to make reliable estimates about travel time saved for each of the types of system identified in Section 8.2 and for each of the types of journey and person types identified in section 8.3. Since the predicted savings in travel time from the use of guidance is in the range of 10 percent (Jeffery et al, 1987), it is desirable to be able to observe differences of the order of 5 percent.

2. The ability to make reliable estimates of safety benefit or disbenefit for the participant group as a whole.

3. The ability to make reliable estimates about the proportion of journeys of different types for which participants seek or accept guidance in whole or in part.

4. The ability to make reliable estimates about proportions on market research issues, such as the proportion of participants being satisfied overall with Autoguide.

5. The ability to detect changes in the level of use of different types of link.

The standard 95 percent confidence interval will be used as the level of precision required.

8.5.3 Travel Time Saved. To establish necessary sample size here, we require some estimate of the variance of travel time. From previous work (Bates, Dix, and May, 1987), we can estimate that the coefficient of variation of travel time is 0.15. If we wish to be able to observe differences of a given size at the .05 significance level and the sample size is large enough to assume a normal sampling distribution, then the difference must be more than twice its standard error. We may assume that the standard error of the difference is the same as the standard error of the mean (remember that we wish to estimate paired differences within a single sample). For a population with a mean of 1 and a standard deviation of 0.15, we obtain the following values:

NUMBER OF PAIRED OBSERVATIONS	2 x STANDARD ERROR
25	0.060
30	0.055
35	0.051
40	0.047
50	0.042
100	0.030
200	0.021
300	0.017
400	0.015

8.5.4 A five percent change here would equal 0.05. As long as that change is greater than twice its standard error, it is statistically significant at the .05 level. Thus, given the estimated inherent variability of travel time, we can estimate a 5 percent difference with the necessary precision when we have a sample size of 40 or more paired observations. It should be emphasized that this sample size represents the number of journeys, not the number of participants.

It therefore appears that the need to estimate travel time saved will not be a critical factor in the choice of sample size. Even quite small samples would provide adequate sizes for this estimation, although if the number of participants were very small, one would have to worry about the effect of the clustering of journeys by participant.

8.5.5 Assuming that, as indicated in 8.2, we wish separately to treat four groups of participants (with and without access to ratruns x with or without access to 'off network' emergency diversion links) and that the participants can be selected so as to require a maximum of only four divisions of the total sample (four is the maximum number of classes in any of the inherent dimensions and, if statistical independence can be assumed, this is therefore the upper bound on the number of subgroups required), there could be as few as 16 groups. Given that 40 observations are required for each group, this implies that at least 640 observations are required.

8.5.6 If the data is to be collected by test vehicles in the manner described in 9.3 this implies 640 (or 1280 if the benefits of'realistic' as well as 'perfect' adherence to advice are to be evaluated) paired journeys. Assuming an experimental period of eight weeks and that no more than two journeys can be covered by the test vehicles in any one day this implies a requirement for eight (or 16) test vehicles.

8.5.7 If the data is to be collected by comparing participants' experiences before and after the introduction of guidance, the requirement would be for each of the 400 drivers to produce data for two journeys during the preguidance phase and a further two during the guidance phase.

8.5.8 If estimates of benefit are required for individual drivers (rather than simply for the group as a whole) then each would have to produce data for 40 journeys during each phase - this amounts to one journey on each day of an eight week period. If estimates for each driver were to be obtained using test vehicles a fleet of 400 would be required!

8.5.9 <u>Safety</u>. Using recent accident statistics to estimate the number of injury accident involvements that a given group of people would have in a given period of time, the injury accident involvement rate for car drivers in built-up areas is 192 involvements per 100 million km (HMSO, 1987). From this we can calculate that 400 car drivers, driving an average of 10,000 miles a year, might be expected to have 12.28 injury accidents in a year, or about 2 in an eight week period. Assuming that there are seven accidents of all types (including damage only) to each reported injury accident (Spicer, Wheeler, and Older, 1980), the same 400 drivers would have about 14 accidents in eight weeks or 56 in 32 weeks. Applying the standard Poisson distribution for small numbers of accidents and the normal distribution for larger numbers (Nicholson, 1987), the minimum change we would be able to observe at the 95 percent confidence level would be a reduction or increase of 10 accidents in eight weeks or 15 in 32 weeks, i.e. we would need a change of at least 27 percent (assuming 32 weeks) or 71 percent (assuming eight weeks) for it to be statistically significant.

8.5.10 The size of the change required means that, even if damage-only accidents are included, the safety effects of Autoguide are unlikely to be proven in a reasonable amount of time using accident data alone. Some other measure of safety is required

8.5.11 The Germans are considering the use of in-car observers to make an assessment of any changes in safety. The observer would count traffic conflicts and unsafe driving actions. This approach seems fraught with bias problems.

8.5.12 The use of conflicts rather than accidents is potentially attractive because the IVUs are apparently capable of recognizing severe decelerations, many of which are likely to be related to traffic conflicts. If the rate of such decelerations per unit of time or distance changed for a group of drivers when they started to use guidance, one would almost certainly have observed a safety effect.

8.5.13 DTp traffic conflict observations have counted three or four grade 3 or greater traffic conflicts a day at a set of sites where five injury accidents involving at least two four-wheeled vehicles were recorded in four years (Spicer, Wheeler, and Older, 1980). Grade 3 or greater traffic conflicts are the ones most likely to be associated with heavy or emergency evasive action (TRRL, 1987). These numbers give a ratio of severe conflicts to injury accidents in the range of 876 to 1168. If we take the ratio to be 1000:1, 400 participants would be involved in about 33 severe traffic conflicts per day.

8.5.14 The same DTp study (Spicer, Wheeler, and Older, 1980) found that small numbers of conflicts have a Poisson distribution, with a variance between days approximately equal to the mean. For numbers of observed conflicts greater than ten per day, however, the normal distribution could be applied, with a variance approximately equal to the mean. The standard error of the mean is then given by:

$$Var(mean) = mean/n$$

where n is the number of days.

8.5.15 For 1320 conflicts over 40 days, the 95 percent confidence interval would therefore be plus or minus 11.5. If the standard error of the difference were equal to the standard error of the mean, then a one percent change (13) in the number of conflicts would be statistically significant at the .05 level. We conclude that, provided that is is technically feasible and that the link between sudden decelerations and conflicts is accepted, the use of IVU data provides a good chance of detecting any objective changes in safety.

8.5.16 The calculations above assume that IVU records are available for all journeys made by all drivers during an eight week period. If this were not practical the precision of the result would be reduced accordingly. Assuming that the data is sampled on a one-day-in-twenty basis, and that the data can be pooled for each of the four eight week periods there would be some 528 conflicts and, using the same assumptions as above, it would be possible to detect a 5% change in the number of conflicts at the .05 level.

8.5.17 Proportion of Journeys for which Users Seek or Accept Guidance. The variance of a proportion for a simple random sample is given by:

$$Var(p) = p(1-p)/(n-1)$$

The 95 percent confidence interval is given by plus and minus two times the standard error of the proportion. The standard error is the square root of the variance. The quantity p(1-p) will be greatest when p is equal to 0.5, and we will therefore set p to this value in deciding on necessary sample size.

SAMPLE SIZE	VARIANCE	2 x STANDARD ERROR (=95% C.I.)
100	0.0025	0.101
200	0.0013	0.071
300	0.0008	0.058
400	0.0006	0.050
500	0.0005	0.045
1000	0.0003	0.032
2000	0.0001	0.022

8.5.18 Thus with 500 journeys for a particular subgroup, one would be able to identify a percentage of 50 as being significantly different from 55 or 45. At 2000 journeys, 50 percent would be significantly different from 52.5 percent or 47.5 percent. At two journeys per week day, 400 people would make a total of 32000 journeys in eight weeks. If data is recorded for all of these journeys it should therefore be possible to make reliable estimates of the proportion of journeys on which guidance is sought or accepted for up to 64 subgroups (32000/500 = 64). If only 16 subgroups are required greater precision will be possible (32000/2000 = 16).

8.5.19 Proportions of the Sample for Market Research Issues. Here once again we apply the standard formula for the variance of a proportion and calculate the standard errors when p is set equal to 0.5. The standard errors are as follows:

SUBGROUP SIZE	VARIANCE	2 x STANDARD ERROR (95% C.I.)
25	0.010	0.204
33	0.008	0.177
44	0.006	0.152
55	0.005	0.136
66	0.004	0.124
77	0.003	0.115
88	0.003	0.107
100	0.003	0.101
111	0.002	0.095

Thus for subgroups of 25, we would be able to identify a proportion of 0.5 as being significantly different from 0.3 or 0.7. We would require subgroups of size 100 (implying only four such subgroups within the 400 participants) to identify a proportion of 0.5 as being significantly different from 0.4 or 0.6. This may well constrain the dissaggregation achievable, but we are fortunate in that this area is the least crucial in terms of assessing the benefits of Autoguide.

8.5.20 Changes in the levels of use of different types of link
The implications of this requirement are difficult to quantify precisely in the absence of information about the proportions of links of different kinds in the area covered by the system. We do not, however, envisage any difficulty in achieving the required sample sizes for links on the guidance network since all such usage can be continuously monitored by the system and this should produce several thousand observations per day throughout the trial period. The problem, if there is one, is likely to lie with links that are not known to the system and for which some form of manual analysis may be necessary (see 11.1.9) and which, for practical reasons, will be possible on a sample basis only. Further work is required to appreciate the precise implications of this.

8.6 Implications for Sample Size and Duration of Experiment.

8.6.1 The preceding discussion has shown that a sample of 400 participants using guidance over eight weeks would provide enough information to make reliable judgments about savings in travel time. Provided severe decelerations can be recorded by the IVUs, the same is true of safety, although here an eight week period would be required prior to guidance, when the participants' decelerations would be monitored. The same period would also be required if journey time savings are to be estimated by the before and after technique rather than via paired test vehicles. The 400 participants using guidance over eight weeks would also yield sufficient data to allow reliable estimates to be made of the proportion of journeys on which guidance was sought or accepted and, probably, of changes in the usage of particular types of link.

8.6.2 The participants' evaluation of guidance would be more problematic. It might not be possible to make reliable judgments about the opinions of 16 subgroups within the total sample. It may be necessary to accept that the precision of estimates will be lower here.

8.6.3 The arguments above relate to the assessment of only one version of the system algorithm. For each additional version to be tested, the experiment will have to run for an additional eight weeks.

8.6.4 Another assumption inherent in the above calculations is that 'novelty' and 'learning' curve effects will have been overcome before the main analysis begins. We suggest that at least four weeks be allowed for this settling down process and that any observations taken during this period could be seen as contributing to an understanding of that process. It should be noted that this settling down period would not need to be repeated for each variant of the system algorithm since the user interface would not alter.

8.6.5 The preceding discussion has concentrated on the sample size required for the participant group and has concluded that 400 should be adequate for most purposes provided that full disaggregation is not required. Similar arguments apply in determining the size of the sample control group and a group of 400 is again recommended. It might, in practice, be possible to get by with a smaller group, but only at the price of requiring more data from each member and sacrificing some of the desired disaggregation.

8.6.6 The required size of the population control group is more difficult to estimate since it depends on the distribution of some fairly obscure characteristics within the population of London drivers. We have not attempted to calculate this precisely, but initial indications are that, to do the job effectively, a sample of several thousand would be required.

8.7 <u>The Timetable for the Evaluation</u>

8.7.1 We suggest that four phases can be identified for the evaluation: A prepilot phase, a non-guidance phase, a guidance phase, and an analysis phase.

8.7.2 During the <u>pre-pilot</u> phase the participant selection software and the simulation model ought to be prepared. Although the model would not be required until the analysis phase it ought ideally to be written well in advance so that its structure can be seen to work and its precise data requirements confirmed prior to the start of the pilot.

8.7.3 Recruitment of the participant and sample control groups ought to commence at least four months before the start of the pilot scheme so that selection of the groups can be completed in good time for cars to be equipped and questionnaires administered before the pilot begins. We note that in Berlin five months was allowed for these processes and this proved barely adequate, because of delays in equipping these vehicles.

8.7.4 The population control data ought to be collected at approximately the same time as the recruitment of the participant and sample control groups.

8.7.5 The <u>non-guidance</u> phase should last at least eight weeks during which data would be collected on technical performance, travel patterns, routes chosen and involvement in incidents and accidents.

8.7.6 The <u>guidance phase</u> should last at least 12 weeks. The first four weeks would allow for driver familiarisation. During the remaining eight weeks the data listed in para. 8.6.5 would be collected. In addition, the test vehicles would be used in their shadow mode to determine the real benefits or disbenefits of guidance. The participants' attitudes to guidance would also be collected during this phase.

8.7.7 We are well aware of the pressures to determine, as soon as possible, whether or not the operators are to receive a licence to continue to operate a scheme beyond the end of the pilot. We are also concerned that the pilot should not be curtailed before it has yielded valuable information about the performance of variants on the basic Autoguide design. We suggest that the pilot could be designed with a breakpoint at twelve weeks, after which, once the data has been analysed, the decision on the licence could be taken. Experimentation could then continue (possibly supported by public funds) for a further 24 weeks. This 24 week period would enable the four variants on the system parameters to be tested.

8.7.8 It follows from the above that we envisage that the analysis phase would be in two parts. The first part would begin after the 12 weeks' experience with guidance, with the emphasis on coming rapidly to a conclusion on whether a continuation licence should be granted. The second part would follow, and would be concerned with a more thorough assessment of the lessons to be learned from the pilot.

9 THE ROLE OF TEST VEHICLES

9.1 The Proposal

9.1.1 We believe that a small fleet of test vehicles, equipped with guidance and driven full-time by employees of the organization conducting the evaluation, should be an integral part of the Autoguide assessment. They would offer the ability to provide systematic monitoring of technical performance; to conduct controlled, as opposed to random, experiments for determining the time value of benefits from guidance; and to enrich the data base of journey times in selected parts of the network in a way that would not otherwise occur in the pilot.

9.1.2 There are thus three main roles for the test vehicles; systematic performance monitoring, benefit assessment and database enhancement. It may be argued that the first of these will not be required since all the system elements will have been proven before the start of the pilot. We suggest, however, that it is necessary to monitor the technical performance of the system throughout the pilot not only to report on the development of any unforeseen problems but to provide a firm basis for interpreting all the other data collected during the evaluation.

9.2 Monitoring the Technical Performance of System Components

9.2.1 The test vehicles would be able to cover the whole guidance network at fixed and repeated intervals to observe the reliability of beacon to vehicle communication. All failures would be logged by the test vehicle driver/passenger and later compared with system information on network communications. Various kinds of in-car electronic equipment (stereos, cellular phones, CB radios, etc.) could be used to ascertain if there were any problems of interference with the in-vehicle guidance equipment.

9.2.2 The same complete coverage of the guidance network would be used to test the system's representation of the actual road network. The test vehicles would evaluate whether, for example, the system was aware of turning restrictions at particular junctions, of the channelling of traffic by lane, or of the times of access restrictions. Similarly, the test vehicles could assess whether advice was given soon enough for a particular junction, and whether the system was responding to exogenous information.

9.2.3 Although the full-scale system will not need to be able to identify guided vehicles, it will need to be able to track them so that it can acquire information on actual link times for updating of route advice. Knowing which test vehicles were at which locations at which times, it will be possible to check from the system log that the system actually obtained the information it was supposed to obtain.

9.2.4 The system will constantly need to compare detected link and journey times with predicted times. An excess of detected over predicted time may be an indication of incidents or congestion. Similarly, the reverse will indicate that traffic is moving more freely than expected. It will be important to know that the detected times are reliable. This could be done by recording them, and subsequently comparing them with the times logged by the test vehicle drivers/passengers.

9.2.5 The test vehicles constitute the best means, during the trial, of assessing the likely impact of a heavy demand on the system in one sector. The chances of two equipped vehicles arriving at a beaconed junction simultaneously are low. One could use one occasion during the trial to test the impact of heavy demand, by routeing a large number of equipped vehicles over an identical route, but this would of necessity be a once-off exercise. It would not permit the testing of hardware or software modifications should problems emerge. On the other hand, a number of test vehicles could be driven simultaneously past certain beacons at various times during the trial, so that any problems could emerge early and there would be time to test modifications. We accept that this test might more properly be conducted before commencement of the pilot.

9.3 Assessment of Benefits

9.3.1 The test vehicles are better suited than the participating vehicles to evaluate the actual journey time saved by using route guidance. Given the problem of estimating the conditions that would have been met on routes that are not actually used on a given day and the extreme unlikelihood of finding a pair of participants attempting to make exactly the same journey at precisely the same time, one under guidance and the other not, it will not be possible to measure time saved directly from observation (system or questionnaire) of the participants. For regular journeys a comparison of the participants' times under guidance with the times experienced on the same routes prior to guidance will provide some information, but it will suffer from problems of seasonality and daily variation.

9.3.2 The test vehicles, on the other hand, could provide the required information in a structured manner. We suggest that, for a sample of participant journeys, the participant vehicle should be shadowed by a test vehicle. This would be done as follows.

9.3.3 On day one the participant vehicle would operate without guidance, alongside a test vehicle using guidance. Since the test vehicle would attempt to follow guidance completely, the pair of trip times obtained would provide the difference between a vehicle obeying all guidance instructions and one not using guidance at all.

9.3.4 On day two the participant vehicle would operate under guidance, while the test vehicle followed the participant's route of the previous day. The participant would be instructed to use his discretion as to when to follow the guidance. He would follow it only when or where it seemed to him to provide the best choice. If a ratrun seemed preferable, it should be used. The pair of trip times obtained would provide a comparison between origin-destination times without guidance and times between the same points under not perfect but real-world use of guidance. The trip by the test vehicle following the intended route might not be a perfect representation of the route a participant would have followed without guidance, because had the participant been driving on that day, he might have amended his route in response to real-time traffic conditions. But it should be possible to confirm with the participant whether or not he would have used the same route given the actual conditions. This is therefore probably a minor flaw.

9.3.5 One could thus obtain both maximum potential time savings from the use of the guidance being offered and realistic estimates of likely savings in normal use. This information is clearly of central significance to the evaluation and we are convinced that it is only by using the test vehicles in this way that accurate estimates can be obtained.

9.3.6 In 8.5.6 we have suggested that the use of test vehicles in this way would require a pool of between 8 and 16 such vehicles depending on whether one were to include assessment of both 'perfect'and 'realistic' adherence to advice.

9.4 Enhancement of the Travel Time Database

9.4.1 Although efforts can be made in selecting the participants (see chapter 10) to ensure that some parts of the network have a concentration of equipped vehicles and thus a higher than average probability of having a relatively rich database of travel times reported by the equipped vehicles, it will not be possible to guarantee that this can be achieved to the required degree.

9.4.2 Given a pool of test vehicles at the disposal of the evaluation team it should be possible, on selected occasions, to deploy them so as to ensure the required frequency of update information on journey times. It is not possible accurately to estimate how many test vehicles would be required for this purpose until the distribution of participant vehicles is known and whether or not, in addition to the 400 participant vehicles, any fleet vehicles are to be equipped to transmit travel times (see para. 12.5.1).

9.4.3 On the pessimistic assumption that no data can be relied on from the participant or fleet vehicles, and assuming that we wish to achieve updates every 5 minutes, we can calculate that for each five mile stretch of road one would need between four and eight vehicles depending on the level of congestion. If, as is likely, updates are required for at least one alternative route in the corridor this implies a pool of at least eight to sixteen test vehicles. If a pool of this size cannot be justified on economic grounds it would be necessary to rely on part of the data coming from the participant or fleet vehicles. The selection of participants and/or fleets, should in any event be designed with this factor in mind.

9.5 Number of Test Vehicles Required

9.5.1 We have calculated that the desirable number of test vehicles would be sixteen but that, at the cost of some loss of data, eight might suffice.

10 THE ROLE AND SELECTION OF PARTICIPANTS AND CONTROL GROUPS

10.1 The Selection of Participants

10.1.1 The discussion in Chapter 8 has indicated that the participants ought to be selected so as to ensure coverage of various dimensions of interest. These include vehicle type, age and sex of the driver, familiarity with the network and the type and location of journeys. It has been argued in 8.3 that we will require the participating population to include sufficient numbers of people in each of: two sexes, three age groups, two journey purposes, two SEGs, five driving styles, four vehicle types, three travel patterns and twelve journey area categories.

10.1.2 While these considerations have focused on the driver, it is important to note that the selection will actually be among vehicles, and the sample will thus provide information on all drivers of those vehicles. This should help to provide further information on use by different types of driver, with differing levels of experience provided, of course, that the driver concerned can be identified.

10.1.3 In addition to the requirements relating to the experimental design the selection of participants will be constrained by practical considerations:

- firstly that they agree to having their vehicles monitored (either via the IVU's regular reports to the beacons or via the IVU record being stored on a cassette tape);

- secondly that they agree, and show ability, to complete journey logs and questionnaires as and when required;

- thirdly that their vehicle is readily compatible with the IVU equipment (it may be advisable to have an approved list of makes, models and vintages);

- fourthly that they expect to use their current vehicle for a significant amount of travel in the scheme area throughout the duration of the pilot.

10.1.4 We anticipate that, as in Berlin, there is likely to be a ready supply of drivers in London who would be keen to participate in the pilot and who would fulfil the criteria set out above. We suggest that they be recruited in the same way as was done in Berlin - via regional press advertisements augmented by appropriate coverage on TV and radio. We further suggest that the required spatial concentration of participants might be achieved by advertising in the local press, garages, car accessory shops and, perhaps, by approaching local firms. A concentration of users of given links might be achieved by a stop-line survey but we doubt that the expense would be necessary or worthwhile. If, as we suggest, a number of commercial and fleet vehicles are to be included in the evaluation then the initial contact would be by letters to targeted firms followed by a direct personal approach.

10.1.5 Having responded to the recruitment campaign, respondents would be sent a screening questionnaire to determine whether they meet the required criteria. We suggest that the results of this questionnaire be processed by pre-prepared software so as to minimise the time taken to select the optimum group of participants. The same software could be used to select the sample control group (see 10.3).

10.1.6 We suggest that, in addition to the 400 vehicles which form the core of the evaluation, it will be worth approaching any people or firms who pay to have their vehicles equipped with a view to their contributing to the evaluation. It may be that some such people would be very willing to contribute information and some might be prepared to allow their IVU records to be analysed. As an incentive to such participation the person or firm might be offered a personal copy of part of the evaluation report including a quantitative assessment of the benefits that Autoguide had provided for his vehicles. Any data so derived would need to be analysed separately, since it would be subject to selection bias.

10.2 Information to be Sought from Participants

10.2.1 We suggest that information be sought from participants in five ways:

- the screening questionnaire
- an initial questionnaire
- logs kept of selected journeys
- information from the IVUs (see chapter 11)
- attitudinal questionnaires.

10.2.2 The screening questionnaire will obtain: personal details of all drivers (name, address, age, sex), vehicle details (make, model, vintage, when to be sold), travel pattern details (monthly mileages in specified parts of London at specified times of the day and by driver, journey purposes, regular journeys, amount of route finding, amount of destination finding) and route finding behaviour (route choice criteria, whether or not they use alternative routes, familiarity with the network). The travel pattern data might best be required in the form of a diary or log in order to ensure that drivers are capable of providing information in such a way.

10.2.3 The initial questionnaire would be conducted after the vehicle is equipped but before any guidance has been received. It would seek details of involvement in traffic accidents during the previous six months, attitudes to new technology and the benefits that they expect to gain from Autoguide (including an estimate of how often individual drivers expect to use it and for what types of journey). None of these questions can realistically be asked in the screening questionnaire since, at that stage, would-be participants will be attempting to impress the organisers.

10.2.4 Driver logs of equipped vehicles would be required. We would want drivers to record vehicle usage via daily odometer readings and to record involvement in any accidents, also on a daily basis. For journeys on selected days only we would require the following to be recorded for all journeys on that day:

- date and time
- driver
- origin, destination and purpose
- familiarity with the route and prior knowledge of the precise location of the destination
- any problems or apparent faults experienced in attempting to follow guidance
- whether guidance was followed and if so, to what extent
- what route would have been followed without guidance (traced on a map)
- estimate of how long that route would have taken
- rating of the guidance given
- location and description of any accidents or near-misses in which the vehicle was involved (useful not only in its own right but also to calibrate the technique for the automatic detection of near misses via records of violent deceleration - Para. 6.8.3).

The journeys for which detailed logs would be required would be prespecified in terms of their date. We consider that sufficient data would be obtained by requiring each participant to provide detailed logs for all journeys on one day in twenty.

10.2.5 The attitudinal questionnaires would be administered a month after the vehicle was first able to receive guidance (i.e. after an initial period of familiarisation) and would be repeated at approximately two-monthly intervals. If the nature of guidance received by a given participant changed during the pilot (as envisaged in 8.2.12) these questionnaires would obviously be phased accordingly. The information sought would include each driver's estimate of:

- the usefulness and reliability of advice received for different types of journey
- the value of the system to them (willingness to pay)
- any resulting changes in travel patterns (other than 'novelty' effects)
- detailed comments on ergonomic aspects of the IVU
- assessments of stress, fatigue and safety (compared to other methods of route finding)
- comments and suggestions on extra features, modifications and enhancements that might be desirable (including the desirability of extending the scheme in London or elsewhere).

10.3 <u>The Sample Control Group</u>

10.3.1 The purpose of the sample control group is described in Section 8.4. It will obviously have to be matched as closely as possible to the participant group and will have to be observed over the same period of time as the participants. If full statistical analysis is required the control group will also have to be equal in size to that of the participant group. These requirements imply that the sample control group should be selected at the same time and from the same population (respondents to the recruitment campaign) as the participant group. Ideally it would be desirable to assign suitable respondents randomly to either the participant group or the control group.

10.3.2 Information required from the sample control group would be of two types; vehicle usage patterns, and involvement in accidents or near-misses. We propose that the information be obtained via an initial questionnaire (which would seek their recent accident history) and travel logs equivalent to those described in 10.2.4 for the participant groups. The information required on a daily basis would be a record of any involvement in accidents and a record of the vehicle's odometer reading. For journeys on selected days only we would require the following information about all journeys on that day:

- date and time
- origin, destination and purpose
- familiarity with the route and prior knowledge of the precise location of the destination
- route followed (traced on a map)
- time taken and distance travelled
- location and description of any accidents or near misses in which the vehicle was involved.

10.3.3 The sample control group cannot be expected to participate in the evaluation without some reward and, perhaps, compensation for not having been chosen to have their vehicles equipped. We suggest that a financial inducement be provided in the form of a payment for data supplied and/or inclusion in some form of lottery or prize draw.

10.4 <u>The Population Control Group</u>

10.4.1 As was explained in Section 8.4., the population control group is necessary to correct any forecasts of the performance of large scale schemes for the fact that the participants in the pilot are not likely to be representative of Londoners at large; most notably in that they are likely to be more technology oriented. The population control group should be drawn at random from London drivers and should ideally be found by random selection from addresses and every care should be taken to maximise the response rate so as to minimise response bias.

10.4.2 From this population control group we would seek information equivalent to that in the screening questionnaire for participants - namely: personal details; vehicle details; travel patterns and route finding behaviour. This would ideally be supplemented by monthly odometer readings and records of involvement in accidents.

10.4.3 We are well aware that this exercise may be regarded as unjustifiably expensive given its apparent peripherality to the main evaluation task. We suggest that it might be acceptable to rely on a scaled down version of the sample albeit with loss of statistical significance but we re-emphasise the importance of producing an unbiased sample for the population control group. As an alternative, estimates of the required data might be made on the basis of existing data sources such as GLTS, various pieces of research on route choice criteria and car use seasonality data but they would not be as reliable as having a proper population control group.

10.4.4 If there is to be a population control group, its members might need to be motivated to make their monthly returns via some form of incentive similar to that in 10.3.3.

11 INFORMATION FROM THE IN VEHICLE UNITS

11.1 Collection of the Information

11.1.1 The attraction of using the IVUs as a source of data is that the process would be more or less automatic and would not require any involvement from the participant drivers. The volume of data from all the equipped vehicles is potentially massive (at least 4000 journeys per week) and the problem is likely to be one of deciding how to select data for processing.

11.1.2 If, however, for reasons discussed in section 4.1.5, it is not thought desirable to allow it to become generally known that the system can identify individual drivers, then it may prove necessary to adopt a very visible method by which to record the IVU data. The best method would probably be to arrange for the IVU to have a data port which could download data to an ordinary cassette tape recorder. This procedure would have obvious cost implications and would require the driver to perform the extra task of downloading the data.

11.1.3 Whether it is recovered via the roadside beacons or via cassette recorders the data which is required from the IVUs will be of 5 main types:

- a record of whether or not guidance was sought for a particular journey;
- a record of the route taken (see 11.1.9);
- a record of the time taken to complete the journey (if possible less the time spent at intermediate destinations such as petrol stations);
- a record of times spent in total and at slow speed on each link;
- a record of locations at which excessive decelerations were detected by the IVU.

11.1.4 In addition to these it would be wise also to record the recommended route devised by the IVU in the light of information from each beacon because, although this information could be reconstructed from the system itself, to do so would incur significant storage and housekeeping problems. We recognise that transmission of the route advice back to the system might overload the vehicle-to-beacon link and suggest that this is a positive reason why IVU records should be stored on tape.

11.1.5 While the cassette may provide an adequate method for storing the journey and person specific data, information will still be required directly from the vehicle. In particular, the system itself will require link journey times to be transmitted in real time.

11.1.6 The journeys for which IVU data is to be stored for subsequent analysis would include all those for which the drivers will be required to provide written logs. We have suggested in paragraph 10.2.4 that such data will be required for all the journeys made on one day in each twenty (=4 weeks). Assuming 400 pilot vehicles this will produce about 1000 such journey records per month of the pilot. In addition to this we envisage a requirement for a sample of journeys from test vehicles and shadowed participants or, as an alternative, and inferior, method of estimating real savings, a sample of regular journeys made before guidance is available and again after it is available. If the test vehicle method is used we would expect this to produce a further 100 or so journey records per month of the pilot. If the before and after method is used we would envisage an extra 2000 or so journey records during the 'before' period.

11.1.7 It will be recalled from section 9.2 that the test vehicles were also to be used to monitor the technical performance of the system, in which case we would be receiving a further 460 or so hours of journey data per month.

11.1.8 A further use for the IVU data would be to provide examples of information for planning purposes which could be made available, or marketed, to government and private concerns. The data in question would, of course, have to be transmitted by the IVUs to the system log. While such data would need to be treated as confidential at the individual level, it could provide useful aggregate indicators of flow and journey patterns by time of day, and locations where congestion was particularly severe or where conflicts were unusually frequent.

11.1.9 It is particularly important to be able to identify usage of links that are not known to the system. This raises technical problems because the IVU cannot log where it has been in terms of link numbers (as it does for links which are known to the system) and would instead have to store vectors to describe its path. This would probably result in greater volumes of data than would be required to describe travel over a similar distance on <u>known</u> links and this may result in more data than can be transmitted back to the beacons. Such data would therefore have to be stored on cassette. Analysis of the data would also be problematic (see 11.2.2).

11.2 <u>Processing of the Information</u>

11.2.1 The handling of data from IVUs and cassettes is not necessarily a straightforward process. We have therefore set out our initial suggestions for ways in which it might be processed.

11.2.2 The data on routes followed by vehicles will be of two types - that which relates to travel on links known to the guidance system (not all of which are necessarily on the guidance network) and travel on links not known to the system. Data on the 'known' links will be more reliable since it can be corrected for accumulated dead reckoning error while data on 'unknown' links cannot be so corrected. We would hope that the 'known' network can be as extensive as possible so as to reduce this problem. Manual analysis of the vectors followed on unknown links might be possible by means of graphical displays of the vectors superimposed on a detailed map.

11.2.3 As was indicated in Section 9.3, we would recommend that some analysis be conducted to compare the routes which the IVUs record (on the known network) with those recorded by the drivers of test vehicles. This would perhaps be most effectively done using computerised network representations and graphics whereby routes recorded manually on maps can be input to the computer via a mouse or digitiser.

11.2.4 If comparisons are to be made of routes actually taken with and without guidance, similar graphics would be a very useful aid during the analysis even though the 'hard data' produced for the evaluation report might relate only to types of link used, distances travelled and time taken.

11.2.5 If estimates are to be made of travel times that would have been experienced by a driver at a given time on a given day (e.g. in the comparison of routes taken for regular journeys before and after the advent of guidance), then additional software will be required. At its simplest this software might assume that historic data for the links in question provided a reasonable estimate and at is most sophisticated it would make use of any journey times reported by other vehicles who happened to use some of the links in question at the time in question. Given the extreme unlikelihood that such data would be available for all the links in question at the appropriate time on the day in question, the software will most probably have to rely on historic profiles updated by such data as is available. Such procedures would presumably be built into the system's own journey time update mechanisms but it must be recognised that to rely on them blindly would be to risk exaggeration of the benefits of Autoguide.

11.2.6 As a check on the accuracy with which the system can estimate link travel times we have suggested, in Para. 9.2.4, that the times recorded by test vehicles be compared with those predicted by the system. We suggest that this test could be further strengthened by comparing, for a sample of journeys, the link times recorded by the participants' IVUs with those previously predicted by the system. This would require that the system should store part of its predicted link travel time table for use in this analysis.

11.2.7 The data on the location of excessive deceleration would be analysed in three ways. Firstly (if further research proves it necessary) it would be compared with the driver's recollection of dangerous situations in order perhaps to 'calibrate' the deceleration threshold appropriate to his driving style. Secondly it would be quantified for input to an analysis of any changes in the amount of such deceleration before and after the introduction of guidance. Thirdly it would be plotted onto a map so as to provide potentially valuable data for use in road safety work.

11.2.8 The data on travel times, time spent at slow speeds, and links followed transmitted by the IVUs to the system log for planning purposes could similarly be plotted on maps to indicate journey patterns, routes taken, and locations of particularly serious congestion.

12 INFORMATION FROM ADMINISTRATIVE AND OTHER SOURCES

12.1 Scope

12.1.1 The previous three chapters have discussed the information required from the test vehicles, from the participants and control groups and from the in-vehicle units. We will now consider the information required from the system itself and from those involved in the establishment and operation of the pilot.

12.2 Information from the System Log

12.2.1 The system log ought obviously to be organised to record information on its performance, in terms of down-time of its components: the central computers, the communications systems and the beacons.

12.2.2 It should also record details of which algorithms were being used (at what time and for which categories of vehicle), to update the travel time files and to calculate optimum paths. A central record should also be kept (for each link) of the rate at which updated link travel times are being received from equipped vehicles. A similar record should be kept of any input of exogenous information (road works, major incidents etc.

12.2.3 Although the record of which links have been followed will also be available in the IVU data (as described in Chapter 11), it may be appropriate to take it from the system record. This is because, whereas the IVU data may for practical reasons be recovered only for a sample of journeys, the system log will be able to log link usage continuously because it will be receiving the data from the beacons along with the link travel times.

12.2.4 As was indicated in para. 11.1.3, we suggest that the record of what guidance was transmitted to each vehicle might be most efficiently stored by the IVUs rather than by the system log. It may, however, be necessary to reconstruct the advice that a driver who did not request advice, would have received, had he requested it. Similarly it may be necessary to reconstruct the advice that a driver, who having requested advice did not follow it completely, would have received had he followed it completely. In either case the IVU will obviously not have the necessary information and it would have to be recovered from the system log. This would, in fact, be a very tedious process since it would involve estimating the time at which the vehicle would have reached various beacons before ascertaining the advice that the beacon would have been transmitting at that time. It would also require the storage of a substantial amount of data, including a record of all advice from all relevant beacons throughout the period of interest.

12.2.5 As was indicated in Section 6.3, one of the measures of the quality of advice given by the system would be a comparison of network times actually experienced with those assumed by the system as the basis of its advice. Records should be kept in the system log of the extent to which link times transmitted by equipped vehicles differ from those previously estimated by the system. Although this data would not be required for all links at all times, it ought to be disaggregated by type of link, time of week, and stage within the pilot.

12.2.6 Although, because of the limited number of equipped vehicles, they will be of limited use during the pilot, various items of information from the system log are of potential value to the network planners, traffic and road safety engineers. At some stage during the pilot the system log ought to be interrogated to show examples of origins and destinations; link flows of equipped vehicles; travel times on links and for journeys; the locations of congestion; and the locations of concentrations of traffic conflicts. Software should be written to allow this data to be effectively presented on a VDU or in hard copy as appropriate.

12.3 Information from the System Operators

12.3.1 The main requirement will be for details on costs. This will obviously require the full co-operation of the system operators. Information would need to be disaggregated into broad tasks such as: establishing and maintaining the network, installing and maintaining the beacons, installing and maintaining the communications links, acquiring and maintaining the central computers, devising and maintaining the various algorithms, inputting exogenous information and producing planning information.

12.3.2 In addition it will be helpful to obtain information on technical performance of the system, the adequacy of the network data base and any adjustments to the network in the light of roadworks and other diversions. It must be accepted that it will be a matter of judgment on the part of the operators as to what items of information are provided.

12.4 Information from Third Parties

12.4.1 A record ought to be kept of any resources expended by such bodies as the Police, highway authorities and the Meterological Office in the establishment and operation of the pilot. Resource inputs for which a payment is made by the system operators (e.g. access to meterological forecasts or advance notice of roadworks) need not be included provided that they are fully and separately identified within the costs in 12.3 above.

12.4.2 Local authorities and others may also be able to comment on the technical performance of the network database and of the procedures for incorporating information on roadworks, diversions and other emergencies. They would also be asked to provide accident data for evaluation purposes.

12.4.3 We would also expect local authorities to comment on any unexpected side effects of the autoguide system (interference with communications systems etc.) were they to arise.

12.4.4 The evaluation of the pilot should allow for the commercial value of planning data from the system log (as described in para. 12.2.6) to be assessed. This will require examples of what could be produced to be shown to potential buyers. These may include highway authorities, the police and also agencies with a commercial interest in flow information for purposes such as deciding on the siting of advertisement hoardings or petrol stations.

12.5 <u>Information from Equipped Fleets</u>

12.5.1 Chapter 7 raised the possibility of equipping one or more fleets of vehicles which make intensive use of the network with transmit-only devices to provide enhanced travel time data. The Germans are using taxis in this way.

12.5.2 We doubt that taxis would be appropriate in London given their frequent stops, U-turns and use of minor roads. We also anticipate that a substantial fleet would be needed to provide the enhancement in travel time data which would be required. Nevertheless, we consider that it would be worthwhile to investigate the use of large fleets, such as post office vehicles, as a potential data source. We have not ourselves been able to pursue this issue.

13 EVALUATION PRIORITIES AND COSTS

13.1 Background

13.1.1 The last seven chapters have reviewed the possible information sources for meeting the data requirements identified in Part I. As noted in Section 4.4, there will inevitably be pressures to reduce the costs and the timetable of the evaluation to a minimum.

13.1.2 In this chapter we address the problem by assigning priorities to these data items and estimating costs for those which we argue should be of high priority. This costing exercise is inevitably incomplete and approximate. In particular, we are unable to provide reliable estimates of costs for many of the hardware requirements.

13.1.3 As noted in Chapter 1, a prime focus of our brief was to develop a common methodology for the pilot projects in Berlin and in London. We have taken this as a principal justification for assigning data items a high priority, and start, in Section 13.2, by considering the opportunities for comparability with Berlin.

13.1.4 The remaining justifications for high priority items are given in Section 13.3. Cost estimates follow in Section 13.4.

13.2 Comparability with Berlin

13.2.1 Comparability of evaluation does not require or imply that the route guidance systems themselves, the manner of their implementation, or the locations in which they are installed, should be identical. Indeed, diversity in these respects can increase the value of the comparison. However, comparability of evaluation does imply that, where possible, similar indicators should be produced.

13.2.2 London covers a larger area than does West Berlin, it has a larger core area with a denser network of streets and a less regular network pattern. London experiences more severe congestion than does Berlin. Another major difference is that, at any given time, London has a larger number of drivers who are unfamiliar with the part of the network that they are using. West Berlin, by contrast, has a closed network and a more experienced pool of drivers.

13.2.3 All these differences are likely to affect the operation and usefulness of route guidance and it will be of value in the design of future schemes to note any such differences and ascertain their cause. Unfortunately, it will not be possible to be sure which of the various network conditions or driver characteristics has contributed the most to any difference in the performance of Autoguide and LISB. However, provided that the guidance networks and beacon locations are carefully chosen, the participants carefully selected and the evaluation data sufficiently disaggregate, some conclusions may be able to be drawn.

13.2.4 The specification of LISB is now complete. That of Autoguide is still under development. It is quite conceivable that differences will exist in terms, for example, of the travel-time data-base up-date procedures, the route choice criteria and the nature of any secondary network.

13.2.5 Any such differences would, of course, affect the performance of the route guidance system and, as with network conditions and driver characteristics, these differences in performance will be of great interest provided that they can be attributed. Such attribution will be difficult unless the Autoguide scheme is designed with comparability in mind. This might mean, for example, ensuring that the LISB route choice algorithms are incorporated in Autoguide for at least a part of the time for some of the respondents.

13.2.6 In the interests of comparability we would, of course, ideally ensure that precisely the same measurements were made using precisely the same procedures in London/Autoguide and Berlin/LISB.

13.2.7 Unfortunately, this is not achievable in practice. There are three main reasons for this: cultural differences, administrative/political differences and professional/technical differences. Taken together they result in it being impractical to specify the same evaluation scheme for London as has already been decided for Berlin.

13.2.8 The cultural differences are perhaps the least significant. The Berlin evaluation places a considerable emphasis on the measurement of attitudes using techniques which are specific to a given language or culture and there would be little point in attempting to reproduce them precisely in London.

13.2.9 The administrative/political differences result from the decision to operate the London Autoguide system in the private sector. The consequence of a reliance on the private sector for the finance to support the pilot scheme will be a quite understandable reluctance on the part of the operators to allow their competitors to gain commercial advantage from the open publication of the results of the pilot scheme. Thus, London evaluation has to be planned against a background of anticipated reluctance of the operators to allow the release of sensitive information. For example, while the government may wish to know whether the system produces real benefits for its users, the operator may not wish to divulge information if it suggests that the benefits are meagre.

13.2.10 One implication of private funding in London is concern by would-be operators that the potential market for route guidance could be damaged if it became widely known that the system could track individual drivers. If such concerns were effectively to rule out the kind of centralised monitoring that is proposed as a central component of the LISB trial's evaluation, this would obviously reduce the opportunities for comparison between London and Berlin. We have suggested an alternative method of storing IVU data which may overcome this problem.

13.2.11 The professional/technical differences between the evaluations arise largely from the opportunities which our present study has provided to reassess the requirements for evaluation. This has enabled us to develop more fully the concepts for the use of IVUs (for safety monitoring), test vehicles, driver logs and a sample control group. Although our technical proposals differ in these ways from those for Berlin, most of the data sources will be similar in nature in the two evaluations.

13.2.12 We conclude that, despite these differences, the pilot route guidance systems in Berlin and in London provide substantial opportunities for co-operation in the evaluation of the systems and for the consequent improvement in our knowledge, not only of route guidance but of driver behaviour and network performance generally.

13.3 Evaluation Priorities

13.3.1 This analysis of the opportunities for comparability suggests that, at the very least, the London pilot evaluation should include information from system logs, IVU data, driver comments, operators and third parties, all of which will be used as information sources in Berlin.

13.3.2 We are satisified that the provision of test vehicles is an essential element in our strategy for estimating benefits from the guidance provided, and will offer a substantial improvement on the method planned for Berlin. We consider, therefore, that they should be a high priority element. We do, however, acept that it might not be possible to justify a pool of sixteen such vehicles - our costings are based on a pool of only eight.

13.3.3 We consider that the use of driver logs is essential if more detailed information is to be obtained on performance of the system, and responses to it, on specific journeys, and we intend as part of our separate study funded by SERC, to obtain such data in Berlin. Our proposals for a device to download journey-specific information should help considerably in providing the necessary information.

13.3.4 Control groups represent a more difficult issue. We are convinced that a sample control group is crucial in assessing benefits, user responses and safety implications, and would allocate it high priority. The population control group could be considered to be more peripheral; however, without it it will be difficult to ensure that the effects of route guidance are not being exaggerated by an undue emphasis on those who are attracted to new technology. We accept, however, that this could be given low priority, and we have therefore not costed it.

13.3.5 An equipped fleet making intensive use of particular parts of the network to provide additional travel time data would be a valuable addition to the evaluation. However, it would be likely to be expensive, and for this reason we accept that it should not be given high priority. We have therefore not costed it.

13.3.6 The other dimension in which costs could be saved is in timescale. We have argued in section 8.6 that the minimum requirement is for eight weeks without guidance and twelve weeks with guidance, together with the time required for the preliminaries of sample selection and equipping vehicles. We suggest that it should be possible to make a decision on a full scale system after twelve weeks' guidance and the analysis of 12 weeks' data, but that it would be highly desirable to obtain information for a further 24 weeks on the effects of varying the basis for guidance. We consider this important enough to have costed it separately.

13.4 Costs

13.4.1 In estimating costs, we have only provided estimates for the work involved in setting up the evaluation. We have no basis for assessing the costs of computer storage, nor have we attempted to estimate the costs of data analysis and interpretation. We consider other costs for each of the data sources in turn, where relevant for both a 20 and a 44 week equipped period.

13.4.2 System logs We have not been able to estimate the costs of the necessary software development or data storage for this. It would probably be necessary for one person to be employed half time in collating this information for a period of either 25 or 49 weeks (including 5 weeks for induction/familiarisation). A technician would cost £6,600 for 25 weeks plus an additional £6,400 for the extra 24 weeks.

13.4.3 IVU Data We have assumed that 400 vehicles would anyway be equipped as part of the Pilot. There would be an additional cost of equipping them with cassette recorders to download journey information. We have assumed a cost of £50 per vehicle to include tapes, on the understanding that the technology for downloading is readily available. We have made no allowance for the developments necessary to record sudden decelerations. We have again assumed that a technician would be required half time to manage this data and again that 5 weeks would be required for induction/training. This gives an overall cost for the twenty week experiment of £26,600 plus an additional £6,400 for the extra 24 weeks.

13.4.4 Test Vehicles The high priority requirement is for eight test vehicles for either 12 or 36 weeks; they would not be required in the non-guidance phase. We have assumed that each vehicle would be hired at £100 per week, and equipped with an IVU at a cost of £500. Drivers would cost £250 per week and fuel £50 per week. Thus, each vehicle would cost £5,300 for 12 weeks, or £14,900 for 36 weeks. Eight vehicles would cost £42,400 for the 12 week period plus an extra £76,800 for the additional 24 weeks.

13.4.5 Participant Selection Participant recruitment advertisements might be expected to cost £2,000, and an initial screening of perhaps 5,000 applicants £10,000 in surveys. Two person months would be required in developing the necessary software and managing the selection process and a further two person months in processing the data. We have assumed a total staff cost of £5,000, giving a total cost of £17,000.

13.4.6 <u>Driver Logs</u> One log would be required from each driver every 20 days. This implies 800 (2x400) logs during the pre-guidance phase, 1200 (3x400) during the 12 week period and an additional 2400 (6x400) during the 24 week period. Design and piloting of these might cost £2,000, and administration £5 per log. In addition, one person would be required full time to manage the participant group, at a cost of £350 per week. Including a provision for induction/training the cost for the 20 week period would be £20,750 plus £20,400 for the additional 24 weeks.

13.4.7 <u>Driver Comments</u> Two questionnaires would be administered to each driver for the shorter experimental period; a further three (one for each test) might be required for the longer period. We have assumed a cost of £5,000 for design and piloting and £5 per questionnaire for administration. This gives a cost of £9,000 for the shorter period, plus £6,000 for the additional 24 weeks.

13.4.8 <u>Sample Control Group</u> This would be selected as part of the process in para. 13.4.5 at no extra cost. It would require the driver logs in para. 13.4.6 and the initial questionnaire in para. 13.4.7. There would, however, be an additional cost of perhaps £2,000 in maintaining the interest of the group for 20 weeks and a further £2,400 to maintain it for a further 24 weeks. Assuming that it can be managed within the resources included in 13.4.6, this gives an overall cost of £14,000 for the shorter period with a further £18,000 for the additional 24 weeks.

13.4.9 <u>Operator and Third Party Data</u> We assume that one person would be required part time to maintain contact with the operator and third parties at a cost of £4,000 for the 20 week period plus £4,800 for the additional 24 weeks.

13.4.10 <u>Lower Priority Items</u> We have not estimated for the population control group or the equipping of a fleet of vehicles, which we see as lower priority items. Similarly, we have costed for eight rather than sixteen test vehicles.

13.4.11 For all the above items, it will be necessary to employ a senior project manager full time at perhaps £500 per week. Including provision for induction this would cost £10,600 for the shorter period plus £12,000 for the additional 24 weeks.

13.4.12 These tentative estimates are summarised below. They indicate that, excluding the costs of the IVUs, the software development, data storage and analysis, evaluation would cost £150,950 for a 12 week guidance period, and an additional £150,800 for a further 24 weeks of experimentation.

Equipped Period	20 weeks	Extra cost for a further 24 weeks
Item	£	£
System logs	6,600	6,400
IVU data	26,600	6,400
Test vehicles	42,400	76,800
Participant selection	17,000	0
Driver logs	20,750	20,400
Driver comments	9,000	6,000
Sample control group	14,000	18,000
Operator and third party data	4,000	4,800
Management	10,600	12,000
TOTAL	£150,950	£150,800

14 RECOMMENDATIONS

14.1 Recommendations for Evaluation

14.1.1 The evaluation should be designed to supply the full list of information requirements listed in para 5.2.

14.1.2 Some of these data items may be commercially sensitive. It will be necessary for the Secretary of State to decide, in discussion with the operator, which items should be publicly available. (para. 5.9).

14.1.3 It is also essential that the pilot provide the basis for simulating the effects of various types of full scale system, in London and elsewhere. (para. 5.3).

14.1.4 The prime data sources should be the system log, IVU data, test vehicles, driver logs, driver comments, a sample control group and information from operators and third parties. (Section 13.3).

14.1.5 While of lower priority, consideration should be given to the use of a population control group and an additional fleet of equipped vehicles. (Section 13.3).

14.1.6 A total of 400 participant vehicles will be sufficient, provided that they are carefully selected. The sample control group should be equal in size and selected in the same way. (para. 10.3.1).

14.1.7 If a population control group is used, a very substantial sample will be required. (para 8.4.5).

14.1.8 A minimum of eight equipped test vehicles is required. If resources permit, this should be increased to 16. (para. 9.5.1).

14.1.9 The minimum time table should involve four months for participant selection and equipping vehicles, a period of eight weeks without guidance and a period of twelve weeks with guidance. (Section 8.5).

14.1.10 It should be possible to make a decision on whether to proceed to a full scale system at the end of this period, once the data collected has been analysed. It will be important for the Department to specify, in advance, the basis on which such a decision is to be made. (para. 5.13).

14.1.11 It is highly desirable for the evaluation then to be extended for a further 24 weeks to enable variants of route guidance to be tested. We consider it appropriate for the costs of this stage in the evaluation to be met by the public sector. (para. 8.6.7).

14.1.12 In order to be seen to be objective, it is essential that the evaluation be conducted by the public sector, preferably under the auspices of TRRL. (para. 5.12).

14.2 Specification of the Pilot

14.2.1 The specification should include provision of at least 8 test vehicles. (para. 9.5.1).

14.2.2 The specification should include provision for equipping the pilot vehicles with cassettes and downloading data, unless the issues of privacy and data transfer capacity can be resolved in other ways. (para. 11.1.2).

14.2.3 It is particularly important that the network and beacon locations be designed to provide variety in the density of each within the 'wedge' generally and preferably within the central, inner and outer parts of the 'wedge'. (paras. 8.2.6 - 8)

14.2.4 The design of the guidance algorithm should permit as many as possible of the variants identified in para. 8.2.12 to be tested.

14.3 Research and Development

14.3.1 A key requirement on which work is needed urgently is the development of the simulation model for assessing the effects of full scale schemes. (para. 7.5.1).

14.3.2 Research and development is urgently required to test the ability of the IVUs to record rapid decelerations and to determine appropriate threshholds for the classification of such decelerations. Research may be required to demonstrate the relationship between various levels of deceleration and 'near misses'. (para. 11.2.7).

14.3.3 Development work is required to provide a means of downloading IVU data to cassette if this procedure is selected (para. 11.1.2).

14.3.4 Software will be required for the selection of participants and for the processing of data from the system log and the IVUs.

14.3.5 More generally, further work is required on the development of procedures for analysing and interpreting the data to be collected.

14.4 General

14.4.1 Although there are marked differences between plans for evaluation in Berlin and London, there are several opportunities for comparison of results. Every effort should be made to benefit from these.

14.4.2 While we fully accept the reason for establishing the London pilot in the private sector, there are substantial public benefits to be gained from a thorough evaluation of the pilot and variants on it. The Department should be prepared to finance a large part of the evaluation on this basis, and should not permit commercial interests to restrict unduly the opportunities which the pilot provides.

15 REFERENCES

Bates, J., Dix, M and May, A. D. (1987) Travel time variability and its effects on time of day choice for the journey to work. Proc. Seminar C, PTRC Summer Annual Meeting, Univ. of Bath.

DTp (1988) Autoguide consultation - Pilot stage proposals, DTp, London.

HMSO (1987) Road Accidents Great Britain, 1986.

Jeffery, D. J., Russam, K. and Robertson, D. I. (1987) Electronic Route Guidance by Autoguide: the Research Background. Traffic Engineering and Control, October, 1987.

May, A. D. (1987) Route Guidance in Berlin: The potential for collaborative research. ITS Technical Note 214.

Nicholson, A. J. (1987) The Estimation of Accident Rates and Countermeasure Effectiveness. Traffic Engineering and Control, October, 1987.

Spicer, B. R. et al (1980) Variations in Vehicle Conflicts at a T-junction and Comparison with Recorded Collisions. TRRL, SR 545.

TRRL (1987) Highway Safety: The Traffic Conflict Technique, IHT 1987.

Von Tomkewitsch, R. (1987) LISB: Large Scale Test of Navigation and information System, Berlin. ITTT Seminar, PTRC Summer Annual Meeting, Bath.