Fourth Framework Telematics Applications Programme - Transport Sector

CONVERGE Project TR 1101

Deliverable 2.4.1

Checklist for Preparing a Validation Plan: Updated Version

Issue 3
Deliverable 2.4.1

Checklist for Preparing a Validation Plan: Updated Version

Issue 3 (11 September 1998)

Principal Authors: Dr David Maltby University of Salford
Steve Morello ISIS S.A.
Ken Perrett Transport Research Laboratory
Jean Hopkin Transport Research Laboratory

CONVERGE - Validation Quality Participants:
ISIS S.A., Lyon, France
Salford University Business Services, UK
ERTICO s.c., Brussels, Belgium
CERT-ONERA, Toulouse, France
CSST, Turin, Italy
Heusch Boesefeldt, Hamburg, Germany
SETRA, Bagneux, France
TNO, Delft, The Netherlands
Transport Research Laboratory, Crowthorne, UK
Checklist for Preparing a Validation Plan: Updated Version

CONVERGE TR1101 D2.4.1 Page i September 1998

Table of Contents

- Technical Abstract ................................................................................................1
- Executive Summary ..............................................................................................2
- 1. Introduction .......................................................................................................3
  - 1.1 Background ................................................................................................. 3
  - 1.2 Role of this Document .............................................................................. 3
  - 1.3 Approach ...................................................................................................... 4
  - 1.4 Structure of Deliverable .............................................................................. 5
- 2. Summary of the Validation/Evaluation Process.............................................6
  - 2.1 The Validation Context ............................................................................. 6
  - 2.1.2 Validation within the Project ................................................................ 7
  - 2.2 The Validation Plan ................................................................................... 7
- 3. Preparation of Draft Validation Plan ..............................................................10
  - 3.1 Analysis of user requirements ................................................................... 10
  - 3.2 Definition of functional specifications / system architecture ...................... 11
  - 3.3 Objectives of Application ......................................................................... 11
  - 3.4 Summary review of application and validation sites ................................... 12
  - 3.5 Phases of Development ............................................................................. 13
  - 3.6 Definition of expected impacts and groups of users/non-users affected by
    the application (called Appraisal Groups) ..................................................... 13
  - 3.7 Selection of impacts to be validated and justification of this selection ........ 14
  - 3.8 Confirmation of the adequacy of the demonstration site for impact analysis .... 15
  - 3.9 Confirmation that impacts required for multi- criteria analysis or social-cost
    benefit analysis or financial appraisal can be validated ................................... 15
  - 3.10 Definition and categorisation of assessment objectives at verification stage
    of validation ..................................................................................................... 16
    - 3.10.1 Definition of Verification .................................................................... 16
    - 3.10.2 Categorisation of Assessment Objectives .......................................... 16
  - 3.11 Outline of Validation Methods at Verification Stage for Each Assessment
    Objective ......................................................................................................... 18
    - 3.11.1 Definition of indicators ....................................................................... 18
    - 3.11.2 Definition of reference case ............................................................... 19
    - 3.11.3 Definition of Success .......................................................................... 19
    - 3.11.4 Definition of the methods of quantitative and qualitative assessment ...... 20
    - 3.11.5 Definition of conditions of quantitative and qualitative assessment ...... 20
    - 3.11.6 Statistical Considerations - Representativeness/Sampling/Confidence
      Levels ........................................................................................................... 21
    - 3.11.7 Problems of integrity in quantitative and qualitative assessment .......... 21
3.11.8 Resources for Verification ................................................................. 22
3.12 Definition and categorisation of assessment objectives at demonstration stage of validation ................................................................. 24
  3.12.1 Definition of Demonstration ............................................................ 24
  3.12.2 Categorisation of Assessment Objectives ......................................... 24
3.13 Outline of Validation Methods at Demonstration Stage for Each Assessment Objective ................................................................. 24
  3.13.1 Definition of indicators ................................................................. 25
  3.13.2 Definition of reference case ............................................................ 25
  3.13.3 Definition of success ................................................................. 25
  3.13.4 Definition of the methods of quantitative and qualitative assessment ................................................................. 25
  3.13.5 Definition of conditions of quantitative and qualitative assessment ................................................................. 25
  3.13.6 Statistical Considerations - Representativeness/Sampling/Confidence Level ................................................................. 25
  3.13.7 Problems of integrity of quantitative and qualitative assessment ........... 25
  3.13.8 Resources for Demonstration ........................................................ 26
  3.13.9 Timescale of demonstrator and methods of validation ......................... 26
  3.13.10 Expansion of results of demonstration stage of validation .................. 26
  3.13.11 Identification of Potential for European Added Value ....................... 26
3.14 Template for Draft/Final Validation Plan .................................................. 27
3.15 Review of Draft Validation Plans ............................................................ 27

4. Preparation of Final Validation Plan .......................................................... 28

5. Evaluation of Validation Results ............................................................... 29
  5.1 Data Analysis and Presentation ............................................................ 29
  5.2 Framework for Project Deliverables on 'Validation Results' ...................... 29
  5.3 Evaluation of Verification Results ........................................................ 29
  5.4 Evaluation of Demonstration Results .................................................... 29

6. Model Validation Plan - TRACAR Project ................................................... 30
  6.2.1 TUIP - Transport Unit Identification and Positioning ......................... 33
  6.2.2 TUM - Transport Unit Management .................................................. 35

7. Examples of elements of model validation plans ......................................... 50
  7.1 Introduction ......................................................................................... 50
7.2 Hypothetical Example of Impact Analysis for the Telematics Applications for Travellers Area - Videotex Demand Spreading Application ........................................50
  7.2.1 Background ..............................................................................................50
  7.2.2 Objectives of Hypothetical Application ......................................................51
  7.2.3 Summary Review ........................................................................................51
  7.2.4 Expected Impacts and Appraisal Groups ....................................................52
  7.2.5 Selection of impacts to be validated ...........................................................53
  7.2.6 All Assessment Objectives and Appraisal Groups .......................................53
  7.2.7 Example of Selection of Assessment Objectives and Indicators ..................55
  7.2.8 Experimental Design - An Example ..........................................................55

7.3 Hypothetical Example of Impact Analysis for the Maritime: Waterborne Transport Telematics Application Area - Vessel Traffic Systems Application ........................................59
  7.3.1 Background ..............................................................................................59
  7.3.2 Objectives of Hypothetical Application ......................................................60
  7.3.3 Summary Review ........................................................................................61
  7.3.4 Expected Impacts and Appraisal Groups ....................................................61
  7.3.5 Selection of impacts to be validated ...........................................................62
  7.3.6 All Assessment Objectives and Appraisal Groups .......................................63
  7.3.7 Selection of Assessment Objectives and Indicators ....................................65
  7.3.8 Experimental Design - An Example ..........................................................66

8. Keyword List .....................................................................................................70
9. References .........................................................................................................71
10. Glossary ...........................................................................................................72
    10.1 Standard Terms ............................................................................................72
    10.2 Acronyms and Abbreviations ......................................................................73
Appendix A: Template for Validation Plan ............................................................74
Appendix B: Feedback Report Template ...............................................................96

List of Figures

Figure 1: Process in Defining the Validation Plan ..................................................8
Figure 2: Steps involved in outlining the Methods of Validation in relation to Particular Assessment Objectives in the Draft Validation Plan ...................................18
Figure 3: The TRACAR Project’s Test Corridor .......................................................30
Figure 4: System functioning at the terminals (entrance/exit, storage) ...................34
Figure 5: Impact categorisation and respective evaluation methods .......................46
List of Tables

Table 1: Summary Review of Application and Validation Sites ........................................ 12
Table 2: Beginning and End of Phases of Development .................................................. 13
Table 3: Definition of Expected Impacts of Application for Different Appraisal Groups . 14
Table 4: Selection of Impacts for Validation for Particular Appraisal Groups ................. 15
Table 5: Checking that impacts required for multi-criteria analysis or social-cost benefit analysis or financial appraisal can be validated ...................................... 15
Table 6: Definition of Assessment Objectives with respect to Categories of Assessment and Users Groups or Appraisal Groups involved in Validation... 17
Table 7: Definition of Validation Methods for Each Assessment Objective in the Draft Validation Plan ........................................................................................................ 23
Table 8: Application overview ......................................................................................... 33
Table 9: Impacts expected ............................................................................................... 36
Table 10: Decision makers and assessment objectives ..................................................... 39
Table 11: Objectives of the technical assessment and user groups involved ................. 40
Table 12: Methods for technical assessment ................................................................... 41
Table 13: Assessment objectives and the related user groups ....................................... 42
Table 14: Methods for user acceptance assessment ....................................................... 44
Table 15: Assessment objectives and the affected target groups ................................. 47
Table 16: Methods for impacts and socio-economic assessment ................................. 48
Table 17: Technologies and/or Approach Used, CORD Functions/Sub-Functions and Decision Makers .............................................................. 51
Table 18: Expected Impacts for different Appraisal Groups ......................................... 54
Table 19: Definition of Assessment Objectives and User Groups involved in Validation for Impact Analyses .......................................................... 54
Table 20: Examples of Indicators for Assessment Objective 1 to 3 in Table 19 .......... 55
Table 21: Technologies and/or Approach Used and Decision Makers ......................... 61
Table 22: Expected Impacts for different Appraisal Groups ......................................... 64
Table 23: Definition of Assessment Objectives and User Groups involved in Validation for Impact Analyses .......................................................... 64
Table 24: Indicators for Assessment Objective 1 - To assess the safety impact on ferries ..........................................................
Acknowledgements

The present document is built on Validation Quality Support Work in the 4th Framework Transport and Environment Telematics Applications Programme.

The authors wish to thank the Head of DGXIII-C6 of the European Commission for his support for all this work.

For the work within the CONVERGE Project TR 1101, the authors also wish to acknowledge the support of Ms. Marion Le Louarn, the CONVERGE Project Officer and their colleagues in the CONVERGE Validation Quality Team.

For work within the ANIMATE Project C6 1102, the authors wish to acknowledge the support of Mr. Wolfgang Boch, Co-ordinator of the Environment Sector within DGXIII-C6 of the European Commission and Mr. John Crompton, Managing Director of ARTTIC and Co-ordinator of the ANIMATE Project.

Dr. Maltby wishes to acknowledge the support of his partners in the Environment Sector Validation Quality Support Contracts:

Dr. J.A. Cunge Laboratoire d’Hydraulique de France - LHF, Grenoble (F)
Dipl. Ing. H.-J. Heich TÜV Rheinland, Köln (D).

The authors also wish to acknowledge the contributions of many evaluation managers in 4th Framework Transport and Environment Sector Telematics Applications Validation Projects to their present views about approaches to validation activities in Telematics Applications Projects.
Technical Abstract

This ‘Updated Checklist’ provides guidance on:

• the review process for draft validation plans in the T-TAP
• preparation of draft validation plans
• preparation of final validation plans
• presentation and evaluation of validation results.

Templates are provided for a draft/final validation and for the individual feedback reports produced by CONVERGE for the Commission and also sent to projects. The template for a draft/final validation plan is the major change since publication of the draft version of this ‘Updated Checklist’ in March 1998.

The deliverable provides examples of a model validation plan provided by the TRACAR Project TR 1059 and two other examples for elements of validation plans as concrete examples of the recommended approach.

The deliverable is part of the CONVERGE Project’s continuing efforts to provide support to projects in their evaluation/validation activities.

This Updated Checklist is complementary to the CONVERGE “Guidebook For Assessment of Transport Telematics Applications”, which has also been updated at the same time. Whilst the Guidebook is intended to give more general guidance and recommendations regarding the assessment or validation process, the updated Checklist provides detailed advice on how to produce draft and final validation plans.
Executive Summary

This deliverable builds on Validation Quality Support already provided by the CONVERGE Project in the 4th Framework Transport Telematics Applications Programme and by the ANIMATE Project in the 4th Framework Environment Telematics Applications Programme.

This Updated Checklist is complementary to the CONVERGE “Guidebook For Assessment of Transport Telematics Applications”, which has also been updated at the same time. Whilst the Guidebook is intended to give more general guidance and recommendations regarding the assessment or validation process, the updated Checklist provides detailed advice on how to produce draft and final validation plans.

Chapter 2 summarises the Validation/Evaluation Process as used in Transport-TAP Projects. It discusses the validation process, why it is important, and the role of the validation plan in that process.

Chapters 3 to 5 present the approach as recommended by the CONVERGE Project to validation planning and evaluation of validation results.

Chapter 3 provides guidance on the preparation of the draft validation plan and explains the basis on which the CONVERGE Project reviews draft validation plans for the Commission. Chapter 4 provides guidance on the preparation of the final validation plan. Chapter 5 discusses the work involved in the presentation and evaluation of validation results.

Chapter 6 is a ‘Model Validation Plan’ provided by the TRACAR Project TR 1059. Chapter 7 provides two further examples of elements of model validation plans.

Appendix A provides a template for a draft/final validation plan for all RTD&D (Research, Technological Development and Demonstration) projects to use. This template for a draft/final validation plan is the major change since publication of the draft version of this ‘Updated

Appendix B comprises the template for the feedback report which CONVERGE provides to the Commission and to a project after review of its validation plan.
1. **Introduction**

1.1 **Background**

The main objective within the Transport Sector of the 4th Framework Telematics Applications Programme (TAP) is to validate and demonstrate the transport telematics technologies in terms of technical, socio-economic and political aspects. Each of the around 60 1st Call and a further 30 4th Call research and demonstration projects involved are required to prepare a draft validation plan 6 months after the beginning of the project.

In order to maximise the effectiveness and benefits of the Programme, it is essential for projects to adopt consistent assessment and validation methods. In particular, projects dealing with comparable applications should take a common approach for their respective assessment and validation activities so that the results from different projects are comparable and impacts of comparable applications can be generalised in a consistent and effective way.

Concise and comprehensive guidance was needed for assessment and validation activities at the start of the Fourth Framework Programme. To this end, the CONVERGE Project provided support to 1st Call Transport -TAP Projects, through:

- Deliverable DVQ5.1 - “Guidebook for Assessment of Transport Telematics Applications” (Zhang, *et al.*, 1996)
- Deliverable DVQ3.2 - “Checklist for Preparing a Draft Validation Plan” (Maltby and Morello, 1996)

1.2 **Role of this Document**

This deliverable builds on the support already provided by the CONVERGE Project to 1st Call Transport Telematics Applications Projects by producing an updated version of the “*Checklist for Preparing a Draft Validation Plan*” for use by 4th Call Transport Telematics Projects.

It builds also on the support already given through the ANIMATE Project to 1st Call Environment Telematics Applications Projects.

This Updated Checklist is complementary to the CONVERGE Deliverable D2.3.1, “Guidebook for Assessment of Transport Telematics Applications: Updated Version” (Zhang, *et al.*, 1998b), which has also been updated at the same time. Whilst the Guidebook is intended to give more general guidance and recommendations regarding the assessment or validation process, the updated Checklist provides detailed advice on how to produce draft and final validation plans.

The CONVERGE Deliverable DVQ3.2 - ‘Checklist for Preparing a Draft Validation Plan’ provided standard tables for preparation of a validation plan.
The objective of this ‘Updated Checklist’ is to update Deliverable DVQ3.2 by providing:

- explanatory text for these standard tables
- examples
- a validation plan template
- a feedback report template

The explanatory text is based on a deliverable, ‘Guidelines for Preparation of Validation Plans’ (Maltby, Cunge and Heich 1996), prepared for Fourth Framework 1st Call Environment Telematics Projects as part of validation support provided by the ANIMATE Project C6 1102.

The examples are drawn from the CONVERGE Deliverable DVQ5.2 - ‘Model Validation Plans.’

The validation plan template has been provided as an additional aid to projects in preparation of validation plans. It is based precisely on the advice provided in the ‘Updated Checklist.’

The feedback report template is provided to show 4th Call Projects the basis on which CONVERGE reviews validation plans for the Commission and to illustrate the point that there is a one-to-one relationship between the advice given by CONVERGE for the preparation of validation plans and the structure of the feedback report.

1.3 Approach

Chapters 2 provides an overall view of the Evaluation/Validation Process.

Chapters 3, 4 and 5 are based on the CONVERGE Deliverable DVQ3.2 - ‘Checklist for Preparing a Draft Validation Plan’ and the ANIMATE Deliverable, ‘Guidelines for Preparation of Validation Plans’. The tables of the CONVERGE Deliverable DVQ3.2 are used in Chapter 3 and they are explained and discussed on the basis of slightly re-drafted text from the ANIMATE deliverable.

Section 3.15 (Review of Draft Validation Plans) is included to advise 4th Call Projects about the way in which CONVERGE prepares its review of draft validation plans.

The examples in Chapters 6 and 7 have been taken from the CONVERGE Deliverable DVQ5.2.

References are made to the complementary CONVERGE Deliverable D2.3.1 - “Guidebook For Assessment of Transport Telematics Applications: Updated Version,” at relevant points in the text.
1.4 **Structure of Deliverable**

The deliverable has the following chapters:

Chapter 1 - Introduction

Chapter 2 - Summary of the Validation/Evaluation Process

Chapter 3 - Preparation of Draft Validation Plan

Chapter 4 - Preparation of Final Validation Plan

Chapter 5 - Evaluation of Validation Results

Chapter 6 - Model Validation Plan - TRACAR Project

Chapter 7 - Examples of elements of model validation plans

Additionally, a Keyword List, References and a Glossary are provided.

Appendix A provides a template for a draft/final validation plan for all RTD&D (Research, Technological Development and Demonstration) projects to use. This template for a draft/final validation plan is the major change since publication of the draft version of this ‘Updated

Appendix B comprises the template for the feedback report which CONVERGE provides to the Commission and to a project after review of its validation plan.
2. **Summary of the Validation/Evaluation Process**

2.1 **The Validation Context**

In this chapter we discuss the validation process, why it is important, and the role of the validation plan in that process. Behind every European RTD&D project, there are numerous actors and interested parties who hope to learn from the developments, trials and demonstrations. Such experience will result from simply taking part in the project’s work. However, there is also a need for a more formal assessment of the systems and services in operation, to verify that they meet specifications of performance or that they have the hoped-for impacts on users. Generally, there is also a need to validate the systems/services in real-life service amongst the general public.

In any case, a well thought out validation plan is essential to ensure that:-

- stakeholders’ needs are met
- the right assessments will be properly carried out
- results at the end of the project will be robust and clear.

A good plan should be a powerful tool of management for the project itself, and has already proved its worth as an aid to programme management for the European Commission.

*As a management tool for the project, a validation plan should describe the essential activities in a simple, coherent way related to the final results of the project.*

This will help a project focus on those results and, especially in larger projects with several sites, will also establish a common vocabulary and understanding within the project.

As project client, the EC has found it beneficial to see, at an early stage in each project, a draft validation plan setting out the objectives and methods of assessment. This gives an early indication of the quality of validation results to be expected, and has been an important element in the annual review process.

* A good validation plan sets evaluation goals that can serve as a yardstick to assess a project’s final achievements.*

The structured process a project uses to plan, define and then carry out validation is described below using the model of the TRACAR project draft validation plan. Before entering into the details of the process, we summarise briefly the general framework for validation. This has two levels:

- the project’s surroundings, and
- within the project itself.
the applications in terms of their basic functions. The validation process assesses how those functions perform in practice. The use of a common set of functions (as recommended in the Guidebook) allows better integration of the validation of different sites in a project and between projects.

**Building the application**
Although system development and building is usually the main work of a project, it features too in the validation plan, which must describe the physical application of the telematics systems and services in enough detail as part of a record of all aspects that bear on the validation results and their interpretation. For example, the plan should describe:
- the general operating environment,
- the types of users,
- how elements are integrated,
- etc.
Validation
This is the phase in a project where system operation is tested, either in a restricted trial to verify its performance or in a wider demonstration to validate its use by end users such as system operators, transport businesses, drivers or passengers. We discuss this in more detail below.

Exploitation
It is almost certain that the validation results will be a major input to a sound exploitation plan. So the measurements of system performance, user acceptance, impacts, costs, and benefits should be designed to feed into and complement any market assessment or product marketing. Good evaluation results become good sales material.

The draft validation plan of the TRACAR project (TR1059), on which our example Model Validation Plan is based, contains a diagram showing clearly the steps in validation; we reproduce it below, with grateful acknowledgement to the TRACAR project.

**Figure 1: Process in Defining the Validation Plan**

This framework, based on that contained in the first versions of the CONVERGE ‘Guidebook’ and ‘Checklist,’ is in fact universally valid for any RTD&D project.

In preparing the validation plan for a particular project, the evaluation team should systematically consider each of the steps, and adapt each element as relevant to that project’s nature.

To help readers in adapting the generic approach to their own project, the model validation plan presented in the next chapter is accompanied by guidance at each step.

2.2 The Validation Plan

The creation and use of a validation plan should not be seen as the task of a single individual working apart from the application development team. The project should rather employ an evaluation team representing all the various key interests – test sites, main applications, area of assessment.

The written validation plan should represent a consensus of these interests, and be a realistic reflection of the resources actually available.

The validation plan, although only a general requirement in its early, draft form (at 6 months into each project), should be a living document, and be revised at each stage of planning. Thus, while on one hand the draft plan would not be expected to contain detailed questionnaire design, the final version of the plan produced before validation actually begins should be very specific and provide details of validation practice.

Even if not a contractual obligation, all projects should produce a “final” validation plan once all the details of data collection and measurements have been decided.

Chapter 2 of the CONVERGE Deliverable D2.3.1, ‘Guidebook for Assessment of Transport Telematics Applications: Updated Version,’ also provides an overview of the assessment process.
3. **Preparation of Draft Validation Plan**

The Draft Validation Plan should be comprehensible without the reader having to refer to other documents. This means that it might be necessary to include in it summaries of important sections of related deliverables, such as those on:

- User Needs
- System Architecture.

The Draft Validation Plan should contain the following elements:

- precise statement of the specific objectives of the application
- summary review of application and validation sites
- definition of expected impacts and groups of users/non-users affected
- selection of impacts to be validated and justification of this selection
- confirmation that impacts required for multi-criteria analysis or social cost-benefit analysis or financial appraisal can be validated
- confirmation of the adequacy of the demonstration site for impact analysis
- definition and categorisation of assessment objectives at verification stage of validation
- outline of validation methods at verification stage for each assessment objective
- definition and categorisation of assessment objectives at demonstration stage of validation
- outline of validation methods at demonstration stage of validation for each assessment objective
- identification of the potential for European Added Value on the basis of:
  - comparison of validation results across different sites in the project
  - comparison of validation results with those of other projects.

3.1 **Analysis of user requirements**

Analysis of user requirements will influence the following aspects of a validation plan:

- definition of the objectives of an application
- definition of the system architecture
- definition of the assessment objectives
- definition of the user groups and appraisal groups to be involved in validation
- possibly the actual methods of validation

Reports on analysis of user requirements for Fourth Framework TAP projects usually have to be prepared at the same time as the draft validation plan. This means that inputs from the analysis of user requirements into the draft validation plan have to be made in a tentative way and finalised in a revision of the draft validation plan.

Chapter 3 of the CONVERGE Deliverable D2.3.1, ‘Guidebook for Assessment of Transport Telematics Applications: Updated Version,’ also discusses definition of user needs; in particular it

3.2 Definition of functional specifications / system architecture

Definition of functional specifications / system architecture will influence the following aspects of a validation plan:

- definition of assessment objectives, particularly those connected with:
  - physical functioning
  - HMI aspects
  - possibly impact assessment
- defining methods of validation for the verification stage in particular.

Definition of functional specifications is usually done after preparation of the draft validation plan in Fourth Framework TAP projects. This means that inputs from the definition of functional specifications can only be put into a revision of the draft validation plan.

The need to revise draft validation plans for Fourth Framework TAP projects from the viewpoint of inputs from the analysis of user requirements and the definition of functional specifications highlights the need for a final validation plan.

3.3 Objectives of Application

There is often confusion about the difference between the ‘Objectives of an Application’ and

The objectives of an application describe what it has been designed to do; for example, the main objective of the Videotex Traffic Service in Chapter 7 is to facilitate ‘demand spreading’; that is, to ameliorate demand for travel by car by virtue of car travellers:

- changing route
- rescheduling of journey
- changing mode.

An assessment objective describes one purpose of validation; for example, for the Videotex Application one assessment objective is to assess change of route by car drivers as a consequence of using the Videotex Traffic Service.

So, assessment objectives are usually ‘mirror images’ of the objectives of an application - this is probably the main reason for the confusion between the two concepts.

The objectives of the application should be stated precisely.
3.4 Summary review of application and validation sites

The Application should be summarily reviewed as in Table 1 on the basis of the following information:

- Technologies and Approach Used
  ⇒ it is useful to include a figure describing the architecture of the application in terms of the technologies being used

- Functionalities
  ⇒ a Telematics Functions List should be used if available (CONVERGE Project, 1997)

- Main Decision Makers involved in the validation of the application
  ⇒ these are groups who will be influential in defining the assessment objectives for validation and in determining whether verification results justify proceeding to the demonstration stage of validation

- User Groups
  ⇒ it is important to identify these groups because they should be involved in validating the application and will probably represent the main market for the application
  ⇒ it is likely that these groups have been identified already in the ‘Analysis of User Requirements’ phase of the project
  ⇒ they might include:
    ◊ operators of the application
    ◊ intermediate users of the application like providers of information
    ◊ end-users of the application
  ⇒ in view of this definition of ‘User Groups,’ there is likely to be some overlap between ‘Main Decision Makers’ and ‘User Groups’
  ⇒ a distinction should be made between ‘User Groups within the Project’ and ‘User Groups outside the Project’ involved in validating an application because the latter have an important role in independent assessment of the validation results

- Verification Site:
  ⇒ details of the site should be provided including a plan or map

- Demonstration Site:
  ⇒ details of the site should be provided including a plan or map.

Table 1: Summary Review of Application and Validation Sites

<table>
<thead>
<tr>
<th>Application</th>
<th>Technologies and Approach Used</th>
<th>Functionalities (Telematics Functions if available)</th>
<th>Main Decision Makers</th>
<th>User Groups Within Project</th>
<th>User Groups Outside Project</th>
<th>Verification Site</th>
<th>Demonstration Site</th>
</tr>
</thead>
<tbody>
<tr>
<td>Application 1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Application 2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
3.5 Phases of Development

As already indicated, 4th Framework TAP Projects have the following phases:

1. Analysis of user requirements
2. Definition of functional specifications
3. Building of a demonstrator
4. Validation of a demonstrator in a process covering 2 stages (Verification and Demonstration)
5. Exploitation plan

It is not unusual for the beginning and end of these phases to be revised in the course of a project. As this could have important consequences for the quality of validation if the time for validation is compressed towards the end of a project, it is important to monitor these phases from the beginning of the project.

Additionally, use of common validation plans across sites within a project requires a consideration of the beginning and ends of the verification and demonstration phases for the applications and/or sites featuring in the comparison.

It is, therefore, helpful to list the beginning and end of the five phases for each application, as indicated in Table 2.

<table>
<thead>
<tr>
<th>Application</th>
<th>Analysis of user requirements</th>
<th>Definition of functional specifications</th>
<th>Building of a demonstrator</th>
<th>Verification</th>
<th>Demonstration</th>
<th>Exploitation Plan</th>
</tr>
</thead>
<tbody>
<tr>
<td>Application 1</td>
<td>mmyy-MMYY</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Application 2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

3.6 Definition of expected impacts and groups of users/non-users affected by the application (called Appraisal Groups)

An impact describes the changes or effects brought about by an application resulting from its implementation in an experimental or ‘real-life’ situation.

It is very important to describe the expected impacts precisely because they hold the key for the benefits or disbenefits of the application.

Equally, it is important to define precisely the different groups of people affected by the application because:

- benefits and disbenefits may differ from one group to another in both type and scale
- perception of these benefits and disbenefits may vary from one group to the next
These benefits and disbenefits will be experienced by:

- operators of the application
- intermediate users of the application like providers of information
- end-users of the application
- individuals or groups affected indirectly from the application but who cannot be classed as either operators, end-users or intermediate users

In the case of the last category, groups may be in close association like businesses or families or loose associations like communities, interest groups, society, etc. These different groups of users/non-users affected by the impacts of the application should be referred to as APPRAISAL GROUPS.

As for ‘User Groups’ in Section 3.4, definition of these ‘Appraisal Groups’ should make a clear distinction between those connected with the project and those who are independent of the project.

The expected impacts and their related appraisal groups should be summarised in Table 3, together with a judgement about the anticipated level of impact, using the following scale:

++ very positive impact
+ positive impact
0 neutral/uncertain impact
- negative impact
-- very negative impact

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Impact Appraisal Group 2.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Impact Appraisal Group .</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Notes:
1. The impacts are likely to vary between the Appraisal Groups, while the scale of impacts which are common to different appraisal groups might vary between these appraisal groups.

3.7 Selection of impacts to be validated and justification of this selection

It may not be possible to validate all the expected impacts; for example:

- it may not be possible to quantify the impact by either measurement or simulation
  ⇒ in such cases, at least an expert qualitative assessment of the impact should be provided.
- it may be judged wasteful of resources to validate impacts considered neutral or uncertain in their scale of impact.
- impacts may be considered too indirect or too difficult to define to justify validation.
resources for validation may be too limited for comprehensive validation.

Each expected impact in Table 4 should be categorised in one of the following ways:
1. impact to be validated quantitatively
2. impact to be validated qualitatively
3. impact not to be validated

This categorisation should be summarised as in the following table within its Appraisal Group, with an explanation for these decisions.

Table 4: Selection of Impacts for Validation for Particular Appraisal Groups

<table>
<thead>
<tr>
<th>Particular Appraisal Group</th>
<th>Impact</th>
<th>Impact Category (1, 2 or 3)</th>
<th>Explanation for Category</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1.</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>2.</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

3.8 Confirmation of the adequacy of the demonstration site for impact analysis

In parallel with selection of the impacts to be validated, the proposed demonstration site should be examined to ensure that:
- impacts can be validated
- the extent and configuration of the site should be sufficient to validate the main impacts
- if comparison of validation results across sites is planned in order to demonstrate the European Value of validation results, sites should be compared to ascertain whether significant differences between sites might invalidate such comparisons

3.9 Confirmation that impacts required for multi-criteria analysis or social-cost benefit analysis or financial appraisal can be validated

If conduct of multi-criteria analysis or social-cost-benefit analysis or financial appraisal of the benefits and costs associated with an application are part of the objectives of a project, it is very important at an early stage to establish that it will be possible to validate the impacts required for such analyses.

This could be done as in Table 5.

Table 5: Checking that impacts required for multi-criteria analysis or social-cost benefit analysis or financial appraisal can be validated
### Checklist for Preparing a Validation Plan: Updated Version

#### Application Impacts required for multi-criteria analysis

<table>
<thead>
<tr>
<th>Application</th>
<th>Impacts required for multi-criteria analysis</th>
<th>Can impact be validated? YES / NO</th>
<th>Explanation if NO</th>
</tr>
</thead>
<tbody>
<tr>
<td>Application 1</td>
<td>Impact 1</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Impact 2</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Impact 3</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Impact 4</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### 3.10 Definition and categorisation of assessment objectives at verification stage of validation

#### 3.10.1 Definition of Verification

The European Commission has defined ‘Verification’ as follows:

‘The verification stage of validation will use a small but sufficient sample of users in a real-life situation to test the technical feasibility of the demonstrator and to yield preliminary findings on user acceptance’ ([see EC, 1994b](#)).

#### 3.10.2 Categorisation of Assessment Objectives

Assessment objectives at the verification stage of validation should concentrate on:

- testing the physical functioning of the application
- testing user acceptance of the application

Some limited impact analysis might be conducted at the verification stage if the main decision makers ask for reassurance regarding the main impacts prior to the demonstration stage.

Assessment objectives should be defined and categorised as in Table 6 in association with their related user groups.
Table 6: Definition of Assessment Objectives with respect to Categories of Assessment and Users Groups or Appraisal Groups involved in Validation

<table>
<thead>
<tr>
<th>Assessment Objectives</th>
<th>User Groups or Appraisal Groups as appropriate</th>
</tr>
</thead>
<tbody>
<tr>
<td>Testing Physical Functioning $^1$</td>
<td>User Groups</td>
</tr>
<tr>
<td>1.</td>
<td></td>
</tr>
<tr>
<td>2.</td>
<td></td>
</tr>
<tr>
<td>.</td>
<td></td>
</tr>
<tr>
<td>Testing User Acceptance $^1$</td>
<td>User Groups</td>
</tr>
<tr>
<td>1.</td>
<td></td>
</tr>
<tr>
<td>2.</td>
<td></td>
</tr>
<tr>
<td>.</td>
<td></td>
</tr>
<tr>
<td>Impact Analysis $^2$</td>
<td>Appraisal Groups</td>
</tr>
<tr>
<td>1.</td>
<td></td>
</tr>
<tr>
<td>2.</td>
<td></td>
</tr>
<tr>
<td>.</td>
<td></td>
</tr>
<tr>
<td>Multi-criteria analysis or social-cost benefit analysis or financial appraisal $^3$</td>
<td>Appraisal Groups</td>
</tr>
<tr>
<td>1.</td>
<td></td>
</tr>
<tr>
<td>2.</td>
<td></td>
</tr>
<tr>
<td>.</td>
<td></td>
</tr>
</tbody>
</table>

Notes
1 at both verification and demonstration stages.
2 mainly at demonstration stage but possibly at verification stage.
3 only at demonstration stage.
3.11 Outline of Validation Methods at Verification Stage for Each Assessment Objective

The outline of the validation methods to be used at the verification stage should present sufficient information to permit an independent judgement that it is likely to result in good verification. The validation methods (the experimental design) that will be used in relation to individual assessment objectives should be outlined.

The process of defining the draft verification plan should have the steps defined in Figure 2.

**Figure 2: Steps involved in outlining the Methods of Validation in relation to Particular Assessment Objectives in the Draft Validation Plan**

<table>
<thead>
<tr>
<th>Step</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>Definition of Indicators</td>
</tr>
<tr>
<td>2.</td>
<td>Definition of Reference Case</td>
</tr>
<tr>
<td>3.</td>
<td>Definition of Success</td>
</tr>
<tr>
<td>4.</td>
<td>Definition of Methods of Quantitative or Qualitative Assessment</td>
</tr>
<tr>
<td>5.</td>
<td>Definition of Conditions of Quantitative or Qualitative Assessment (referred to as ‘Blocking’ in the language of experimental design)</td>
</tr>
<tr>
<td>6.</td>
<td>Statistical Considerations - Representativeness/Sampling/Confidence Level</td>
</tr>
<tr>
<td>7.</td>
<td>Problems of Integrity in Quantitative or Qualitative Assessment</td>
</tr>
</tbody>
</table>

This information should be provided as follows.

### 3.11.1 Definition of indicators

An indicator is the main criteria of quantitative or qualitative assessment. It is a parameter indicating the performance or impacts of an application. For quantitative assessment, it is either measured directly or derived from measurement or modelling / simulation.

Quantitative assessment is either objective or ‘hard’ (e.g. physical measurement) or subjective or

Use of indicators for qualitative assessment is more problematic because measurement is replaced by personal judgement. This may mean that qualitative assessment is more applicable to user acceptance than to impact analysis. All projects should be able to test physical functioning of an application by quantitative assessment. Simulation / modelling can make a contribution to both quantitative and qualitative assessment.

An indicator is linked to a specific assessment objective.

Indicators should be defined as precisely as possible.
The main qualification is that precise definition of indicators for testing physical functioning of an application and user acceptance will probably have to await completion of the definition of functional requirements phase. The validation plan should be continually revised in this way up to the validation phase. Indicators for impact analysis are less likely to be affected in this way.

This means that any indicators not defined precisely enough in the draft validation plan will have to be defined precisely in a revised validation plan.

It may be necessary to use more than one indicator for each assessment objective.

3.11.2 Definition of reference case

The performance and impacts of an application are usually compared against some existing situation in order to show that the application meets required standards (for tests of the physical functioning of the application and user acceptance) and that use of the application is an improvement on alternative ways of achieving the objectives of the application (for user acceptance again and impact analysis). ‘Before and After’ studies might feature in assessment of user acceptance and impact analysis.

Reference cases will vary with category of assessment objectives and may be required for individual indicators or be the same over a group of indicators or assessment objectives, for example:

- with testing physical functioning of an application, reference cases may be standards for individual indicators;
- with user acceptance and impact analysis, the reference case may simply be a current application or way of doing things and so be common over a number of user acceptance or impact analysis assessment objectives.

3.11.3 Definition of Success

This defines the expectation about the performance and impacts of the application. The success or failure of validation results are tested against this criteria. Therefore, it has a vital role in the validation methodology.

Definition of success should be given, AS APPROPRIATE, for individual indicators, groups of related indicators, individual assessment objectives or groups of assessment objectives. Examples may be:

- a standard for an individual indicator in relation to testing the physical functioning of the application
- an acceptable level of satisfaction with respect to user acceptance
- accepted ‘best practice’ with respect to testing HMI aspects of user acceptance
- for impact analysis, a reduction in pollution beyond a defined level.

For quantitative assessment, definition of success is linked to statistical considerations; viz. the number of measurements required of an indicator to establish that a change in performance is at least as large as anticipated performance (the definition of success) depends upon:
• the magnitude of this anticipated change in performance
• the amount of variability in the measurements
• the level of statistical confidence which is expected

3.11.4 Definition of the methods of quantitative and qualitative assessment

Methods of quantitative assessment could include:
• physical measurement (called ‘objective’ measurement)
• questionnaire surveys, structured interviews (called ‘subjective’ measurement)
• simulation / modelling

Methods of qualitative assessment could include:
• simulation / modelling
• less structured ways of obtaining opinion
• professional judgement

The difference between simulation/modelling used in quantitative assessment and simulation/modelling used in qualitative assessment lies in the quality of the specification, calibration and validation of simulation/modelling; these are of lower quality when used in qualitative assessment but sufficiently developed to add to professional judgement.

Obviously, in using qualitative assessment, it is desirable to be transparent in explaining how validation results were obtained (for example, who was consulted and how they were consulted) because of the less structured nature of the approach. A greater degree of doubt must also accompany any assessment of validation results by qualitative assessment for the same reason.

It is anticipated that validation of a Transport-TAP Project would be mainly by quantitative assessment.

Even though verification should feature a ‘real-life’ situation (see definition in Section 3.10.1), it may be necessary to simulate this situation to test user acceptance at the verification stage if testing on public facilities is judged inappropriate.

3.11.5 Definition of conditions of quantitative and qualitative assessment

Verification should be conducted for the range of conditions (time and spatial variation, meteorological conditions, for example) likely to be met at the demonstration stage of validation. This range of conditions should be defined as precisely as possible in order to relate performance of the application as directly as possible to the conditions under which it will be verified.
3.11.6 Statistical Considerations - Representativeness/Sampling/Confidence Levels

3.11.6.1 Representativeness

Objective and subjective measurement should be representative of a population. This means, in principle, that measurements should be taken at random from the parent population and respondents in questionnaire surveys should be selected at random.

In practice this requirement should not be a major problem for testing the physical functioning of an application because these tests are usually extend over continuous operation of the application. However, testing the physical functioning of an application at the verification stage of validation over the range of conditions likely to be met during the more comprehensive demonstration stage of validation may be more problematic.

‘Representativeness’ can be expected to be more of a problem in tests of user acceptance if the number of users involved is small and these users are working within the project. As indicated in Section 3.4, this problem can be helped by involvement in tests of user acceptance of users from both inside and outside the project.

3.11.6.2 Sampling

Sampling design and calculations need not be provided in the draft verification plan. However, some measures of sampling should be provided, such as:

- duration of the verification stage;
- exhaustiveness of the testing of the physical functioning of the application;
- number of users involved.

3.11.6.3 Confidence Levels

An estimate based on professional judgement should be given about the statistical confidence level that can be associated with any measurement of indicators.

In cases of a statistical approach to sampling being judged INAPPROPRIATE by a project for measurement of particular indicators, explanations should be provided for such decisions.

3.11.7 Problems of integrity in quantitative and qualitative assessment

Methods of quantitative and qualitative assessment should be chosen to avoid any major problems of integrity. ‘Integrity’ is the language of experimental design but it is really about ‘lack of integrity.’ This will be of particular concern for qualitative assessment.
The main considerations will be:

1. **Insularity:**
   - even though the conditions of measurement of indicators should have been defined as precisely as possible, the performance of the application may be affected by variations in these conditions between measurements.

2. **Disturbance of the Assessment Process:**
   - there may be accidental or intended bias introduced into the assessment process by such factors as:
     - respondent fatigue - subjects become mentally and/or physically tired;
     - policy response bias - subjects wish to influence the results of the validation;
     - justification bias - subjects consciously or subconsciously give answers which they think are more acceptable to an interviewer.

3. **Disturbance of the Simulation / Modelling Process:**
   - simulation / modelling may suffer from poor specification even if seemingly adequately calibrated and validated.

This outline of the validation methods to be used at verification plan should be summarised as in Table 7 or the information in this table should be presented in an alternative structure.

**3.11.8 Resources for Verification**

The draft verification plan should take into account the resources available for verification. This consideration will make necessary a process of iteration and prioritising in preparation of the validation plan.
Table 7: Definition of Validation Methods for Each Assessment Objective in the Draft Validation Plan

<table>
<thead>
<tr>
<th>Indicators</th>
<th>Reference Case</th>
<th>Definition of Success</th>
<th>Methods of Quantitative or Qualitative Assessment</th>
<th>Conditions of Quantitative or Qualitative Assessment</th>
<th>Statistical Considerations</th>
<th>Possible Problems of Integrity</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Indicator 1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Indicator 2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
3.12 Definition and categorisation of assessment objectives at demonstration stage of validation

3.12.1 Definition of Demonstration

The European Commission has defined ‘Demonstration’ as follows:

‘The demonstration stage of validation will use a sufficiently large sample of users in a real-life situation to provide information on cost-effectiveness, user friendliness and similar issues, as well as testing the feasibility of the system when used on a large scale’ (see EC, 1994b).

It may not be possible to make a clear transition from the verification to the demonstration stage; for example, some ‘final tuning’ of the application may be required prior to the start of the demonstration stage.

3.12.2 Categorisation of Assessment Objectives

Assessment objectives at the demonstration stage of validation should concentrate on:

- testing the physical functioning of the application
  ⇒ the emphasis will be on operational reliability
- testing user acceptance of the application over a larger group of users and ideally representing users within the project and users outside the project with the role of the latter group that of providing impartial validation
- impact analysis
  ⇒ related to impact analysis
    ◊ multi-criteria analysis
    ◊ social-cost benefit analysis
    ◊ financial appraisal.

Assessment objectives should be defined and categorised as in Table 6 in association with their related user and appraisal groups.

3.13 Outline of Validation Methods at Demonstration Stage for Each Assessment Objective

The outline of the validation methods to be used at the demonstration stage should present sufficient information to permit an independent judgement that it is likely to result in good demonstration. The validation methods (the experimental design) that will be used in relation to individual assessment objectives should be outlined.
The process of defining the draft demonstration plan should be similar to that already defined for the draft verification plan in Section 3.11 and in Figure 2. The main difference will be that categories of assessment objectives will be extended to impact analysis and related multi-criteria analysis, social-cost benefit analysis, and financial appraisal.

Additionally, the following points are relevant:

3.13.1 Definition of indicators

No further points.

3.13.2 Definition of reference case

For impact analyses the main reference case will be a current application or way of doing things. It may be necessary to develop these reference cases in association with related policies and strategies.

3.13.3 Definition of success

No further points.

3.13.4 Definition of the methods of quantitative and qualitative assessment

Both measurement and simulation / modelling of indicators with respect to impacts may be necessary if measurement alone cannot validate the impact; for example, low market penetration of an application presently may make necessary simulation / modelling of its impact for future conditions of high market penetration.

3.13.5 Definition of conditions of quantitative and qualitative assessment

No further points.

3.13.6 Statistical Considerations - Representativeness/Sampling/Confidence Level

No further comment.

3.13.7 Problems of integrity of quantitative and qualitative assessment

The main additional consideration will be:

- completeness
- in the selection of impacts to be validated, care should be taken to ensure that limited resources do not result in non-selection of impacts which might have significant impacts.
This outline of the validation methods to be used at demonstration plan should be summarised as in Table 7: Definition of Validation Methods for Each Assessment Objective in the Draft Validation Plan

3.13.8 Resources for Demonstration

The draft demonstration plan should take into account the resources available for demonstration. This consideration will make necessary a process of iteration and prioritising in preparation of the plan.

3.13.9 Timescale of demonstrator and methods of validation

The details of the measurement plan should be considered in preparation of the Final Validation Plan (see Chapter 4). The measurement plan is concerned with the timing and sequence of measurement in validating the application - it is designed to minimise the possibility of bias entering measurements because the application under test has been tested under different conditions than the application with which it is being compared.

The timescale of a demonstrator should be considered in preparation of the draft validation plan from two viewpoints:

- whether the duration of demonstration is sufficient to validate impacts adequately
- whether the proposed validation methods will capture impacts adequately within the demonstration stage.

3.13.10 Expansion of results of demonstration stage of validation

If it is necessary to expand the results of the demonstration stage of validation as part of the evaluation of a demonstrator, the expansion methodology (e.g. modelling, simulation, extrapolation, professional judgement) should be outlined; for example, estimating the effects of greater market penetration of an application.

3.13.11 Identification of Potential for European Added Value

Finally, the potential for adding European Added Value to the Project should be identified in terms of comparing validation results:

- across sites within the project
- across sites within the project and other projects.

This means that the project should first of all identify similar applications (applications with comparable functionalities) within the project.

It should, then, develop a common validation plan for use on all sites within the project where such applications are being tested.
Opportunities for the project to compare its validation results with those of other projects testing similar applications may arise through the cross-project collaborative studies programme of the CONVERGE Project (Zhang, et al, 1998a).

### 3.14 Template for Draft/Final Validation Plan

A template for a draft/final validation plan is provided in Appendix A. Its structure is consistent with the advice given in this Chapter.

### 3.15 Review of Draft Validation Plans

A European Commission requirement is that a project with a validation phase submits its draft validation plan for review at 6 months after the start of the project.

This review will be conducted by the CONVERGE Project on behalf of the Commission. The procedure is to prepare a feedback report which is sent to the Commission and to the project. The template used for this feedback report is in Appendix B. This template is consistent with the guidance on preparation of draft validation plans provided in this chapter.

4. Preparation of Final Validation Plan

As already indicated, the draft validation plan should be revised as more complete information becomes available within a project and completed as a ‘Final Validation Plan.’

The Final Validation Plan should be completed sufficiently in advance of verification starting on any of the projects' sites. It may still be necessary to revise this ‘final validation plan’ to take account of verification results.

The final validation plan should cover both the verification and demonstration stages of validation. It should involve the following work:

• an update of the draft validation plan

• design of the methods of quantitative and qualitative assessment, including:

  ◊ definition of the criteria against which the validation results will be evaluated - these will be based on the definitions of success for particular assessment objectives and indicators previously addressed in the draft validation plan

  ◊ finalisation of statistical considerations for quantitative assessment - confirmation of representativeness of samples and preparation of sample calculations including confirmation of estimates of confidence levels presented previously in the draft validation plan

  ◊ definition of the methods of expansion of results at the demonstration stage of validation (see Section 3.13.10)

  ◊ presentation of the basis of analysis to be used in any multi-criteria analysis, social-cost benefit analysis, and financial appraisal, etc.

  ◊ preparation for data analysis and presentation in relation to the design of the methods of quantitative and qualitative assessment.

• definition of the measurement plan (see Section 3.13.9) - it will be important to plan the timing of measurement of indicators for the application and the reference application in such a way that the measurement plan will not itself be a contribution to bias in the comparison of the performance of the application and the reference application

  ◊ this will be a particularly important consideration if the number of measurements involved is small

• confirmation of the potential for European Added Value in the project (see Section 3.13.11)

• preparation of a detailed workplan for the organisation for the management of validation.

Section 5.2.4 of the CONVERGE Deliverable D2.3.1, ‘Guidebook for Assessment of Transport Telematics Applications: Updated Version,’ also provides advice on socio-economic assessment.
5. **Evaluation of Validation Results**

5.1 **Data Analysis and Presentation**

Chapter 8 of the CONVERGE Deliverable D2.3.1, ‘Guidebook for Assessment of Transport Telematics Applications: Updated Version,’ also provides advice on data analysis and presentation.

5.2 **Framework for Project Deliverables on Validation Results**

CONVERGE has also provided advice on the content and structure of project ‘Validation Results’ deliverables (Maltby, et al, 1998).

5.3 **Evaluation of Verification Results**

The following work should be involved:

- an explanation should be provided for any deviations at verification from the verification plan and an evaluation presented of the effects of such deviations on the verification results;
- evaluation of the verification results against the criteria confirmed in the final validation plan;
- confirmation of the assessment objectives and validation methods for the demonstration stage.

5.4 **Evaluation of Demonstration Results**

The following work should be involved:

- an explanation should be provided for any deviations at demonstration from the demonstration plan and an evaluation presented of the effects of such deviations on the demonstration results
- evaluation of the demonstration results against the criteria confirmed in the final validation plan
- estimation of benefits and costs associated with the application by multi-criteria analysis or socio-economic analysis or financial appraisal
- contribution of validation results to the exploitation plan.
6. Model Validation Plan - TRACAR Project

We are grateful to the TRACAR Project for permitting us to use their Draft Validation Plan for this example and also agreeing to our making constructive comments on it. These comments are in boxes.

### Objectives

The Project aims to track, monitor and manage vehicles, containers/swap bodies and their contents along a corridor from an originating site in Denmark to a destination site in Italy, crossing three borders (DK/DE, DE/AU, AU/IT) (see Figure 3) The journey will involve road and rail transportation, with the project monitoring the change of mode, as well as the status of the freight, within the container/swap body.

Clear maps or plans of both verification and demonstration sites is an essential part of a validation plan.

![Figure 3: The TRACAR Project’s Test Corridor](image)

The Project will concentrate on two Freight and Fleet Management (FFM) applications:

- Identification and positioning of transporting units (containers and swap bodies) along combined transport chains (TUIP Application - Transport Unit Identification and Positioning);
- Management of transporting units’ movement along combined transport chains (TUM Application - Transport Units Management).
Sufficient details of the User Needs Analysis should be presented with the Validation Plan to provide a context for:

- the characteristics of the proposed applications; e.g. objectives, technologies and approaches used, functionalities, main decision makers
- determination of assessment objectives for validation.

TRACAR undertook an analysis of user requirements related to the management and supervision of a multi-modal transport system. The analysis focused particularly on sensitive goods transportation, such as frozen food and fresh products, as this was identified as a very appropriate freight sector in a preliminary user survey at the beginning of the project. The TRACAR user needs analysis resulted in the following conclusions:

- The user groups in the total transport market look for advanced electronic, datasafe, reliable, flexible and user-friendly transport monitoring and management systems.
- The SMEs within the transport market need these systems at low cost to be able to adopt them and become more effective, and thus more competitive.
- The ideal system must be able to interface to existing technologies, i.e. radio- and tele-communication, and to existing infrastructures.
- Transport modes other than road transport require an improved supervision system for intermodal transport if the needs of sensitive cargo are to be met. If such improved facilities were available, user groups would be motivated to use combined transport modes more in the future.
- If a system can be used/exploited in other public areas such as AVI, public transport, emergency transport and hazardous transport, etc., it will increase the interest for many different user groups, public as well as private, thus speeding up and easing its introduction in the market place.
- The TRACAR project will address the needs of end-users who want to make use of any combination of road, rail, sea and inland waterways transport. The end-users need a solution as to how, at any time during the journey, they can obtain reliable information about the geographical position and condition of their freight. Thus, a total logging of the entire freight operation is required for providing all parties involved (i.e. end-users, forwarding agents, transporters and customs) better information quality.

The above analysis has determined that the following conditions and measurements should be automatically recorded and secured for any combined transport journey:-

- Information concerning the forwarding/shipping agent, sender/receiver, loading place, freight, temperature and humidity measurement during loading.
- The transporter’s identity, arrival and departure time to/from loading places.
- The conditions during accompanied journeys, including temperature and humidity measurements, essential vibrations, transport route and transport time.
- The transporter’s arrival time for reloading to another transporter or terminal.
- The conditions at reloading from one transporter to another.
- The temperature and humidity at the acceptance of the freight by the receiver/terminal.
• The automatic electronic alarms at changes in the freight unit’s conditions during transport or storage, including changes in temperature and/or humidity, or by non-authorised entry to, or movement of the freight unit from, the transporting unit or from the terminal.

• Automatic electronic positioning of the freight unit during transport or storage.

• Electronic remote controlled adjustment of temperature and humidity in transported or stored freight unit.

• Electronic access to reading and treatment of forwarding documents, references, etc. and recorded data.

The TRACAR project will provide the necessary system and tools for all transport companies, especially for small and medium-sized companies where cost is a factor, thus improving the entire combined freight transport sector’s ability to react to changes in market demands.

Corresponding to the user needs identified, the objectives of the two applications are to:-

• provide data safe, reliable, flexible and user-friendly tools for transport monitoring and management;

• enable users become more effective, and thus more competitive, through low cost of the equipment and with an interface to existing technologies, i.e. radio- and tele-communication as well as to existing infrastructures;

• improved supervision of intermodal transport to meet the needs of sensitive cargo;

• support other public areas such as AVI, public transport, emergency transport and hazardous goods transport.

In Table 8, the key characteristics of the two applications are summarised in a tabular format, under the following headings:

• application name or type

• major technologies whose application is going to be tested

• functionality of the application to be tested (based on CORD’s definitions of functions and sub-functions)

• test or demonstration site name.
### Table 8: Application overview

<table>
<thead>
<tr>
<th>Application</th>
<th>CORD Functions/ Sub-function ¹</th>
<th>Technologies and/or Approach Used</th>
<th>Decision makers</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. TUIP</td>
<td>• Logistics &amp; Freight Management Operation Control (SF 8.1.4) Fleet/Resource Management Operation Control (SF 8.2.3)</td>
<td>• Low frequency tag/antenna • On-board Computer • GPS • Existing Communication System</td>
<td>• shippers • forwarders • terminal operators • rail transport operators • transport operators • technology providers</td>
</tr>
<tr>
<td>2. TUM</td>
<td>• Vehicle/Cargo Management Preparation (SF 8.3.1) • Vehicle/Cargo Management Operation (SF 8.3.2)</td>
<td>• Memorising Tag, • Upgrade L/F technology • Sensors &amp; Controllers • DSRC • Existing Communication systems</td>
<td>• shippers • forwarders • terminal operators • rail transport operators • transport operators • technology providers</td>
</tr>
</tbody>
</table>

Note:

1. CORD Functions/Subfunctions have been revised (CONVERGE Project, 1997).

### 6.2.1 TUIP - Transport Unit Identification and Positioning

The Project will use low radio frequency, non-battery powered, tag/transponder technology for locating and positioning of vehicles and swap bodies locally, and GPS for wide area positioning.

The transporting units will be equipped with controllers, short range radio equipment and with interfaces to various sensors for temperature supervision. The transporting units will also be equipped with positioning and identification equipment.

For the purpose of identification, a tag will be mounted at the top of the swap body unit. When stacked, an antenna placed at the bottom of the swap body can read the tag at the top of the unit below it, thus establishing its position in the terminal and in the stack. The antenna at the bottom of the unit will also be able to communicate with tags placed on the road or terminal surface - see Figure 4: System functioning at the terminals (entrance/exit, storage).
A figure (not a system architecture diagram) showing the technology and approaches being used and the physical linkages between different elements of the technology is an essential part of a validation plan.

**Figure 4: System functioning at the terminals (entrance/exit, storage)**

By equipping the vehicles with tag-readers and by implementing tags on the surface of the road at the terminal entrances and exits, automatic registration of all conditions during arrival, hand-over and departure of the cargo can take place.

With full utilisation of the concept, all immobile sites (factories, terminals, depots, etc.) and the transporting/pulling units (trucks, locomotives or ship/river boats) will be equipped with controllers for interfacing to the equipment in the transporting units by short range radio and the existing communication systems.

If a suitable communication technology is not available in the pulling units, portable controllers can be fitted and interfaced to all existing standard communication systems, and the necessary communication will be based on one or more of these systems. Applied tags, which are inductive, can in addition be combined with the systems for automatic vehicle identification (AVI) as well as the systems for precise positioning. In order to comply with future demands for memory capacity, the read/write speed, a broader applicability of a common technology in the transport sector as a whole, and non-battery products, the existing Low Frequency Communication technology have to be upgraded. All other technologies and applications used in the demonstrator, are based on existing products and know-how.
6.2.2 TUM - Transport Unit Management

For the purpose of information storage and in order to secure the cargo documentation against manipulation, a non-erasable memory-tag logging all events deviating from the pre-set transport conditions, and an antenna are to be placed beneath the swap body. The equipment will be linked to the controller within the unit. The memorising tag can also be accessed from the pulling, or terminal equipment, via the interface as well as with external portable equipment. On arrival at the destination, a complete log of all relevant aspects of the journey is available and can provide evidence in the event of disputes or insurance claims.

In order to avoid any problems concerning the different infrastructure administrations in the Test Corridor and in the absence of a common standard of communication equipment, and similar absence of equipment for positioning in the railway locomotive for the actual demonstration, only the trucks used and not the railway locomotive, will be equipped with controller systems. Instead, a controller system will be installed in each block-transport on the rail journey.

These controllers will contain the necessary equipment for positioning of the block of units being transported, for internal communication with a single unit in the block and for communicating with the surrounding world including the consignor, forwarding agent, operator, etc. In the demonstrator, all communication with the mobile unit (truck, train) or immobile unit (terminal, depot) is based on the use of the GSM system. The controllers on the trucks can also be connected to the existing GPS systems or other systems for positioning of transport units.

Any authorised user involved in the combi-trans process can, on a remote basis, using radio, GSM or satellite communications, access the data collected and effect changes as allowed and/or required (e.g. temperature, humidity regulation).

The expected result is the development of a standard for supervision/management of combined freight transport, and a precise data logging of the total journey, including the relevant forwarding data. The system is expected to be especially attractive for SME users.

Expected impacts should be examined to determine which can be validated. An explanation should be provided for decisions not to validate an impact. The critical consideration is that it may not be able to validate particular impacts for sound methodological reasons but validation of such impacts might be crucial in terms of assessing the benefits of the application.

Any project must have objectives, leading to intended impacts, those which the telematics applications should produce. The assessment should verify these expected impacts. In fact, the project cannot be properly designed without a good idea of the nature and scale of likely impacts. If the system would affect certain categories of users or non-users, this would suggest the groups to be targeted in any measurements or data collection, and the respective sample size (for measurements, questionnaires, etc.) required.
Thus it is necessary to identify system impacts expected in principle and assess their approximate magnitude and target groups, before deciding which impacts should be assessed in the project. Where it is neither necessary nor reasonable to assess all possible system impacts, such a process can help focus the evaluation activities on the most important issues. In Table 9, an overall picture of the pre-assessment is depicted providing for each of the applications and the affected target groups, an estimation of the expected impacts’ magnitude.

**Table 9: Impacts expected**

<table>
<thead>
<tr>
<th>Impacts expected</th>
<th>Target groups</th>
<th>Applications</th>
<th>Impact / Magnitude</th>
</tr>
</thead>
<tbody>
<tr>
<td>Improved operational efficiency</td>
<td>shippers</td>
<td>TUIG</td>
<td>++</td>
</tr>
<tr>
<td></td>
<td></td>
<td>TUM</td>
<td>++</td>
</tr>
<tr>
<td></td>
<td>combined transport operators</td>
<td>TUIG</td>
<td>++</td>
</tr>
<tr>
<td></td>
<td></td>
<td>TUM</td>
<td>++</td>
</tr>
<tr>
<td></td>
<td>drivers</td>
<td>TUIG</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>TUM</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>terminal operators</td>
<td>TUIG</td>
<td>++</td>
</tr>
<tr>
<td></td>
<td></td>
<td>TUM</td>
<td>++</td>
</tr>
<tr>
<td></td>
<td>insurance companies</td>
<td>TUIG</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>TUM</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>public authorities</td>
<td>TUIG</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>TUM</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>society</td>
<td>TUIG</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>TUM</td>
<td>0</td>
</tr>
<tr>
<td>Improved service quality provision</td>
<td>shippers</td>
<td>TUIG</td>
<td>++</td>
</tr>
<tr>
<td></td>
<td></td>
<td>TUM</td>
<td>++</td>
</tr>
<tr>
<td></td>
<td>combined transport operators</td>
<td>TUIG</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td></td>
<td>TUM</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>drivers</td>
<td>TUIG</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>TUM</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>terminal operators</td>
<td>TUIG</td>
<td>++</td>
</tr>
<tr>
<td></td>
<td></td>
<td>TUM</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>insurance companies</td>
<td>TUIG</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>TUM</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>public authorities</td>
<td>TUIG</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>TUM</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>society</td>
<td>TUIG</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td></td>
<td>TUM</td>
<td>0</td>
</tr>
</tbody>
</table>
## Table 9: Impacts expected (cont’d)

<table>
<thead>
<tr>
<th>Impacts expected</th>
<th>Target groups</th>
<th>Applications</th>
<th>Impact Magnitude</th>
</tr>
</thead>
<tbody>
<tr>
<td>Improved working conditions</td>
<td>shippers</td>
<td>TUIP</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>TUM</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>combined transport</td>
<td>TUIP</td>
<td>++</td>
</tr>
<tr>
<td>operators</td>
<td></td>
<td>TUM</td>
<td>++</td>
</tr>
<tr>
<td></td>
<td>drivers</td>
<td>TUIP</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td></td>
<td>TUM</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>terminal operators</td>
<td>TUIP</td>
<td>++</td>
</tr>
<tr>
<td></td>
<td></td>
<td>TUM</td>
<td>++</td>
</tr>
<tr>
<td></td>
<td>insurance companies</td>
<td>TUIP</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>TUM</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>public authorities</td>
<td>TUIP</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>TUM</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>society</td>
<td>TUIP</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>TUM</td>
<td>0</td>
</tr>
<tr>
<td>Increased market opportunities</td>
<td>shippers</td>
<td>TUIP</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td></td>
<td>TUM</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>combined transport</td>
<td>TUIP</td>
<td>++</td>
</tr>
<tr>
<td>operators</td>
<td></td>
<td>TUM</td>
<td>++</td>
</tr>
<tr>
<td></td>
<td>drivers</td>
<td>TUIP</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>TUM</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>terminal operators</td>
<td>TUIP</td>
<td>++</td>
</tr>
<tr>
<td></td>
<td></td>
<td>TUM</td>
<td>++</td>
</tr>
<tr>
<td></td>
<td>insurance companies</td>
<td>TUIP</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>TUM</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>public authorities</td>
<td>TUIP</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>TUM</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>society</td>
<td>TUIP</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>TUM</td>
<td>0</td>
</tr>
<tr>
<td>Reduced environmental pollution</td>
<td>shippers</td>
<td>TUIP</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>TUM</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>combined transport</td>
<td>TUIP</td>
<td>0</td>
</tr>
<tr>
<td>operators</td>
<td></td>
<td>TUM</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>drivers</td>
<td>TUIP</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>TUM</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>terminal operators</td>
<td>TUIP</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>TUM</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>insurance companies</td>
<td>TUIP</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>TUM</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>public authorities</td>
<td>TUIP</td>
<td>++</td>
</tr>
<tr>
<td></td>
<td></td>
<td>TUM</td>
<td>++</td>
</tr>
<tr>
<td></td>
<td>society</td>
<td>TUIP</td>
<td>++</td>
</tr>
<tr>
<td></td>
<td></td>
<td>TUM</td>
<td>++</td>
</tr>
</tbody>
</table>
Table 9: Impacts expected (cont’d)

<table>
<thead>
<tr>
<th>Impacts expected</th>
<th>Target groups</th>
<th>Applications</th>
<th>Impact Magnitude</th>
</tr>
</thead>
<tbody>
<tr>
<td>Improved road safety</td>
<td>shippers</td>
<td>TU</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>combined transport operators</td>
<td>TU</td>
<td>++</td>
</tr>
<tr>
<td></td>
<td>drivers</td>
<td>TU</td>
<td>++</td>
</tr>
<tr>
<td></td>
<td>terminal operators</td>
<td>TU</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>insurance companies</td>
<td>TU</td>
<td>++</td>
</tr>
<tr>
<td></td>
<td>public authorities</td>
<td>TU</td>
<td>++</td>
</tr>
<tr>
<td></td>
<td>society</td>
<td>TU</td>
<td>++</td>
</tr>
<tr>
<td>Reduced transport costs</td>
<td>shippers</td>
<td>TU</td>
<td>++</td>
</tr>
<tr>
<td></td>
<td>combined transport operators</td>
<td>TU</td>
<td>++</td>
</tr>
<tr>
<td></td>
<td>drivers</td>
<td>TU</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>terminal operators</td>
<td>TU</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>insurance companies</td>
<td>TU</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>public authorities</td>
<td>TU</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>society</td>
<td>TU</td>
<td>++</td>
</tr>
</tbody>
</table>

* ++ very positive; + positive; 0 neutral; - negative; -- very negative
Objectives

Assessment objectives reflect judgements made about what expected impacts can be validated. This means that determination of assessment objectives is an iterative process which reflects what decision makers would like assessed in principle and is constrained by what assessment objectives, particularly those related to impacts, can be validated in practice.

The assessment process is to start by identifying assessment objectives on the basis of the, requirements and preferences of the decision makers and other stakeholders concerned in the project. The Assessment objectives should correspond to the criteria for making judgements and choices and a hierarchy needs to be established from the general to the specific.

The Project has established assessment objectives on the basis of the user needs analysis. To ensure a consistent approach it was thought helpful to depict the relationships between the:
- transport telematics applications to be tested,
- relevant decision makers, and
- assessment objectives
- in an overall picture, as shown in Table 10.

### Table 10: Decision makers and assessment objectives

<table>
<thead>
<tr>
<th>Applications</th>
<th>Decision Makers</th>
<th>First Level Assessment Objectives</th>
<th>Second Level Assessment Objectives</th>
</tr>
</thead>
<tbody>
<tr>
<td>TUIP</td>
<td>shipper</td>
<td>to determine the extent to which the application can</td>
<td></td>
</tr>
<tr>
<td></td>
<td>forwarder</td>
<td>improve the transparency of combined transport operations</td>
<td></td>
</tr>
<tr>
<td></td>
<td>terminal operators</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>rail transport operator</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>transport operator</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>technology provider</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>terminal operators</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>rail transport operator</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>transport operator</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>technology provider</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>insurance organisation</td>
<td></td>
<td></td>
</tr>
<tr>
<td>TUM</td>
<td>shipper</td>
<td>to determine the extent to which the application can</td>
<td></td>
</tr>
<tr>
<td></td>
<td>forwarder</td>
<td>improve the control over the total combined transport chain</td>
<td></td>
</tr>
<tr>
<td></td>
<td>terminal operators</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>rail transport operator</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>transport operator</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>technology provider</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>insurance organisation</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

- to assess the system’s
  - reliability
  - user friendliness
  - market opportunities

- to determine the extent to which the application can:
  - improve working conditions
  - improve the level of services provided
  - improve the efficiency of operations
  - contribute to reduced environmental pollution and improved safety
  - reduce transportation costs for the society
### Verification

Verification is the first stage of validation. Technical assessment and tests of user acceptance also feature at the second stage of validation – the demonstration stage. The emphasis on technical assessment at the verification stage is on whether the application meets its physical performance standards in terms of correct functioning and meeting standards; the emphasis at the demonstration stage is on reliability of physical functioning because the demonstration stage is usually longer and under more varied conditions. The main tests of user acceptance are at the demonstration stage when a larger and more varied sample of users should be involved. The objective of user acceptance tests at the verification stage is to determine that the main decision makers, at least, are satisfied that the application is performing well enough to justify proceeding to demonstration stage.

In this example, technical assessment and tests of user acceptance are only illustrated at the verification stage.

#### Technical assessment

Technical assessment focuses on the question whether the tested applications fulfil their pre-specified technical requirements and objectives. Technical assessment is usually the first assessment to be performed in the research and development process of an ATT system. It is used as a tool to manage the technical development, that is to establish if the system is reliable and effective from the technical point of view. The indicators (i.e. evaluation criteria) which are to be used for this assessment, are technical indices of the functions of the tested equipment. The objective of the present technical assessment and the related user groups are presented below:

<table>
<thead>
<tr>
<th>Assessment objectives</th>
<th>User groups involved</th>
</tr>
</thead>
</table>
| Assess system reliability | • shippers  
|                        | • combined transport operators  
|                        | • drivers  
|                        | • terminal operators |

The detailed approach for the technical assessment is summarised in Table 12.
Table 12 is fairly typical of experimental design for technical assessment – physical measurement is generally straightforward given adequate instrumentation, transparent performance indicators, no major constraints on securing an adequate number of measurements and, as a consequence, tests can usually be carried out at a high level of statistical accuracy – of course, the corollary is that these conditions are a pre-requisite for demonstrating the required high levels of reliability of the physical functioning of the system.

Table 12: Methods for technical assessment

<table>
<thead>
<tr>
<th>Application</th>
<th>Assessment Objective</th>
<th>Indicators</th>
<th>Reference Case</th>
<th>Methods of Measurement or Simulation</th>
<th>Measurement Conditions or Conditions to be Simulated</th>
<th>Statistical Considerations</th>
<th>Measurement Plan</th>
<th>Integrity of Measurement or Simulation</th>
</tr>
</thead>
<tbody>
<tr>
<td>TUIP</td>
<td>• System reliability</td>
<td>• number of breakdowns • Successful position polling (%) • Successful ID requests (%)</td>
<td>pre-defined criteria</td>
<td>manual recording</td>
<td>Real-life operation</td>
<td>● 1000 ● 500 ● 1000</td>
<td>95% ● 0.3% ● 95 % ● 95%</td>
<td>1 July - 31 Dec. 1997</td>
</tr>
<tr>
<td>TUM</td>
<td>• System reliability</td>
<td>• number of breakdowns • Successful remote intervention (%)</td>
<td>pre-defined criteria</td>
<td>manual recording</td>
<td>Real-life operation</td>
<td>● 1000 ● 100</td>
<td>95% ● 0.3% ● 95 %</td>
<td>1 July - 31 Dec. 1997</td>
</tr>
</tbody>
</table>
Tests of User Acceptance tend to concentrate on the HMI of the workstation. This would certainly be the case at the verification stage of validation. However, User Acceptance encompasses all the performance characteristics of the application. So, particularly at the demonstration stage of validation, the concern is whether the application meets the expectations of the users in totality. This means that the results of technical assessment and of impact analyses (and related social cost benefit analyses, financial assessments, etc.) also contribute to tests of user acceptance.

User acceptance assessment aims at estimating users’ attitudes to and perception of application(s) investigated. Here the users may be the shippers, transport operators or the authorities who will operate the system, etc.

For the user groups that actually operate the system, user acceptance is often estimated in terms of perceived benefits, e.g. improvement in safety, reduction of transport costs, HMI friendliness, etc.

The Assessment will be performed by using interviews, user forums etc. In Table 13 user acceptance assessment objectives are listed, together with relevant users.

### Table 13: Assessment objectives and the related user groups

<table>
<thead>
<tr>
<th>Assessment objectives</th>
<th>User groups involved in validation</th>
</tr>
</thead>
<tbody>
<tr>
<td>friendliness</td>
<td>• shippers</td>
</tr>
<tr>
<td></td>
<td>• combined transport operators</td>
</tr>
<tr>
<td></td>
<td>• drivers</td>
</tr>
<tr>
<td></td>
<td>• terminal operators</td>
</tr>
<tr>
<td>Assess impact on working conditions</td>
<td>• shippers</td>
</tr>
<tr>
<td></td>
<td>• combined transport operators</td>
</tr>
<tr>
<td></td>
<td>• drivers</td>
</tr>
<tr>
<td></td>
<td>• terminal operators</td>
</tr>
<tr>
<td>Assess impact on service level provision</td>
<td>• shippers</td>
</tr>
</tbody>
</table>
The detailed approach for the user acceptance assessment is summarised in Table 14.

Table 14 is based on TRACAR’s draft validation plan – the following points are relevant to development of the final validation plan:

- **assessment objectives and related indicators should be defined more precisely; for example, ‘User Friendliness’ will need to be defined in terms of the characteristics of the workstation and these characteristics would feature explicitly in questions in the interviews and/or questionnaires**

- **measurement methods will vary between verification and demonstration – the main difference will be participation of fewer users, probably just representatives of the main decision makers, less intensive measurement, and probably use of less formal methods of measurement at the verification stage – however, the outcome of verification must still be results in which confidence can be placed**

- **measurement conditions will require specification for both verification and demonstration stages – the main difference will be more limited use of the application at the verification stage,**

- **it is necessary to keep a clear view about the number of questionnaires/interviews required – it is related to questionnaire design and results analysis – the more it is necessary to disaggregate results in term of the characteristics of respondents on one hand, and those of performance of the application on the other, the greater the number of questionnaires that will be required to satisfy a certain level of statistical confidence – put another way, if resources control the number of interviews/questionnaires, the degree of disaggregation of the results analysis will control the statistical confidence level**

- **overall definition of success will have to be related more precisely to precise definitions of the indicators for ‘user friendliness, improved working conditions and improved service level’ – see above**

- **the measurement plan will need to be defined more precisely; for example, will interviews/questionnaires be used only at the end of the demonstration period for validation of the demonstration stage?**

- **there are usually considerations of integrity with use of interviews and questionnaires such as:**
  - the representativeness and partiality of the sample of respondents
  - the confidence that can be placed in self-completed questionnaires.
### Table 14: Methods for user acceptance assessment

<table>
<thead>
<tr>
<th>Application</th>
<th>Assessment Objective</th>
<th>Indicators</th>
<th>Reference Case</th>
<th>Methods of Measurement or Simulation</th>
<th>Measurement Conditions or Conditions to be Simulated</th>
<th>Statistical Considerations</th>
<th>Measurement Plan</th>
<th>Integrity of Measurement or Simulation</th>
</tr>
</thead>
<tbody>
<tr>
<td>TUIP</td>
<td>Ease of use (ordinal scale)</td>
<td>pre-defined criteria</td>
<td>interviews, questionnaires</td>
<td>-</td>
<td>100</td>
<td>95%</td>
<td>85%</td>
<td>1 Jan. - 28 Feb. 1998</td>
</tr>
<tr>
<td></td>
<td>Improved working conditions</td>
<td>decreased paper handling (ordinal scale)</td>
<td>existing situation</td>
<td>interviews, questionnaires</td>
<td>-</td>
<td>100</td>
<td>95%</td>
<td>85%</td>
</tr>
<tr>
<td></td>
<td>Improved service level</td>
<td>improved information quality provision (ordinal scale)</td>
<td>existing situation</td>
<td>interviews, questionnaires, user forums</td>
<td>-</td>
<td>100</td>
<td>95%</td>
<td>85%</td>
</tr>
<tr>
<td>TUM</td>
<td>User friendliness</td>
<td>Ease of use (ordinal scale)</td>
<td>pre-defined criteria</td>
<td>interviews, questionnaires</td>
<td>-</td>
<td>100</td>
<td>95%</td>
<td>85%</td>
</tr>
<tr>
<td></td>
<td>Improved working conditions</td>
<td>decreased paper handling (ordinal scale)</td>
<td>existing situation</td>
<td>interviews, questionnaires</td>
<td>-</td>
<td>100</td>
<td>95%</td>
<td>85%</td>
</tr>
<tr>
<td></td>
<td>Improved service level</td>
<td>improved exception management (ordinal scale)</td>
<td>existing situation</td>
<td>interviews, questionnaires, user forums, manual recording</td>
<td>-</td>
<td>100</td>
<td>95%</td>
<td>85%</td>
</tr>
</tbody>
</table>
### Demonstration

Technical assessment (tests of physical functioning) will continue in the demonstration stage of validation with an emphasis on monitoring the reliability of the applications. The main tests of user acceptance will be done at the demonstration stage. This section focuses on impact analysis which generally features mainly at the demonstration stage.

#### 6.6.1 Impact assessment

Impact assessment is the measurement or estimation of the impacts (effects) of an application, (such as those on transport efficiency, user behaviour, etc.) for the affected target groups (shippers, system operators, society as a whole etc.).

Within the Project a socio-economic evaluation will be carried out. In order to relate impact assessment to the socio-economic evaluation methods to be used, impacts have been classified into three broad categories (see Figure 5):

- impacts which can be expressed directly in monetary terms
- impacts which can be expressed indirectly in monetary terms
- impacts which can only be expressed only in qualitative terms.

The reason for the existence of this evaluation is that, the development of such advanced systems should also provide significant benefits for the society as a whole, in order to justify the spending of public money (through EU or national governments). Thus, an assessment of the social impact of the applications, will be performed. The socio-economic indicators will derive from the society’s objectives. The possible effects of the pilots on the society as a whole, can be summarised in two broad categories:

- increased transport efficiency (leading to the reduction of the cost of transport operations and the enhancement of the service level enjoyed by the simple citizen)
- increased road safety and environmental protection (leading to the improvement of life quality).

It has been estimated that, if only 5% of the total road cargo transport (31.5 thousand million tonne-km) was transferred to rail transport, the net saving of CO2 emissions (road minus rail = 0.093 kg per tonne-km) would be 5,900 million tonnes CO2 emissions. It is further anticipated that, if 50% of all food currently transported by road was transferred to rail (quite possible with reliable and safe supervision) the net saving of CO2 emissions in the EU would be approximately 7,000 million tonnes.

Apart from the environmental benefits and savings in this area, it is estimated that the annual transport costs would be reduced by 2,000 MECU if 50% of road food cargo was transferred to rail. At the same time, the railway infrastructure would receive an extra turnover of approximately 4,400 MECU, which would help the future development of this important sector.
The main issues to be investigated are:

- Firstly, what are these potential cost reductions?
- Secondly, how may the benefits from cost reductions and greater efficiency in utilisation of resources be distributed?
- Thirdly, what are the related overall benefits to society?

In order for the society’s economic benefits to be estimated, the following procedure will be followed:

- during the demonstrator, the cost reductions due to ATT use, will be measured
- these reductions will be extrapolated to the European road freight market
- an estimation of the consumer’s consequent economic benefits will be made, according to economic theory.

Also, contribution of the applications to safety and environmental protection is to be estimated. This will be done by estimating the reduction of accidents and emissions due to cargo shifted from road to combined transport, taking into account existing transportation statistics.

The evaluation methods to be used have been selected for different impact categories (Figure 5).

**Figure 5: Impact categorisation and respective evaluation methods**

![Diagram of impact categorisation and evaluation methods]

- Impacts
  - monetary expressed
    - measurable but with no agreed money values (e.g. time savings)
      - Cost-Benefit Analysis
      - Experts opinion or literature
  - non-monetary expressed
    - purely qualitative (e.g. customer service enhancement)
      - Medium-term impacts
      - Long-term impacts
        - Multi-Criteria Analysis
        - SWOT Analysis

- Synthesis
  - Multi-Criteria Analysis
In Table 15 impacts and socio-economic assessment objectives are listed, together with affected target groups.

### Table 15: Assessment objectives and the affected target groups

<table>
<thead>
<tr>
<th>Assessment objectives</th>
<th>User groups involved in validation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Assess impact on environmental pollution</td>
<td>public authorities</td>
</tr>
<tr>
<td></td>
<td>society</td>
</tr>
<tr>
<td>Assess impact on safety</td>
<td>public authorities</td>
</tr>
<tr>
<td></td>
<td>society</td>
</tr>
<tr>
<td>Assess impact on transport cost reductions for the simple citizen</td>
<td>society</td>
</tr>
</tbody>
</table>

Table 16 is based on TRACAR’s draft validation plan. The following are considerations for the final validation plan:

- As efficiency of operations features as a second level assessment objective (see Table 10), it would be reasonable to include an analysis of the impact of the application on the methods of operation of the main users of the applications and a related financial assessment – as this is related to the assessment objectives for user acceptance concerning ‘improved working conditions’ and ‘improved service level,’ (Table 14) these objectives should be made part of this impact analysis – a complication for such an analysis might be the operation of the applications in parallel with current systems and applications during the demonstration – this situation would obviously make it difficult to associate unambiguously the impacts caused by the applications with the applications themselves – however, investment is unlikely to take place in these applications unless they generate financial benefits for the freight operators concerned – it is, therefore, paramount that every effort is made to achieve results from validation which assist decisions regarding investment by freight operators, even if such results are constrained by practical difficulties to be the outcome of ‘partial analysis.’

- The indicators provided for ‘reduced environmental pollution’ and ‘improved safety’ are very aggregate – in reality both these criteria will be sensitive to the types of roads used and other characteristics of routes for road freight and combined transport – details of these routes and the method of calculation of such impacts should be provided in the final validation plan.

- Figure 5 shows, in principle, the way of estimating cost reductions – the final validation plan should lay out the detailed method of calculation.

- The number of measurements is probably the number of operations of the trains in the combined transport operation during the demonstration stage of validation – this should be made clearer.

- Considerations of integrity (of estimation in this case) should include the robustness of the assumptions which will be necessary to expand the results of a very specific freight operation to the potential effects on EU combined transport – the method of expansion and the associated assumptions should be provided in the final validation plan.
**Table 16: Methods for impacts and socio-economic assessment**

<table>
<thead>
<tr>
<th>Application</th>
<th>Assessment Objective</th>
<th>Indicators</th>
<th>Reference Case</th>
<th>Methods of Measurement or Conditions to be Simulated</th>
<th>Statistical Considerations</th>
<th>Measurement Plan</th>
<th>Integrity of Measurement or Simulation</th>
</tr>
</thead>
<tbody>
<tr>
<td>TUIP and TUM Reduced environmental pollution</td>
<td>reduction of driven kms on the roads existing situation</td>
<td>• direct measurement and model-based analysis • extrapolation of demonstrator results to EU level</td>
<td>Real-life operation</td>
<td>200 95%</td>
<td>Reduciton of 10,000 veh-mill. Km in EU</td>
<td>1 April - 31 Aug. 1997</td>
<td>-</td>
</tr>
<tr>
<td>Improved safety</td>
<td>reduction of driven kms on the roads existing situation</td>
<td>• direct measurement and model-based analysis • extrapolation of demonstrator results to EU level</td>
<td>Real-life operation</td>
<td>200 95%</td>
<td>Reduction of 10,000 veh-mill. Km in EU</td>
<td>1 April - 31 Aug. 1997</td>
<td>-</td>
</tr>
<tr>
<td>Cost reductions and more efficient use of resources</td>
<td>transport cost reduction (%) existing situation</td>
<td>• model-based analysis • extrapolation of demonstrator results to EU level</td>
<td>Real-life operation</td>
<td>200 95%</td>
<td>Cost reduction by 2000 MECU in EU</td>
<td>1 April - 31 Aug. 1997</td>
<td>-</td>
</tr>
</tbody>
</table>
In the context of validation, the issue regarding ‘European Added Value’ is making the validation results more valuable.

This can be done by combining results with those from comparable projects. This requires in order of importance:

i) comparable functionalities
ii) comparable assessment objectives
iii) comparable demonstration sites
iv) comparable experimental design.

The most comprehensive comparison of results can be done when all four conditions are met but results can be compared, albeit in more general terms, if only the first condition is met.

Obviously, common validation plans facilitate this process.

It should be borne in mind that the Commission may oblige projects to co-operate in this way and may consider providing additional resources to permit projects to add value to their results.

The Project is keen to demonstrate both an awareness and confidence to ensure that a “European Added Value” dimension be incorporated within the Project.

The Transport Telematics Programme includes a large number of projects whose objectives and content overlap or are shared. The real “customers” for the results of projects include as well as the direct participants also a broad audience of potential users, purchasers, manufacturers and operators. For them it is important to know how the various systems and applications perform, how much they cost to buy and to operate, and what impacts they might produce if implemented.

The Project partners appreciate that this sort of knowledge is not directly available from any one project alone. Without appropriate syntheses of project results, decision makers (those who will manufacture, own, operate or use transport telematics systems) will have little evidence of the relative advantages of comparable systems, and which may be the most appropriate for their particular environmental and traffic conditions and for their own policies/purposes.

The Consortium will consider the interoperability aspects of its solution with other existing solutions and projects running in the 4th Framework Programme. A first approach with two similar projects, has already been made and further co-operation is to be expected at the following stages of the project.
7. **Examples of elements of model validation plans**

7.1 **Introduction**

In this chapter, two hypothetical examples of impact analyses are presented using the framework to preparation of a validation plan recommended by the CONVERGE Project (Zhang, *et al.*, 1996; Maltby and Morello, 1996).

The reasons for focusing on impact analyses are:

- it is reasonable to expect that validation approaches for assessment objectives dealing with impacts will vary more between applications, and particularly between application areas, than those for tests of physical functioning or tests of user acceptance
- the CONVERGE review of draft validation plans (Maltby, *et al.*, 1996) indicated that impact analysis was an area of major difficulty for projects.

A discursive approach has been used in order to emphasise the professional judgements involved in preparing a validation plan because of practical difficulties in formulating perfect plans and the importance of limiting validation to assessment objectives which are likely to lead to useful validation results.

7.2 **Hypothetical Example of Impact Analysis for the Telematics Applications for Travellers Area - Videotex Demand Spreading Application**

7.2.1 **Background**

In France, there are about 5 million Videotex (or ‘Minitel’) terminals, mainly in domestic use but also available in offices, hotels and supervised public locations. The cost of their use is low and they can be used to obtain information on services generally, weather conditions, and road travel conditions on the major roads (motorways and national routes) of the national road network (called 3615 ROUTE). Traffic conditions are described in fairly general terms such as ‘heavy congestion,’ ‘heavy traffic but fluid conditions,’ and ‘light traffic and no congestion.’

France also has a national ‘real-time’ database (the TIGRE system) of traffic conditions on the national road network. This information is supplied by motorway and other traffic information centres to the regional CRIRCs (Regional Centres for Traffic Information and Communications) and thence to the CNIR (National Centre for Traffic Information) near Paris. By ‘real-time,’ we mean traffic conditions in the immediate past few minutes – as long as it takes to collect the information automatically and transfer it to the CNIR or any end-user. Traffic conditions are also available on a historical basis for forecast periods beyond ‘real-time.’ These take into account seasonal factors, day of the week, time of day, planned maintenance and special events.

The national road network is available on the Michelin GDF 2.0 cartographical database. So traffic conditions, road events, and trip origins and destinations can be described with reference to this database.
An itinerary calculator is available. Its algorithm uses ‘real-time’ or ‘forecast’ traffic information to estimate trip duration and to propose alternative itineraries.

7.2.2 Objectives of Hypothetical Application

It is proposed to extend the videotex traffic service by providing route itinerary advice on the basis of either ‘real-time’ or forecast traffic conditions. The motorway toll cost of alternative itineraries will also be provided.

Additionally, a public transport database for services by train and air will be added to the traffic service.

The main objective of the demonstration from the viewpoint of impact analyses will be to examine the contribution of the application to ‘demand spreading;’ that is, what is the contribution of the application to ameliorating demand for travel by car by virtue of

- changing route
- rescheduling of journey
- changing mode.

It is also possible that the application might be used by road freight operators from the viewpoint of choosing the best route and rescheduling of journeys.

7.2.3 Summary Review

Table 17: Technologies and/or Approach Used, CORD Functions/Sub-Functions and Decision Makers

<table>
<thead>
<tr>
<th>Technologies and/or Approach Used</th>
<th>CORD Functions/Subfunctions ¹</th>
<th>Decision Makers</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Videotex (Minitel) Network</td>
<td>SF7.1.1 transport mode selection/planning</td>
<td>• Videotex operator</td>
</tr>
<tr>
<td>• ISDN network for PCs with a ‘Minitel’ interface</td>
<td></td>
<td>• National road administration</td>
</tr>
<tr>
<td>• Videotex server</td>
<td>SF7.1.2 route planning</td>
<td>• Road operators</td>
</tr>
<tr>
<td>• Itinerary Calculator</td>
<td>SF7.2.1 road characteristics</td>
<td>• Rail operators</td>
</tr>
<tr>
<td>• TIGRE ‘real-time’ database</td>
<td></td>
<td>• Air operators</td>
</tr>
<tr>
<td>• ‘Static’ databases</td>
<td></td>
<td>• Road freight operators</td>
</tr>
<tr>
<td>GDF database</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note:
1. CORD Functions/Subfunctions have been revised (CONVERGE Project, 1997)
7.2.4 Expected Impacts and Appraisal Groups

7.2.4.1 Expected Impacts

The main expected impacts are:
- change of route by car driver
- rescheduling of journey by car driver
- change of mode by car driver
- change of route for lorry
- rescheduling of route for lorry.

For rescheduling of journeys, the possibilities will range from changing the hour but not the day to changing both the hour and day of departure. These changes might be associated with or without changes of route.

In principle, one should also be concerned about the impact of the new service on the videotex operator, in particular the financial effects in terms of investment costs, running costs and revenue. The extent to which it is possible to do this will depend upon the transparency of agreements about cost sharing and payment for services between the major decision-makers, the extent to which the market will determine the level of charge to the end-user, and the existence of a legal entity which has responsibility for providing the videotex service.

The general situation at the moment is that such conditions do not exist. Therefore, the minimum validation of this impact should consist of monitoring of the functioning of the application, estimation of costs which can be clearly attributed to the application, and estimating the level of charge which end-users might pay for the service if it is not provided on a commercial basis.

7.2.4.2 Appraisal Groups

Impacts depend upon changes in human behaviour. Appraisal groups are the groups changing their behaviour or the ways they do things; they enjoy the benefits of suffer from the disbenefits of the application.

Appraisal groups clearly involved are:
- car drivers
- lorry operators/drivers.

Appraisal groups involved less tangibly because it might be difficult validating the impacts of the application on them are:
- videotex operator
- national road administration
- road operators
- rail operators
- air operators.
Table 18 shows the expected impacts for the different appraisal groups. It is anticipated that car drivers are quite likely to change route or reschedule their journeys as a result of using the application but it is much less likely that they will change mode. Lorry operators/drivers will not change mode as a result of using the application because the public transport database is designed for passenger traffic. It is anticipated that the impact of the application on lorry operators/drivers changing route or rescheduling journeys will be less than for car drivers because:

- other systems with such objectives will already be in fairly common use
- there are greater constraints on rescheduling commercial journeys.

As already discussed the likely impact on the operations of all the parties contributing to the operation of the videotex application will be difficult to estimate.

### 7.2.5 Selection of impacts to be validated

In view of the above considerations, it has been decided to validate the following impacts:

- change route by car drivers
- rescheduling journey by car drivers
- change mode by car drivers
- change route by lorry operators/drivers
- reschedule of journey by lorry operators/car drivers.

It has also been decided to monitor all the transparent effects on the operations of the videotex operator. This should identify any significant effects at the interfaces of the videotex operator with the national road administration, road operators, rail operators and air operators.

### 7.2.6 All Assessment Objectives and Appraisal Groups

Table 19 shows the assessment objectives and groups involved in validation for impact analysis on the basis of the considerations discussed in Section 7.2.5
Table 18: Expected Impacts for different Appraisal Groups

<table>
<thead>
<tr>
<th></th>
<th>Change route</th>
<th>Reschedule journey</th>
<th>Change mode</th>
<th>Impact on operations</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Car drivers</td>
<td>++</td>
<td>++</td>
<td>-/0/+</td>
<td>N/A</td>
</tr>
<tr>
<td>2. Lorry operators/drivers</td>
<td>+</td>
<td>+</td>
<td>0</td>
<td>N/A</td>
</tr>
<tr>
<td>3. Videotex operator</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>-/0/+</td>
</tr>
<tr>
<td>4. National road administration</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>-/0/+</td>
</tr>
<tr>
<td>5. Road operator</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>-/0/+</td>
</tr>
<tr>
<td>6. Rail operators</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>-/0/+</td>
</tr>
<tr>
<td>7. Air operators</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>-/0/+</td>
</tr>
</tbody>
</table>

Table 19: Definition of Assessment Objectives and User Groups involved in Validation for Impact Analyses

<table>
<thead>
<tr>
<th>Category of Assessment</th>
<th>Assessment Objectives</th>
<th>Groups involved in Validation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Impact Analysis</td>
<td>1. To assess change of route by car drivers</td>
<td>• car drivers</td>
</tr>
<tr>
<td></td>
<td>2. To assess rescheduling of journeys by car drivers</td>
<td>• lorry drivers/operators</td>
</tr>
<tr>
<td></td>
<td>3. To assess change of mode by car drivers</td>
<td>• videotex operator</td>
</tr>
<tr>
<td></td>
<td>4. To assess change of route by lorry operators/drivers</td>
<td>• national road administration</td>
</tr>
<tr>
<td></td>
<td>5. To assess rescheduling of journeys by lorry operators/drivers</td>
<td>• road operators</td>
</tr>
<tr>
<td></td>
<td>6. To monitor impact on the operations of the videotex operator</td>
<td>• rail operators</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• air operators</td>
</tr>
</tbody>
</table>
7.2.7 Example of Selection of Assessment Objectives and Indicators

Table 20 shows the indicators for Assessment Objectives 1 to 3 in Table 19.

Table 20: Examples of Indicators for Assessment Objective 1 to 3 in Table 19

<table>
<thead>
<tr>
<th>Assessment Objective</th>
<th>Indicators</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. To assess change of route by car driver</td>
<td>Evidence of change of route</td>
</tr>
<tr>
<td>2. To assess rescheduling of journeys by car driver</td>
<td>Evidence of rescheduling of journeys</td>
</tr>
<tr>
<td>3. To assess change of mode by car drivers</td>
<td>Evidence of change of mode by car drivers</td>
</tr>
</tbody>
</table>

7.2.8 Experimental Design - An Example

This section presents an example for the Experimental Design comprising:

- Assessment Objectives
- Indicators
- Reference Case
- Method of Measurement or Simulation
- Conditions of Measurement or Simulation
- Statistical Considerations
  - Number of Measurements
  - Confidence Level
  - Overall Definition of Success
- Measurement Plan
- Integrity of Measurement or Simulation for Impact Analysis

Consider now the experimental design for the Assessment Objectives and indicators in Table 20.

7.2.8.1 Reference Case:

The reference case will be the bases on which car drivers would have chosen the routes and/or scheduling of journeys, or decided to use rail or air services instead, if they had not used the application.
This introduces a potentially great source of variation between end-users. Such alternative information sources will include:

- advice from other people
- maps
- radio information service at home
- radio service in car
- telephone calls to information centres
- information sources for public transport
- itineraries provided by automobile associations
- route planning software
- teletext service on TV
- Minitel 3615 Route Service.

### 7.2.8.2 Method of Measurement

It is unlikely to be practicable to obtain evidence of ‘demand spreading’ by monitoring changes in traffic flows and use of rail and air services in aggregate. The only possible methods of measurement will be:

- electronic questionnaires
- questionnaires sent out to respondents who indicate a willingness (perhaps by short electronic questionnaire which is part of the Minitel service) to co-operate and returned by post as a self-completed questionnaire
- structured interviews completed by an interviewer at a pre-arranged time with respondents who have indicated a willingness to co-operate
- panel surveys.

### 7.2.8.3 Conditions of Measurement

The main considerations will be:

- timing of the demonstration
- duration of the demonstration.

Timing will depend upon the type of traffic at which the application is aimed:

- holiday traffic will require a holiday period when traffic congestion is common and which is of sufficient duration to let end-users familiarise themselves with the system and to conduct surveys as well as possible
- business and commuting traffic will permit validation at any time of the year except the main holiday periods.
A complication will be that uses of the application will probably have quite significant regional differences in terms of the type of traffic even though a nationally – based database underpins the application.

The main consideration regarding duration will be that end-users should have sufficient time to learn about the existence of the application, and to familiarise themselves with it if they use it, and that the surveys can be conducted as well as possible.

**7.2.8.4 Statistical Considerations**

**7.2.8.4.1 Number of Measurements**

The results of the questionnaire surveys or structured interviews would be used to estimate the contribution of the application to demand spreading in terms of changing route, rescheduling journeys, or changing mode. Other changes in travel behaviour such as deciding not to make the journey or deciding to make another journey should not be overlooked.

So at the most aggregate level, end-users would fall into the following groups:

- those who changed their travel behaviour as a consequence of using the application
- those who did not change their travel behaviour as a consequence of using the application.

The number of surveys required depends upon how close an estimate to the statistical population is required and upon the relative importance of the different indicators of demand spreading. At this highest level of aggregation, the number of surveys required for a reasonable level of statistical accuracy would probably be in the low hundreds.

However, an analysis at this level is not very insightful because many factors will affect the results. Examples are:

- type of traveller in terms of such factors as socio-economic status, sex, amount of car travel per year, usual means of trip planning
- travel purpose
- anticipated level of congestion on the road network at the planned time of travel
- origin and destination of journey
- number of times application used – for example, an end-user might change his trip planning as a consequence of some uses of the application but not of others.

It is necessary, therefore, to determine the lowest level of disaggregation of the analysis which is necessary to obtain sound policy guidance from the validation. At such a level of disaggregation, the number of surveys required for a reasonable level of statistical accuracy would probably be in the low thousands.

**7.2.8.4.2 Overall Definition of Success**

The overall definition of success would be measured in terms of statistically significant evidence that use of the application had resulted in demand spreading.
7.2.8.5 Measurement Plan

The main considerations are:

- use of a representative sample of videotex users – a list of videotex private subscribers would provide a sample frame – residential and business subscribers should be included to investigate differences in use between different types of subscribers – use of videotex in supervised public places should probably be omitted because of the complications of conducting surveys of public use

- the new videotex service should be introduced sufficiently in advance of the demonstration stage of validation to ensure that it is known about – advertising in the media could be used – the important consideration about the effects of the intensity of advertising on the results of the validation is whether it simply generates more use of the application or whether the relative changes in travel behaviour which take place as a consequence of using the application vary with different levels of advertising

- the duration of validation should be sufficient to ensure good survey practice

7.2.8.6 Integrity of Measurement

The main considerations are:

- use of a representative sample of subscribers

- whether it is possible to have confidence in results from self-completed questionnaires – sounder results would probably be obtained by using more expensive structured interviews – if resource constraints mean that fewer interviews could be conducted, it is probably better to have results in which more confidence can be placed even though a lower statistical accuracy would be attached to them.
7.3 Hypothetical Example of Impact Analysis for the Maritime:
Waterborne Transport Telematics Application Area - Vessel Traffic
Systems Application

7.3.1 Background

Conventional vessel traffic systems (VTS) track vessels by radar. Radar has the disadvantages of:

- limited range
- shadowing effects of irregular hilly coastlines and islands
- inability to identify vessels automatically
- inability to classify automatically vessel cargoes, particularly in the special case of hazardous cargoes
- requires voice contact to enquire about details of vessel
- cannot confirm details of vessel.

The tracking capability of conventional VTS can be increased by installing transponders to vessels so that vessels so fitted can be tracked beyond the range of radar. Where vessels are beyond the range of shore-based VHF interrogators of transponders, positional data can be supplied to the VTS centre via satellite. Other advantages of using transponders are, depending on the system employed:

- fitted vessels can be identified more clearly on the graphical VTS traffic image display within the area of radar coverage
- short messages can be transmitted between VTS centre and fitted vessel and vice versa
- they offer the potential for ship - ship communication.

Use of transponders may also facilitate:

- information about a vessel such as ship identity, ship ETA at destination port, ship characteristics, hazardous cargo information etc.
- new contributions to automatic control of a vessel through more accurate path prediction
- on-shore piloting as a result of such improvements in path prediction.

Implementation of transponder technology to date in waterborne transport is sporadic and, so far, uncoordinated. It is also necessary for VTS authorities to recognise the benefits of investment in shore-based VHF interrogators and use of satellites in order to track ships fitted with transponders. VTS are considered to fall into three categories; coastal, port / harbour and inland waterways. The first and last tend to be in the public sector but the middle category – probably the largest group – is mainly in private hands. The perception of benefits of tracking by transponder technology will vary between these three groups of VTS authorities because the objectives of applications will vary. Coastal VTS authorities will perceive benefits in terms of the lower likelihood of grounding (and possibly collision, depending upon the responsibility of the VTS authority to avoid this happening) of ships which they can track using transponder technology, and the related lower incidence of marine and coastal pollution and lower dependence upon emergency services.
Port and harbour VTS authorities are likely to perceive benefits in terms of the impact of the technology on their financial efficiency. In this case, the transponder technology provides additionally the benefits of an automatic record of use of berth facilities and an electronic link with port services such as ship agents and pilotage. In this context, application of the technology will be particularly useful when it is fitted to ships which account for a significant proportion of either the navigational and/or berthing capacity of a port or which carry hazardous cargo.

This variability in the application of tracking by transponder in terms of the characteristics of the demonstration site and the type of shipping involved means that impacts vary with the main objectives of an application.

This is a technology with benefits for shipping companies, shipping services, VTS services, pilotage services, emergency services, marine and coastal pollution services, port authorities and port services. The main problem for investment in the technology is obtaining distributions of investment and operating costs between all these parties which reflect as fairly as possible the distribution of benefits between them.

### 7.3.2 Objectives of Hypothetical Application

An existing international channel is partly covered by radar-based VTS services on either side of the channel leaving a central section not covered by radar. Dense and varied traffic passes at all times of the day through the channel while hourly ferry services operate across it. A significant proportion of the traffic operating through the channel carries hazardous cargo. Shipping through the channel is directed along seaways in each direction, one within radar coverage of one national VTS centre and the other within the radar coverage of the other national VTS centre. There are hazards to navigation in both seaways in the form of sandbanks.

As part of a demonstration, it is proposed to fit all the ferries with transponders as well as the ships of a company specialising in hazardous cargo which pass through the channel on average once a day in each direction. One of the two national VTS authorities will invest in a shore-based VHF interrogator and an object orientated graphical VTS traffic image display. This workstation will have facilities for identifying vessels fitted with transponders, special recognition of vessels carrying hazardous cargo, for fusing tracking information from radar sources, and for differentiating the sources of tracking information.

The main objectives of the application are to:

- continuously track ferries by radar when within range, and by transponder, with a consistent correlation of the radar track data with the transponder track data
- continuously track ships carrying hazardous cargo by radar when within range, and by transponder within the limits of the shore-based interrogator, with a consistent correlation of the radar track data with the transponder data
- demonstrate the transmission of messages between the VTS centre and vessels fitted with transponders and vice versa
- demonstrate the contribution to safety of shipping
- demonstrate the contribution to efficiency of working methods in VTS centre.
7.3.3 Summary Review

Table 21: Technologies and/or Approach Used and Decision Makers

<table>
<thead>
<tr>
<th>Technologies and/or Approach Used</th>
<th>Decision Makers</th>
</tr>
</thead>
<tbody>
<tr>
<td>VTS</td>
<td>Competent Authority - Coast Authorities</td>
</tr>
<tr>
<td>Radar</td>
<td>VTS Authorities</td>
</tr>
<tr>
<td>Transponders</td>
<td>Regulatory Authorities</td>
</tr>
<tr>
<td>VHF Interrogator</td>
<td>Emergency Authorities</td>
</tr>
<tr>
<td>LAN/WAN</td>
<td>Ships’ Masters</td>
</tr>
<tr>
<td></td>
<td>Vessel Operators</td>
</tr>
</tbody>
</table>

Table 21 shows the technologies and approach used, and the decision makers.

7.3.4 Expected Impacts and Appraisal Groups

7.3.4.1 Expected Impacts

The main expected impacts are:

- increased safety for ferries
- increased safety for ships with transponders carrying hazardous cargo
- increased safety for shipping generally
- possibly greater efficiency of working methods in the VTS centre.

7.3.4.2 Appraisal Groups

Impacts depend upon changes in human behaviour. Appraisal groups are the groups changing their behaviour or the ways they do things; they enjoy the benefits of suffer from the disbenefits of the application.

For increased safety, the appraisal groups are:

- VTS Authorities
- VTS Employees
- Coastal Authorities
- Emergency Services
- Regulatory Authorities
- Ships’ Masters
- Vessel Operators
- Insurance Companies
• Classification Societies
• P&I (Protection and Indemnity) Clubs.

For greater efficiency of working methods in the VTS centre, the appraisal groups are:
• VTS Authorities
• VTS Employees.

Table 22 shows the variation in expected impacts with the different appraisal groups, and an indication of the scale of the expected impact. The likely impact of increased safety for ferries on the coastal authorities is uncertain because the impact of any grounding or sinking as a result of an accident would fall mainly upon the emergency services and there would be insignificant coastal and marine pollution. Ships masters would experience the main impact of increased safety for vessels fitted with transponders because VTS services will be able to give advice if necessary beyond the area of radar coverage, particularly with respect to potential groundings or collisions as ships enter the area of radar coverage and to times of poor visibility and bad weather conditions. Increased safety is likely to have the main impact on coastal authorities with respect to ships carrying hazardous cargoes.

There is uncertainty about the impact of the application on working methods in the VTS centre; the application should permit operators to control shipping more efficiently by virtue of the higher quality of tracking but the workload can be expected to increase because of the greater coverage of tracking and the increased data flow. Automatic identification of all ferries will lead to more efficient communication with them in the event of their needing navigational advice with respect to potential collisions. In principle, if there is a significant improvement in safety, insurance premiums should decline. However, this depends upon the relationship between the penetration of transponder technology and the characteristics of a particular geographical area of navigation in terms of type and intensity of traffic, weather and sea conditions, and other natural hazards.

### 7.3.5 Selection of impacts to be validated

It would be reasonable to validate increased safety for vessels fitted with transponders. The methodological problem is that this will have to be based on potential grounding or collision analysis as the frequency of actual groundings or collisions will be too low to provide a sound basis for examining the impact of the application.

It is debatable whether an attempt should be made to validate the impact on working methods in the VTS centre for two reasons:

- the changes in working practices will be modest bearing in mind that only a small number of vessels will be fitted with transponders
- the VTS centre will probably be operating the current work station in parallel with the demonstration work station and, therefore, a transparent comparison of current with new working practices will not be possible.

However as a minimum, in parallel with the potential grounding or collision analysis used to examine the safety impact, a record should be kept of significant potential effects on working practices, and structured interviews conducted with VTS Authorities and employees.
It will be pointless validating any impact on insurance companies, Classification Societies and P&I Clubs because of the low penetration of transponders in the merchant fleet.

Therefore the following impact alone should be validated:

- the safety impact for vessels fitted with transponders.

This impact analysis will give insights into the safety impact for vessels not fitted with transponders by virtue of any potential collision analyses.

Additionally, in parallel with the potential collision analyses for vessels fitted with transponders, a record should be kept of significant potential effects on working practices in the VTS centre.

**7.3.6 All Assessment Objectives and Appraisal Groups**

Table 23 shows the assessment objectives and groups involved in validation for impact analysis on the basis of the considerations discussed in Section 7.3.5
**Table 22: Expected Impacts for different Appraisal Groups**

<table>
<thead>
<tr>
<th>Category of Assessment</th>
<th>increased safety for ferries</th>
<th>increased safety for ships with transponders carrying hazardous cargo</th>
<th>increased safety for shipping generally</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. VTS Authorities (*)</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>2. VTS Employees</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>3. Coastal Authorities</td>
<td>-/+/0/+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>4. Emergency Services</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>5. Ships Masters</td>
<td>++</td>
<td>++</td>
<td>+</td>
</tr>
<tr>
<td>6. Vessel Operators</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>7. Insurance Companies</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
</tbody>
</table>

(*) No impact is shown for VTS Authorities but it could be expected that the validation results would influence its responsibilities and how these should be carried out.

**Table 23: Definition of Assessment Objectives and User Groups involved in Validation**

<table>
<thead>
<tr>
<th>Category of Assessment</th>
<th>Assessment Objectives</th>
<th>Groups involved in Validation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Impact Analysis</td>
<td>1. To assess the safety impact on ferries</td>
<td>• VTS authorities</td>
</tr>
<tr>
<td></td>
<td>2. To assess the safety impact on vessels fitted with transponders and carrying hazardous cargoes</td>
<td>• VTS employees</td>
</tr>
<tr>
<td></td>
<td>3. To assess the safety impact on shipping generally</td>
<td>• ships’ masters</td>
</tr>
<tr>
<td></td>
<td>4. To assess the effect on working practices in the VTS centre</td>
<td></td>
</tr>
</tbody>
</table>
7.3.7 Selection of Assessment Objectives and Indicators

Table 24 shows the indicators for Assessment Objective 1 in Table 23– to assess the safety impact on ferries.

Table 24: Indicators for Assessment Objective 1 - To assess the safety impact on ferries

<table>
<thead>
<tr>
<th>Assessment Objective</th>
<th>Indicators</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. To assess the safety impact on ferries</td>
<td>1. Number of navigational interventions to ferries by VTS with respect to potential groundings</td>
</tr>
<tr>
<td></td>
<td>2. Time to potential grounding for vessel which is the subject of the intervention at time of intervention</td>
</tr>
<tr>
<td></td>
<td>3. Number of navigational interventions to ferries by VTS with respect to potential collisions</td>
</tr>
<tr>
<td></td>
<td>4. Number of navigational interventions to other vessels by VTS with respect to potential collisions</td>
</tr>
<tr>
<td></td>
<td>5. Time to closest point of approach (TCPA) to potential collision for vessel which is the subject of the intervention at time of intervention</td>
</tr>
</tbody>
</table>

As already indicated in Section 7.3.5, it will be necessary to base the assessment of safety impact on potential groundings or collision analysis.

Whether the impact analysis is based on potential groundings, or potential collisions, or both will depend upon the statutory responsibility of VTS to intervene in these situations. Coastal VTS has responsibility for preventing groundings but responsibility for avoiding collisions might be left by VTS to ships’ masters to avoid VTS being blamed for collisions as a result of distracting a master from the navigation of his ship.

Let us assume that VTS has only responsibility for warning ships’ masters in the event of a potential grounding.

Two indicators are proposed:

- number of navigational interventions by VTS
- time to potential grounding for vessel which is the subject of the intervention at time of intervention.

A distinction has to be made between interventions outside the area of radar coverage and interventions within the area of radar coverage as a result of using the application.

Outside the area of radar coverage, it is reasonable to anticipate that the application will have the most significant effect on time to potential grounding at time of intervention because VTS will be able to track ferries in the central section of the channel which does not have radar coverage.
Within the area of radar coverage, the application has the advantage over conventional radar of clearer identification of vessels fitted with transponders on the VTS traffic image display. This may raise the perception of VTS operators about the need for intervention but it can be anticipated that improvements in performance on the basis of the indicators will be small and therefore more difficult to validate.

A possible weakness in the number of navigational interventions as an indicator outside the area of radar coverage is that the application will encourage VTS to intervene as soon as a potential grounding course is identified. Such understandable action would overlook the possibility that the vessel might have taken avoiding action without the need for intervention prior to what would have been the time of intervention if the application were absent. The number of navigational interventions is differentiated between ferries and other ships to show any change in the proportions of interventions by type of vessel.

**7.3.8 Experimental Design - An Example**

This section presents an example for the Experimental Design comprising:

- Assessment Objectives
- Indicators
- Reference Case
- Method of Measurement or Simulation
- Conditions of Measurement or Simulation
- Statistical Considerations
  - Number of Measurements
  - Confidence Level
  - Overall Definition of Success
- Measurement Plan
- Integrity of Measurement or Simulation for Impact Analysis

Consider now the experimental design for the assessment objective and indicators of Section 7.3.7.

**7.3.8.1 Reference Case:**

The reference case will be the current radar based workstation. A complication is that the demonstration workstation will probably have to be operated in parallel with the current system for two reasons:

- the demonstration system will be in the user acceptance test stage
- it might also be subject to certification procedures by the regulatory authorities.
As indicated in Section 7.3.5, this complicates the task of examining the performance of the application in comparison with the current system. It is likely to have three important effects on the design and practice of validation:

- the ‘double-manning’ implicit in running two parallel systems will probably put a resource limitation on the duration of parallel operation
- having different personnel operating the two systems in parallel will be a possible source of bias in the results of the validation
- the two systems will not be able to operate truly in parallel because the existing system will probably be paramount as an accepted basis for decision-making – the main role of the demonstration system will be to give earlier warnings to the staff operating the existing system of the need for intervention.

### 7.3.8.2 Method of Measurement

The main methods of measurement will be:

- automatic timing of intervention, record of type/identity/position/speed/direction of vessel which was subject of intervention, and record of type/identity/position/speed/direction of other vessels involved
- time to potential grounding for vessel which was subject of intervention at time of intervention.

The main consideration for measurement is:

- completing paired measurements (i.e. measurements using the demonstrator and measurements using the existing system) accurately.

### 7.3.8.3 Conditions of Measurement

The main considerations are:

- potential groundings will probably be relatively infrequent
- measurement should take place at times which are representative of traffic, seasonal, meteorological, tidal, and visibility conditions, etc.
- the need to distinguish between interventions outside the area of radar coverage and interventions within the area of radar coverage as discussed in Section 7.3.7.

It would be useful to examine historical records of potential groundings or, if these are not available, to obtain the views of VTS operators to establish if there are any patterns to the occurrence of potential groundings in order to establish the most fruitful times for measurement.

If the pattern of potential groundings is random in time, it will be necessary to reconsider the method of measurement because it will be inefficient of resources to have additional staff committed to the demonstration system on any continuous basis.
As the demonstration system is more likely to be a ‘faster trigger’ of navigational intervention by VTS operators than the current system, a possible solution to this problem of infrequent, events would be to use the demonstration system operationally during the demonstration stage of validation. As the demonstration system will identify tracking by radar and transponder sources, it will be possible to obtain actual times of intervention and estimate the time of intervention had the current system have been used. However, this solution would require acceptance by the regulatory authorities of the demonstration system.

7.3.8.4 Statistical Considerations

7.3.8.4.1 Number of Measurements

Intervention outside the area of radar coverage

As indicated in Section 7.3.7, the time to potential grounding at time of intervention is the indicator most likely to show an improvement of performance with the demonstration system. As this is inherent in the demonstrator, it should not take too many events to show a significant improvement. The main problem will lie in testing the application over different kinds of conditions, over which there might be a considerable variation in the frequency of potential groundings, and this frequency will probably be low on average.

Intervention inside the area of radar coverage

As also indicated in Section 7.3.7, differences in performance between the demonstrator and the existing system are likely to be much smaller. It will be, therefore quite difficult to validate the indicators. It may not be possible to validate the safety impact in terms of particular indicators and it may be necessary to limit the validation to obtaining the views of the VTS operators on use in general of the demonstrator. Depending upon the duration of validation, they may be able to identify safety-related advantages of the demonstrator.

7.3.8.4.2 Confidence Level

Intervention outside the area of radar coverage

Testing of the time to potential grounding at time of intervention should be done at the 95% level. However, the low frequency of such events will mean that it will not be practicable to take measurements at the verification stage to provide a basis for statistical design of the number of measurements required to achieve this level of statistical accuracy. The confidence level will be dictated by the number of events recorded during the demonstration.

7.3.8.4.3 Overall Definition of Success

Intervention outside the area of radar coverage

Statistically, the overall definition of success will be a significant increase in the time to a potential grounding for a vessel which was the subject of intervention at time of intervention, in all conditions of operation.
However, as this result is inherent in the demonstrator, it is at least equally important that the demonstrator shows extremely high levels of functional reliability, consistency between tracking from transponder and radar sources, reliable and useful transmission of short messages between the VTS centre and vessels fitted with transponders, and the workstation display is judged as user-friendly.

This statement suggests that:

- the key issue for quantitative validation of the safety impact is what is the minimum measurement required to be satisfied about the safety advantage of the demonstrator
- continuous logging of events which highlight the safety impact of the application and interviews with VTS operators about such events will also play an important role in validation of safety impact.

### 7.3.8.5 Measurement Plan

The measurement plan will be dictated by the frequency of potential groundings, as already discussed under ‘Conditions of Measurement.’ The other main issue, already discussed, is whether the demonstrator can be tested in parallel with the existing system, or whether the VTS centre can replace the existing system by the demonstrator during the demonstration stage of validation.

### 7.3.8.6 Integrity of Measurement

The following points have already been discussed:

- VTS operators using the application outside the area of radar coverage may intervene sooner than is necessary from the viewpoint that vessels might have taken avoiding action before it would have been possible to intervene using the existing system
- parallel operation of existing and demonstration applications complicates comparison of performance and makes qualification necessary about any effects on working practices.
8. **Keyword List**

9. References


## 10. Glossary

### 10.1 Standard Terms

<table>
<thead>
<tr>
<th>Term</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>Application</td>
<td>A telematics system or service as installed and operating in a real-life environment.</td>
</tr>
<tr>
<td>Appraisal groups</td>
<td>Different groups of users/non-users affected by the impacts of an application. Benefits and disbenefits are estimated for these impacts.</td>
</tr>
<tr>
<td>Assessment</td>
<td>The process of determining the performance and/or impacts of a candidate application, usually in comparison to a reference case (existing situation or alternative applications), and usually including an experimental process based on real-life or other trials, often involving users.</td>
</tr>
<tr>
<td>Assessment Objective</td>
<td>A precise statement of an individual objective of validation - it should be associated with a precise definition of the associated indicator(s) and definition of success.</td>
</tr>
<tr>
<td>Definition of Success</td>
<td>This defines the expectation about the performance and impacts of the application. The success or failure of validation results are tested against this criteria. So it has a vital role in the validation methodology. It is most exact when it is defined for a single indicator.</td>
</tr>
<tr>
<td>Demonstration</td>
<td>The demonstration stage of validation will use a sufficiently large sample of users in a real-life situation to provide information on cost-effectiveness, user friendliness and similar issues, as well as testing the feasibility of the system when used on a large scale.</td>
</tr>
<tr>
<td>Evaluation</td>
<td>The process of determining the value of an application in comparison to alternative applications and/or to a “base case”, and deriving recommendations for decision makers based on identifying requirements on and analysing results of related experiments.</td>
</tr>
<tr>
<td>Impact</td>
<td>An effect of an application on the people who are affected by the application.</td>
</tr>
<tr>
<td>Indicator</td>
<td>An indicator is the main criteria of quantitative or qualitative assessment. It is a parameter indicating the performance or impacts of an application. For quantitative assessment, it is either measured directly or derived from measurement or simulation. Use of indicators for qualitative assessment is more problematic because measurement is replaced by personal judgement.</td>
</tr>
<tr>
<td>Main Decision Makers</td>
<td>People or groups who will be influential in defining the assessment objectives for validation and in determining whether verification results justify proceeding to the demonstration stage of validation.</td>
</tr>
</tbody>
</table>
Reference Case  The performance and impacts of an application are usually compared against some existing situation in order to show that the application meets required standards (for tests of the physical functioning of the application and user acceptance) and that use of the application is an improvement on alternative ways of achieving the objectives of the application (for user acceptance again and impact analysis). ‘Before and After’ studies might feature in assessment of user acceptance and impact analysis.

User Groups  Groups involved in validating the application - they will probably represent the main market for the application - they might include:
• operators of the application
• intermediate users of the application like providers of information

Validation  Validation is the process of testing how an application performs with respect to the assessment objectives.

Verification  The verification stage of validation will use a small but sufficient sample of users in a real-life situation to test the technical feasibility of the demonstrator and to yield preliminary findings on user acceptance.

10.2 Acronyms and Abbreviations

<table>
<thead>
<tr>
<th>Acronym / Abbreviation</th>
<th>Definition</th>
<th>Remarks</th>
</tr>
</thead>
<tbody>
<tr>
<td>ATT</td>
<td>Advanced Road Transport Telematics</td>
<td></td>
</tr>
<tr>
<td>AVI</td>
<td>Automatic Vehicle Identification</td>
<td></td>
</tr>
<tr>
<td>CONVERGE</td>
<td>Fourth Framework Transport Telematics Support Project</td>
<td>TR1101</td>
</tr>
<tr>
<td>CORD</td>
<td>Co-ordination project for R&amp;D</td>
<td>DRIVE II V2056</td>
</tr>
<tr>
<td>DRIVE I</td>
<td>Dedicated Road Infrastructure for Vehicle Safety in Europe</td>
<td>2nd Framework Programme</td>
</tr>
<tr>
<td>DRIVE II</td>
<td>Dedicated Road Infrastructure for Vehicle Safety in Europe</td>
<td>3rd Framework Programme</td>
</tr>
<tr>
<td>RTD&amp;D</td>
<td>Research, Technological Development and Demonstration</td>
<td></td>
</tr>
<tr>
<td>TTAP</td>
<td>Transport Telematics Applications Programme</td>
<td></td>
</tr>
<tr>
<td>TTCO</td>
<td>Transport Telematics Central Office</td>
<td></td>
</tr>
<tr>
<td>VQ</td>
<td>Validation Quality</td>
<td>Work area of the CONVERGE Project</td>
</tr>
<tr>
<td>VTS</td>
<td>Vessel Traffic Systems</td>
<td>TR1056 TRACAR</td>
</tr>
</tbody>
</table>
Appendix A: Template for Validation Plan
Appendix B: Feedback Report Template
Telematics Applications Programme - Transport Sector
CONVERGE Project TR 1101
Validation Quality Review

Feedback Report on the Draft Validation Plan for

Project name
TR0000

Reviewer: e.g. Ken Perrett
Transport Research Laboratory
Tel: + 44 1344 770 027
Fax: + 44 1344 770 913
E-mail: KenPerrett@compuserve.com

Version 1.0 (FINAL Version)

<table>
<thead>
<tr>
<th>Approved by:</th>
<th>Dr David Maltby</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Validation Quality Co-ordinator</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Signature:</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Date:</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
</tr>
</tbody>
</table>
1. Introduction

1.1 Role of this document

This document contains the results of the review of your Draft Validation Plan.

Section 1 provides an brief overview of the project.

Section 2 contains general comments on the validation methodology used and summarises the results of the review for the project as a whole.

Section 3 contains recommended actions regarding your future validation/evaluation work.

Section 4 contains details of the review process. This is in the form of a review checklist, results for each site/application combination, specific remarks and recommendations for internal use within your project.

Please note that Sections 1-3 for all projects will be included as an annex in our report to the Commission on project Draft Validation Plans (CONVERGE Validation Quality Deliverable DVQ3.3/1- ‘Summary Report on Project Draft Validation Plans.’).

1.2 Overview of the project

General overview of the project for the purposes of the summary management report.

2. General comments

2.1
2.2
2.3
2.4
Table 1: Summary of review for the project as a whole

<table>
<thead>
<tr>
<th>Element</th>
<th>Yes/No</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Evaluation Methodology Framework</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Are you confident at this stage that the validation methods are likely</td>
<td></td>
<td></td>
</tr>
<tr>
<td>to result in well-founded results?</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Is the role of the draft validation plan in the assessment process as</td>
<td></td>
<td></td>
</tr>
<tr>
<td>a whole described clearly?</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Is there a clear management structure for evaluation / validation?</td>
<td></td>
<td></td>
</tr>
<tr>
<td>**Listing of All Applications being Validated and their Phases of</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Development**</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Is the time scale reasonable for validation activities?</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Identification of Potential for European Added Value</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Does the draft validation plan address explicitly the issue of European</td>
<td></td>
<td></td>
</tr>
<tr>
<td>added Value?</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Is any attempt made to group applications on the basis of common</td>
<td></td>
<td></td>
</tr>
<tr>
<td>functionality?</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Are common validation plans proposed?</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Is validation integrated across sites?</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
3. **Recommended actions**

Summary of recommended action to be presented to the Commission. Where necessary provide further details in Section 4.3.1 for use by the project evaluation manager.

3.1
3.2
3.3
3.4

4. **Comments specific to the project**

4.1 **Detailed review**

The Draft Validation Plan was reviewed against a checklist devised by the CONVERGE Validation Quality team and agreed with the European Commission. The review questions and the completed answers for each of the applications are summarised in Table 2b.

The specific remarks which follow provide more information on the highlighted aspects within Table 2b.

4.2 **Specific remarks**

Refer to copies of Tables 2b in annex

4.3 **Summary of advice to the evaluation manager**

Summary of advice to be taken into account by the evaluation manager in the continuing process of evaluation.

4.3.1 **Recommended actions**

Explanation of the necessary actions, which are summarised in Section 3, with further details for use by the evaluation manager where necessary.

4.3.2 **Additional comments and advice**

5. **Annexes**
Include any additional material, particularly completed versions of Table 2b as appropriate. A copy of Table 2b follows. Further copies are available as “Table2b.doc” for insertion.
### Table 2b: Detailed review

**Application:** _____________________________________________  
**Site:** _____________________________________________

<table>
<thead>
<tr>
<th>Review questions</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>General issues for Draft Validation Plan</strong></td>
<td></td>
</tr>
<tr>
<td>Summary Review of Application</td>
<td></td>
</tr>
<tr>
<td>Are all applications described adequately</td>
<td></td>
</tr>
<tr>
<td>Are the objectives of the application stated clearly</td>
<td></td>
</tr>
<tr>
<td>Is the verification site described adequately (figure, plan or map, etc.)?</td>
<td></td>
</tr>
<tr>
<td>Is the demonstration site described adequately (figure, plan or map, etc.)?</td>
<td></td>
</tr>
<tr>
<td><strong>Definition of expected impacts</strong></td>
<td></td>
</tr>
<tr>
<td>Are any significant impacts omitted?</td>
<td></td>
</tr>
<tr>
<td>Are impacts defined precisely enough?</td>
<td></td>
</tr>
<tr>
<td>Do you agree with the suggested scale of impacts?</td>
<td></td>
</tr>
<tr>
<td><strong>Selection of impacts to be validated</strong></td>
<td></td>
</tr>
<tr>
<td>Is this considered?</td>
<td></td>
</tr>
<tr>
<td>Do you agree with the selection?</td>
<td></td>
</tr>
<tr>
<td><strong>Adequacy of the demonstration site for impact analysis</strong></td>
<td></td>
</tr>
<tr>
<td>Is this considered?</td>
<td></td>
</tr>
<tr>
<td>Do you agree that the demonstration site is adequate?</td>
<td></td>
</tr>
<tr>
<td><strong>Assessment objectives at verification</strong></td>
<td></td>
</tr>
<tr>
<td>Are the assessment objectives stated clearly?</td>
<td></td>
</tr>
<tr>
<td>Are the user groups defined adequately?</td>
<td></td>
</tr>
<tr>
<td><strong>Outline of Verification Methods</strong></td>
<td></td>
</tr>
<tr>
<td><strong>Indicators</strong></td>
<td></td>
</tr>
<tr>
<td>Are they defined reasonably precisely, bearing in mind that functional specification may not have been completed?</td>
<td></td>
</tr>
<tr>
<td>Are they compatible with their associated assessment objectives?</td>
<td></td>
</tr>
<tr>
<td><strong>Reference Cases</strong></td>
<td></td>
</tr>
<tr>
<td>Are these outlined?</td>
<td></td>
</tr>
<tr>
<td>Are they valid?</td>
<td></td>
</tr>
<tr>
<td><strong>Methods of Measurement or/and Simulation</strong></td>
<td></td>
</tr>
<tr>
<td>---------------------------------------------</td>
<td>---</td>
</tr>
<tr>
<td>Are these outlined?</td>
<td></td>
</tr>
<tr>
<td>Are they defined adequately?</td>
<td></td>
</tr>
<tr>
<td><strong>Conditions of Measurement or/and Simulation</strong></td>
<td></td>
</tr>
<tr>
<td>Are these outlined?</td>
<td></td>
</tr>
<tr>
<td><strong>Overall Definition of Success</strong></td>
<td></td>
</tr>
<tr>
<td>Is this defined with associated indicators?</td>
<td></td>
</tr>
<tr>
<td><strong>Sampling</strong></td>
<td></td>
</tr>
<tr>
<td>Do they show concern that their measurements should be representative?</td>
<td></td>
</tr>
<tr>
<td>Do they provide some measures of sample size?</td>
<td></td>
</tr>
<tr>
<td><strong>Statistical Considerations</strong></td>
<td></td>
</tr>
<tr>
<td>Do they provide any estimates of confidence levels that might be associated with their measurements?</td>
<td></td>
</tr>
<tr>
<td><strong>Integrity of Measurement or/and Simulation</strong></td>
<td></td>
</tr>
<tr>
<td>Do they show any awareness of this issue?</td>
<td></td>
</tr>
<tr>
<td><strong>Overall issues for Verification Methods</strong></td>
<td></td>
</tr>
<tr>
<td><em>Is the Outline of Verification Methods satisfactory overall?</em></td>
<td></td>
</tr>
<tr>
<td><strong>Assessment Objectives at Demonstration</strong></td>
<td></td>
</tr>
<tr>
<td>Are the assessment objectives stated clearly?</td>
<td></td>
</tr>
<tr>
<td>Are the user groups defined adequately?</td>
<td></td>
</tr>
<tr>
<td><strong>Outline of Demonstration Methods</strong></td>
<td></td>
</tr>
<tr>
<td><strong>Indicators</strong></td>
<td></td>
</tr>
<tr>
<td>Are they defined reasonably precisely, bearing in mind that functional specification may not have been completed?</td>
<td></td>
</tr>
<tr>
<td>Are they compatible with their associated assessment objectives?</td>
<td></td>
</tr>
<tr>
<td><strong>Reference Cases</strong></td>
<td></td>
</tr>
<tr>
<td>Are these outlined?</td>
<td></td>
</tr>
<tr>
<td>Are they valid?</td>
<td></td>
</tr>
<tr>
<td><strong>Methods of Measurement or/and Simulation</strong></td>
<td></td>
</tr>
<tr>
<td>---------------------------------------------</td>
<td>---</td>
</tr>
<tr>
<td>Are these outlined?</td>
<td></td>
</tr>
<tr>
<td>Are they defined adequately?</td>
<td></td>
</tr>
<tr>
<td><strong>Conditions of Measurement or/and Simulation</strong></td>
<td></td>
</tr>
<tr>
<td>Are these outlined?</td>
<td></td>
</tr>
<tr>
<td><strong>Overall Definition of Success</strong></td>
<td></td>
</tr>
<tr>
<td>Is this defined with associated indicators?</td>
<td></td>
</tr>
<tr>
<td><strong>Sampling</strong></td>
<td></td>
</tr>
<tr>
<td>Do they show concern that their measurements should be representative?</td>
<td></td>
</tr>
<tr>
<td>Do they provide some measures of sample size?</td>
<td></td>
</tr>
<tr>
<td><strong>Statistical Considerations</strong></td>
<td></td>
</tr>
<tr>
<td>Do they provide any estimates of confidence levels that might be associated with their measurements?</td>
<td></td>
</tr>
<tr>
<td><strong>Integrity of Measurement or/and Simulation</strong></td>
<td></td>
</tr>
<tr>
<td>Do they show any awareness of this issue?</td>
<td></td>
</tr>
<tr>
<td><strong>Overall issues for Demonstration Methods</strong></td>
<td></td>
</tr>
<tr>
<td>Will the timing of the demonstration stage capture the major impacts?</td>
<td></td>
</tr>
<tr>
<td>Is the scale of the demonstration sufficient to estimate impacts?</td>
<td></td>
</tr>
<tr>
<td>Are there plans to expand results?</td>
<td></td>
</tr>
<tr>
<td>Are they proposing to estimate benefits by socio-economic analysis or financial assessment, etc?</td>
<td></td>
</tr>
<tr>
<td>At this stage are you confident that they will be successful with socio-economic analysis or financial assessment, etc?</td>
<td></td>
</tr>
<tr>
<td><strong>Is the Outline of Demonstration Methods satisfactory overall?</strong></td>
<td></td>
</tr>
</tbody>
</table>
Abstract:
This document provides the Draft/ Final Validation Plan for the validation activities within the ____________project. Insert details of sites and outline of applications.

Keyword list:
_____________________________________________, Validation
TABLE OF CONTENTS

1. INTRODUCTION .......................................................................................................... 4
   1.1 OBJECTIVES FOR THE PROJECT ........................................................................... 4
   1.2 OVERVIEW OF APPLICATIONS AND SITES .......................................................... 4
   1.3 PHASES OF DEVELOPMENT .................................................................................. 4
   1.4 USER NEEDS ........................................................................................................... 4

2. COMMON APPROACH .................................................................................................... 5
   2.1 THE EVALUATION FRAMEWORK .......................................................................... 5
   2.2 MANAGEMENT OF EVALUATION .......................................................................... 5
   2.3 OBJECTIVES ........................................................................................................... 5
   2.4 SYSTEM ARCHITECTURE ...................................................................................... 5
   2.5 ANALYSIS OF IMPACTS ....................................................................................... 5
   2.6 COMMON INDICATORS ......................................................................................... 5
   2.7 POTENTIAL FOR EUROPEAN-ADDED VALUE THROUGH COMPARATIVE ASSESSMENT .. 5

3. SITE 1 .............................................................................................................................. 6
   3.1 DESCRIPTION OF APPLICATION .......................................................................... 6
   3.2 DESCRIPTION OF VERIFICATION SITE ................................................................. 6
   3.3 DESCRIPTION OF DEMONSTRATION SITE ............................................................ 6
   3.4 SYSTEM DESIGN OBJECTIVES ............................................................................. 6
   3.5 ASPECTS OF SYSTEM ARCHITECTURE TO BE VALIDATED ................................. 6
   3.6 DESCRIPTION OF THE TECHNOLOGIES .............................................................. 6
   3.7 DECISION MAKERS .............................................................................................. 6
   3.8 EXPECTED IMPACTS AND USER GROUPS AFFECTED ........................................ 6
   3.9 SELECTION OF IMPACTS TO BE VALIDATED ..................................................... 6
   3.10 OUTLINE OF VALIDATION METHODS AT VERIFICATION STAGE ...................... 7
   3.11 OUTLINE OF VALIDATION METHODS AT DEMONSTRATION STAGE ................. 7

4. SITE 2 .............................................................................................................................. 9
   4.1 DESCRIPTION OF APPLICATION .......................................................................... 9
   4.2 DESCRIPTION OF VERIFICATION SITE ................................................................. 9
   4.3 DESCRIPTION OF DEMONSTRATION SITE ............................................................ 9
   4.4 SYSTEM DESIGN OBJECTIVES ............................................................................. 9
   4.5 ASPECTS OF SYSTEM ARCHITECTURE TO BE VALIDATED ................................. 9
   4.6 DESCRIPTION OF THE TECHNOLOGIES .............................................................. 9
   4.7 DECISION MAKERS .............................................................................................. 9
   4.8 EXPECTED IMPACTS AND USER GROUPS AFFECTED ........................................ 9
   4.9 SELECTION OF IMPACTS TO BE VALIDATED ..................................................... 9
   4.10 OUTLINE OF VALIDATION METHODS AT VERIFICATION STAGE ...................... 9
   4.11 OUTLINE OF VALIDATION METHODS AT DEMONSTRATION STAGE ................. 10

5. REFERENCES/BIBLIOGRAPHY ..................................................................................... 12

6. GLOSSARY ..................................................................................................................... 12

APPENDICES .................................................................................................................. 13

A. CHECKLIST TABLES .................................................................................................... 13

B. <OTHER APPENDICES ETC.> .................................................................................... 21
SUMMARY
Text (max. 2 pages).

Notes about this template

The file includes some simple instructions in italic text. There are also more detailed instructions and annotations which can be read on screen by double clicking on the annotation mark.
1. INTRODUCTION

This section refers to checklist tables 1 and 2

1.1 OBJECTIVES FOR THE PROJECT

1.2 OVERVIEW OF APPLICATIONS AND SITES

1.3 PHASES OF DEVELOPMENT

1.4 USER NEEDS

These will be provisional in the draft validation plan, based on previous experience, other sites, etc. The results of the user needs analysis for this project will be included in the final validation plan (also, in the draft validation if user needs analysis is well advanced or completed at this stage).
2. COMMON APPROACH

This section should be used to show the features in the approach to validation in the project which are common to the sites or applications to be validated. The individual sections of the validation plan may also include other aspects which are not common across the project as a whole.

2.1 THE EVALUATION FRAMEWORK

2.2 MANAGEMENT OF EVALUATION

2.3 OBJECTIVES OF THE DEMONSTRATION

2.4 SYSTEM ARCHITECTURE

Include simple one page diagram showing the main functions to be validated.

2.5 ANALYSIS OF IMPACTS

(if impact assessment is part of the validation objectives)

2.6 COMMON INDICATORS

2.7 POTENTIAL FOR EUROPEAN-ADDED VALUE THROUGH COMPARATIVE ASSESSMENT

This section should show how the project will add value to its site-specific validation results by comparing validation results across sites. It may be based on a summary of more detailed information presented in the Site-Specific Sections. The topics covered should include the following:

2.7.1 Applications to be compared

2.7.2 Sites to be compared

2.7.3 Expected Common Impacts

2.7.4 Validation methods for comparisons between sites

The emphasis should be on common validation methods; e.g. indicators, reference cases, methods of assessment, conditions of measurement, analysis of data, etc.
3. **SITE 1**

This outline assumes that each site has a separate validation plan, and therefore separate sections of the plan for each site. Some projects may have a common validation plan across all sites, in which case it will be appropriate to have a set of site descriptions followed by a common plan for all sites.

3.1 **DESCRIPTION OF APPLICATION**

*Include objectives of the application*

3.2 **DESCRIPTION OF VERIFICATION SITE**

*Include maps and diagrams*

3.3 **DESCRIPTION OF DEMONSTRATION SITE**

*Include maps and diagrams*

*Include confirmation that site is adequate for impact analysis*

3.4 **SYSTEM DESIGN OBJECTIVES**

3.5 **ASPECTS OF SYSTEM ARCHITECTURE TO BE VALIDATED**

3.6 **DESCRIPTION OF THE TECHNOLOGIES**

3.7 **DECISION MAKERS**

3.8 **EXPECTED IMPACTS AND USER GROUPS AFFECTED**

*This section refers to checklist table 3.*

3.9 **SELECTION OF IMPACTS TO BE VALIDATED**

*This section refers to checklist tables 4 and 5.*
3.10  **OUTLINE OF VALIDATION METHODS AT VERIFICATION STAGE**

3.10.1  Assessment objectives

*This section refers to checklist table 6.*

3.10.2  Outline of validation methods

*This section refers to checklist table 7.*

3.10.2.1  Indicators

3.10.2.2  Reference case

3.10.2.3  Definition of success

3.10.2.4  Definition of methods of quantitative and qualitative assessment

3.10.2.5  Measurement conditions

3.10.2.6  Statistical considerations

*Include representativeness, sampling, confidence levels*

3.10.2.7  Integrity of measurement

3.10.3  Resources for verification

3.11  **OUTLINE OF VALIDATION METHODS AT DEMONSTRATION STAGE**

3.11.1  Assessment objectives

*This section refers to checklist table 6.*

3.11.2  Outline of validation methods

*This section refers to checklist table 7.*

3.11.2.1  Indicators

3.11.2.2  Reference case
3.11.2.3 Definition of success

3.11.2.4 Definition of methods of quantitative of qualitative assessment

3.11.2.5 Measurement conditions

3.11.2.6 Statistical considerations
   Include representativeness, sampling, confidence levels

3.11.2.7 Integrity of measurement

3.11.3 Resources for demonstration

3.11.4 Timescale of demonstrator and methods of validation

3.11.5 Expansion of results of demonstration

3.11.6 Identification of potential for European Added Value
4. **SITE 2**

4.1 **DESCRIPTION OF APPLICATION**

Include objectives of the application

4.2 **DESCRIPTION OF VERIFICATION SITE**

Include maps and diagrams

4.3 **DESCRIPTION OF DEMONSTRATION SITE**

Include maps and diagrams
Include confirmation that site is adequate for impact analysis

4.4 **SYSTEM DESIGN OBJECTIVES**

4.5 **ASPECTS OF SYSTEM ARCHITECTURE TO BE VALIDATED**

4.6 **DESCRIPTION OF THE TECHNOLOGIES**

4.7 **DECISION MAKERS**

4.8 **EXPECTED IMPACTS AND USER GROUPS AFFECTED**

*This section refers to checklist table 3.*

4.9 **SELECTION OF IMPACTS TO BE VALIDATED**

*This section refers to checklist tables 4 and 5.*

4.10 **OUTLINE OF VALIDATION METHODS AT VERIFICATION STAGE**

4.10.1 **Assessment objectives**

*This section refers to checklist table 6.*
4.10.2 Outline of validation methods

This section refers to checklist table 7.

4.10.2.1 Indicators

4.10.2.2 Reference case

4.10.2.3 Definition of success

4.10.2.4 Definition of methods of quantitative and qualitative assessment

4.10.2.5 Measurement conditions

4.10.2.6 Statistical considerations

Include representativeness, sampling, confidence levels

4.10.2.7 Integrity of measurement

4.10.3 Resources for verification

4.11 OUTLINE OF VALIDATION METHODS AT DEMONSTRATION STAGE

4.11.1 Assessment objectives

This section refers to checklist table 6.

4.11.2 Outline of validation methods

This section refers to checklist table 7.

4.11.2.1 Indicators

4.11.2.2 Reference case

4.11.2.3 Definition of success

4.11.2.4 Definition of methods of quantitative of qualitative assessment
4.11.2.5 Measurement conditions

4.11.2.6 Statistical considerations
*Include representativeness, sampling, confidence levels*

4.11.2.7 Integrity of measurement

4.11.3 Resources for demonstration

4.11.4 Timescale of demonstrator and methods of validation

4.11.5 Expansion of results of demonstration

4.11.6 Identification of potential for European Added Value
5. REFERENCES/BIBLIOGRAPHY

6. GLOSSARY

Abbreviations Explanation
APPENDICES

A. CHECKLIST TABLES

Introduction

This Appendix contains the framework for the checklist tables. These are designed to be copied and edited to produce the individual validation plans for each project in a common format. However the structure of the projects varies and evaluation managers for individual projects will need to consider how best the checklist information can be laid out for the particular combination of applications and sites.

If there are a considerable number of site/application combinations, it may be more appropriate to produce the checklist tables as a separate volume to assist the reader who will need to relate the tables to the text in the main body of the validation plan.

It is suggested that projects retain the essential numbering of the tables as this will aid comparison across projects. Where there is a requirement for the tables to be replicated, projects will need to extend the numbering system. For example, the information in table 7 is required for each indicator, for each assessment objective, for each category of assessment and for each application. It is suggested that the following form is used:-

```
    table 7-n-m etc
    where n is the application/site combination
    and m is the category of assessment/assessment objective
```

The precise number of levels and appropriate grouping (ie indicators, objectives, categories and applications) is left to projects to define.

WORD annotations have been used to provide advice about the content of the tables and how they could be adapted to suit the needs of projects. The annotations and an indication of their content are indicated by text 'Further information ....'. To show the position of the annotations on the screen, you will need have selected Non-printing characters - All 'in the View page of the Tools - Options menu. Once this is selected, you should see [CONVERGE] on the document window. To view the text of the annotation, double click on the Annotation mark in WORD6 or in WORD7, put the cursor on the annotation mark. Having read the text of the annotation, you may wish to Close the annotation pane.
### Table 1: Summary Review of Applications and Validation Sites

<table>
<thead>
<tr>
<th>Application</th>
<th>Application 1</th>
</tr>
</thead>
<tbody>
<tr>
<td>Technologies and approach used</td>
<td></td>
</tr>
<tr>
<td>Functionalities (Telematics Functions if available)</td>
<td></td>
</tr>
<tr>
<td>Main Decision Makers</td>
<td></td>
</tr>
<tr>
<td>User Groups Within Project</td>
<td></td>
</tr>
<tr>
<td>User Groups Outside Project</td>
<td></td>
</tr>
<tr>
<td>Verification Site</td>
<td></td>
</tr>
<tr>
<td>Demonstration Site</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Application</th>
<th>Application 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Technologies and approach used</td>
<td></td>
</tr>
<tr>
<td>Functionalities (Telematics Functions if available)</td>
<td></td>
</tr>
<tr>
<td>Main Decision Makers</td>
<td></td>
</tr>
<tr>
<td>User Groups Within Project</td>
<td></td>
</tr>
<tr>
<td>User Groups Outside Project</td>
<td></td>
</tr>
<tr>
<td>Verification Site</td>
<td></td>
</tr>
<tr>
<td>Demonstration Site</td>
<td></td>
</tr>
</tbody>
</table>

**Further information**
- **Treatment of sites**
Table 2: Beginning and End of Phases of Development

<table>
<thead>
<tr>
<th>Application 1</th>
<th>Dates</th>
<th>Explanatory notes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Analysis of user requirements</td>
<td>mmyy - mmyy</td>
<td></td>
</tr>
<tr>
<td>Definition of functional requirements</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Building of a demonstrator</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Verification</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Demonstration</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Exploitation Plan</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Application 2</th>
<th>Dates</th>
<th>Explanatory notes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Analysis of user requirements</td>
<td>mmyy - mmyy</td>
<td></td>
</tr>
<tr>
<td>Definition of functional requirements</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Building of a demonstrator</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Verification</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Demonstration</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Exploitation Plan</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Further information
- Explanatory notes
- Treatment of sites
Key to impacts in Table 3

<table>
<thead>
<tr>
<th>Impact</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Further information
• Explanation of description of impacts

Table 3: Definition of Expected Impacts for Different Appraisal Groups

<table>
<thead>
<tr>
<th>Application 1</th>
<th>Impact 1</th>
<th>Impact 2</th>
<th>Impact 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Impact Appraisal Group 1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Impact Appraisal Group 2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Impact Appraisal Group 3</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Application 2

| Impact Appraisal Group 1 |          |          |          |
| Impact Appraisal Group 2 |          |          |          |
| Impact Appraisal Group 3 |          |          |          |

Key:
•++ very positive impact
•+ positive impact
•0 neutral or uncertain impact
•- negative impact
•-- very negative impact

Further information
• Treatment of sites
Table 4: Selection of Impacts for Validation for Particular Appraisal Groups

Application 1

<table>
<thead>
<tr>
<th>Impact</th>
<th>Impact category (validation to be quantitative/ qualitative/ not to be validated)</th>
<th>Explanation for category</th>
</tr>
</thead>
<tbody>
<tr>
<td>Appraisal Group 1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Impact 1.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Impact 2.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Impact .</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Appraisal Group 2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Impact 1.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Impact 2.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Impact .</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Appraisal Group .</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Impact 1.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Impact 2.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Impact .</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Application 2

<table>
<thead>
<tr>
<th>Impact</th>
<th>Impact category (validation to be quantitative/ qualitative/ not to be validated)</th>
<th>Explanation for category</th>
</tr>
</thead>
<tbody>
<tr>
<td>Appraisal Group 1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Impact 1.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Impact 2.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Impact .</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Appraisal Group 2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Impact 1.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Impact 2.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Impact .</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Appraisal Group .</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Impact 1.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Impact 2.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Impact .</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Further information
- Treatment of sites
- Impact category
Table 5: Checking that impacts required for multi-criteria analysis or social cost-benefit analysis or financial appraisal can be validated

<table>
<thead>
<tr>
<th>Application</th>
<th>Impacts required for multi-criteria analysis etc</th>
<th>Can impact be validated?</th>
<th>Explanation if No</th>
</tr>
</thead>
<tbody>
<tr>
<td>Application 1</td>
<td>Multi-criteria analysis etc 1</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Impact 1.</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Impact 2.</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Impact .</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Multi-criteria analysis etc 2</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Impact 1.</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Impact 2.</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Impact .</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Multi-criteria analysis etc</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Impact 1.</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Impact 2.</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Impact .</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Application 2 etc.</td>
<td>Multi-criteria analysis etc 1 etc.</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Impact 1 etc.</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Table 6: Definition of Assessment Objectives with respect to Categories of Assessment and Users Groups or Appraisal Groups involved in Validation

<table>
<thead>
<tr>
<th>Application 1</th>
<th>Assessment Objectives</th>
<th>User Groups or Appraisal Groups</th>
</tr>
</thead>
<tbody>
<tr>
<td>Testing Physical Functioning</td>
<td>1.</td>
<td></td>
</tr>
<tr>
<td></td>
<td>2.</td>
<td></td>
</tr>
<tr>
<td>Testing User Acceptance</td>
<td>1.</td>
<td></td>
</tr>
<tr>
<td></td>
<td>2.</td>
<td></td>
</tr>
<tr>
<td>Impact Analysis</td>
<td>1.</td>
<td></td>
</tr>
<tr>
<td></td>
<td>2.</td>
<td></td>
</tr>
<tr>
<td>Multi-criteria analysis or social cost-benefit analysis or financial appraisal</td>
<td>1.</td>
<td></td>
</tr>
<tr>
<td></td>
<td>2.</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Application 2</th>
<th>Assessment Objectives</th>
<th>User Groups or Appraisal Groups</th>
</tr>
</thead>
<tbody>
<tr>
<td>Testing Physical Functioning</td>
<td>1.</td>
<td></td>
</tr>
<tr>
<td></td>
<td>2.</td>
<td></td>
</tr>
<tr>
<td>Testing User Acceptance</td>
<td>1.</td>
<td></td>
</tr>
<tr>
<td></td>
<td>2.</td>
<td></td>
</tr>
<tr>
<td>Impact Analysis</td>
<td>1.</td>
<td></td>
</tr>
<tr>
<td></td>
<td>2.</td>
<td></td>
</tr>
<tr>
<td>Multi-criteria analysis or social cost-benefit analysis or financial appraisal</td>
<td>1.</td>
<td></td>
</tr>
<tr>
<td></td>
<td>2.</td>
<td></td>
</tr>
</tbody>
</table>

Further information
- Treatment of sites
Table 7: Definition of Validation Methods for Each Assessment Objective

<table>
<thead>
<tr>
<th>Application 1</th>
<th>Category of Assessment</th>
<th>Physical functioning/ User Acceptance/ Impact Analysis etc</th>
</tr>
</thead>
<tbody>
<tr>
<td>Assessment Objective 1</td>
<td>Indicator 1</td>
<td>Indicator 2</td>
</tr>
<tr>
<td>Reference Case</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Definition of Success</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Methods of Quantitative or Qualitative Assessment</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Conditions of Quantitative or Qualitative Assessment</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Statistical Considerations</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Representativeness</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Measure of Sampling</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Estimate of Statistical Confidence</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Possible Problems of Integrity</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Assessment Objective 2</td>
<td>Indicator 1</td>
<td>Indicator 2</td>
</tr>
<tr>
<td>Reference Case</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Methods of Quantitative or Qualitative Assessment</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Conditions of Quantitative or Qualitative Assessment</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Statistical Considerations</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Representativeness</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Measure of Sampling</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Estimate of Statistical Confidence</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Possible Problems of Integrity</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Further information
- assessment objectives
- treatment of sites
- treatment of indicators
B. <OTHER APPENDICES ETC.>