Deliverable 4.2 Final Impact Evaluation Framework

Project:	STTRIDE
Project full title:	**S**MARTER **T**RAVEL **T**ECHNOLOGY **R**EVIEW FOR **I**NVESTMENT **DE**CISIONS
Research Services Agreement:	CEDR TRANSNATIONAL ROAD RESEARCH PROGRAMME - Agreement for CEDR Call 2015: User Needs: Smarter Travel Technology Review for Investment Decisions (STTRIDE)
Date:	5 June 2018
Version:	4

• TRL UK • VTT Finland • RISE Sweden •

Supported by

The sole responsibility for the content of this agreement lies with the authors. It does not necessarily reflect the opinion of CEDR. CEDR is not responsible for any use that may be made of the information contained therein

Document Information

Contractual date of delivery to CEDR	30 June 2018
Actual date of delivery to CEDR	5 June 2018
Work package contributing	WP 4
Circulation	TBC
Abstract	TBC
Author(s)	Jean Hopkin, TRL, Marcus Jones, TRL, Jesse Fahnestock RISE, Erkki Siira, VTT
Reviewer	Marcus Jones

Document Status

Version	Date	Contributor	Reason for change/remarks
1	04.04.18	Jean Hopkin	Updated D4.1 to include points from pilot
2	12.04.18	Jesse Fahnstock	Additional points from Sweden
3	23.05.18	Jean Hopkin	Harmonisation of content with modules in D5.0
4	05.06.18	Jean Hopkin	Added hyperlinks and minor editing

CONTENTS

1 ABOUT THIS FRAMEWORK 1

2 EVALUATION IN THE CONTEXT OF STTRIDE 2

2.1 Background to STTRIDE 2
2.2 This document 2
2.3 What is evaluation? 4
2.4 Why evaluate? 5
2.5 How does evaluation planning fit in the life cycle of a project? 5
2.6 Why use a common framework? 6
2.7 What is the most suitable evaluation approach to use? 6
3 EVALUATION FRAMEWORK 8

3.1 Overview of stages in impact evaluation 8
3.2 Define user needs for the results 9
3.3 Describe the intervention 12
3.4 Describe the intervention logic 13
3.5 Define evaluation objectives 16
3.6 Frame the research questions 19
3.7 Pre-assessment of outcomes and impacts 23
3.8 Define assessment methods and write evaluation plan 26
3.9 Data collection and analysis 39
3.10 Report results 42
4 CONTENTS OF AN EVALUATION PLAN 45

5 BIBLIOGRAPHY 46

APPENDIX 1 GLOSSARY 47

1 About this framework

STTRIDE is funded by the Conference of European Directors of Roads. It is addressing how best to use technological advances to deliver positive modal shift towards sustainable travel, with emphasis on the inter-urban network managed by National Road Authorities.

Technology innovation in the mobility sector is moving at a rapid pace. Many emerging technologies are having or could have a significant impact on people's preferred mode of transport over the next twenty years. This technology-driven paradigm shift provides an opportunity for significant change in traveller behaviour without necessarily requiring major infrastructure investment or legislative intervention. Indeed, this could result in a substantial difference in future transport network demands, emissions and contribute to healthy lifestyles.

Harnessing the potential of technological development can make more efficient use of existing transport infrastructure and services, as well as facilitating the introduction of new and improved ones. For example, passenger information systems could increase bus occupancy, improving the case for investing in service improvements, thereby encouraging further modal shift.

The pace of change is such that it could be a challenge for road and transport authorities to understand the potential impacts and timescales associated with a wide range of technologies. Once a new technology has appeared, it can also be difficult to assess its impacts. As a result there is a knowledge gap for authorities wishing to understand how to support, respond to or invest in the technologies that will deliver their preferred outcomes.

STTRIDE has identified and analysed technologies which can be seen to affect positive modal change over the next 20 years and to support CEDR authorities in two ways: by providing a toolkit for selecting investment options for new technologies and a common evaluation framework (summarised in the diagram) for assessing their impacts.

This document provides a common framework for CEDR authorities and their consultants to use when evaluating the impact of technology-based interventions aimed at encouraging modal shift. It is intended to provide a consistent basis for planning, conducting, analysing and reporting on such interventions that will enable the various authorities to share and compare their results, learn from the lessons of others and build an evidence base for decision-making. It provides a 'menu' from which authorities can plan their evaluations within an overall project management programme, taking account of their own national and local objectives and priorities and national guidance on evaluation, potentially incorporating this into their own guidance documents.

It has been developed on the basis of experience and best practice in evaluation guidance for transport interventions. It makes the case for evaluation, sets out the role of evaluation in the project lifecycle and considers the most appropriate approach to use for evaluations within the scope of the STTRIDE project.

The evaluation framework provides guidance on each stage in the process of evaluating interventions involving new technologies to achieve modal shift. It also provides common frameworks for writing evaluation plans and reporting results to enable road authorities to compare their results.

2 Evaluation in the context of STTRIDE

2.1 Background to STTRIDE

STTRIDE is a European project funded by the Conference of European Directors of Roads. The project is addressing how best to use technological advances to deliver positive modal shift towards sustainable travel, with a particular emphasis on the inter-urban network managed by National Road Authorities.

Technology innovation in the mobility sector is moving at a rapid pace. Many emerging technologies are having or could have a significant impact on people's preferred mode of transport over the next twenty years. This technology-driven paradigm shift provides an opportunity for significant change in traveller behaviour without necessarily requiring major infrastructure investment or legislative intervention. Indeed, this could result in a substantial difference in future transport network demands, emissions and the contribution to healthy lifestyles.

Harnessing the potential of technological development can make more efficient use of existing transport infrastructure and services, as well as facilitating the introduction of new and improved ones. For example, passenger information systems could increase the occupancy of buses, which improves the business case for investing in improvements to the service, thereby encouraging further modal shift.

The pace of change is such that it could be a challenge for road and transport authorities to understand the potential impacts and timescales associated with a wide range of technologies. Once a new technology has appeared, it can also be difficult to assess its impacts. As a result there is a knowledge gap for authorities wishing to understand how to support, respond to, or invest in, the technologies that will deliver their preferred outcomes.

STTRIDE has identified and analysed technologies which can be seen making impact on positive modal change over the next 20 years and used this analysis to support CEDR authorities by providing a toolkit for selecting investment options for new technologies and a common evaluation framework that can be used to assess the impacts of implementing such technologies. These comprise the STTRIDE Evaluation Process Guidelines.

2.2 This document

This document provides a common framework for CEDR authorities and their consultants to use when evaluating the impact of technology-based interventions aimed at encouraging modal shift. These might include the transport impacts of 'non-transport' interventions and the potential benefits of the transport industry implementing technologies from other industries to achieve transport objectives, as well as applying technology to transport interventions. It is intended to provide a consistent basis for planning, analysing and reporting evaluation of such interventions within an overall project management programme that will enable the various authorities to share and compare their results, learn from the lessons of others and build an evidence base for decision-making. It is however important to note that objectives and priorities may vary from one country to another, so the scope of evaluations may vary within the overall 'menu' provided in this framework. Countries may also need to take account of national guidance on conducting evaluation and may potentially incorporate this document into their own guidance. The remainder of this introductory section makes the case for evaluation, sets out the role of evaluation in the project lifecycle and considers the most appropriate approach to use for evaluations within the scope of the STTRIDE project.

This evaluation framework provides guidance on each stage in the process of evaluating interventions involving new technologies to achieve modal shift in Section 3.1 to 3.9 of this document; the main points at each stage are summarised in green text boxes and examples are provided in blue text within the smaller text boxes. Section 3.10 provides a common framework

for reporting results to enable road authorities to compare their results, while Section 4 presents a common structure for writing an evaluation plan. A glossary of terms is included in the Appendix.

This document has been developed on the basis of experience and best practice in evaluation guidance for transport interventions. The main sources used are listed in the bibliography in Section 5.

A draft version of the framework was piloted by reviewing it with volunteer road authorities in Ireland and Sweden to understand the feasibility of applying it, both in terms of the data available or potentially available and the processes involved. The draft was then refined in the light of this experience and comments from road authorities, before making this final version available for use.

The evaluation framework is also available as a set of modules and templates in the STTRIDE Evaluation Process Guidelines. These can be downloaded from the STTRIDE web site. To navigate between different modules of these guidelines, users may click on the rows within Figure 2.1 which contain hyperlinks to the modules on the web site. Hyperlinks to the templates associated with tables, diagrams and report outlines are provided in the relevant sections of this document.

Figure 2.1 Navigation within STTRIDE Evaluation Process Guidelines

2.3 What is evaluation?

Evaluation is a planned and structured assessment of the extent to which an intervention has met its objectives after it has been implemented; this is sometimes termed 'ex poste' evaluation. It assesses the benefits, financial costs and negative consequences.

Before making investment decisions for new interventions, an appraisal will often be carried out to identify the potential benefits, costs and impacts; such an appraisal (or pre-implementation evaluation or 'ex ante' evaluation) can be used both to justify the investment and to plan how and where the intervention will be implemented in more detail. This appraisal will also identify the aspects of the intervention to be evaluated and the feasibility of doing this can be taken into account in the decision to invest in the intervention.

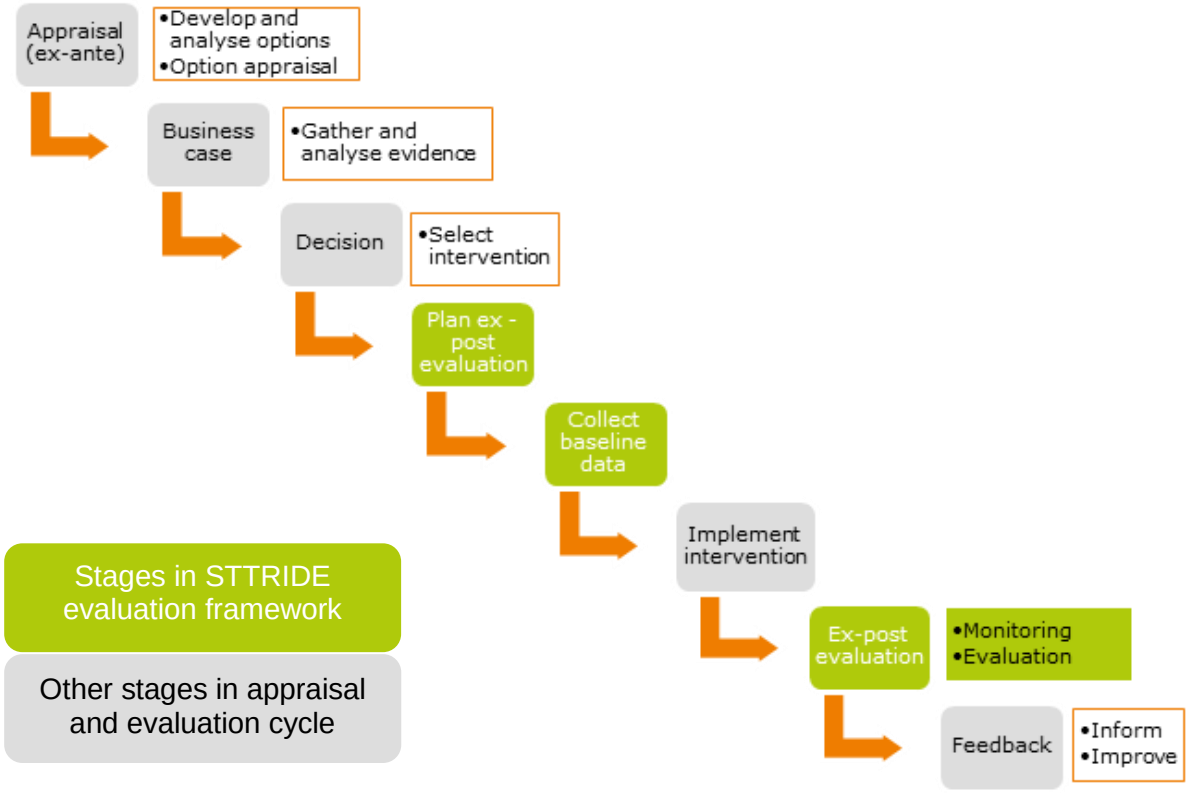
Usually in the case of interventions involving new technologies, a trial is undertaken before full scale implementation. Evaluation of the trial will then inform the decision on whether or not to proceed with further implementation. The approach to evaluation set out in this framework is appropriate for both trials and full scale implementations.

In the case of STTRIDE, the focus is on assessment of impacts after implementation (ex post evaluation). Although the focus in this case is on what happens after an intervention has been implemented, it is important that it is planned before implementation so that the current situation can be accurately monitored to provide a baseline against which to compare the intervention's impacts. Monitoring and evaluation are an important part of project management, and the approach presented here can be incorporated within project management processes.

Evaluation is often carried out by an organisation that is independent of those involved in implementing the intervention so that it can be clearly seen as an unbiased assessment.

The diagram below summarises the stages in the evaluation cycle from ex ante appraisal of potential interventions, to ex post evaluation of intervention and feedback to inform improvements to the intervention and future decisions. It indicates the timing of the stages which are covered by the STTRIDE evaluation framework within the overall process.

Figure 2.2 Ex-post evaluation within the appraisal and evaluation cycle



2.4 Why evaluate?

Evaluation after an intervention has been implemented has three main purposes:

- To understand the impacts of the intervention on travellers and other stakeholders and the extent to which it has contributed to its objectives; in the case of STTRIDE interventions reducing single occupancy car use on the inter-urban network and thereby cutting congestion and achieving wider policy impacts such as improving safety, environment, health and well-being
- To improve performance and learn lessons that will benefit other schemes in future
- To check that the investment made in the intervention can be justified by the benefits achieved, thereby supporting the business case for similar schemes elsewhere.

Thus evaluation is a key element of implementing any scheme involving new technology where the impacts and consequences are not well known and understood.

2.5 How does evaluation planning fit in the life cycle of a project?

A key document for evaluation is the evaluation plan. This document is prepared by the evaluation team and agreed by the relevant stakeholders in the intervention, initially in draft form but then refined at intervals during a project as scheme details are finalised and detailed plans for evaluation are defined. This evaluation team may consist of in-house staff or external consultants or a mixture of the two (for example by contracting out the specialist work involved in designing and conducting surveys and other data collection).

Evaluation planning should begin early in the life of a project. This enables the requirements for baseline and long term monitoring data to be defined early enough for any equipment and resources needed to be procured and operational in good time; procurement specialists will need to be involved at this stage. This early start to evaluation planning is indicated in Figure

2.1 above, in which planning the evaluation and collecting baseline data take place before the intervention is implemented. Early evaluation planning also helps to manage risks before investment decisions are made; mapping out the logic of the intervention and its likely impacts enables unintended consequences to be identified and addressed at the outset.

When implementing new technologies aimed at achieving modal shift, it is important to take account of the long term nature of some impacts. The lag time for some technologies is long so evaluation in the first year or two after the intervention may need to focus on what the intervention has delivered (outputs) and the short term outcomes. It will be important to consider whether to plan for evaluation resources to be made available for a longer term impact assessment to be carried out in future, in addition to the evaluation of short term outcomes.

2.6 Why use a common framework?

The STTRIDE evaluation framework provides a common, consistent basis for CEDR members to use when evaluating new technologies to encourage mode shift away from single occupancy car journeys and then reporting on the results. This consistency is intended to help CEDR members to compare different interventions and benefit from a pool of evaluation results focused on interventions addressing this specific objective.

Additionally, in cases where trial interventions are managed via external research or consulting engagements, the framework can provide guidance on evaluation *from* a CEDR member *to* its contracting partner.

The framework is intended to be sufficiently generic that it can help users to think through the issues relevant to evaluating this type of intervention and then prepare an evaluation plan that is tailored to their own context and type of intervention. Thus not all of the details set out in the framework will be relevant to all contexts and types of intervention.

Within the context of this common framework, there will be some national differences in the way that it can be applied, for two reasons:

- The roles of NRAs will differ between countries – for example in the extent to which they are involved in policy and operational roles; this may affect the scope of evaluation required in particular countries
- Some countries provide guidance on how evaluation of transport or other investment should be carried out, and it will be important to take this into account.

2.7 What is the most suitable evaluation approach to use?

Various approaches to evaluation are available. When considering the influence of technologies within the domain of the STTRIDE project, there are several features which are relevant when deciding which approach to adopt:

- The influence of new technologies on modal shift will often be indirect, with the result that attribution of impacts to the introduction of those technologies (rather than other factors) will be difficult
- The nature and scale of the likely impacts is uncertain, so it will be important to gain some understanding of why any changes took place, rather than just identifying that they happened
- Connected journeys and multi-modal journeys do not lend themselves to traditional data collection and measurement; for example transport services usually collect data on their own users, not users of other modes. Therefore it is likely that suitable baseline data will not be readily available and bespoke baseline data collection will be needed, covering the interfaces or interchanges between modes

- Changes in other parts of the network or travel demands (e.g. from new housing, workplace or retail developments) may obfuscate changes resulting from specific interventions
- As mentioned earlier, some new technologies will have impacts several years into the future, so these will not be apparent within a normal evaluation time frame of 1-2 years.

The most robust approach to assessing the level of impact of an intervention is to use an 'experimental' approach by analysing two comparable areas, one of which receives the intervention and the other does not. However this approach does not provide explanations for any changes and relies on the premise that the area without the intervention is not influenced by other factors.

Taking into account the advantages and disadvantages of various approaches to evaluation, a combination of approaches is recommended for the STTRIDE context.

- For some elements, an 'outcome' approach is sufficient, looking at the short and medium term changes following the intervention – for example to identify the change in number and proportion of journeys made by different modes.
- Some aspects of a theory-based approach are recommended, working within the overall framework of an 'intervention logic map' that sets out the theoretical connections between an intervention and its impacts. This approach makes it possible to investigate why and under what conditions any changes occurred, using qualitative and quantitative research methods, in cases where it is difficult to attribute changes specifically to the intervention. Gathering data from different sources and then comparing them to assess how they complement and support each other will provide confidence in the results.

In addition, it is important to understand not just the context for implementation but also the conditions set by the actors involved and their relationship to the technology and data about impacts that may be generated. While guiding principles for data collection and analysis are useful, different actor configurations and choices about technology will affect what data can and will be generated, how it can be used, etc. The evaluating authority may in some cases be dependent on other actors (e.g. research partners, local authorities, technology providers) who define and control data collection and analysis, and should adjust its guidance accordingly.

3 Evaluation framework

3.1 Overview of stages in impact evaluation

Following best practice principles, this evaluation framework involves working through a series of stages in the process of evaluating the impacts of an intervention, from the initial investigation of what users need from the evaluation to the reporting of evaluation results. These stages are set out in the diagram below. The following sub-sections (3.2 to 3.10) outline each of these stages.

Figure 3.1: Stages in the evaluation process

Stages in impact evaluation

In each section of this document, an example is provided to indicate how the information summarising each stage can be set out systematically in a table or diagram. These tables and diagrams are compiled in an evaluation plan. The evaluation plan is a living document, built up and agreed by the various members of the evaluation team. It provides a single reference source for use throughout the evaluation process. It is intended to be stand-alone, but may need to summarise and refer to other key documents. The proposed content of an evaluation plan is set out in Section 4.

3.2 Define user needs for the results

User needs

A clear understanding of who will use the results of the evaluation and how they will use them is an important starting point for designing the evaluation approach. Each of these stakeholders will have their own objectives and priorities which will determine the types of results they need from the evaluation. Such objectives and priorities may also vary between stakeholders with the same role but in different countries.

The resources available for the evaluation from each stakeholder should be identified here.

User needs: Smart Infrastructure

The implementation of smart infrastructure will involve a considerable amount of cooperation between the infrastructure owners, service providers and transport operators. Each party has their own objectives, and these need to be discovered. In joint investments, who pays what also needs to be considered. Infrastructure owners, such as municipalities, are interested in safety or providing a better environment for their residents. Businesses will need a positive return on their investment even if their work on the smart infrastructure is technical.

User needs: Voice Recognition

Speech recognition promises easier user interaction in situations where complex choices are being made. It is especially helpful in opening up interactions with foreign users who do not know the local language well – they can speak in their native language, with real time translation into the local language.

For transport operators, better service for their customers should be a top priority.

User needs: Advanced Fare Management

Public transport operators see Advanced Fare Management Systems as a way of optimising occupancy, reducing boarding times and operating costs and increasing revenues.

Travellers see such systems as a way of making public transport more attractive and accessible, especially for those who are not used to travelling by public transport.

User needs: V2X technologies

V2X communications technologies are already associated with defined lists of services that should be available at different stages ('Day One' and 'Day Two'). These objectives will be the starting point for road authorities and service providers, with specific objectives related to testing, verification, and upscaling possible additions.

Table 3.1: Example of Definition of user needs

Type of user/ stakeholder	User	Typical objectives	Areas of interest
National government	National transport authority	Improve road safety Reduce harmful emissions from transport Foster economic growth Improve health and well-being Improve accessibility Foster social inclusion	Contributing to national policy objectives
	National road authority	Reduce casualties Reduce congestion Improve journey reliability	Improving network performance
Local government	Local authority 1	Reduce vehicle emissions Reduce noise levels/ noise pollution	Improve quality of life for residents
	Local authority 2	Improve journey times for through traffic	Reduce delays
Transport operators	Road operator	Reduce casualties Reduce congestion Improve user satisfaction Reduce road maintenance requirements	Meeting performance targets
	Public transport operator 1	Maintain/ increase customer base Reduce operating costs Increase revenues Improve quality of service	Increasing profitability
	Public transport operator 2	Improve adherence to timetable Improve/ maintain accessibility	Increase customer base
	Freight	Improve rate of on-time delivery Reduce wasted driver hours	Improve profitability
Service providers	Technology infrastructure	Maintain or increase customer base Improve quality and range of services delivered	Improve profitability Extend market
	Communications	Maintain or increase customer base Improve quality and range of services delivered	Improve profitability Extend market

Type of user/ stakeholder	User	Typical objectives	Areas of interest
	Travel information & navigation	Maintain or increase customer base Improve quality and range of services delivered	Improve profitability Extend market
	Traffic management	Reduce congestion Improve journey reliability	Improve network performance
	Payment	Improve efficiency of payments Reduce fraud/ missing payments	Improve profitability Increase customer base
	Shared transport	Maintain or increase customer base Improve quality of service delivered	Improve profitability Extend market
	Emergency services	Meet service targets	Improve efficiency
	Infotainment	Maintain or increase customer base Improve quality of service delivered	Improve profitability Extend market
Travellers	Private motorists	Improve journey time reliability Improve ease/ comfort of journey	Journey quality
	Freight drivers	Improve journey time reliability Decrease vehicle-km per unit	Meet delivery and performance targets
	Pedestrians	Improved safety and efficiency of walk journeys	Improved quality of walking environment
	Cyclists	Improved safety and efficiency of walk journeys	Improved quality of cycling environment
	Public transport users	Improved connections and interchanges with other modes Improved journey information/ support Improved journey reliability Improve accessibility and attractiveness of service	Improved level of service Improved quality of service
Residents		Reduce adverse impacts of road traffic on the local area	Improved quality of life

A template for this table is available in the file Templates 5 of the Evaluation Process Guidelines on the STTRIDE web site.

At this stage it is also useful to identify the financial and staff resources that are available for the

evaluation and which stakeholders are providing them – for example as in the table below. This initial plan will need to be reviewed once the metrics and data needs are clear.

A template for this table is available in the file Templates 5 of the Evaluation Process Guidelines on the STTRIDE web site.

Table 3.2: Resource plan for evaluation

Role	Organisation	Year 1		Year 2		Year 3	
		Staff	€	Staff	€	Staff	€
Project management							
Evaluation planning							
Data collection							
Data analysis							
Report writing							
Data capture equipment							
Other							

3.3 Describe the intervention

Description

Describing the intervention is the next step in defining the evaluation. The technologies and measures to be implemented, where and when, along with the objectives of the intervention for the various users/ stakeholders, are the main elements of the description in this stage. This will help to identify the scope and focus of the evaluation.

The description may draw on information assembled during the ex ante appraisal before deciding to implement the intervention.

Description of an Open Data intervention

An intervention could involve the development of an API that makes a publicly held transport/ traffic database accessible to third-party application developers.

Objectives would include supporting the development of new mobile transport applications, improving the usability of public transport services, and driving increases in shared and/or multi-modal transport options.

Table 3.3: Description of the intervention

	User/ Stakeholder A Road operator	User/ Stakeholder B City authority	User/ Stakeholder C Service provider
Technologies and measures to be implemented	New technology providing a service that will encourage a reduction in single car occupancy at congested sites on the inter-urban network		
Objectives of the intervention	A.1 Reduce peak hour congestion A.2 Improve customer experience	B.1 Reduce the impact of interurban congestion on local roads	C.1 Increase take-up C.2 Extend customer base
Site (s)	Roads/ towns Include a map		
Dates	Month and year for start and completion of implementation		

A template for this table is available in Template 6 of the Evaluation Process Guidelines on the STTRIDE web site.

3.4 Describe the intervention logic

Intervention logic

The intervention logic summarises the main components that are needed to enable the intervention to deliver its intended impacts, including the type of change in mode use that the technology is expected to bring about (which may include unintended changes). It also indicates how these components are connected. In this stage of the evaluation, a logic map diagram is created which summarises these links.

The intervention logic may be based on the information gathered during the ex ante appraisal and preparation of the business case for the intervention, but may also draw on evaluation evidence from similar or related interventions elsewhere.

As a first stage in summarising the intervention logic for a scheme within the STTRIDE framework, it is useful to understand the types of change in mode use that are expected to be brought about by the intervention. There are broadly five types of mode change that may be brought about by new technology:

- Increased sharing/ occupancy for existing trips, with reduced vehicle km for the same person km
- Shift to a different mode for existing trips, with reduced vehicle km for the same person km
- Replace existing journeys with shorter ones serving the same purpose (for example by using local pick-up points or local shared remote working office), with reduced vehicle km and reduced person km
- Replace journeys and shift mode, with a greater reduction in vehicle km for a reduction in person km
- Avoid travel (for example by home working, home delivery, or teleconferencing), with no vehicle km or person km for these activities.

It is important to bear in mind that with the exception of avoiding travel, all of these types of change could increase the multi-modal nature of journeys as people walk or cycle to access other modes.

The intervention logic map helps to identify whether there are any links between the components that are unclear, the evidence that is required from the evaluation, and to highlight any gaps in the evidence which will determine the focus of the evaluation effort and what will need to be measured. It will also help to ascertain which evaluation approach will make it possible to attribute the changes that are measured to the intervention. This logic map is then used to frame the research questions and decide on the evaluation approach.

If an appraisal was carried out before deciding to invest in the scheme, this is likely to have considered the nature of the expected impacts and the mechanisms by which these might be achieved. Such appraisals provide a useful starting point for developing the logic map. However other sources of information that were not part of the appraisal or which have become available since the appraisal (such as evaluation of similar interventions elsewhere) should also be considered. Consulting stakeholders will also provide useful insights into the intervention logic.

An example logic map is shown overleaf for introduction of smart ticketing technology (Ball S et al, 2015).

The first part of the logic map summarises the context for the intervention from the point of view of the different types of stakeholder, for example:

- National transport policies supported by the intervention
- Regional or local issues and priorities addressed by the intervention
- Other contextual factors that may influence the ability of the intervention to achieve its outcomes and impacts; these may be associated with different stakeholder groups.

The intervention logic then considers the inputs, which can be measured quantitatively or quantitatively in the evaluation:

- Financial resources invested to implement the intervention
- Other resources invested – for example staff, skills, equipment, research.

The outputs are the next stage in the intervention logic. The outputs are what the intervention is going to produce and they will be monitored during the evaluation. They may include:

- Physical products of the intervention such as a new database of public transport timetables and interchanges or a web site providing the booking service for a shared car scheme.
- Activities which result directly from the intervention such as promotion events or services for users
- Participation which results directly from the intervention – i.e. the types of stakeholder or geographic areas that will be influenced or affected.

In the example logic map shown, the evaluation objective index numbers are referenced in brackets in the relevant components of the intervention logic.

The next stage of the logic map identifies the outcomes that the intervention is aiming to achieve, including changes in quality of journeys (such as security or ease of interchange between modes), mode use and traffic but should also identify potential unintended outcomes. These will be monitored during the evaluation. Outcomes are often separated into:

- Short term outcomes
- Medium term outcomes (1 – 2 years).

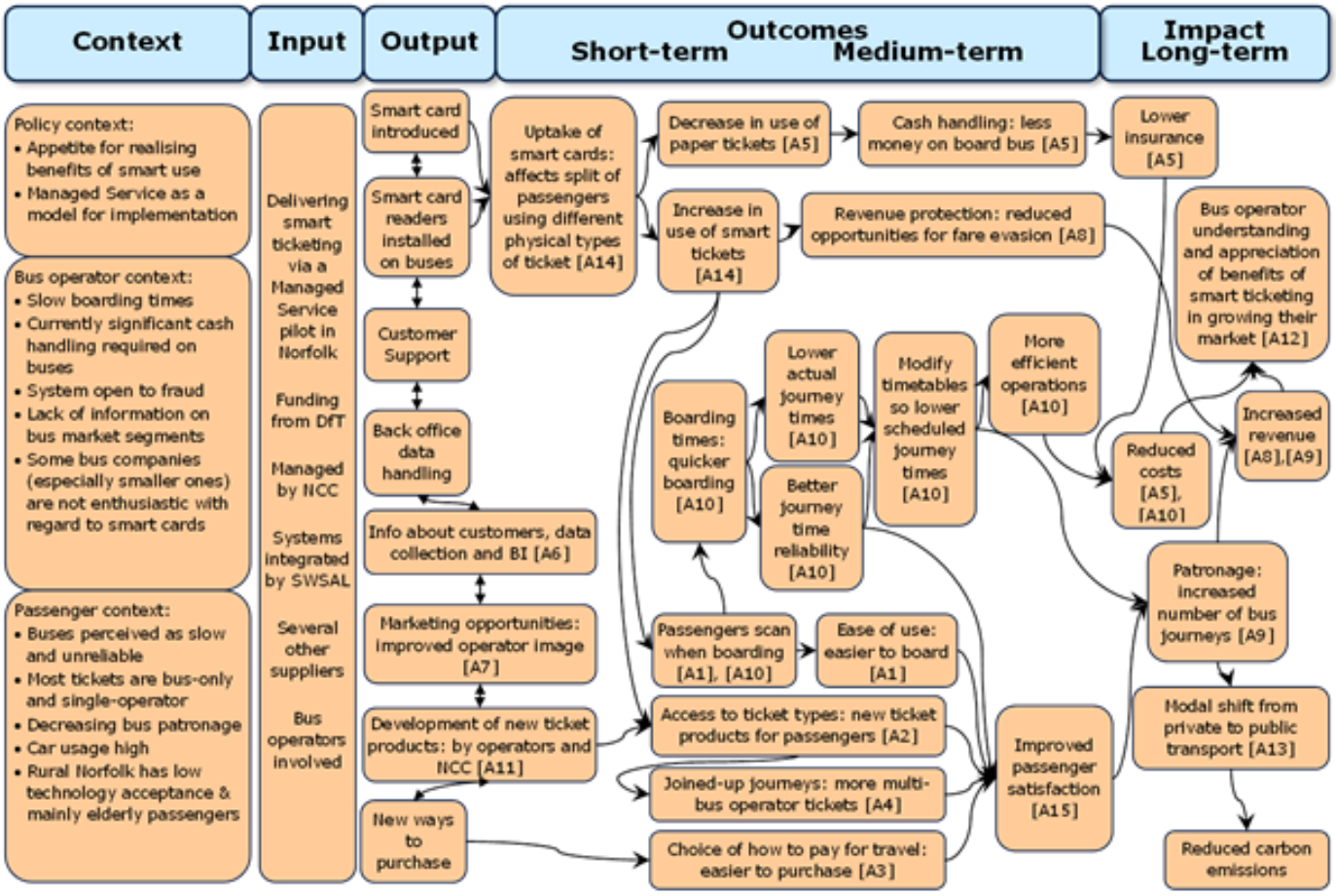
The final stage of the logic map is the long term impacts of the intervention, including the societal consequences of the changes in mode use, which will be measured during the

evaluation. Again potential unintended as well as intended impacts should be considered. These long term impacts may be quite different from the short term impacts and may include impacts on:

- Environment (noise, air quality etc.)
- Safety
- Health
- Well-being
- Accessibility
- Social inclusion
- Economy.

In the case of new technologies, as mentioned in Section 2.7, these impacts may become evident several years into the future.

Figure 3.2: Example intervention logic map



Source: Ball S, 2015

A template for creating an intervention logic map is available in Template 7 of the Evaluation Process Guidelines on the STTRIDE web site.

Note that before setting out the logic map in diagrammatic form, it may be helpful to create a table which can be used to identify and agree on the components to be included.

3.5 Define evaluation objectives

Evaluation objectives

To meet the needs of the various users of the evaluation results (defined in the first stage of setting out the evaluation framework), different types of assessment will be appropriate. These may be:

- Performance assessment (such as technical performance, reliability)
- User acceptance assessment (such as users' opinions, preferences, take-up rates, willingness to pay)
- Impact assessment (such as safety, environment, economy, user behaviour, mode use)
- Socio-economic evaluation (benefits and costs of the intervention)
- Financial assessment (costs of setting up and running the scheme, rate of return on investment, payback period).

The core of a STTRIDE evaluation of introducing new technologies to encourage changes in mode use is likely to be an impact assessment after the intervention, but other types of assessment will also be important. For example the user acceptance and impacts may be affected by system performance, so the results of a performance assessment will help to inform the results of the user acceptance and impact assessment. Public sector stakeholders who have invested in the intervention such as the local authority, or in some cases the National Road Authority, are likely to need a socio-economic evaluation to justify the investment and to inform future investment decisions. Some stakeholders, such as those operating a commercial service, will need a financial assessment to understand the effect on the costs of a service.

Evaluation objectives: Advanced Fare Management

Public transport operators would usually need to see a financial assessment of Fare Management Systems, while representatives of users would expect to see a user acceptance assessment.

Evaluation objectives: Wearable Technologies

Certain categories of wearable technology have already been accepted by users. For example activity bands and smart watches have been successful and have promoted a healthier lifestyle, with steps or exercising being rewarded and leading to people walking or cycling more. But there have been several unsuccessful entries to the market where the technology has not been mature enough or the users have not seen the value of the device.

Table 3.4: Example of evaluation objectives and user groups

Type of assessment	Evaluation objective	Stakeholder/ user groups involved
Performance	Assess the scale of service delivery/ amount of service provided	Transport operators Service providers
	Assess reliability of the service	Transport operators Service providers
	Assess availability of the service	Transport operators Service providers
	Assess the interoperability of the service	Government Road authority Transport operators Service providers
User acceptance	Assess awareness of the service	Travellers Transport operators Service providers
	Assess accessibility of the service	Travellers Transport operators Service providers
	Assess ease of use of the service	Travellers Transport operators Service providers
	Assess preferences for the service vs other options	Travellers
	Assess willingness to pay for the service	Travellers
Impact assessment	Assess the impact on provision of new products/ services	Local authority Road authority Travellers
	Assess the impact on the need for changes to other systems/ services	Transport operators Service providers
	Assess the impact on introducing new 'pain points' or 'bottlenecks' – i.e. the next area for action after introducing the intervention	Local authority Road authority Transport operators Service providers
	Assess the impact on mode use	Local authority Road authority Travellers

Type of assessment	Evaluation objective	Stakeholder/ user groups involved
	Assess the impact on single car use on the inter-urban network	Road authority Travellers
	Assess the impact on journey efficiency for different modes	Local authority Road authority Travellers
	Assess the impact on journey quality	Local authority Road authority Transport operators Service providers Travellers
	Assess the impact on accessibility	Government Local authority Road authority Transport operators Travellers
	Assess the impact on the environment	Government Local authority Road authority
	Assess the impact on safety	Government Local authority Road authority Travellers
	Assess the impact on health and well-being	Government Travellers
	Assess the impact on social inclusion	Government Local authority Travellers
Socio-economic evaluation	Assess the societal gains and losses	Government Local authority Road authority
Financial assessment	Assess the financial impacts	Transport operators Service providers

A template for this table is available in Template 8 of the Evaluation Process Guidelines on the STTRIDE web site.

3.6 Frame the research questions

Research questions

The research questions define the questions to be answered during the evaluation. A series of research questions will be defined to address each assessment objective. These questions can also be used to structure the analysis and reporting.

Sub-questions are defined and a check is made to ensure that all elements of the logic map are included in the research questions.

Two types of research question may be relevant: those which inform the evaluation objectives, and those which help provide an understanding of the impacts.

3.6.1 Research questions to inform evaluation objectives

Research questions may be of various types: open, specific, or exploratory. The table below sets out example research questions to address evaluation objectives. For simplicity of presentation, the research questions shown in the examples cover several aspects of each issue; in practice these would be separated out into sub-questions.

Table 3.5: Example of evaluation objectives and research questions

Type of assessment	Evaluation objective	Example research questions
Performance	Scale of service delivery/ amount of service provided	What is the scale of service delivery or amount of service provided, where, and when?
	Reliability of the service	How often does the service fail to operate as planned, and how long do these incidents last?
	Availability of the service	For what percentage of each day or week is the service fully operational/ what proportion of potential users have access to the service?
	Interoperability of the service	To what extent is the new system/ service interoperable with others/ existing services?
User acceptance	Awareness of the service	Are users aware of the service and informed about what it offers them?
	Accessibility of the service	What proportion of attempted uses fail because the user does not have the equipment/ service needed to access it and how many potential users does this represent?

Type of assessment	Evaluation objective	Example research questions
	Ease of use of the service	Have users found the service easy to use, and if so to what extent/ how? How long does it take users to complete actions required to use the service? How frequently is the service used by individuals? How is use of the service shared between different user groups?
	Preferences for the service vs other options	How do users think the new service compares with what they did before in terms of cost, quality, overall journey time etc.?
	Willingness to pay for the service	How much were users willing to pay for the service, and how did this vary between types of user, types of journey etc.?
Impact assessment	Provision of new products/ services	Has there been a change in the number and/ or scale of delivery new products/ services that have been provided, and what has changed as a result?
	Introduction of new 'pain points' or 'bottlenecks'	Has the new technology led to a new 'pain point' in the transport system?
	Changes to other systems/ services	Has there been a need to change other systems/ services as a result of introducing the intervention, and what is the impact of those changes?
	Mode use	Has the number of journeys made by each mode changed on the inter-urban network and or other roads, and if so in what way? Has the number of shared mode journeys changed on the inter-urban network and or on other roads, and if so in what way? Has the number of multi-modal or connected journeys changed on the inter-urban network and or other roads, and if so in what way? Have existing journeys been replaced with shorter ones that serve the same purpose, and if so which ones, where, and how? Have journeys been avoided by home working, home delivery, teleconferencing etc., and if so which ones, where and how?

Type of assessment	Evaluation objective	Example research questions
	Single car use on the inter-urban network	Has the number of single car occupant journeys on the inter-urban network changed, and if so, at what times, and by how much?
	Journey efficiency	Have journey times changed on the interurban network and/ or other roads, for car and other modes, and if so, in what way? Have journey speeds changed on the interurban network and/ or other roads, for car and other modes, and if so, in what way?
	Journey quality	Has the quality of journeys changed on the interurban network and/ or other roads, and if so, in what way?
	Accessibility	Has the number of people with access to key facilities changed in the area, and if so, in what way?
	Environment	Has fuel consumption changed on the interurban network and/ or other roads, and if so, in what way?
	Safety	Has the number of casualties changed on the interurban network and/ or other roads, and if so, in what way?
	Health	Has the amount of active travel changed, and if so in what way? Have vehicle emissions changed on the interurban network and/ or other roads, and if so, in what way?
	Well-being	Has the overall level of well-being changed in the area and if so, in what way?
	Social inclusion	Has the overall level of social inclusion changed in the area and if so, in what way?
Socio-economic evaluation	Societal gains and losses	What is the monetary value of changes in safety, economy, health and the environment? What were the overall costs of setting up and operating the scheme and over what timescale are these incurred?
Financial assessment	Financial impacts	What were the additional costs for service providers and operators of setting up and operating the scheme over what timescale?

Type of assessment	Evaluation objective	Example research questions
		What were the additional revenues accrued by service providers over what timescale? Has the National Road Authority made cost savings, and if so how and over what timescale?

A template for this table is available in Template 9 of the Evaluation Process Guidelines on the STTRIDE web site.

3.6.2 Research questions to understand impacts

In addition to the research questions associated with the evaluation objectives, some stakeholders are likely to require different types of evidence: some will wish to demonstrate that the investment has delivered the impacts that were anticipated (accountability evaluation), while others will wish to gain an understanding of which interventions work, under what circumstances and why (knowledge based evaluation). These requirements may generate additional research questions aimed at further understanding of how and why the outcomes and impacts were achieved, and whether the anticipated outcomes and impacts were realised.

For example research questions in a knowledge-based evaluation might include:

- Were there any unanticipated impacts or displacement effects?
- To what extent were the observed changes in mode use additional to what would have happened in the absence of the scheme?
- What are the main factors or mechanisms that led to the intervention achieving its impacts?
- How were the impacts distributed between different groups of travellers, types of road, area?
- Which target groups was the intervention most effective for?
- What lessons can be learned for development of future interventions?

An accountability evaluation might include questions such as:

- Did the intervention result in the anticipated outcomes and impacts, including change in traffic flow, journey time, journey quality, cost of travel, ease of use of services, mode use, social, economic or environmental factors?
- To what extent were the planned outputs delivered?
- To what extent has the change in mode use anticipated from the intervention been achieved?
- To what extent did anticipated costs and benefits match those actually incurred or realised?

Having set out the broad types of question relevant to the intervention covering the evaluation objectives and the types of evaluation, the next stage is to develop specific sub-questions under each of the research questions. These can be reviewed against the intervention logic map to check that the questions match the logic map, and that all elements of the logic map are covered. If there are gaps, a decision will need to be taken on whether the scope of the evaluation should be extended to cover the gaps, or whether these gaps should be excluded from the scope of the evaluation.

The research questions related to changes associated with the intervention should be used to develop hypotheses for statistical testing during the analysis. A hypothesis is a statement linking a cause to an effect and predicts the expected direction of any change or difference.

3.7 Pre-assessment of outcomes and impacts

Pre-assessment

In order to plan an evaluation it is important to obtain the best possible understanding of the nature and scale of likely outcomes and impacts before designing the evaluation in detail. This information can then be used to determine which measurements are taken and how many, during the data collection phase of the evaluation.

There are some fundamental uncertainties in the domain of influencing single car use on inter-urban networks by introducing new technologies. Primarily these concern the extent and nature of the impact of new technologies on modal shift and the potential for connected and multi-modal journeys.

The logic map created earlier in the evaluation process will have set out the types of short and medium term outcomes and long term impacts that are expected and this information can be used as the starting point here. The potential unintended consequences of the intervention should be included among the likely impacts considered, as well as the intended impacts.

It is helpful to summarise the outcomes and impacts which are expected for each type of user or stakeholder, and the likely qualitative or quantitative magnitude of the impact in a table. Evidence from other cases of similar interventions could be reviewed to inform this. An example is shown overleaf. A table such as this can be used to make a final selection of the outcomes and impacts that will be included in the assessment.

Pre-assessment: Augmented Reality

A predicted outcome for augmented reality (AR) is that it provides rich content and visualisation for travellers on foot and in vehicles. Vehicle-based AR is provided in a controlled environment and the content it can provide is easier to access and manage by the content provider. Using mobile devices or special AR equipment, people may access this rich content during their journey.

The most probable outcome is for navigation to become easier as routes can be shown through the AR, as well as service locations and points of interest. Especially in the case of tourists and visitors, AR promises to provide help for people who are confused or lost.

Pre-assessment: Electric Vehicles

Wider availability of electric bicycles and infrastructure for using them may encourage multi-modal journeys; cycling to public transport stops or stations and then using public transport for the longer inter-urban leg of the journey.

Pre-assessment: Traffic Management Systems

Traffic management systems can benefit all modes by smoothing traffic flow and making journeys shorter and more predictable. But if private car journeys benefit as much or more than other modes, the impact on mode shift away from single occupancy car use can be neutral or even negative.

Table 3.6: Example expected outcomes and impacts

Outcomes and impacts expected	Type of user/ stakeholder	Scale and direction of impact
Increase in number of products/ services available	Local authority Road authority Travellers	++
Change in other systems/ services	Local authority Road authority Transport operator Service provider Travellers	?
Increase in level of sharing/ occupancy for existing trips	Local authority Road authority Travellers	+
Reduction in vehicle use	Local authority Road authority Travellers	-
Reduction in distance travelled in vehicles on existing journeys	Local authority Road authority Travellers	-
Increase in journeys by cycle/ bus/ walk	Local authority Road authority Travellers	+
Change in number of multi-modal or connected journeys	Road authority Local authority Transport operators Service providers Travellers	++
Increase in frequency of service use	Road authority Local authority Transport operators Service providers Travellers	++
Improved journey efficiency by car and/ other modes	Local authority Road authority Travellers	+

Outcomes and impacts expected	Type of user/ stakeholder	Scale and direction of impact
Improved journey quality	Local authority Road authority Transport operators Service providers Travellers	+
Improved accessibility to facilities	Government Local authority Road authority Transport operators Travellers	+
Reduction in environmental impacts	Government Local authority Road authority	-
Improvement in safety	Government Local authority Road authority Travellers	+
Improved health	Government Travellers	+
Improved well-being	Government Travellers	+
Improved social inclusion	Government Local authority Travellers	+
Increased profitability of transport services	Local authority Transport operators Service providers	++
Potential unintended impacts	Government Local authority Road authority Service providers Travellers	?

A template for this table is available in Template 10 of the Evaluation Process Guidelines on the STTRIDE web site.

3.8 Define assessment methods and write evaluation plan

Assessment methods

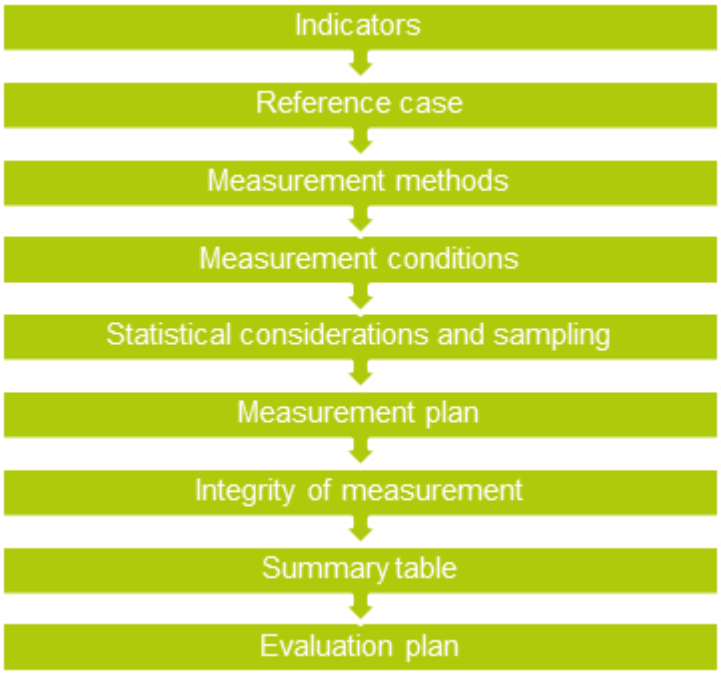
This step defines the assessment methods in more detail by determining which methods are to be used to meet each of the evaluation objectives defined earlier.

The assessment methods can be summarised in a table as shown in Section 3.8.8 below. A flow chart is also recommended.

Once they are defined, the assessment methods are recorded in an evaluation plan.

Defining the assessment methods involves a series of stages as shown in the diagram below and described in Sections 3.8.1 to 3.8.8. They lead to the content of the evaluation plan as outlined in Section 4.

Figure 3.3: Stages in defining assessment methods



3.8.1 Define indicators

The indicators are the variables used to estimate the impacts and thus the extent to which the objectives have been met. They are relative measures, with a denominator that makes it possible to compare them per unit of distance, time, area, user etc. Indicators should be defined for each research question (and therefore linked to an evaluation objective), and in each case there should be a clear definition of how the indicators are to be measured or derived. Linking the indicators with evaluation objectives means that the data collection is focused on the main areas that have

Define indicators: HD Maps

The impact of and investment in HD Maps and road databases can be indicated at the output level in terms of the number of square miles mapped, or the frequency of updates related to roadworks, diversions etc.

At the outcome level, indicators might include the number of calls made by onboard technologies to these databases.

Impact indicators would relate to the rate of accidents and delays associated with roadworks, diversions etc. that have been added to HD Maps and databases.

been identified for assessment, and less relevant data is not collected.

The indicators used should be able to meet two criteria:

- Clearly reflect the impact
- Can be assessed reliably using the methods available.

When defining the indicators, the availability of new types of data should be considered. It is often the case that new types of data are generated when new technologies are implemented; such data may be useful, if not central, to the evaluation. For example advanced fare management systems collect data on customers' use and habits, while cooperative ITS (C-ITS) services can collect data from vehicles to produce information on vehicle flows on routes at different times. In addition, technology advances are making new data sources available through 'Open Data' initiatives which can be useful for monitoring and evaluation.

The indicators will also vary depending on where the technologies are deployed; they may be in vehicles, carried or worn by travellers, or be part of a service delivery 'infrastructure' such as an operations room, a data 'hub' on or near roads, interchanges, etc.

With the focus on connected and multimodal journeys in STTRIDE, it will be important to include indicators reflecting use of the interfaces between modes; for example cycle parking at public transport stations, use of public transport interchanges and park and ride services.

Bearing in mind the STTRIDE focus on sustainable travel, it is worth noting that in some countries, specific indicators have been developed for assessing impacts associated with active travel (walking, cycling etc.) such as journey quality, physical activity, absenteeism, safety, environment, and time savings. Techniques have been developed to take account of the likely decay in impacts of some types of scheme over time and the variation in response to interventions between different types of user (such as commuters, leisure users and utility users).

It is recommended that a set of common indicators be defined for use in different National Road Authorities for assessing similar interventions. This will make it possible to make comparisons between areas and interventions, enabling road authorities to learn from each other.

Two types of indicator should be defined: those that are generic to any investigation of the impact of technologies on mode use, and those that are specific to certain technologies.

Suggested indicators that could be used to address the example research questions listed in Section 3.6 above are shown in the table below.

Table 3.7: Example indicators

Example research questions	Example indicators
What is the scale of service delivery or amount of service provided, where, and when?	**Outputs** • Number of items of equipment/ vehicles sold/ provided/ equipped in different areas and times • Number of applications downloaded/ server calls • Number of km of network equipped with technology • Square km of communications coverage • Speed/ latency of communications • Number of leaflets distributed/ hits on information web site/ enquiries answered • Number of service providers engaged • Number of safety/ information messages delivered to different user groups
	Outcomes • Number of subscribers/ users • Number of new services established • Number of requests/ notifications during service delivery • Number of safety/ information messages received by different user groups • Number of incidences of users responding to information provided • Availability of service at planned time • Availability of service for on-demand use
Has there been a need to change other systems/ services as a result of introducing the intervention, and what is the impact of those changes?	• Number of systems/ services changed • Degree of impact of these changes on: operators, service providers, users
Has the new technology led to a new 'pain point' in the transport system?	• Features of pain point and number of users/ trips affected
How often does the service fail to operate as planned, and how long do these incidents last?	• Number of incidents in operating week when service fails to operate as planned • Mean, minimum and maximum duration of incidents in operating week when service fails to operate as planned

Example research questions	Example indicators
For what percentage of each day or week is the service fully operational/ what proportion of potential users have access to the service?	• Percentage of operating week when service is available for, and has capacity for, users • Percentage of all potential users who have subscribed to the service • Minimum time required between booking service and travelling
To what extent is the new system/ service interoperable with others/ existing services?	• Number/ proportion of existing/ other services that the new service is interoperable with
Are users aware of the service and informed about what it offers them?	• Number/ proportion of potential users who know about the service • Number/ proportion of potential users who have an accurate understanding of its key features
What proportion of attempted uses fail because the user does not have the equipment/ service needed to access it and how many potential users does this represent?	• Percentage of attempts to use the service which fail and why (because the user does not have smartphone, credit card, internet etc.) • Percentage of potential users who are unable to use for these reasons
Have users found the service easy to use, and if so to what extent/ how?	• Percentage of users rating the service as easy/ very easy to use • Shares/ likes/ ratings (where applicable)
How long does it take users to complete actions required to use the service?	• Time taken to complete actions to use the service
How frequently is the service used by individuals?	• Time elapsed between uses of the service
How is use of the service shared between different user groups?	• Percentage distribution of uses of service by different user groups

Example research questions	Example indicators
How do users think the new service compares with what they did before in terms of cost, quality, overall journey time etc.?	• Time before first use • Time between uses • Cost • Waiting time • Time spent using service • Number of links/ modes • Journey time • Number of referrals/ shares made to other potential users • Compromises made
How much were users willing to pay for the service, and how did this vary between types of user, types of journey etc.?	• Willingness to pay for service on commute journeys • Willingness to pay for service on business journeys (in course of work) • Willingness to pay for service on leisure journeys
Has there been a change in the number and/ or scale of delivery new products/ services that have been provided, and what has changed as a result?	• Number of products/ services available
Has the number of journeys made by each mode changed on the inter-urban network and or other roads, and if so in what way?	Data on indicators for all users and each user group: • Number of vehicles of each type on interurban and other roads • Number of trips by each mode on interurban and other roads • Number of connected/ multi-modal trips on interurban and other roads • Distance travelled by each mode on interurban and other roads • Percentage of trips/ distance by each mode on interurban and other roads • Percentage of trips on interurban and other roads that are connected/ multimodal • Correlation between distribution of messages and use of inter-urban network, such as percentage change in vehicles in response to messages about better routes, delays, alternative modes
Has the number of shared mode journeys changed on the inter-urban network and or on other roads, and if so in what way?	• Number of journeys by public transport or shared car on interurban and other roads/ routes

Example research questions	Example indicators
Has the number of multi-modal or connected journeys changed on the inter-urban network and or other roads, and if so in what way?	• Number of journeys involving two modes or more (excluding walk to access to public transport) • Number of journeys involving interchange within modes e.g. multi-stage public transport journeys • Number of journeys involving relevant combinations of modes • Number of journeys involving relevant interchanges e.g. use of cycle parking at public transport stations, park and ride, hire of bicycles/ e-bikes
Have existing journeys been replaced with shorter ones that serve the same purpose, and if so which ones, where, and how?	• Number of journeys shifted to a closer destination
Have journeys been avoided by home working, home delivery, teleconferencing etc., and if so which ones, where and how?	• Number of journeys avoided
Has the number of single car occupant journeys on the inter-urban network changed, and if so, at what times, and by how much?	• Number of single occupant car journeys on interurban and other roads
Have journey times changed on the interurban network and/ or other roads, for car and other modes, and if so, in what way?	• Mean journey time on interurban and other roads for each relevant mode
Have journey speeds changed on the interurban network and/ or other roads, for car and other modes, and if so, in what way?	• Mean journey speed on interurban and other roads for each relevant mode
Has the quality of journeys changed on the interurban network and/ or other roads, and if so, in what way?	• Standard deviation of journey time on interurban and other roads for each relevant mode • Standard deviation of journey speed on interurban and other roads for each relevant mode • Number of vehicles parked at the roadside

Example research questions	Example indicators
Has fuel consumption changed on the interurban network and/ or other roads, and if so, in what way?	• Estimated fuel consumption (petrol, diesel, electricity, other) by mode on interurban and other roads derived from number of vehicles, types, distance, speed
Has the number of casualties changed on the interurban network and/ or other roads, and if so, in what way?	• Number of fatalities by mode • Number of serious injuries by mode
Has the amount of active travel changed, and if so in what way?	• Number of trips by walk/ cycle/ access to public/ shared transport • Distance travelled on walk/ cycle/ access to public/ shared transport trips • Proportion of trips by walk/ cycle/ access to public/ shared transport • Proportion of distance by walk/ cycle/ access to public/ shared transport
Have vehicle emissions changed on the interurban network and/ or other roads, and if so, in what way?	• Estimated emissions of CO_2 derived from vehicle numbers, types, distance, speed
What is the monetary value of changes in safety, economy, health and the environment?	• Monetised valuations of changes in safety, journey time, levels of active travel, emissions
What were the overall costs of setting up and operating the scheme and over what timescale are these incurred?	• Investment costs for different organisations • Operational costs for different organisations • Maintenance costs for different organisations
What were the additional costs for service providers and operators of setting up and operating the scheme over what timescale?	• Difference between investment costs and usual costs for different organisations • Difference between operational costs and usual costs for different organisations • Difference between maintenance costs and usual costs for different organisations
What were the additional revenues accrued by service providers over what timescale?	• Revenues for transport operators/ service providers (may be derived from data on trips/ subscriptions etc.)

Example research questions	Example indicators
Has the National Road Authority made cost savings, and if so how and over what timescale?	• Costs incurred by the NRA to operate and maintain the network

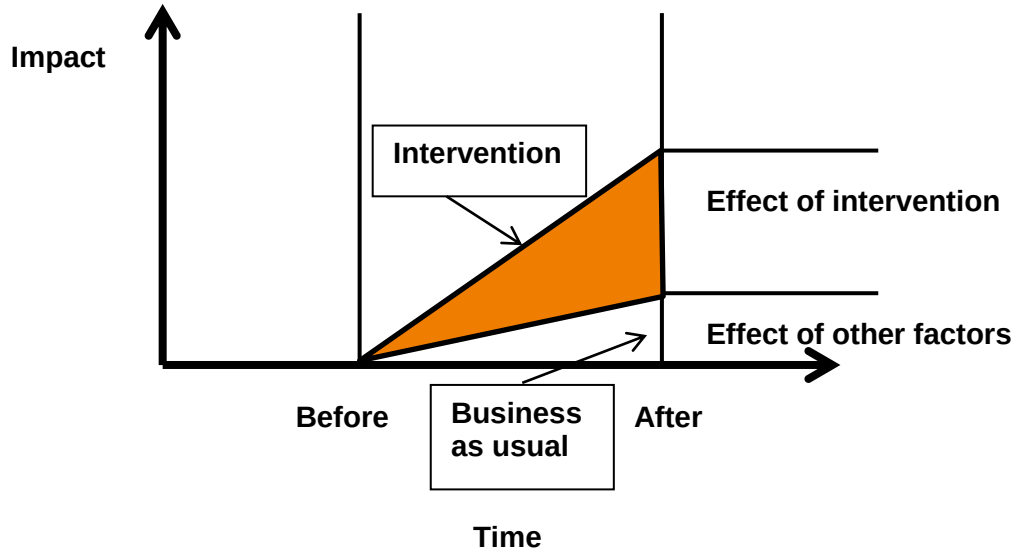
A template for this table is available in Template File 11a of the Evaluation Process Guidelines on the STTRIDE web site.

3.8.2 Define the reference case

The reference case, or baseline, defines the existing situation without the intervention; this is sometimes known as the 'do nothing' scenario, in which no other investment in the transport network takes place. Comparison of the situation as a result of the intervention with this reference case is what determines the impact of the intervention.

Given that the impacts of new technologies will often take some time to work through into behavioural changes affecting mode use, it is important to consider whether the reference case or baseline is likely to change during this time. If so, then the reference case is actually an estimate of the likely situation if existing policies continue without the intervention or other planned transport investments take place, rather than a reflection of the 'before' situation; this is often described as the 'business as usual' scenario or the 'do minimum' scenario. For example road collisions are tending to reduce across Europe as a result of sustained policies to reduce casualties; an evaluation of the safety impacts of new technology would take into account these forecast changes in casualties; a downturn in the economy could also have an influence on the demand for travel and levels of car use.

Figure 3.4: Before (baseline), business as usual and after scenarios



When assessing the impacts of new technologies on mode use, the reference case that is relevant to most of the evaluation objectives is likely to be the 'before' or 'business as usual' situation in the area where the intervention is to be implemented. However in the case of an

assessment of the performance of the technologies, the reference case may be defined in industry technical standards.

Another factor to consider in defining the reference case for interventions involving new technologies is that if the intervention generates data that can be used in the evaluation, it will be necessary to obtain equivalent data from other sources for the reference case, rather than relying on the intervention itself to provide all of the required data.

3.8.3 Define measurement methods

Investigations to assess the impact of new technologies on mode use will largely be based on measurements made during real life conditions. The measurement methods will vary depending on the intervention that is being evaluated; they may include:

- Questionnaire surveys of users and other stakeholders (such as household or 'destination based' surveys, roadside surveys, surveys on public transport vehicles or at public transport stops)
- Counts (automatic or manual) of levels of use at key points in the area influenced by the intervention, using continuous or periodic monitoring
- Collection of data from the technology or service delivered in the intervention
- Collection of operational and usage data from services
- Extraction of data from secondary and 'open data' sources such as mapping, population, census, local plans and routine monitoring data on traffic and travel; this can reduce the need to collect new data and reduce costs.

The limitations of these various sources of data will need to be taken into account when defining the measurement methods to be used. For example when considering the use of secondary and 'open' data, the timeliness, sample sizes, data quality, ownership and access rights may limit the suitability of the data for the evaluation. Collaboration with owners of such data may overcome these issues and by obtaining anonymised raw data through such collaboration, some of the limitations associated with using the processed data may be overcome.

A specific example of data quality considerations may arise when assessing impacts on cycling; it is important to bear in mind the limitations of using automatic traffic counting technologies to accurately detect the number of cyclists, while the location of automatic traffic counts for vehicles should be selected to avoid sites where queues are expected to form.

When planning real life data collection, it is important to consider legal, ethical issues and safety issues and other risks associated with collecting data from individuals and organisations, and to plan for how to deal with them. Legal and ethical issues will also need to be considered when extracting data from secondary sources.

One factor to be considered is the availability of data from commercial sources. For example transport operators are often unwilling to release data on levels of use and revenues for evaluation purposes so it may be necessary to use proxy measures, observations or surveys.

There are other reasons why measurements in real life conditions are not feasible, due to the scale or nature of the measurements involved. In such cases simulation or modelling may be used if there is sufficient data available on previous patterns of traffic and travel and other secondary data.

For example it may be necessary to use a modelling approach to assess the impacts of new technologies on travel behaviour across the network if it is not feasible to collect sufficient original data on traffic and travel behaviour for this purpose.

Another case when modelling is used is for measuring changes in vehicle emissions as a result of changes in the composition or speed of traffic; an evaluation will often estimate this change on the basis of known relations between vehicle types, flows and speed.

3.8.4 Define measurement conditions

The conditions under which data collection is carried out should be controlled and as stable as possible. It is well known that weather conditions, time of day, day of week and season of the year all influence patterns of travel, mode use and traffic levels, while traffic volumes and flows may influence the performance of some interventions, such as traffic management. Thus the difference between the reference case and the intervention may vary under different measurement conditions. A 'neutral' or representative data collection period should be defined, avoiding public holidays, school holidays and the period immediately before and after these, and also large scale local events; in cases where education institutions are an important feature it is advisable to avoid exam periods as well as holiday periods.

Levels of cycling and to some extent walking can vary considerably with weather conditions, between winter and summer and between holiday and periods and normal working weeks. It is therefore recommended that where new technologies are designed to influence walking or cycling, data collection on use of these modes takes the form of continuous monitoring from automatic count sites at key points around the area expected to be affected by the intervention. If such count sites are not already established, it is advantageous to set them up as early as possible in the project – ideally at least a year before the intervention is implemented and preferably earlier. This will reduce uncertainty in interpreting the data on impacts on levels of walking and cycling after the intervention has been implemented. If it is not possible to set up automatic count sites to monitor walking and cycling, regular manual counts can be used but unless these are carried out frequently and at consistent times over an extended period, it will not be possible to ascertain with any certainty whether any changes can be attributed to the intervention rather than seasonal or weather-related factors. Established best practice for such manual counts is that they cover a 12 hour period on each day.

If a simulation model is being used, it may represent some situations more accurately than others, leading to differences in the apparent impacts of the intervention that reflect differences within the simulation model rather than differences in impacts.

It is also important to bear in mind that some of the indicators measured may have a strong correlation with parameters which describe the measurement conditions. For example travel time on a road network is closely associated with the traffic level. The measurements made will need to be carried out in a way that allows for this 'confounding factor' by taking account of variations in traffic levels in designing the data collection and analysis. This will have implications for the resources required for the evaluation.

There may also be confounding factors that need to be taken into account if a change in the transport network or services available in a neighbouring area has knock-on effects in the area where the intervention is planned.

3.8.5 Define statistical considerations and sampling

Where possible, a statistical approach should be used to define the number of measurements required to determine the impacts of the intervention in order to ensure that the appropriate quantity and quality of data is obtained in order to be able to attribute the impacts accurately. Depending on the nature of each indicator, the number of measurements required may be defined by one or more of the following: number of days, duration of measurements each day, number of units (vehicles, people), number of sites. It is recommended that a statistician is involved in this process.

For each indicator, the level of change that is expected to be brought about by the intervention should be estimated, using expert judgement and any available evaluation results from similar interventions. This is known as the 'overall definition of success'.

If a statistical approach is possible, the level of statistical confidence associated with this level of change that is acceptable or required should then be defined. For example it might be expected

that a 5% reduction in single car use on motorways in the study area could be achieved, and that it is desirable to collect enough data to be able to state that there is 95% confidence that this level of change has taken place as a result of the intervention.

However even if a statistical approach is not possible, it is still important to ensure that there is a link between evaluation objectives and the overall definition of success through the definition of success for the relevant indicators.

For individual indicators, having determined the definition of success and level of statistical confidence required, the sample size should be defined – i.e. the number of measurements that are to be taken to represent the 'population' of all possible measurements. The following considerations should be borne in mind:

- Larger samples are usually needed for questionnaire surveys than simple counts because disaggregation of responses into sub-groups during the analysis means that fewer responses of any one type are available
- Larger samples are also needed as:
    - The expected level of change becomes smaller
    - The variation between individual measurements becomes larger
    - The level of statistical accuracy becomes greater
    - The number of sites that are to be compared increases (to more than one)
- Results that are based on objective or 'hard' measurements such as automatic vehicle counts can be treated as being more 'credible' than those based on subjective or 'soft' measurements such as questionnaire surveys of reported travel patterns.
- When planning counts of traffic, cycling and walking, the sample size is defined both by the times of day and number of days in the week when data are collected and by the number of sites where count data are collected. Automatic traffic counts would usually be designed to cover periods of two weeks or more in order to take account of variations from day to day. Particularly for cycling and walking, the number of sites needs to be calculated carefully. For example, experience of interventions to encourage cycling shows that in a medium-sized town, at least 15 automatic count sites are needed before a reasonable picture of changes in cycling can be obtained.

There will however be some situations when it is not possible to define the number of measurements statistically because the 'population' from which the sample could be drawn is too small. For example if assessing the impact of an intervention on the operators providing a service, there may only be a few operators available for inclusion in the assessment.

3.8.6 Define the measurement plan

Some data will be available from operational, monitoring and routine statistical sources and 'open data' – for example traffic flows and casualty numbers; in these cases the measurement plan defines the scale, scope and timing of the measurements that will be used in the evaluation. Other data such as user acceptance, number of walk trips and use of travel interchanges will usually need to be collected specifically for the evaluation.

One factor to consider when evaluating the impact of new technologies is who owns the data, which affects who has access to it for evaluating the impacts. It may be that only one party has access to the data, and that they may need to anonymise and pre-process the data before it can be used by others in the evaluation team.

Another factor which will be important in some cases is that data may migrate from legacy systems to new systems introducing the new technology. This may affect the availability of the data, its quality and scope, which could in turn affect the extent to which it is possible to compare between the baseline and the situation after the intervention has been implemented.

The duration of the measurements will need to take account of the fact that changes in travel patterns take time. The impacts of new technologies may also take some years to become evident. People are reluctant to change to new modes of travel or new ways of arranging their travel because this involves additional effort to gather the information they need to make the switch and potentially adapt other elements of their life to make such changes possible. It may be necessary to collect data at different time intervals to capture short term and longer term impacts.

The timing of the measurements for indicators of impacts of the intervention and for the reference case should be defined carefully so that this does not introduce any bias in the assessment. For example traffic and travel patterns should be measured at similar times (of day, week and year). It is helpful to produce a chart setting out time schedule for the various phases of the evaluation.

The area in which the measurements are to take place will need to be defined carefully to ensure that the evaluation objectives are met and the anticipated outcomes and impacts are captured. For at least some interventions designed to reduce single car use on the inter-urban network, this area is likely to include the surrounding road network and routes feeding into the inter-urban network. The location(s) will need to be clearly described when specifying the measurement plan, with maps, geographic coordinates and photographs as appropriate.

Once the measurement plan has been defined, the resource plan will need to be reviewed and it may be necessary to adjust either or both plans, to ensure that the measurement plan is appropriate for the resources.

3.8.7 Identify integrity of measurement

To ensure integrity of measurement, three types of factor should be considered.

- Completeness of the coverage to ensure that the measurements cover all of the significant impacts (not just those which are easiest or cheapest to measure)
- The scope of the measurements includes all of the factors which might influence the impact of the intervention or the characteristics of the reference case
- Accidental or intended bias to measurements may occur through factors such as respondent fatigue, policy response bias (respondents or participants who wish to influence the results), and justification bias (respondents give answers – knowingly or unknowingly - which they think the interviewer will find more acceptable); another form of bias may arise if the investigation assumes that users have received the necessary information to make rational decisions and informed comments relating to the intervention, when this is not the case.

3.8.8 Summary table

The following table provides a framework for summarising the details of the assessment method as described in Section 3.5 to 3.8.7. A template is available in Template File 11a of the Evaluation Process Guidelines on the STTRIDE web site.

In a complex evaluation, it may also be helpful to create a chart of information flows informing the assessment objectives and research questions.

Table 3.8: Summary of assessment methods for each type of assessment, evaluation objective and research question

Type of assessment	Evaluat-ion objective	Research question	Indicators	Reference case	Methods of measurement	Measurement conditions	Statistical aspects			Measurement plan	Integrity of measurement
							Sampling	Statistical confidence	Definition of success		
Performance											
User acceptance											
Impact											
Socio-economic											
Financial											

3.9 Data collection and analysis

Data

The main considerations for planning data collection and analysis are summarised here. These cover:

- Collecting baseline data
- Checking and analysing baseline data
- Collecting data after the intervention has been implemented
- Checking and analysing the 'after' data and comparing with the baseline in order to answer the research questions.

The design and implementation of data collection are specialist activities which need careful specification and will often involve skills that need to be bought in from external organisations; procurement specialists will provide guidance to the evaluation team at this stage. For the detailed planning of statistical aspects of data analysis, it is recommended that a statistician is involved in the evaluation team.

3.9.1 Collect baseline data

The baseline data collection should be of a scale, scope and level of detail necessary to assess the changes resulting from the new technology in a statistically robust manner where at all possible. Thus it should provide data on all of the indicators defined in the evaluation plan, following the statistical methods and measurement conditions set out there.

It is helpful to document any special events or unusual activities that take place while the baseline data collection is taking place that may influence the comparison with the situation after the intervention; for example large scale events, road works, weather-related incidents or service disruptions.

Considering the long term nature of the impacts of some new technologies, the baseline data may also need to cover a time period that makes it possible to make predictions of the 'business as usual' scenario. This will enable the changes following the intervention which are attributable to the intervention to be separated from other changes that would have happened as a result of other trends.

As mentioned earlier, the baseline data collection will involve a combination of pulling together data from existing sources and carrying out surveys and observations designed in the evaluation plan.

A pilot test of the data collection methods is useful to identify any problems and to ensure that any questionnaires used are understood correctly before the main baseline data collection is carried out. This pilot is also an opportunity to check the quality of the data and ensure that the training and supervision for data collection are sufficient.

Ideally the baseline data collection should take place early enough to enable the data analysis to be carried out before the intervention is implemented, so that there is a chance to fill any gaps in the baseline data that become apparent during the analysis, before the implementation.

3.9.2 Analyse baseline data

The initial analysis of the baseline data is designed primarily to check that the data collection methods have been successful in obtaining the planned scale, scope and detail necessary to

provide a robust baseline, against which the data collected after implementation of the intervention can be compared. The following types of check are recommended to identify errors in data collection and recording:

- Sample sizes are as planned, ensuring statistical validity of results
- Missing data
- Composition of samples – population characteristics, vehicle characteristics etc.
- Maximum and minimum values on each indicator are within reasonable values
- Unexpected 'clusters' of values for indicators
- Range of dates and times covered.

These checks will involve both internal comparisons within the data and comparisons with relevant comparator data such as population data and automatic traffic monitoring data.

Any adjustments or estimations to fill gaps in the data should be carefully documented; this information will be needed when comparing with data collected after the intervention.

3.9.3 Collect data after the intervention has been implemented

The data collection after the intervention has been implemented should be of a scale, duration, scope and level of detail necessary to assess the changes resulting from the new technology in a statistically robust manner where at all possible. Thus it should provide data on all of the indicators defined in the evaluation plan, following the statistical methods and measurement conditions set out there. Where the data is designed for comparison with the baseline, it should be gathered under the same measurement conditions as the baseline data collection.

As for the baseline, the 'after' data collection will involve a combination of pulling together data from existing sources and carrying out the surveys and observations set out in the evaluation plan. Again, a record of any special events or unusual activities should be maintained to inform the comparisons with the baseline data.

Considering the long term nature of the impacts of some new technologies, the 'after' data collection may need to take place in more than one 'wave', spread out over time in order to capture both the immediate outcomes and the short and longer term impacts.

If any additional data collection is carried out that does not match baseline data, a pilot test of the data collection methods is useful to identify any problems and to check the quality of the data.

3.9.4 Data analysis

Before carrying out the data analysis, it is important to carry out systematic checks on the data quality, as noted for the baseline data in Section 3.9.2. Any missing data or sources of bias or error in the data that are identified should be controlled or corrected in a 'data cleaning' phase before the analysis begins; any such cleaning and correction should be documented.

Further checks should be carried out during the analysis as the results are being interpreted, to question and investigate unexpected results and to ensure that the design and analysis has not been compromised, for example by the influence of confounding factors or the spread of the sample data collected being greater than expected.

The data analysis should be designed to provide answers to the research questions and test the hypotheses posed in the evaluation plan. Targeting the analysis in this way ensures a focused investigation and avoids wasting time on interesting avenues of analysis that do not inform these questions.

Three main types of analysis are likely to be needed in STTRIDE evaluations:

- Descriptive data – primarily for performance assessment and user acceptance assessment

- Assessment of scale and nature of change – primarily for impact assessment, socio-economic evaluation and financial assessment
- Attribution of change to the intervention including statistical comparison of outputs and outcomes.

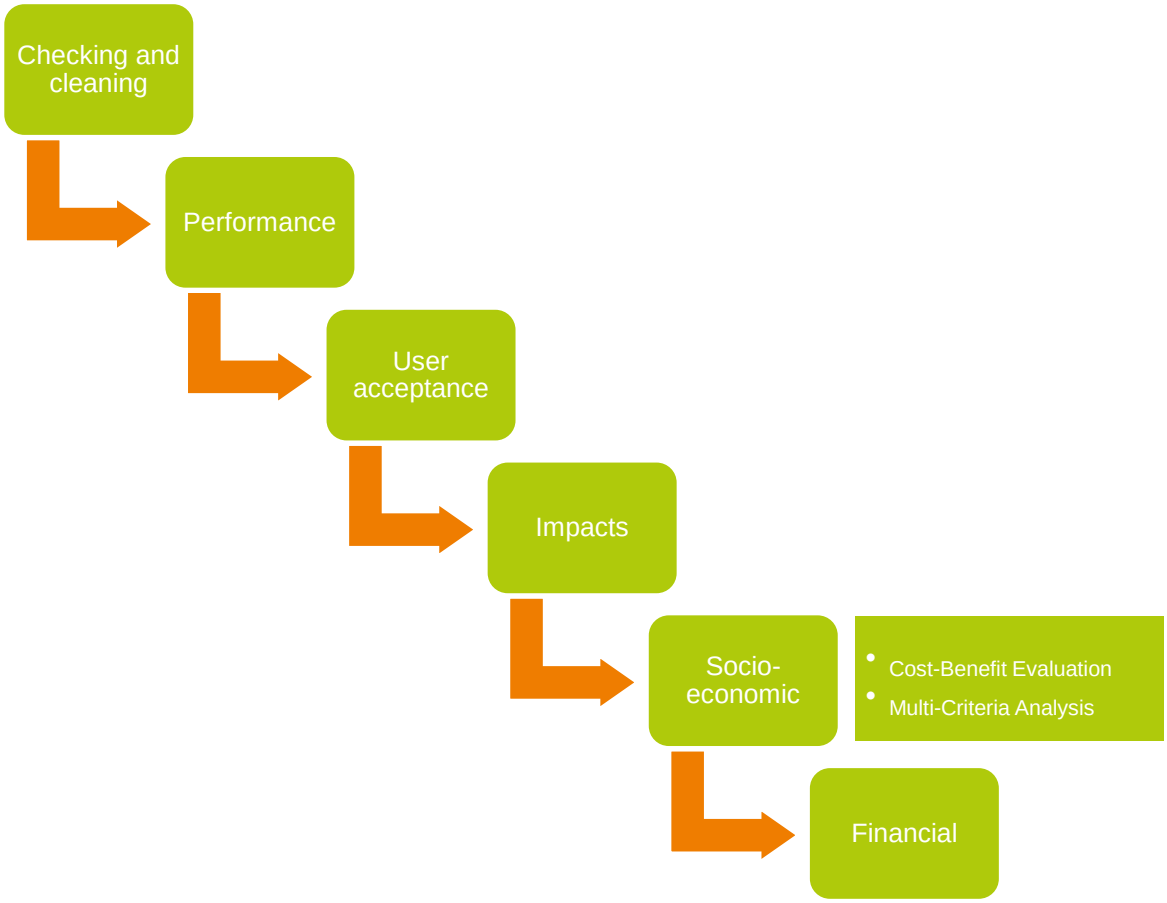
In addition, some evaluations may also need further analysis to understand:

- Mechanisms to explain impacts
- Distribution of impacts.

The statistical methods used depend on the type of data and probability distribution; it is recommended that a statistician is involved in selecting the appropriate statistical techniques and overseeing the analysis.

It is recommended that the analysis is carried out in stages corresponding with the types of assessment, because the results of each stage will inform the next. Also, if the results of the first stages do not identify significant impacts, it may not be considered worthwhile to carry out socio-economic and financial assessment.

Figure 3.5: Stages in the analysis process



Apart from the socio-economic assessment, the analysis undertaken in each stage is outlined in the research questions as set out in Section 3.6. In the case of socio-economic assessment, two different types of analysis may be carried out:

- Cost-Benefit evaluation in which the impacts are all given a monetary value and the ratio of benefits to costs is calculated over a defined future period, discounting future costs to present values using an appropriate discount rate (often defined in national guidance on appraisal of schemes). Note that different countries use their own standard values for the monetisation of impacts on factors such as safety, travel time and emissions.

- Multi-Criteria Analysis, in which some impacts (or criteria) cannot be given a monetary value but quantitative indicators can be defined and if appropriate combined into an overall index.

3.10 Report results

Reporting

Key principles are set out for reporting evaluation results.

A common approach to structuring reports of evaluation results is recommended, to enable National Road Authorities to compare results of different interventions and of similar interventions implemented in different areas.

The structure is in two parts:

- An overview aimed at policy makers and key decision makers
- Detailed results, aimed at a technical audience.

The detailed structure follows that of the evaluation framework, so that the same process can be followed for organising the results as for preparing the evaluation plan.

The following main principles should be followed when reporting results:

- Audience – decision makers and technical teams or practitioners will want different styles and levels of detail
- Transparency – reports should be easy to understand, and the context of the intervention and source of the results should be clear
- Balance – include positive and negative results, qualitative as well as quantitative impacts
- Publicity – publicise the results so that others can learn from them – through journals, conferences, networks and web sites which publish case studies.

A common structure for reporting the evaluation results for projects using the STTRIDE framework is recommended below, following the structure of the evaluation framework. In addition, consideration should be given to creating a leaflet, video or slide presentation for a non-technical audience.

A template for the headings in the report of results is available in Template 13 of the Evaluation Process Guidelines on the STTRIDE web site.

Recommended outline of report on evaluation results

1. Overview of Key Results or Executive Summary
2. Description of the problem
   - 2.1. Area
   - 2.2. Issues to be addressed
3. Description of the intervention
   - 3.1. Objectives
   - 3.2. Technologies
   - 3.3. Users

- 3.4. Stakeholders and user needs for results
- 3.5. Timing of implementation and current status

4. Description of the evaluation
   - 4.1. Area covered by the evaluation
   - 4.2. Timing and type of evaluation
   - 4.3. Intervention logic
   - 4.4. Evaluation objectives
   - 4.5. Research questions
   - 4.6. Expected impacts and impacts included in the evaluation
   - 4.7. Assessment methods – summary of methods for each evaluation objective and research question including reference case, manipulation during analysis and a summary table as set out in Section 3.8.8; refer to Technical Appendix for details.
5. Results (presented as answers to the research questions, with details of indicators, comparison with reference case, statistical significance and bias)
   - 5.1. Performance assessment
   - 5.2. User acceptance
   - 5.3. Impact assessment
   - 5.4. Socio-economic assessment
   - 5.5. Financial assessment
   - 5.6. Overall assessment of impacts
     - 5.6.1. Single-use car journeys on the inter-urban network
     - 5.6.2. Connected and multi-modal journeys
     - 5.6.3. Mode shift
     - 5.6.4. Mode shift for groups of users
     - 5.6.5. Quality and characteristics of journeys (e.g. travel time, reliability, quality, vehicle utilisation over time, waiting time, availability, distance to access service, cost)
     - 5.6.6. Environment
     - 5.6.7. Economy
     - 5.6.8. Society (e.g. safety, accessibility, health, well-being, social inclusion)
6. Comparison with other similar interventions (if available)
7. Transferability of results (summary of local issues and factors which may affect whether similar results could be achieved elsewhere)
8. Lessons learned
9. Glossary
10. Technical appendices
    - 10.1. Evaluation methods
      - 10.1.1. Evaluation objective 1, Research question 1
        - 10.1.1.1. Indicators

10.1.1.2. Reference case
10.1.1.3. Measurement methods
10.1.1.4. Measurement conditions
10.1.1.5. Statistical considerations and sampling
10.1.1.6. Measurement plan
10.1.1.7. Integrity of measurement
10.1.1.8. Data selection, cleaning, manipulation
10.1.2. Evaluation objective 1, Research question 2
(….etc. sub headings as above)
10.2. Key results for each indicator
10.3. Other technical aspects e.g. modelling

4 Contents of an evaluation plan

An Evaluation Plan

An evaluation plan is a living document, built up by the evaluation team and agreed by the stakeholders involved. It provides a single source of reference for the evaluation.

A common structure is recommended for evaluation plans for interventions using the STTRIDE framework.

A template for the headings in the Evaluation Plan is available in Template 11b of the Evaluation Process Guidelines on the STTRIDE web site. Templates for the tables and diagrams in the plan are available using the links associated with the examples in Section 3 of this document.

Recommended outline of an evaluation plan

1. Description of the intervention
    - 1.1. Objectives
    - 1.2. Technologies
    - 1.3. Area/ sites
    - 1.4. Timing of implementation
2. User needs for results
3. Description and mapping of intervention logic
    - 3.1. Expected impacts
    - 3.2. Types of change in mode use
    - 3.3. Intervention logic map
4. Definition of evaluation objectives
5. Definition of research questions
6. Scale and nature of expected outcomes and impacts
7. Selection of impacts to be evaluated
8. Assessment methods
    - 8.1. Definition of indicators
        - 8.1.1. Common core indicators
        - 8.1.2. Local indicators
    - 8.2. Definition of reference case
    - 8.3. Definition of measurement methods
    - 8.4. Definition of measurement conditions
    - 8.5. Statistical aspects
    - 8.6. Measurement plan and time schedule
    - 8.7. Integrity
9. Resources for evaluation and roles and responsibilities
10. Use of common approach to reporting results

5 Bibliography

Ball S, Tindall D, York I, Lloyd L and Torkington C (2015). Delivering smart ticketing via a managed service pilot in Norfolk – Impact evaluation methodology, baseline data and data collection strategy. TRL Report CPR 2073.

Department for Transport, Monitoring local cycle use http://webarchive.nationalarchives.gov.uk/20081022192428/http://www.dft.gov.uk/pgr/roads/tpm/tal/cyclefacilities/monitoringlocalcycleuse

Department for Transport. Monitoring Walking http://webarchive.nationalarchives.gov.uk/20081023012253/http://www.dft.gov.uk/pgr/roads/tpm/tal/walking/monitoringwalking?page=1#1001

Department for Transport (2017). Transport Analysis Guidance: WebTAG. https://www.gov.uk/guidance/transport-analysis-guidance-webtag#guidance-for-the-appraisal-practitioner

Department for Transport (2017), TAG Unit A5.1 Active mode appraisal. London. https://www.gov.uk/government/publications/webtag-tag-unit-a5-1-active-mode-appraisal-december-2017

Hills D and Junge K (2010). Guidance for transport impact evaluations. Tavistock Institute, London

Hopkin J (2015). Evaluation. Road Network Operations and Intelligent Transport Systems A Guide for Practitioners. World Road Association (PIARC)

SUSTRANS (2017). Appraising, monitoring and evaluating cycle schemes. SUSTRANS Research and Monitoring Unit's Recommendations to Highways England, SUSTRANS

Transport Infrastructure Ireland (2016). Project Appraisal Guidelines for National Roads Unit 13.0 – Pedestrian and Cyclist Facilities PE-PAG-02036. http://www.tiipublications.ie/advanced-search/results/document/?id=1468

Transport Infrastructure Ireland (2016). Project Appraisal Guidelines for National Roads Unit 2.1 – Project Appraisal Plan. PR-PAG-02011. http://www.tiipublications.ie/library/PE-PAG-02011-01.pdf

Transport Infrastructure Ireland (2017). Pilot and Trial Projects – Guidelines and Online User Manual. RE-PTP-07003. http://www.tiipublications.ie/library/RE-PTP-07003-02.pdf

Van Rooijen T and Nesterova N (2013). Applied framework for evaluation in CIVITAS Plus II. CIVITAS WIKI project – coordination, evaluation and dissemination of CIVITAS Plus II

Zhang X, Kompfner P, White C, Sexton B, Maltby D and Morello S (1998). Guidebook for assessment of transport telematics applications: updated version. European Commission Telematics Applications Programme – Transport Sector CONVERGE Project D2.3.1, Brussels.

Appendix 1 Glossary

Accountability evaluation	Aimed at demonstrating that the intervention has delivered the impacts that were anticipated in the appraisal
Appraisal	The process of defining objectives, investigating options and weighing up their costs, benefits and risks before making a decision on investment – also known as 'ex ante' evaluation
Attribution	A causal link between changes and an intervention that is credited to that intervention (rather than confounding or external factors)
Business as usual	The situation if existing policies and trends continue without the intervention
Discounting	Method of comparing costs and benefits that occur in different time periods, based on the principle that people prefer to receive goods and services now rather than in the future
Ex post evaluation	An evaluation that is carried out after in intervention has been implemented
Experimental method	A theoretical way of ascertaining the impact of an intervention by comparing two situations which are identical except that the intervention has been applied to one of them
Hypothesis	A statement linking a cause to an effect and predicting the expected direction of any change or difference
Impact	The effects of an intervention which can be seen in the long term – these may be primary or secondary, positive or negative, intended or unintended
Indicator	Parameter for qualitative or quantitative assessment that is either measured directly or derived from a measurement or simulation
Intervention	Project, scheme or programme
Intervention logic	The links between an intervention's inputs and the outputs, short term outcomes and longer term impacts on society
Knowledge-based evaluation	Aimed at increasing understanding of which interventions work in, in what circumstances, and why
Outcome	The short and medium term effects of an intervention
Output	The activities, goods and services produced by an intervention
Reference case	The existing situation without the intervention – also known as the baseline
Theory-based evaluation	Provides systematic articulation and testing of theoretical connections between an intervention and its expected impacts