

# Guidelines for Evaluating Road Safety Education Interventions



# Guidelines for Evaluating Road Safety Education Interventions



August 2004

Department for Transport

This report has been produced by Joanne Sentinella at TRL Limited, as part of a contract placed by the Department for Transport.

Department for Transport  
Great Minster House  
76 Marsham Street  
London SW1P 4DR  
Telephone 020 7944 8300  
Internet service: [www.dft.gov.uk](http://www.dft.gov.uk)

© Crown Copyright 2004.

Copyright in the typographical arrangement and design vests in the Crown.

*This publication (excluding the Royal Arms and logos) may be reproduced free of charge in any format or medium provided that it is reproduced accurately and not used in a misleading context. The material must be acknowledged as Crown copyright with the title and source of the publication specified.*

This document is available on the DfT website: [www.dft.gov.uk](http://www.dft.gov.uk)

Published by the Department for Transport.

Printed in the UK August 2004 on paper containing 80 per cent post consumer waste and 20 per cent TCF pulp.

Product code TINF937

# CONTENTS

<b>1. Introduction</b>	5
Purpose of guidelines	5
Why evaluate?	5
Integrating evaluation into programme design	6
Cost	6
How the guide is organised	7
<b>2. Terminology</b>	8
Process	8
Outcomes	8
Summative evaluation	8
Formative evaluation	9
Terms used to describe different data collection methods	9
Qualitative	9
Quantitative	10
Randomised controlled trial	10
Quasi-experimental design	10
Cross-sectional survey	10
Longitudinal survey	10
<b>3. Review past evaluations</b>	11
<b>4. Choosing an evaluator</b>	12
External evaluator	12
Internal evaluator	12
Employing an external evaluator	13
Working with children	13
Ethical issues	13

<b>5. Doing the evaluation</b>	14
1. Define the objective(s) of the evaluation	15
2. Identify your target group	16
3. Develop evaluation measures	18
Process measures	18
Outcome measures	18
4. Select methods to collect information on evaluation measures	20
Process methodology	20
Interviews	20
Focus groups	23
On-line focus groups	24
Observation	26
Document analysis	28
Questionnaires	29
Outcome methodology	29
Experimental designs	29
Surveys	33
Data collection methods	33
Suitability of different methods for different age groups	48
5. Design and test the materials	49
6. Collect the data/information	51
7. Analyse and interpret the results	52
8. Writing an evaluation report	52
<b>6. General guidance</b>	54
Working with members of the programme team	54
<b>7. Cost effective evaluations</b>	55
<b>8. Feedback</b>	57
<b>REFERENCES</b>	58
<b>APPENDIX A: Case study profiles</b>	60
<b>APPENDIX B: Pawson and Myhill’s Evaluation lessons</b>	68
<b>APPENDIX C: Useful websites</b>	72
<b>APPENDIX D: Feedback Form</b>	73

# 1. Introduction

## Purpose of guidelines

In 2000, the Government introduced a strategy to reduce the number of people killed or seriously injured in a road accident by 2010 (DETR, 2000). The document also set a number of casualty reduction targets, including a target to halve the number of children killed or seriously injured by 2010.

Education measures form part of this strategy to improve the safety of road users. It is therefore important that good practice in the delivery of road safety education is promoted. Evaluation is essential to establish whether interventions that have been implemented are effective at improving road user safety and can contribute towards Best Value Indicators set for Local Authorities. Evidence based practice is also an increasing requirement for road safety practitioners.

The outcome of an evaluation will be influenced by the techniques used to undertake such evaluations. These guidelines offer guidance on the appropriate types of evaluation and the methods to be employed when evaluating a programme.

The guidelines are based upon a critical review of recent developments in evaluation techniques and those already in use over the whole field of education, health and safety research undertaken by an expert in the field of evaluation (see Pawson and Myhill, 2001, for more details). They also include advice from evaluators who have tested the techniques on a number of innovative Road Safety Education (RSE) programmes. An outline of these evaluations can be found in Appendix A and brief 'case study' examples are given in the text. A summary of Pawson and Myhill's Evaluation Lessons is given in Appendix B.

The guidelines aim to assist road safety officers and other practitioners to conduct their own evaluations and to be better informed when commissioning an evaluation.

With this guide we hope to encourage you to conduct an evaluation of your RSE programme by providing practical advice based upon the experiences of evaluators. It should give you a general overview of evaluation and some examples of successful evaluation techniques that could be used.

## Why evaluate?

Evaluation can be used to demonstrate the success of an intervention. It can be used to find out:

- If a programme is effective;
- Why it is effective or ineffective; and
- What can be learned from the successes and mistakes that have occurred.

An evaluation of a programme can help inform policy decisions. It can also help you:

- Decide if the programme is an efficient use of resources;
- Decide if funding should continue; and
- Publicise the programme to gain additional funding and support from outside agencies, if required.

Evaluation can also be used during the development of a programme. It can identify:

- The strengths and weaknesses of a programme;
- Subsequent improvements that can be made; and
- If the materials or method of programme delivery are appropriate.

An evaluation also enables you to share information with others about the programme and its effectiveness. It can:

- Highlight which programmes are effective at improving road safety and can help meet the Government's casualty reduction targets;
- Provide evidence for use in injury prevention; and
- Offer the people taking part in the programme a chance to comment on the programme and share their experiences.

The result you would hope to achieve at the end of a thorough evaluation is a genuinely effective programme of road safety education.

## Integrating evaluation into programme design

It is a good idea to plan your evaluation into the programme design so that it is an ongoing process. This gives you a clearer idea about the aims and objectives of the programme itself and enables you to put procedures in place for the collection of data for the evaluation e.g. before and after measures, number of schools taking part. It also allows you to plan the timetable for the evaluation.

## Cost

Evaluation is often resource intensive. This guide promotes the use of cost-effective methods. It highlights the advantages and disadvantages of the various choices you need to make.

It is best practice to build the cost of an evaluation into the design of a programme. When applying for funding remember to include the cost of an evaluation in your proposal. In general, around 10% of the total programme costs (including staff time) should be budgeted for evaluation.

## How the guide is organised

Section 2 of the guide provides background information about evaluation including the terminology used and description of different types of evaluation.

Section 3 discusses how to find information about published evaluations, which can help you evaluate your own programme.

Section 4 provides guidance on choosing an evaluator to conduct the evaluation.

Section 5 is a practical step-by-step guide to doing an evaluation. It includes information on:

- Defining an evaluation objective
- Identifying the target group
- Developing evaluation measures
- Selecting evaluation methods
- Designing and testing materials
- Collecting the data/information
- Analysing and interpreting the results
- Writing an evaluation report

Section 6 provides general guidance on how to work with the programme developers and managers.

Section 7 discusses the cost-effectiveness of different evaluations.

Much of the work in this guide is based upon the publications listed in the Reference Section on page 58.

This guide is not prescriptive. You may find that other methods are more suitable for certain groups of people or to provide different types of information. Be creative and let us know what your experience has been.

## 2. Terminology

You may encounter a number of new terms when reading this guide or reports of evaluations.

### Process

The term 'process' refers to how a programme operates and is perceived by the people involved in the programme. Interviews and focus groups are usually used to gain information on programme activities and people's perceptions of the programme.

Process evaluation is concerned with:

“seeking to understand what the programme actually does to bring about change.”  
(Pawson and Myhill, 2001: 9)

An evaluation that only focused upon processes could examine various practices and opinion among participants to identify and develop best practice from their viewpoint.

### Outcomes

Outcomes are the changes that result from the programme. They are often related to the programme goals. An example of an outcome would be a change in behaviour as a result of the programme. Measurable indicators are required to examine the extent to which the programme meets its objectives. Questionnaires, quizzes or tests are often used to collect information on indicators in RSE programmes.

Sometimes the word 'impact' is used to describe changes in attitudes, knowledge, skills or behaviour and 'outcome' is limited to changes in casualty rates or accidents (longer term impacts). In this guide they are all described as outcomes.

An evaluation that focuses on outcome measures will inform you whether the programme works but not how or why. An evaluation of process measures can be used to answer these questions.

### Summative evaluation

A summative evaluation is carried out to examine the extent to which a programme meets its stated objectives. It assesses the outcomes of the programme to judge its effectiveness. Often, policy makers or funders of a programme will commission this type of evaluation to decide whether or not a programme should continue.

The evaluator is independent in this type of evaluation and will not provide feedback to the programme staff during the evaluation. Normally, a formal evaluation report will be the only output from the evaluation. The evaluation will only indirectly affect the delivery of the programme.

Data collection is focused on the implementation of the programme and outcome measures. Quantitative experimental designs tend to be favoured in this type of evaluation (see Section 5.5). An important part of the evaluation may be to understand why the programme is or is not working. In this case the processes will also be investigated.

## Formative evaluation

This type of evaluation is carried out during the development, or redevelopment, of a programme. It is done to give feedback to people who are trying to improve something.

The aim of a formative evaluation is to identify strengths and weaknesses of programme design and implementation. It investigates whether anything needs to be done to improve the programme.

Formative evaluation differs from summative evaluations in that feedback is provided throughout the evaluation by the evaluator (the person carrying out the evaluation) to the programme developers and programme managers. This will often result in changes being made to the programme during the evaluation to address problems as they arise. The role of evaluator is more interactive than in a summative evaluation.

The emphasis in this type of evaluation is on programme processes. The evaluator seeks to understand how the programme actually operates and gain an understanding of:

“why certain things are happening, how the parts of the programme fit together, and how people perceive the programme” (Patton, 1986: 139).

This understanding should enable the evaluator to identify which activities are more successful in reaching the programme goals. This type of evaluation can also be used to clarify goals and identify programme outcomes.

The study of process is not limited to formative evaluation, nor are outcomes limited to summative evaluations. In both types of evaluation, process and/or outcome measures can be examined.

## Terms used to describe different data collection methods

An evaluation will involve the collection of data either using qualitative or quantitative research methods.

### *Qualitative*

Qualitative methods ask open-ended questions and include interviews, focus groups, observation and document analysis. The data collected are non-numerical and related to categories.

## *Quantitative*

Quantitative methods collect numerical data that can be used in statistical analysis. Data will be collected as an integral part of an overall research design: normally an experimental research design or survey.

### *Randomised controlled trial*

In a randomised control trial, each participant (individual, school, community) is randomly assigned to an experimental or control group.

### *Quasi-experimental design*

In this experimental design, the experimental and control groups are matched on the characteristics that may be expected to produce a difference in the effects of the intervention. The matching process should ensure that the overall distribution of variables is equivalent within each group.

### *Cross-sectional survey*

A cross-sectional survey involves collecting data from a group of people (a sample of a chosen population) at one point in time.

### *Longitudinal survey*

In a longitudinal survey data will be collected from the same group of people more than once.

### 3. Review past evaluations

Before carrying out an evaluation of a programme it is worth examining whether similar programmes have been evaluated in the past. This can provide an insight into how you might need to evaluate your programme. It may also provide copies of the questionnaires or other materials used to collect data. Potential sources of reports, articles or systematic reviews of RSE interventions are listed here and also in Appendix C:

- Department for Transport (DfT) Road Safety: <http://www.dft.gov.uk>  
Click on Road Safety > Research > Road safety research reports
- Scottish Executive: <http://www.scotland.gov.uk>  
Click on Publications > Catalogue by topic > Transport
- TRL Information Centre: <http://www.trl.co.uk/>  
Click on Publications
- Royal Society for the Prevention of Accidents (RoSPA):  
<http://www.rospace.com/CMS/index.asp> Click on Road Safety > Road Safety Education
- The AA Foundation for Road Safety Research:  
<http://www.aanewsroom.com/aafoundation/reports.cfm>
- Injury Prevention Journal: <http://ip.bmjournals.com> (other health journals can be searched through the Medline link)
- The Cochrane Collaboration: <http://www.cochrane.org/>

If you are implementing a programme that has already been evaluated you may only need to examine whether your programme works as effectively as those previously evaluated.

## 4. Choosing an evaluator

An external evaluator or an internal evaluator can be used to carry out the evaluation. Both will employ the same methods.

### External evaluator

An external evaluator is an independent consultant commissioned to undertake the evaluation. They are independent of the organisation and can often provide a new perspective on the programme. In some cases an external evaluator can act as a facilitator, helping to co-ordinate and train internal evaluators conducting their own evaluation.

### Internal evaluator

Evaluations can be conducted internally by a member of staff, for example a Road Safety Officer. It is not necessarily inappropriate to conduct the evaluation yourself. Other internal evaluators could include 'participants', such as teachers or pupils.

In practice, the decision on who should evaluate a programme will often be pragmatic, dictated by the funding available. Employing an external evaluator can be expensive. There may be insufficient funds available to employ an external evaluator but the resources of a member of staff can be provided in kind to conduct the evaluation.

For the results of the evaluation to be taken seriously, the evaluation needs to be unbiased. It is helpful to use an evaluator who:

- Has not been directly involved with the development or running of the programme;
- Is impartial about the results;
- Will not be pressurised to find certain findings;
- Will report all the findings and not gloss over or fail to report negatives; and
- Can communicate effectively with key personnel (Clarke, 1999).

If the programme team is evaluating itself, steps should be taken to minimise potential biases, for example by issuing self-completion questionnaires that are returned anonymously by post instead of carrying out interviews with other team members or participants.

Whoever is chosen to evaluate the programme needs to be competent and reliable. The evaluator must have the respect of the evaluated staff because a successful evaluation needs the co-operation of all participants.

## Employing an external evaluator

If an external evaluator is used, check that they have relevant experience in using the correct methods (such as focus groups and/or experimental design) and in evaluating programmes similar to yours. If the programme is aimed at children, the evaluator should also have experience of working with children (see below). The evaluator should consider the budget available and other constraints when designing the evaluation.

The role of the evaluator should be clearly defined in the terms of reference. The Social Research Association has produced advice for commissioners of research, which can be found at <http://www.the-sra.org.uk>.

## Working with children

If the evaluation involves working with children or vulnerable adults, ensure that the evaluators have undergone a police check (Disclosure) from the Criminal Records Bureau.

The evaluator needs to be sensitive to the needs and abilities of children at different developmental stages. Adults need to be careful that children's participation is voluntary and ensure children's views are given equal weight to those of other people.

## Ethical issues

Anyone carrying out research needs to make sure that their activities are ethical. It is the responsibility of the evaluator to ensure:

- People are not coerced to take part. Evaluators should obtain informed consent from participants;
- That people are treated with respect when information is collected and used;
- That people know how the information is to be used and they are happy for it to be used in that way; and
- That the report does not identify individuals (McNeish and Downie, 2002).

A code of conduct for researchers and commissioners of research has been produced by the Social Research Association and can be found at <http://www.the-sra.org.uk/Ethicals.htm> (Last accessed February 2004).

The UK Evaluation Society has also issued guidelines for good practice in evaluation ([http://www.evaluation.org.uk/Pub\\_library/Good\\_Practice.htm](http://www.evaluation.org.uk/Pub_library/Good_Practice.htm) Last accessed February 2004). This document contains advice for evaluators, commissioners of evaluations, self-evaluators and participants.

## 5. Doing the evaluation

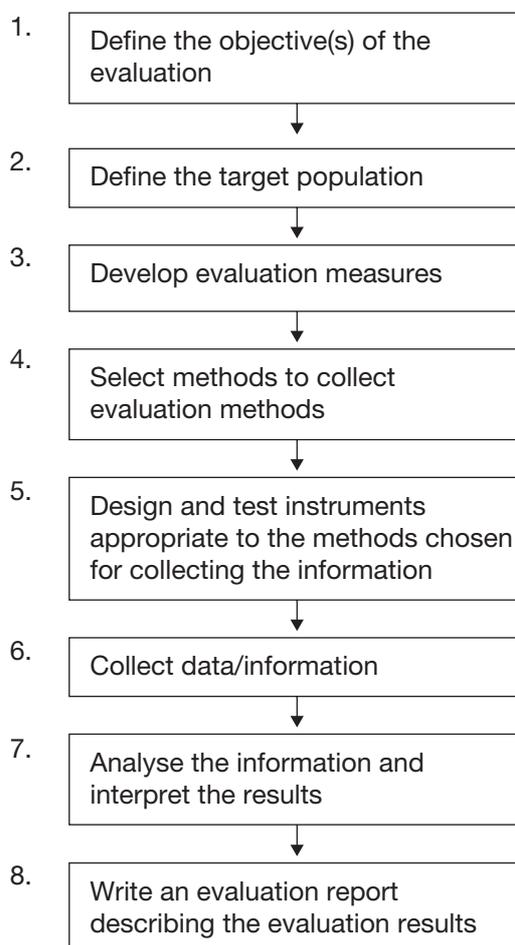
When designing an evaluation you need to consider:

- What aspects of the programme you are evaluating?
- Why are you evaluating the programme? and
- Who is the evaluation for?

With this information you can develop evaluation objectives and focus the evaluation on the questions that matter. Figure 1 shows the steps you need to follow in an evaluation.

The person commissioning the evaluation will often undertake the first three steps. The others will be the responsibility of the evaluator sometimes in collaboration with the commissioner of the evaluation.

Figure 1: Steps in an evaluation



(Adapted from Thompson and McClintock, 1998)

## 1. Define the objective(s) of the evaluation

An evaluation objective is based upon what you want to find out about the programme. The question that is usually asked in a summative evaluation is ‘Does it work?’. An evaluation of a programme’s outcomes can tell you whether it works in terms of meeting its stated objectives. Pawson and Myhill (2001), however, warn that studies, which focus exclusively on the question, ‘Does it work?’ may not be particularly useful when considering more general application of the programme. What works (or fails) in one context may be ineffective (or more effective) in another. Instead, it is recommended that evaluation research should ask the broader question:

“What is it about the programme which works, for whom, in what circumstances, and in what respects?” (Pawson and Myhill, 2001:16).

To answer this broader question you will need to consider:

- *What are the outcomes?* The programme should be assessed on a wide range of criteria and use multiple measures. Intended and unintended outcomes should be examined. An evaluation should find out if there are any unexpected benefits or problems.
- *Who does it work for?* Who is using the programme? Is the programme reaching the target group? What outcome does it produce among this group? Does it work for some groups better than others?
- *How does it work?* How is the programme being used? Is it being implemented as intended? In what circumstances is it being used?
- *Why does it work?* What are the key factors that make it work? What aspects of the programme produced the change? Was it the delivery of the programme and/or the programme content? Were there other influences?

Both the programme outcomes and processes need to be examined to answer these questions. Evaluations should use a combination of methods to explore processes within education programmes and to shed light on how they affect programme outcomes. In practice, the extent of the evaluation will be determined by the resources available (time, budget and skills).

### **Case Study: Management of Gwent's Under Sevens Organisers**

An evaluation of the Under Sevens scheme focused on its use of part-time staff who deliver road safety education specifically to the under sevens age group. The evaluation was limited to the management of the Under Sevens Organisers by the Road Safety Officers. Training and recruitment issues were also explored. By examining the strengths and weaknesses of the scheme management, recommendations for improvement and transferability could be made.

When you are considering the objective of the evaluation it is a good idea to write down the theory behind the intervention: what the intervention intended to change (its aims and objectives), how the change would happen (for example, safer behaviour through attitude change) and the anticipated changes. This information will influence the evaluation measures and research design (see steps 3 and 4).

Typical questions asked in a formative evaluation concern the effectiveness or usefulness of different materials or programme delivery techniques. The evaluation examines whether the programme activities are suitable for the target audience. It can also ask:

- When is the best time to introduce the programme to the target group? or
- How much staff training is required?
- What resources are needed to implement the programme?

All evaluation objectives should be SMART – Specific, Measurable, Agreed (by all involved), Realistic and Time-limited.

## **2. Identify your target group**

Road safety education interventions may be aimed at:

- Children;
- Adults;
- The general public; or
- Companies and/or institutions.

The target group is not necessarily the people who received the intervention but the people you wanted to reach. An evaluation can tell you whether or not you reached this group.

When studying processes the evaluation will go beyond the target group and also examine the perceptions of other participants – people who have an interest in the outcome of the programme (stakeholders). These are likely to include:

- The Government;
- The Local Authority;
- Road Safety Officers (RSOs);
- Those delivering the programme;
- Schools – head teachers, school governors;
- Class teachers;
- Parents/carers;
- Children;
- Other people living in the local community; and
- Those delivering any other road safety programmes to the same target group e.g. the police.

The time and resources available will determine the extent to which each stakeholder group can be involved in the evaluation.

#### **Case Study: Right Start Stakeholders**

Right Start is a pedestrian training programme developed by Lancashire County Council. It teaches basic pedestrian skills to infant aged children in school. The teacher introduces the skills in the classroom, which are then practised at the roadside with parent trainers and reinforced back in the classroom. Stage 1 of the programme is targeted at Reception aged children and Stage 2 Year 1 pupils.

In addition to the collection of outcome data from children, an evaluation of Right Start also gathered information from:

- Parents;
- Parent trainers;
- Area co-ordinators;
- Teachers;
- Head teachers;
- School governors; and
- Road safety officers.

This information indicated how the people involved in the programme perceived Right Start. It also showed how different people used the programme.

### 3. Develop evaluation measures

#### *Process measures*

An evaluation of the programme processes might involve:

- Assessing management of the programme, including delivery and cost efficiency;
- Assessing staffing requirements, and the training of programme staff;
- Examining how and to what extent the programme was implemented;
- Investigating to what extent the target group was reached; and
- Assessing acceptability of the programme to the target group.

Process measures include the suitability of the materials for the target group, the acceptability of the deliverers of the programme to the target group and participants' opinions and satisfaction with the programme. The way the programme is used and is received by participants can also be measured.

Information collected to monitor the progress of the programme can be used as process measures, such as:

- Number and demographics of individuals/groups (e.g. schools) taking part;
- Number of courses/events held; and
- The cost of running the programme.

#### *Outcome measures*

The overall aim of a road safety education programme may be a reduction in road casualty or accident rates. However, these indicators are unlikely to be usable as an outcome measure in local studies. Most studies of road safety education programmes have found little or no change in accident rates for a number of reasons including variable accident reporting systems, timescale and the influence of other factors. The number of accidents in a local area is likely to be too small to detect any significant differences when comparing one year with another. A large sample needs to be monitored over a long period to find a reduction in accident or casualty rates and this may not be practical for a programme that builds up slowly, a few schools at a time.

A reduction in casualty or accident rates may be anticipated from behaviour change; road users will behave more safely as a result of the RSE programme. Road Safety Education aims to change the behaviour of road users by:

- Developing safer attitudes e.g. against drink-driving;
- Improving knowledge and understanding of road safety; and
- Teaching people skills, such as cycle training or hazard awareness.

Changes in behaviour, attitudes, knowledge or skills are used as outcome measures for road safety education programmes. These measures should be specific and reflect the programme's educational objectives, for example an evaluation of a pedestrian training programme which teaches safe routes would measure children's ability to find a safe route as an outcome. An evaluation of a RSE programme should demonstrate whether the programme's educational objectives (changing knowledge, skills or attitudes) lead to safer behaviour. Safer behaviour is usually the primary outcome sought from a RSE programme and should be measured in the evaluation.

It can be difficult to establish that a change in behaviour is a result of your programme and not other factors. Influencing knowledge, skills or attitudes does not necessarily lead to a change in behaviour. Other factors should be acknowledged. Other road safety programmes, run nationally and locally, such as engineering or enforcement measures might influence the outcomes. It is unusual for a road safety education programme to operate in isolation. Schools and parents might teach other road safety lessons to children in addition to the programme being investigated. A control group is unlikely to receive no road safety education.

The results of the evaluation should be assessed carefully. The methods selected will invariably follow very closely the instruction given during the intervention. This means that the evaluation may assess the participant's ability to complete the task rather than demonstrate a clear change in knowledge or behaviour.

Multiple outcome measures will increase the reliability of the findings and should be used to overcome these difficulties.

Outcome measures should normally be measured against a baseline. A baseline is the existing level of safe behaviour, attitudes, knowledge or skills before the intervention is implemented. The amount of change after the intervention is delivered is measured against this baseline. Baseline information can also include local context data to describe what the conditions are like in the area where the intervention is being implemented, such as the demographics of the area, type of environment and engineering or enforcement measures in place. Baseline data may already have been collected for another purpose and should be considered before collecting new data.

**Key points** about evaluation measures:

Process measures

- Measures how the programme works in practice
- Measures acceptability of the programme
- Monitoring data can be used

Outcome measures

- Accident or injury rates are unlikely to be useful in local studies
- Safer behaviour should normally be measured as the primary outcome
- Changes in attitude, knowledge or skills should be measured to demonstrate the programme's educational objectives that lead to safer behaviour
- Measures should be specific and reflect the project's educational objectives
- Other factors which may influence behaviour should be acknowledged
- Multiple measures should be used to increase reliability of the findings
- Outcome measures should be measured against a baseline

## 4. Select methods to collect information on evaluation measures

The methods selected should balance what is most desirable with what is feasible within the timescale and resources available. The scale of the evaluation should also be in proportion to the size of the programme.

### Process methodology

Qualitative methods are usually used to collect information on process measures. These methods involve the collection and analysis of narrative rather than numbers. They include interviews, focus groups, observation and document analysis. Quantitative approaches such as questionnaires may also be used to collect process measures.

#### *Interviews*

In-depth interviews are often carried out with participants to explore their views of the programme and how they used it. They use open-ended questions, which do not have any pre-coded response categories, to generate as much information as possible. Open-ended questions are used because they encourage a fuller response and cannot be answered with a 'yes' or 'no'. Generally, they ask:

- What?
- How?
- Why?

Simply removing the response categories does not make a closed question open, the phrasing of a question is also important. The response should not be implied in the question. An example of the phrasing of a closed and an open question are given below:

- How satisfied would you say you are with the programme? (closed)
- How do you feel about the programme? (open).

In-depth interviews allow the interviewer to increase their understanding of the issues involved by probing each interviewee about his or her answers. A topic guide is used to steer the discussion and ensure the major interests are covered in each interview. However, it must not be so restrictive to discourage the interviewee from raising issues that may not have occurred to the interviewer.

Unstructured interviews aim to resemble a natural conversation; questions and follow-up probes are generated during the interview itself. A semi-structured interview technique may be more useful for a less experienced interviewer as it asks standardised questions about demographics and pre-determined open-ended questions. The interviewer can vary the order and phrasing of the questions and probe for more information.

In-depth interviews are time consuming but only require a small sample to gain a range of views. Interviews can be carried out by telephone or face-to-face. Telephone interviews are less time consuming than face-to-face interviews and generally have high response rates.

Interviews can be tape-recorded and transcribed verbatim. Transcription is time consuming and expensive but provides an accurate record of the conversation that can be used in the analysis. It also allows the interviewer to concentrate on what is being said. A tape-recorder can be off-putting or distracting to some interviewees so permission should always be sought before using it. If it is used it is still a good idea to make notes during the interview as a backup. Selective note taking may be sufficient to generate data or the tape could also be analysed directly by replaying the interview and taking notes. These methods are less time consuming and cheaper than transcription analysis but do risk missing data out.

#### **Case Study: Interviews with Gwent's Under Sevens Organisers (USOs)**

Both the telephone and face-to-face interviews yielded good detail. USOs seemed relaxed and happy to discuss their ideas and concerns using either approach.

The main value of the face-to-face sessions was that it seemed to allow a more free ranging discussion. This meant that items not on the topic guide were more likely to be discussed. The topic guide for the telephone interviews was therefore slightly modified using information gained from the initial face-to-face interviews.

Interviews are useful at the beginning of a study to explore the issues and develop survey materials. Interviews with programme staff and participants enable the evaluator to get a feel for the programme and their views about the programme. They are also useful with groups who have difficulty reading or writing English.

An interview can be biased if the interviewee views the evaluation as an assessment rather than a learning experience. The interviewee may have a vested interest in a positive outcome for the evaluation and feel they have to provide the ‘right’ answers. They may also be reluctant to discuss areas of the programme that might need improvement. It is easier to deal with these issues in a face-to-face interview, when rapport can be established relatively quickly.

**Table 1: Advantages and drawbacks of interviews**

<b>Interviews</b>	
<b>Advantages</b>	<ul style="list-style-type: none"> <li>✓ Generates a greater range and depth of response than other methods, especially if a rapport exists between interviewer and interviewee</li> <li>✓ Can raise issues of which the interviewer was previously unaware, as the topic guide is often very flexible</li> <li>✓ Small samples, if interviewed in-depth, can provide a large range of views</li> <li>✓ Flexibility of conducting interviews face-to-face or by telephone</li> <li>✓ Higher response rates than questionnaires</li> <li>✓ Valuable for developing more effective survey materials for use in an evaluation</li> <li>✓ Useful for evaluating respondents with low levels of literacy</li> <li>✓ The interviewer can rephrase the question if the interviewee does not understand it</li> </ul>
<b>Drawbacks</b>	<ul style="list-style-type: none"> <li>▼ Questions must be skilfully phrased so as to avoid leading the interviewee towards a particular response</li> <li>▼ Experience of interviewing is required to avoid using very restrictive topic guides</li> <li>▼ Interviewees may try to provide the ‘right’ answers, rather than their actual opinion</li> <li>▼ Transcription provides the most accurate results for analysis but presents a considerable cost in terms of the time required</li> <li>▼ The less structured the interview, the more difficult and time consuming it is to analyse</li> <li>▼ The less structured the interview, the more opportunity for bias to creep in to the questioning or interpretation of the answers</li> </ul>

## Focus groups

Focus groups work on the same basis as interviews but ask open-ended questions to a group of people. The advantage of a focus group is that the comments of one participant may stimulate the ideas of others. They are useful when examining the attitudes and opinions of groups.

The composition of the group should be considered carefully. To encourage discussion the members of the group should be similar, in terms of level of involvement in the programme or demographics, and regard each other as equals. If members of the group regard a participant as having 'expert' or greater knowledge this may hinder the discussion. Several focus groups should be run with different groups of people to gain information on different perspectives. In some cases a mix of participants with a range of views can make a successful group. In these cases the group dynamics should be considered carefully.

### Case Study: Right Start Parent Focus Groups

The number of parents who were Approved Trainers within the programme was limited to two in the focus groups as it was felt that other parents may be more reticent in the presence of 'experts'. The views of parents who were less involved in the programme were considered important and so the composition of the groups (6-8 parents) was designed to encourage all parents to participate fully. This strategy meant fewer groups needed to be conducted to gain a range of views and was less expensive.

Ideally each group should comprise between 6 to 8 people. A group moderator or facilitator should encourage all members of the group to take part in the discussion and maintain the focus of the group. As with interviews a topic guide should be used to ensure that the major topics are covered.

The group should be held in a location, which the participants find easily accessible, and which is comfortable and welcoming. Seating should be arranged in a circle and refreshments provided. The group should be held at a time convenient to the majority of participants and will normally last around an hour. Incentives, such as travel expenses, can be offered to encourage people to attend. The provision of childcare for parents could also be considered. It is worth inviting more people than required (say 8 to 10 people) to allow for drop-outs.

Group or class discussions with children work on the same basis. The group should comprise children of the same age or year group. Small focus groups can be conducted with children from the age of 4 (Borgers et al., 2000), although children under the age of 8 are very suggestible and have limited language skills. Interviews with pairs of friends can work well with this age group and could be used in substitute. The children's familiarity with each other can stimulate ideas and discussion.

Ideally two researchers should attend each focus group: one to facilitate the discussion and the other to take notes and monitor the tape-recorder. As with interviews the discussion can be tape-recorded and transcribed. When analysing the data, the evaluator looks for common themes within and between groups. Differences are also noted.

Focus groups are relatively inexpensive to run and can generate information fairly quickly. The analysis, however, can be time-consuming and a skilled facilitator should be used to maintain the focus of the discussion.

**Table 2: Advantages and drawbacks of focus groups**

<b>Focus groups</b>	
<b>Advantages</b>	<ul style="list-style-type: none"> <li>✓ Participants' comments often stimulate a wide variety of ideas amongst the group</li> <li>✓ Can explore the attitudes and opinions that groups have about road safety, rather than just those of individuals</li> <li>✓ For children under eight years of age, interviewing pairs of friends is successful as their familiarity stimulates ideas</li> <li>✓ Process of direct involvement can have a positive effect on how participants perceive the programme</li> <li>✓ Quick and relatively inexpensive to run compared to an experimental study</li> </ul>
<b>Drawbacks</b>	<ul style="list-style-type: none"> <li>▼ Discussion can be hindered if some participants are seen to be 'experts'</li> <li>▼ Recruitment of those who willingly volunteer can bias the discussion group by excluding those people who do not usually like to participate in groups</li> <li>▼ Requires a skilled facilitator to keep discussions on topic</li> <li>▼ Several separate focus groups are necessary to explore significantly different samples (e.g. young children and adolescents)</li> <li>▼ Peer pressure can introduce conformity to the opinions of the group, making some participants reluctant to offer their genuine views</li> <li>▼ Relatively expensive and time consuming to analyse the discussions</li> </ul>

### *On-line focus groups*

On-line focus groups use an Internet chat room as the venue for the discussion. It enables people in different locations to take part in the same discussion without travelling to a central location.

The security of Internet chat rooms involving children is especially important. The chat room should be set up specifically for the purpose immediately before the discussion, closed immediately after, and accessible only by password, which is faxed or phoned directly to the school staff member involved on the day of the discussion. This means the chances of anyone else gaining access are remote. Schools are welcome to provide a teacher to supervise the pupil while he or she is online.

A facilitator and technical support person monitor the discussion in the chat room and deal with any technical problems encountered whilst on-line. A transcript of the discussion can be printed off the facilitator's computer for analysis.

One advantage of an on-line discussion is that it conceals the personal characteristics of the interviewer removing potential interview bias. This can also be a disadvantage as participants can hide details about themselves and present an image which is not accurate. However it does provide respondents with anonymity, which can be useful when discussing a sensitive topic.

#### **Case Study: Junior Road Safety Officers (JRSOs) On-line Focus Groups**

The online focus groups for JRSOs provided useful information about the ways in which the scheme operated in different schools and different local authority areas. This method was used on a small scale in an evaluation of the Leicestershire and Newham JRSO schemes. It had a number of advantages; not least that it avoided the obvious logistical difficulties and expense of transporting school children from London and Leicester to a common location for a discussion. Because the children were not leaving the school grounds, it was easier to obtain permission from schools and parents for them to take part, and less disruptive to their school day. Use of Internet chat rooms may be a new and educational experience for some of the children, fulfilling part of the ICT curriculum. It was also exciting and enjoyable for the pupils and they were generally very well behaved.

There were also some disadvantages. Some children became distracted by the technology, particularly the function that enables participants to 'whisper' (write private messages) to each other or to choose 'faces' to appear next to their written contributions on the screen. 'Whispered' messages that do not involve the technical support person do not appear in the transcript and so not all interactions could be monitored. Other pupils took a while to grasp the way the chat room works and were slow to begin responding to questions. Some seemed to prefer and seek one-to-one interaction with the moderator or technical support person, which was distracting. Part of this problem may stem from classroom training: children are used to raising their hands when they want to say something and speaking one at a time. It may be harder for the quieter individuals to adapt to a situation where everyone is encouraged to contribute ideas at the same time – and quickly. Participants responded to questions at different speeds and therefore the conversation and the transcript was hard to follow. Most interaction took place between the moderator and individual pupils rather than among the pupils themselves. Chat rooms within schools seem to work better when the participants are classmates and so know each other quite well. A few pupils had problems with their computers crashing or losing their Internet connection repeatedly.

Nevertheless, the findings from the online focus groups were remarkably consistent with those from the face-to-face interviews, at least in regard to the kinds of activities JRSOs were involved in and how they felt about their jobs. The fact that individual answers could not be explored in depth did reduce the amount of detail available, but researchers may be able to compensate for this somewhat by running more online focus groups.

**Table 3: Advantages and drawbacks of on-line focus groups**

<b>On-line focus groups</b>	
<b>Advantages</b>	<ul style="list-style-type: none"> <li>✓ Using the internet can overcome the expense and logistical complications of participants from different areas travelling to a single location</li> <li>✓ On-screen displays lessen the influence of the interviewer’s personal characteristics</li> <li>✓ Transcripts of the discussion can be automatically printed to vastly reduce the time needed for analysis</li> <li>✓ Novelty of the experience will, for some, stimulate both participation and further IT skills</li> <li>✓ Relative ease of administration allows several groups to be run in quick succession</li> </ul>
<b>Drawbacks</b>	<ul style="list-style-type: none"> <li>▼ Allowing children access to on-line discussion groups requires careful consideration of all the security implications</li> <li>▼ Anonymity of method can allow participants to create a false impression of themselves and their views</li> <li>▼ Technology may overshadow the purpose of the discussion</li> <li>▼ Direct, secret messages between participants cannot always be monitored</li> <li>▼ Restricted participation for those who cannot type quickly (or at all)</li> <li>▼ Discussions can be slow to begin</li> <li>▼ Content can be difficult to follow due to the variation in the speed of participants’ responses</li> <li>▼ Requires a skilled facilitator to keep discussions on topic</li> <li>▼ Participants (particularly younger groups) sometimes see the facilitator as an authority figure to whom they should direct their comments, rather than interacting as a group</li> <li>▼ Technical problems can interrupt the discussion</li> <li>▼ Difficult to explore individual comments in more detail</li> </ul>

### *Observation*

Observation of programme activities is often used to assess the delivery of the programme. Participant observation involves a member of the evaluation team participating in the activity. This enables them to view the activity from a participant’s perspective. It also provides an opportunity to gain informal opinions of the programme from other participants. The observer keeps a note of his or her experiences and observations of how participants interact with each other. Participant observers can be unobtrusive and should not affect the running of the programme.

Non-participant observation involves the video recording or observation of activities by an evaluator. The disadvantage of this type of observation is that the presence of the observer may change the way people behave and so not reflect normal activities.

Observation can be carried out covertly without telling the participants the identity of the observer or using video. Whilst this should avoid affecting the activity observed it may be considered unethical and therefore unacceptable in some circumstances. Some groups such as teachers and pupils are used to being observed by other adults and observation is often used successfully. The observer can use a checklist to record whether the programme is being delivered as intended.

Observation methods are particularly useful when working with young children or people who have difficulty communicating.

#### **Case Study: 'The Walk' Drama Workshops**

'The Walk' is a drama based programme aimed at Year 5 pupils. The programme provides training for teachers in drama techniques, a resource pack and support from a dramatist, with the aim of producing a play with the pupils based upon walking safely to school. Observation was used extensively in a process evaluation of the programme:

- The training course for teachers was observed to evaluate its effectiveness in engaging and motivating teachers, in teaching drama and improvisation skills and in equipping teachers to undertake the project
- Observations were carried out in schools of the in-class support provided by the dramatist working with the pupils alongside the teacher. These observations sought to assess the support given by the dramatist, the role of the teacher in implementing the project, the response of the pupils, the effectiveness of the medium of drama, and the impact of the messages about walking to school and road safety
- Performances of 'The Walk' by pupils for other children and parents were also observed. The observations sought to assess the impact of the road safety messages on children and parents, the impact of the project and the extent of any associated road safety initiatives.

Table 4: Advantages and drawbacks of observation

Observation	
<b>Advantages</b>	<ul style="list-style-type: none"> <li>✓ Effective for samples who have difficulty directly communicating their views</li> <li>✓ Can assess the delivery of a programme as it happens</li> <li>✓ Using observers who are also participants, it is possible to gain a participant’s perspective</li> <li>✓ Using observers who are also participants, it is possible to gain the informal opinions of the group</li> <li>✓ Using observers who are also participants, it is possible to gain insight into the dynamics and characteristics of the groups receiving the RSE programme</li> <li>✓ Covert observation or video recording avoids the problem of influencing the normal activity of a group</li> </ul>
<b>Drawbacks</b>	<ul style="list-style-type: none"> <li>▼ Open observation or video recording can influence the normal activity of a group</li> <li>▼ Covert observation may be regarded as unethical in some circumstances (e.g. observing illegal activity).</li> </ul>

### *Document analysis*

It is useful to consider existing records before collecting data from scratch. Existing records are often a good source of information on the number of people participating in a programme or the budget, for example. These may be routinely collected to monitor the programme and will be readily available. It can highlight if any baseline data already exists.

Other documents related to the programme such as the minutes of meetings and correspondence can also be studied. These give an insight into decisions that affected the programme. The documents may provide information on how the programme developers intended the programme to be implemented and how and why these intentions were modified during and after the development. The content of materials produced for the programme can be assessed for their suitability for the target group.

Table 5: Advantages and drawbacks of document analysis

Document analysis	
<b>Advantages</b>	<ul style="list-style-type: none"> <li>✓ Using existing data sources saves time and cost of data collection</li> <li>✓ Can be used to collect baseline data</li> <li>✓ Can provide an insight into the programme and the decisions that have been made</li> </ul>
<b>Drawbacks</b>	<ul style="list-style-type: none"> <li>▼ Data may have been collected for another purpose so may not be in the format required or in sufficient detail</li> </ul>

## Questionnaires

User satisfaction surveys are used to collect information from programme participants about their opinions and views of the programme. These surveys usually involve the distribution of self-completion questionnaires, which ask participants about their opinions of the programme and measure the level of satisfaction. Questionnaires can be paper or computer based. An interviewer, on the telephone or in person, can also administer the questionnaire. Using an interviewer will be more expensive but is likely to get a better response rate than a postal survey. A postal survey, however, will provide greater anonymity.

Questions on process measures can be included in the same questionnaire used to collect outcome measures. An example would be to ask pupils what road safety activities they took part in at school. For a summary of the key advantages and disadvantages of using questionnaires, please refer to the next section on Data Collection Methods.

## Outcome methodology

Quantitative methods collect data expressed in terms of numbers. They are used to examine whether interventions have had any detectable effect. Data will usually be collected within some well-defined research design: the style of research design will depend on the nature of the intervention. It will also depend on whether the evaluation is intended as an independent measure of its effectiveness or alternatively that the intervention is performing as effectively or better than was shown in some previous evaluation.

## Experimental designs

Randomised control trials or quasi-experimental research designs should ideally be used to evaluate safety education interventions.

### *Randomised control trial*

In a randomised control trial, each participant (individual, school, community) is randomly assigned to an experimental or control group. Each participant has an equal chance of being allocated to either group. Random assignment should remove differences between experimental and control groups by distributing outside influences equally between the two. It enables the researcher to conclude that any statistically significant differences that arise between the groups following an intervention are due to the intervention and not to pre-existing systematic differences between those receiving the intervention and the control sample.

In practice this method is unlikely to be feasible for a road safety education programme. Groups, schools and/or individuals usually choose whether or not they want to take part in the programme, they self-select, and will have reasons for choosing either way. It is not likely that you will be allowed to assign the sample randomly between the treated (experimental) and untreated (control) groups.

### *Quasi-experimental design*

Quasi-experimental design is used when random allocation is not possible.

You can try to reduce differences between groups by sampling two matched groups, an experimental group and a control group. This involves matching the groups on the characteristics that may be expected to produce a difference in the effects of the intervention. Matching can be done at an individual or aggregate (group) level. Aggregate level matching is used more often because it is cheaper than matching individuals. It should ensure that the overall distribution of variables is equivalent within each group. In a small evaluation participants could be matched at an individual level one for one. Typical matching criteria for a children's road safety education programme include:

Aggregate level (e.g. school)

- Age
- Sex
- Ethnicity
- Proportion of children with English as an additional language
- Deprivation indicators (eligibility for free school meals, unauthorised absence)
- Academic ability (published school performance data, statement of special educational needs (SEN))
- Area.

Variables are selected because they are likely to produce a difference in the outcome measure and/or have been associated with increased risk of a child being involved in a road accident.

In practice, matching experimental and control groups is not always easy. If the programme is well established its coverage may be so extensive that the number of participants who can act as controls is limited. Some participants may not want or be able to take part in the evaluation, especially if they are only to act as controls. This means that compromises may need to be made in the matching process. It is also possible that the experimental and control groups may start the evaluation as equivalent groups but will not remain the same. Participants may drop out of the programme or withdraw from the evaluation. When using quasi-experimental designs evaluators need to note how similar the groups are and also describe how they are not similar. Some differences between the experimental and control groups can be allowed for in the statistical analysis. However, these allowances reduce the precision of the analysis. Using a larger sample may overcome this loss of precision but will increase the cost of the evaluation.

The sample size required depends on the level of accuracy required and the amount of variation in the sample. It is harder to find a statistical difference between two groups in a small sample. It is a good rule to aim for a sample that is at least five times greater than the number of cells in the most elaborate table you are likely to produce. For example, if you are carrying out a controlled before and after study and one or more of your closed questions has five responses then you need a sample of  $(2 \times 2 \times 5) \times 5 = 100$ . If you wanted to look at sex differences and three age groups as well this increases the minimum sample by  $(2 \times 3)$  to give a total of 600 children. For a survey you should also consider the level of non-response you may encounter and increase your sample size accordingly.

It is advisable to seek advice on a suitable sample size from a statistician. Local Authorities often employ statisticians or you may find a local University can offer assistance. Information on quantitative sampling can be found in de Vaus (1996), Kalton (1983), Florey (1993) or Campbell et al. (1995). The latter article is available on the British Medical Journal (BMJ) website <http://www.bmj.com>.

### *Controlled before and after study*

In a before and after study involving both experimental and control groups, both groups are tested before the intervention is given (pre-test) and again soon after it has been delivered (post-test). The before measure is used to obtain a baseline measure of skill (knowledge, behaviour or attitude) and to demonstrate the equivalence of the groups before the intervention. Analysis of the after data should then show whether there has been any change in both groups and whether the change in the experimental group is significantly different from the change in the control group.

A second post-test measure can also be taken some time after the intervention has been delivered (for example 3 months). This will provide a more accurate measure of the maturation effect and demonstrate whether the change is maintained over time. However, the increased amount of data collection will be reflected in the cost of the evaluation.

### *Controlled after study*

In this design no pre-test measures are collected from the experimental and control group and so the comparability of the untreated groups cannot be demonstrated. Where the evaluator is confident that there are no significant differences between the groups this approach has the advantage that it is cheaper and less time consuming because it requires less data collection.

### *Uncontrolled before and after studies*

This design is useful where you want to check that your administration of an intervention is as effective as that of the programme developers or of other Road Safety Units. Assuming that the developers or other units have conducted a satisfactory controlled test, you need only copy their methods but test an experimental sample only. You should compare your results with the reported results and if they are not significantly different then your administration of the intervention is as effective as the published administration.

This sort of study, copying a developed evaluation method, saves money for both the development of evaluation tools and measurement and analysis time. It also avoids the need to involve schools or groups, who have not taken the intervention up.

#### **Key points about experimental designs:**

##### Randomised control trial:

- Evenly distributes outside differences by randomly assigning individuals to the experimental or control groups
- Randomised trials are likely to be inappropriate for most evaluations of RSE as the decision to take part is usually made by the participant

##### Quasi-experimental design:

- Controls differences that may effect the results of the intervention by matching the experimental and control groups
- Matching can take place at an individual or aggregate (group) level
- Individual matching suited to smaller groups
- Aggregate matching is more cost-effective for larger evaluations
- Finding and recruiting participants for control groups can be problematic
- Participants may opt out of the evaluation before it is complete so it is important to use a sample that is large enough to account for this
- A large sample is essential if there are many criteria to evaluate

### Key points about experimental designs: *(continued)*

#### *Controlled before and after study:*

- Measurements are taken before the intervention is delivered to obtain baseline figures and assess whether the experimental and control groups are equivalent
- The same measurements are taken soon after the intervention has been delivered to find out if there are differences between the experimental and control groups
- The same measurement can be taken again some time later to establish if there are any long term effects

#### *Controlled after study*

- Only applicable when the experimental and control groups are known to be equivalent
- Less data collection means this approach is cheaper and less time consuming

#### *Uncontrolled before and after study*

- Can be used to compare your administration of a programme against the results obtained by the developers or other implementers of the same programme
- If your results are broadly comparable, the administration was effective

## Surveys

Surveys are a systematic way of collecting numeric data, usually in the form of questionnaires. They differ from experimental designs in that data may only be collected once and there is no control group. Surveys are useful in a formative evaluation to collect baseline data on knowledge, attitudes and behaviour of the target group. Questionnaires are also useful for gathering demographic information about participants such as age, sex and ethnicity.

### Key points about surveys:

- Data is collected once and there is no control group
- Used in formative evaluations to collect baseline data on knowledge, attitudes and behaviour of the target group
- Useful for collecting demographic information about participants

## Data collection methods

Data on behaviour change should be collected in the most natural setting for the behaviour, for example observing behaviour at the roadside, but this is not always possible. Actual behaviour is difficult to measure so tasks are often set to measure specific behaviours relevant to the programme. Data on specific behaviours will be collected in practical tests, using table-top models or self-report. A variety of methods can be used to collect data on behaviour, attitudes, knowledge or skills, some of which are described below.

- *Questionnaires*

Information on attitudes and self-reported behaviour can be collected using questionnaires. Questionnaires use closed or open-ended questions to gather information.

Closed questions have pre-determined response categories, which are usually developed using the results of in-depth interviews and a pilot survey. It is important to provide responses for all possible answers to closed questions, including 'don't know' or 'not applicable'. It is also good practice to include an 'other, please specify' category. For example:

**How do you usually get to school in the morning?**  
 (For the main part of the journey. Please tick one box only)

Walk	<input type="checkbox"/>
Ride a bike	<input type="checkbox"/>
Bus	<input type="checkbox"/>
Car, van or taxi	<input type="checkbox"/>
Train	<input type="checkbox"/>
Other (please write here):	_____

Standard response categories and rating scales can also be used. An example of a closed question seeking information about self-reported behaviour is given below.

**How often do you do the following things?**  
 (Please tick one box only on each line)

	Never	Hardly ever	Sometimes	Fairly often	Very often
Run across the road without looking	<input type="checkbox"/>				
Look both ways when crossing the road	<input type="checkbox"/>				

Attitude measures can consist of a series of statements that collectively reflect the respondent's attitude. For each statement the respondent is asked to state the extent they agree or disagree with it, for example:

**How do you feel about the following statements?**  
 (Please tick one box only on each line)

	Strongly agree	Agree	Neither agree or disagree	Disagree	Strongly disagree
RSE is important	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
RSE develops essential life skills	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

The response categories can be simplified for 8-11 year old children to 'agree, not sure and disagree'.

Attitudes and opinions can also be measured using a 5 or 7 point scale between bipolar adjectives, as illustrated below:

Please tick one box on each line to show your answers to the following statement.

The course was:

Interesting							Boring
	<input type="checkbox"/>						
Easy							Difficult
	<input type="checkbox"/>						

For questions asking an opinion, it is important to emphasise that there is no right or wrong answer. Developing attitude measures is a specialised task so it is better to use an established scale wherever possible.

Ranking questions ask the respondent to rank items in a list in order of importance or preference. They are used to gain information on how certain topics are viewed in relation to others. For example:

**Please rank the following activities in order of preference**  
(1 = liked the most, 4 = liked the least)

Walk to school week	<input type="checkbox"/>
Road safety competition	<input type="checkbox"/>
Cycle training	<input type="checkbox"/>
Video about road safety	<input type="checkbox"/>

Open-ended questions do not restrict the respondent to choosing from a list of pre-coded responses. They allow the respondent to elaborate and are better at getting detailed information. Responses to open-ended questions can be difficult to categorise so coding is a time-consuming process. Questionnaires using closed questions are quicker to administer and data enter. In a self-completion questionnaire open questions should be kept to a minimum.

A quiz can be included as part of the questionnaire, to collect information on the participant's road safety knowledge.

Table 6: Advantages and drawbacks of questionnaires

<b>Questionnaires</b>	
<b>Advantages</b>	<ul style="list-style-type: none"> <li>✓ Quick, simple and relatively cheap to administer</li> <li>✓ Closed questions are easy to collect, compare and analyse the results</li> <li>✓ Open questions can be used to collect more detailed information</li> <li>✓ Can collect the views of a large number and range of participants</li> <li>✓ Anonymity of questionnaires may improve response rate for sensitive questions</li> <li>✓ Anonymity of questionnaires may improve the response from shy individuals who may not respond to other collection methods</li> <li>✓ Minority groups can be represented fairly</li> <li>✓ Analysis of coded responses is relatively straightforward and can provide powerful statistical evidence for the effectiveness of the programme</li> </ul>
<b>Drawbacks</b>	<ul style="list-style-type: none"> <li>▼ A low response rate can result in a biased sample (although this can be minimised)</li> <li>▼ Closed questions can restrict response categories and limit the depth of the evaluation</li> <li>▼ Pilot surveys are required to develop appropriate response categories for closed questions</li> <li>▼ Categorising and coding responses to open ended questions is time consuming</li> <li>▼ Development of effective questionnaires requires specialist knowledge and can incur high costs if attempted from scratch</li> <li>▼ Questionnaires can limit response rates by being perceived as 'boring'</li> <li>▼ No cues available as to the honesty of respondents' answers</li> <li>▼ Literacy skills required for responding can exclude some individuals</li> <li>▼ Lack of concentration and/or understanding in younger respondents can produce flawed data</li> </ul>

## • Quizzes

A quiz consists of a series of questions about road safety. It is easiest to analyse if it uses multiple choice, true/false statements or questions that can be answered with 'yes' or 'no'. The respondent's knowledge of road safety is measured by summing the total number of correct answers given.

Visual materials such as photographs can be used in the quiz to test the knowledge of vocabulary or ability to identify safe places. Signs, symbols and pictures can replace words to represent different response categories, for example:

✓ true

? not sure

✗ false

### **Case Study: Measuring Knowledge and Attitudes Among Children in the Leicestershire JRSO Evaluation**

A questionnaire was designed for pupils in Years 4, 5 and 6. A road safety quiz was used to obtain an index of pupils' road safety knowledge. The quiz consisted of three sections: 'True or false?' (eight items), 'Which one is the right answer?' (two items) and 'Safe and unsafe places' (10 items). Responses to all quiz items were recoded to 1=correct, 0=incorrect, and then summed. The maximum possible individual score for knowledge was 20. Some quiz questions were based directly on the content of JRSO newsletters, and others were based on general road safety guides such as the *Highway Code for Young Road Users*.

A 12-item scale, 'Thinking about road safety', was used to measure road-safety-related attitudes. For each attitude statement, the pupil was asked to indicate whether they agreed or disagreed with each statement or were not sure.

A self-completion questionnaire was considered unsuitable for Year 3 pupils so discussions were conducted with groups of 10 Year 3 pupils instead. The researcher introduced the discussion by inviting participants to talk first about safety in general, in order to assess children's awareness of road safety. The groups' first spontaneous responses were an indicator of the prominence of road safety compared with other safety issues of concern to children.

After the initial open question about general safety, aspects of road safety that had been mentioned spontaneously were explored in greater detail in order to assess children's understanding of road safety. Themes discussed with the pupils included conspicuity, cycle helmets and cycling in general, seatbelts and safe behaviour when driving, safe and unsafe places to play, and how to walk near the road and cross the road safely.

Table 7: Advantages and drawbacks of quizzes

Quizzes	
<b>Advantages</b>	<ul style="list-style-type: none"> <li>✓ Easy to collect and analyse data if response categories are multiple choice, true/false or yes/no answers</li> <li>✓ Can incorporate visual materials</li> <li>✓ Can be combined with other forms of evaluation such as questionnaires</li> </ul>
<b>Drawbacks</b>	<ul style="list-style-type: none"> <li>▼ Self-completion quizzes can be unsuitable for younger respondents who may not be able to answer on their own</li> <li>▼ Forced choice response categories can provide a misleading evaluation if respondents guess at answers</li> </ul>

## • Card tests

Card tests are useful for younger children to test their awareness of danger or understanding of safe and unsafe situations. The test involves the use of cards depicting a safe or dangerous situation in a photograph or line drawing. Photographs are used in preference to line drawings because they have higher face validity, in that the image closely resembles a real situation. The test is scored according to certain tasks such as identification of whether the situation is safe or dangerous and providing reasons for that judgement.

### Case Study: Right Start Card Test

A card test was developed for use in evaluation of the Right Start pedestrian training programme, targeting Reception and Year 1 pupils, aged 4-6. It used eight photographs, carefully posed according to the content of the Right Start programme, which were shown one at a time to the child. The child's responses to the photographs were scored according to the tasks described below.

The interviewer tells the participant that they are going to look at some photographs. Some of the photos show children doing safe things, and others show children doing things that are less safe or even dangerous.

Task 1: The interviewer shows a photo and says to the participant, "What do you think about this picture? How is the child acting?" (If necessary, prompt: safely or not safely?) The child decides whether the photo depicts safe or unsafe behaviour.

Task 2: The interviewer asks, "What is safe/dangerous about the way the child is acting?" The child explains his or her reasons for the decision. These are scored according to mentions of Right Start content on a checklist.

This test was used as an indicator of understanding rather than awareness of danger. It examined differences in the decision-making of trained and untrained children.

In addition to the test described in the case study, different card tests have been developed by Warden et al. 1997, Coppens (1986), Lewis et al., (1997) and Whitebread and Neilson (1998).

**Table 8: Advantages and drawbacks of card tests**

Card tests	
<b>Advantages</b>	<ul style="list-style-type: none"> <li>✓ Effective for younger age groups</li> <li>✓ Can be used to assess respondent’s awareness of dangers and their understanding of safe practice on the road</li> <li>✓ Photographs can be used which provide a realistic image</li> <li>✓ Can examine the decision making process respondents use</li> </ul>
<b>Drawbacks</b>	<ul style="list-style-type: none"> <li>▼ Photographs need to be carefully staged to create a realistic image of a road safety situation. They should also be taken in a setting respondents can relate to (e.g. reflect the type of housing in the area the respondents live).</li> <li>▼ Prompts are often necessary to stimulate respondents’ thinking making it harder to evaluate actual awareness of road safety</li> </ul>

• **Draw and write**

Draw and write is a qualitative technique that has been used to examine children’s understanding of ideas. It involves inviting children to draw a picture of a situation, such as a safe place to play, and describe what is happening. Content analysis is used to quantify the number of themes depicted in the pictures and stories.

This technique is useful with primary school aged pupils. The content of the pictures/stories can be compared before and after the programme has been delivered. In other studies the technique has been used to compare the content of pictures/stories of pupils who received the intervention and those who have not.

Researchers have found that responses lack variation in classes where a large amount of conferring between pupils takes place (MacGregor et al., 1998). Another potential problem would be the teacher or researcher influencing the pupils’ responses. This bias could be minimised by briefing teachers before the exercise.

### Case study: An Evaluation of the “S’Cool to be Safe: On the Road” Theatre-in-Health Education (TIHE) Project

McWhirter et al. (2002) used the draw and write technique with 9-11 year old pupils in eight schools. The pupils were asked to draw and write about someone their own age doing something risky and what they would do or say if they were there. Road safety was the most frequently depicted topic for both experienced and inexperienced pupils. The results differed from previous studies of pupils with similar age groups with more pupils depicting a ‘road safety’ category in this study.

The author explains that the technique is sensitive to the context in which it is carried out and in this study the inexperienced pupils (from the same school who had not yet taken part in the programme) knew they would be taking part in a TIHE, perceived by teachers to be related to road safety. The authors suggest that it was, perhaps, not unsurprising that pupils related the task to road safety.

Table 9: Advantages and drawbacks of draw and write

Draw and write	
<b>Advantages</b>	<ul style="list-style-type: none"><li>✓ Useful method for primary school age children</li><li>✓ Content of drawings and descriptions can be compared pre/post road safety education to measure the extent to which the programme has had an effect</li><li>✓ Content of drawings and descriptions can be compared to a control group without any educational intervention to measure the effect of the programme</li></ul>
<b>Drawbacks</b>	<ul style="list-style-type: none"><li>▼ Respondents may confer whilst completing the task limiting the variation of responses</li><li>▼ How the task is presented may bias the results (e.g. if it is associated with road safety in some way)</li></ul>

### • Story bag

Another qualitative method used to examine children’s understanding of ideas is the story bag narrative. This involves the child telling a story in response to cues or prompts pulled from ‘the bag’. The child’s narrative will be tape recorded and transcribed for analysis.

Table 10: Advantages and drawbacks of story bag

Story bag	
<b>Advantages</b>	✓ Story prompts can creatively evaluate awareness of road safety and understanding of appropriate behaviour
<b>Drawbacks</b>	▼ Transcription and content analysis of narratives can be time consuming

## • Travel diaries

Self-reported behaviour can be collected using travel diaries. These can also be used to provide a measure of exposure. From the age of 7 children should be able to complete a simple diary, recording their journey to and from school and who accompanied them on the journey. This data can also be collected as a group activity with a teacher reading out the different options and asking the children to raise their hands if they travelled that way. This method is often used with younger children. The data can then be recorded on a graph or database. Older children (aged over 10) should be able to complete an individual diary on all the journeys they make in a week.

Examples of the questions used to collect information about travel to school can be found at:

- Department for Transport School Travel website: Click on Local Transport > Local transport plan – process and initiatives > Service provision and initiatives > School travel > Guidance for schools  
[http://www.dft.gov.uk/stellent/groups/dft\\_localtrans/documents/page/dft\\_localtrans\\_504070.hcsp](http://www.dft.gov.uk/stellent/groups/dft_localtrans/documents/page/dft_localtrans_504070.hcsp)
- Young TransNet: <http://www.youngtransnet.org.uk>
- Sustrans Safe Routes to School questionnaires:  
[http://www.saferoutestoschools.org.uk/html/pub\\_info.htm](http://www.saferoutestoschools.org.uk/html/pub_info.htm). The pupil and sixth form questionnaires include questions about the route to school and the number of roads crossed.

The National Travel Survey also collects data about journeys from children using a travel diary.

Travel diaries are time consuming to complete and normally have a low response rate. Therefore, the data obtained may not be representative of the sample.

Table 11: Advantages and drawbacks of travel diaries

Travel diaries	
<b>Advantages</b>	<ul style="list-style-type: none"> <li>✓ Provides evidence of actual road safety behaviour over a specified period of time</li> <li>✓ Measures respondents' exposure to road safety situations</li> <li>✓ Travel behaviour of younger children can be measured with group polls for each day of the evaluation period</li> <li>✓ Older children (over 10) can usually complete their own diaries</li> </ul>
<b>Drawbacks</b>	<ul style="list-style-type: none"> <li>▼ Diaries are time consuming to complete and this can often mean the response rate is low</li> <li>▼ Data might not fully represent sample due to low response rate</li> </ul>

## • Observed behaviour in the road environment

Observation methods are used to assess behavioural change. The behaviour of children outside school could be observed, for example, to assess safe pedestrian behaviour. This method is used to collect data on actual behaviour in a natural setting.

Video is often used to record pedestrian behaviour in the road environment. If video cameras need to be installed outside a school, permission should be sought from the relevant Local Authority department with responsibility for the area. It is good practice to contact the school, who may want to seek the permission of parents or governors before approving the installation of cameras. The process of gaining approval from the people being observed may affect their behaviour.

Recording behaviour that has been observed can be problematic. It is often difficult to determine the age or sex of participants being observed and the observer's view may also be restricted. It should also be remembered that the people being observed have not necessarily taken part in the programme.

**Table 12: Advantages and drawbacks of observed behaviour in the road environment**

Observed behaviour in the road environment	
<b>Advantages</b>	✓ Can observe and evaluate actual behaviour of participants in a natural setting
<b>Drawbacks</b>	<ul style="list-style-type: none"> <li>▼ Observation is time consuming</li> <li>▼ Being observed may affect a participant’s behaviour, if they are aware they are being observed</li> <li>▼ Observers may encounter difficulties determining gender, age and behaviour of participants. The data recorded is subjective.</li> <li>▼ Restricted viewing angles may obscure important details</li> <li>▼ Cannot identify those who have/have not received road safety education</li> </ul>

• **Practical testing in the road environment**

Road user behaviour can also be observed in the road environment by setting participants a number of tasks to demonstrate their skills. Their performance is then judged against pre-determined criteria to obtain a score for the task. In addition to performing the task the participants could be asked to explain their choice of route for example. This is used to provide a measure of understanding. The roadside setting provides situational cues, which is useful for younger children who have difficulty abstracting ideas.

A risk assessment should be undertaken before testing. The site should be in reasonable proximity to the school and be roughly comparable between schools, covering the same specified range of features. Parental permission is needed before taking children to the roadside.

The risk assessment may highlight a number of health and safety procedures that should be followed during testing such as: pedestrian skills testing should be conducted on a one-to-one basis; the child’s hand should be held at the roadside; and both the evaluator and child should wear reflective vests. Cyclists should wear a cycle helmet during an assessment of their cycling skills, and, before testing drivers on the road, assessors should ensure that participants have a valid driving licence. The Kerbcraft pedestrian training scheme web site: [www.kerbcraft.org.uk](http://www.kerbcraft.org.uk) contains examples of risk assessments undertaken by a number of Local Authorities.

Roadside assessments are time-consuming and expensive due to the high levels of staffing required.

### Case Study: The Effectiveness of Child Cycle Training Schemes (Savill et al., 1996)

This study examined whether cycle training had an effect on the road safety knowledge and skills of children aged around twelve. It included an on-road test including various manoeuvres at a quiet 'T' junction near the school. These manoeuvres included turning left, right and overtaking a parked car. The assessor used a pre-coded form to record how well each child performed each manoeuvre and gave a safety rating (safe/unsafe) overall and for each manoeuvre.

Parental consent was obtained for each child taking part in the on-road assessment. Where possible the child rode their own bicycle or one provided by the researchers. All the children who took part were required to wear a fluorescent belt and cycle helmet, which was supplied by the researchers where necessary.

Table 13: Advantages and drawbacks of practical testing in the road environment

Observed behaviour in the road environment	
<b>Advantages</b>	<ul style="list-style-type: none"><li>✓ Tasks can be evaluated to assess skill and understanding of specific elements of the road safety programme</li><li>✓ Roadside setting provides natural situational cues, especially for younger children, making the task more realistic</li></ul>
<b>Drawbacks</b>	<ul style="list-style-type: none"><li>▼ Must assess risk and gain parental consent where necessary prior to the evaluation</li><li>▼ Staffing requirements make practical evaluations time consuming and expensive</li></ul>

### • Table-top model

Pedestrian skills can also be measured using a table-top model, which simulates the road environment. The model consists of an aerial view of a road layout and includes buildings and trees. Three-dimensional models work better than two-dimensional models. Toy cars and dolls are used to create traffic scenarios. The model should have realistic dimensions: the pavement should be at least three times wider than the feet of the doll pedestrian so it can be positioned unambiguously at either side of the pavement.

The child is asked to move a doll representing a pedestrian around the model following a story, which is used to introduce tasks. The task is scored against criteria on a score sheet. To gain a measure of understanding the child is asked to narrate the task. However, some children have difficulty doing this, especially if they are shy. The child can be probed about their actions but only generally ('what is X doing now?') to avoid leading the child. It can be difficult to assess the child's actions as they move around the model.

Ampofo-Boateng and Thomson (1991) found that their model produced comparable results to those obtained at the roadside in assessing safe places and routes. A model is less time-consuming to administer than a roadside test and allows strict comparability across different children.

#### **Case Study: Right Start Model Test**

The model test involved asking each child individually to move a doll representing themselves and one representing an adult around a board depicting the pavement, a zebra crossing, a junction with a driveway, a park and shops. During the test, children were asked to walk the dolls along the pavement to the 'cake shop' past an alleyway/driveway and narrate what they were doing. They were then asked to move the dolls to the 'park' (which involved crossing the road).

The test was designed to assess the children's:

- Understanding of pavement, kerb and road and their functions
- The need to hold hands and walk on the inside of the accompanying adult
- The need to watch out for drive and alleyway entrances
- Where to stop before crossing a road
- What to do when waiting to cross
- What to do while crossing

And for year 1 children:

- To cross straight over
- Where to cross at a zebra crossing

Researchers found scoring the model test problematic. It was difficult to establish whether the dolls were holding hands without prompting the child. The children were also distracted by the toy cars on the model. The model test may have been too abstract for some 5-6 year olds. A three-dimensional model might have been more successful.

Trained reception year children performed better on the model test than untrained reception year children. There was no difference in the scores of year 1 trained and untrained children.

Table 14: Advantages and drawbacks of table-top models

Table-top model	
<b>Advantages</b>	<ul style="list-style-type: none"> <li>✓ Cheaper and less time consuming than a roadside study</li> <li>✓ Lower risk than roadside testing</li> <li>✓ Comparable across different participants</li> <li>✓ A range of road safety behaviours can be evaluated</li> <li>✓ Training on how to use the model can improve understanding and subsequent performance in the task</li> </ul>
<b>Drawbacks</b>	<ul style="list-style-type: none"> <li>▼ Scale of model must be proportional otherwise participants' actions may seem ambiguous</li> <li>▼ Children can be reluctant to describe their actions unless prompted, which raises the danger of leading the child's response</li> <li>▼ Two-dimensional models may be too abstract for young children to apply their road safety knowledge. A three-dimensional model could be used</li> <li>▼ Can be difficult to assess the participants' actions as they navigate the model</li> <li>▼ Young children can be distracted by the presentation of the task</li> </ul>

## • Computer and video based tests

Computer and video-based tests simulate the road environment. The test normally involves the participant clicking a button, which is linked to the video or computer, in response to a task.

An example of a computer-based test is the hazard awareness test for drivers. Drivers are shown video clips recorded from the driver's view and are asked to press a button when they spot a hazard. The time taken to identify the hazard is recorded.

The advantage of a computer or video-based test compared to a roadside assessment is that a set of identical traffic scenarios can be presented to each participant and it is easy to administer. The test conditions are held constant so the results are also comparable between participants. Various factors can be controlled for in a computer or video-based test such as speed of traffic and pedestrians.

There is also a concern that the participant's ability to use the evaluation tool is measured rather than their skill in these tests. Additional time and resources are required to train the children how to use the computer package. Computer assisted data collection is particularly useful for ages 8-11 as it can help maintain their attention.

Unless a validated computer or video-based test is available, the cost of using this type of test is high due to development costs.

**Case Study: University of Strathclyde’s Basic Performance Task**

This test has been developed by researchers at University of Strathclyde to measure visual timing and gap selection. The child is presented with a number of locations and is asked to indicate when they would cross the road by clicking the computer mouse. The computer scores the test automatically.

The test was validated against roadside tests in a recent study carried out by TRL of children living in traffic-calmed areas. This study found that the results of the two tests were closely comparable when used with 7-9 year old pupils.

**Table 15: Advantages and drawbacks of computer and video based tasks**

Computer and video based tests	
<b>Advantages</b>	<ul style="list-style-type: none"> <li>✓ Easy to administer</li> <li>✓ Enables rapid data collection and analysis</li> <li>✓ Easy to replicate and repeat specific road safety scenarios and hold constant traffic speed, traffic density and other variables</li> <li>✓ Can measure exact reaction times to evaluate spontaneous awareness of road safety issues</li> <li>✓ Novelty of task can help focus younger people</li> <li>✓ Motivation can be enhanced with special graphics, music or games between test items</li> </ul>
<b>Drawbacks</b>	<ul style="list-style-type: none"> <li>▼ Should only use validated tools to ensure ease of use is not a factor in the results</li> <li>▼ Participants should be trained how to use the programmes which is costly and time consuming</li> <li>▼ Expensive to develop</li> </ul>

**Suitability of different methods for different age groups**

Table 1 gives an indication of the suitability of different methods for different age groups. Most methods will need to be adapted for different age groups. As Pawson and Myhill (2001) warn, “knowledge of the ‘age-appropriateness’ of research instruments is still at the trial and error stage” (Pawson and Myhill, 2001:27). The methods selected should be appropriate for the development stage of the participant. Further information about working with children can be found in Grieg and Taylor (1999).

Table 16: Suitability of different methods for different age groups

Method	Age Groups			
	Infant	Children Junior	Secondary	Adults
Individual interviews			✓	✓
Friendship pair interviews	✓	✓		
Focus group/group discussion		✓	✓	✓
On-line focus group		✓	✓	✓
Self-completion questionnaire		✓*	✓	✓
Interviewer questionnaire	✓	✓	✓	✓
Quiz	✓*	✓*	✓	✓
Card test	✓	✓		
Draw and write	✓	✓		
Observation	✓	✓	✓	✓
Travel diaries – individual		✓	✓	✓
Travel diaries – group	✓	✓	✓	✓
Roadside test	✓	✓	✓	✓
Model test	✓	✓		
Computer based tests		✓	✓	✓

\*May need to be administered by an interviewer or teacher, one-to-one with each child

## 5. Design and test the materials

All materials should be designed with the age of the participants in mind.

Focus groups and other qualitative methods are useful in providing information for the topics and possible response categories of questionnaires. Language and, in the case of the questionnaire, layout and font size should be chosen carefully to be as clear and straightforward as possible. Consideration should be given as to whether materials should be translated into other languages or provided in alternative formats for people with English as an additional language or a disability. A large font (14 point Arial) should be used in self-completion questionnaires and quizzes aimed at children. The length of a self-completion questionnaire should be kept to a minimum. Even for older children, a self-completion questionnaire should not exceed 8 pages.

Children can participate in the development of questionnaires, coming up with questions or helping to word or phrase a question in the language they use. Open questions at the end of a questionnaire give the child opportunity to freely respond. Advice should also be sought from teachers on the suitability of the questionnaire or materials. Methods and materials should be adapted for people with special needs, disabilities or English as an additional language.

Thorough piloting of the materials should be undertaken, preferably in the geographic location of the main study. This enables the evaluator to consider local terminology and ensure that any images are representative of the area in which they are to be used. Piloting can highlight whether children use 'path' or 'pavement'. It can also show what they understand are 'speed humps' or 'speed bumps'. A pilot study only needs to be carried out with a small number of people and can save time by highlighting problems early.

It is important to use the same language as the participants in an interview or focus group.

#### **Case Study – Piloting of the JRSO Child Materials**

The JRSO interview topic guide and questionnaire (including questions on self-reported behaviour, attitudes, knowledge and process) were piloted at eight schools in Leicestershire. The control version of the questionnaire and the group discussion topic guide were piloted in two schools in Berkshire: one suburban school and one inner-city school with a high proportion of children from non-English-speaking backgrounds. These pilot schools represented the opposite ends of the spectrum in terms of average Key Stage 2 scores, and English was not the first language for many of the children at the inner-city school. Feedback from the teachers, pupils and teaching assistants and results of the questionnaires enabled the evaluators to judge whether the language and complexity of the questions were at an appropriate level for pupils in the full range of abilities. This was an extremely important part of the design process, as the aim of the questionnaire was to provide an index of road safety knowledge and attitudes, not reading ability.

The following points should be considered when designing questions for a questionnaire:

- Question wording should be clear and unambiguous;
- Each question should only contain one idea;
- Use simple language;
- Avoid the use of double negatives; and
- Do not lead the respondent.

Questions should follow a logical sequence and use filters to enable the respondent to skip questions not relevant to them. Personal questions are normally asked at the end of a questionnaire. Children, however, are used to writing their name at the top of school exercises so they can be asked demographic questions at the beginning. If the questionnaire is completed as a class activity it is advisable to leave any open questions until the end so that pupils can complete them at their own speed.

Children will try to answer questions even if they do not understand them. It can be more difficult to spot an inaccurate answer to a closed question requiring a yes/no response than to an open question (Waterman et al., 2001). Children are more likely to say they do not understand or don't know if asked an open question whilst a closed question will prompt a response even if they think the question is 'silly'. When interviewing children open questions should be used as much as possible.

Quiz questions should not be too difficult or too easy: questions that almost everyone or no-one can answer will not allow the detection of small differences between groups.

Using a published and generally accepted test measure may compromise the precise assessment of programme objectives but allows comparison between studies. New tests and attitude measures must be tested for reliability and validity.

## 6. Collect the data/information

Multiple test measures are more reliable than using a single measure, although they can introduce a learning effect whereby completion of one test improves performance in another. To avoid a learning effect the test order should be randomised. This can be done using a Latin square, which allocates an equal number of participants to each possible test order, or by shuffling the cards in a card test.

The amount of time each child spends on a set of tests should be kept to a minimum. Ideally, this time should not exceed 30 minutes. Testing or interviewing individual or small groups of children should be carried out away from the class in a quiet room such as a library. The pupils' ordinary timetable should be obtained so that the tests can be scheduled around class activities such as Physical Education (PE).

When working with children in Special schools it is advisable to have an adult helper accompany the child at all times, especially if the evaluator is unused to working with children with behavioural difficulties.

Before administering a self-completion questionnaire or quiz to a class it is advisable to seek advice from the class teacher. The class teacher will be able to advise on the best way to approach his or her class. Year 6 pupils and more able Year 5 pupils are normally able to work through a self-completion questionnaire individually. Other Year 5 and younger pupils will need the researcher or teacher to read all the questions and response options out aloud. This can be done as a class activity or individually in a structured interview. Teaching assistants, teachers and the evaluator should be available to work with individual pupils including those with special educational needs as they complete a self-completion questionnaire or quiz.

Response rates for self-completion questionnaires can be maximised by attending the school to distribute and collect the questionnaires in person. There can be relatively poor returns from teacher questionnaires left at the schools, compared with those completed while the evaluator is present.

A covering letter should accompany a postal questionnaire, which explains the purpose of the survey. Respondents should be reassured that any information they provide is confidential and anonymous. A pre-paid self-addressed envelope should also be provided for return of the questionnaire. Issuing reminders can increase response rates.

## 7. Analyse and interpret the results

Following collection and analysis of the evaluation measures, the results will require interpretation and need to be placed into context. This means clarifying the original objectives of the evaluation and relating the findings to the theory behind the programme.

It can be difficult to establish that a change in knowledge, attitudes or behaviour is a result of your programme and not other factors. Influencing knowledge, skills or attitudes does not necessarily lead to a change in behaviour. Other factors should be acknowledged. Other road safety programmes, run nationally and locally, such as engineering or enforcement measures might influence the outcomes. It is unusual for a road safety education programme to operate in isolation. Schools and parents might teach other road safety lessons to children in addition to the programme being investigated. A control group is unlikely to receive no road safety education.

The results of the evaluation should be assessed carefully. The methods selected will invariably follow very closely the instruction given during the intervention. This means that the evaluation may evaluate the participant's ability to complete the task rather than demonstrate a clear change in knowledge or behaviour.

## 8. Writing an evaluation report

Evaluation reports can take various forms such as an article for publication in a scientific journal or a report for a committee. A written summary or presentation to a group may be all that is required. The type of report will depend upon whom it is being written for – the audience. These can include:

- The Government
- The Local Authority
- Programme managers
- Programme deliverers
- Schools
- Parents
- Other Road Safety Officers
- Academics

Several reports may need to be written from the evaluation, presenting the results in a variety of formats. The amount of background and detail that needs to be included should be considered. Often the Executive Summary is distributed and a full version made available on request. The evaluation report should be brief and concise. Typical contents of an evaluation report include:

<b>Executive Summary</b>	Summary of the report, detailing the main points from each section. This should be no longer than two sides of A4.
<b>Background</b>	Describe the background to the programme: when and why the programme was developed; who it is targeted at; how it is delivered; how long it has been running for; and the coverage of the programme. What are the aims and objectives of the programme? This section may include a short literature review of relevant material.
<b>Aims</b>	State the objectives of the evaluation. Who commissioned the evaluation and why? What is the scope of the evaluation? What are the evaluation measures?
<b>Methods</b>	How was the programme evaluated? Describe the research design and methods used to collect data; describe the materials used; sampling strategy (including sample sizes and response rates); and the methods used to analyse the data.
<b>Findings</b>	What were the results of the evaluation measures? What has been learnt? Summarise the results under headings. Where possible use tables, graphs and diagrams to illustrate the results. Include examples of what people actually said in interviews.
<b>Discussion</b>	Did the programme meet its objectives? What improvements could be made to meet them? What were the constraints of the research design or difficulties of the evaluation? How may these have influenced the results?
<b>Recommendations</b>	Recommendations for action. These should be clearly linked to the outcomes of the evaluation. This may be the only section some people read.
<b>Appendices</b>	Include copies of the evaluation materials and supplementary information relevant to the evaluation.

It is a good idea for evaluators to critically assess their research and acknowledge ways in which flaws with the research design or difficulties encountered during the evaluation may have affected the results.

It is important that whatever the outcome of the evaluation, the findings are disseminated.

## 6. General guidance

### Working with members of the programme team

Programme staff may view the evaluation as an assessment of their work and, consequently, can be quite protective of it. This is more likely to be the case in summative evaluations where contact with the evaluator is limited and this can lead to biases in the data collection. In addition, the programme team may be concerned about the evaluation damaging existing relationships with participants, which can lead to delays in gaining access to participants. Programme staff need to have confidence that the evaluators will represent the programme fairly and not generate difficulties for the future.

#### **Case Study: Right Start Testing in Schools**

On each of the days testing in Lancashire a member of the road safety team from Lancashire County Council was available to introduce the evaluators to the school. The evaluators found this helped to ensure the smooth running of the testing.

## 7. Cost effective evaluations

It is important to consider the scale of the evaluation. It is not always necessary to carry out a large, full-scale evaluation. The evaluation should be in proportion to the type of programme being implemented.

Evaluations may vary considerably in cost according to the methodology selected. Different methodologies will yield different types of information and answer different questions. Part of the evaluator's task is to choose the type of evaluation and the type of tools most appropriate for the question to be answered.

For example, the evaluation might be needed to ensure that all the participants are happy with the way an intervention is working. This could be done through interviews, focus groups or questionnaires. A simple answer could be obtained using questionnaires but this may give only a limited picture if the responses are too general and only obtained from some of the participants. A better understanding of how participants interact on the programme is likely to be obtained through focus groups, but these will be more expensive as a range of different participant types will need to be involved in the groups to give a full answer.

Or the need may be to confirm whether the programme should continue. In this case, outcome will need to be measured in a form that is acceptable to the funding agencies. It may be sufficient to demonstrate that the programme is producing general benefits for the target group. If however the need is to assess how to improve the programme, or whether increased funding could improve its effectiveness, a more detailed evaluation will probably be necessary to identify its effect for specific groups, and to identify groups for which it is not effective.

Costs may be kept low by limiting numbers of participants surveyed or tested, but the resulting conclusions will be less reliable.

Using less experienced staff to design questionnaires or conduct interviews or focus groups will also reduce costs, but while this may be sufficient to provide broad indications, or identify major shortcomings in programmes, the results are likely to be less informative, than if more expert staff are used.

One way of reducing costs is to use standard test materials. Not only does this remove the costs of developing the materials but also, potentially reduces the number of times tests need to be applied, if normalised responses are available from other previous users of the materials. For example, instead of doing before and after tests for control samples, it may be possible only to test those who have received the intervention, and compare their responses with those from other areas where the intervention (or one with similar objectives) had been implemented successfully. This approach however provides no opportunity to tailor the test materials to question the participants against the specific messages they have been given in the intervention. It may therefore need to rely on tests of general road safety knowledge and behaviour.

A different approach may be needed when implementing programmes known to be effective than when developing new types of programme in a new area. For an evaluation of a known programme it may be sufficient to examine the implementation and delivery of the programme to ensure that it is consistent with practice found effective elsewhere.

Evaluation can enhance a programme by providing information on how to improve it. It provides the evidence base for future programmes and can be used to support claims for future funding.

## 8. Feedback

We expect to review and update this guide. If you would like to comment or make any suggestions on the guidelines, please contact us using the feedback form in Appendix D.

## References

- Ampofo-Boateng, K. and Thomson, J.A., (1991) 'Children's perception of safety and danger on the road', *British Journal of Psychology*, 82, pp 487-505.
- Borgers, N., de Leeuw, E., and Hox, J., (2000) 'Children as respondents in survey research: Cognitive development and response quality', *Bulletin de Méthodologie Sociologique*, April 2000, N.66, pp 60-75.
- Campbell M.J., Julious S.A. and Altman D.G. (1995) 'Estimating Sample sizes for binary, ordered categorical, and continuous outcomes in two group comparisons.' *British Medical Journal*, 311: 1145-8.
- Coppens, N., (1986) 'Cognitive characteristics as predictors of children's understanding of safety and prevention.' *Journal of Pediatric Psychology*, 11, pp 1898-195.
- Clarke, A., with Dawson, R., (1999) *Evaluation research: An introduction to principles, methods and practice*. London: Sage.
- Florey C du V (1993) 'Sample size for beginners'. *British Medical Journal*, 306: 1181-4.
- Greig, A., and Taylor, J., (1999) *Doing research with children*. London: Sage.
- Kalton, G. (1983) *Introduction to survey sampling*. Series: Quantitative applications in the Social Sciences, No. 35, Thousand Oaks, CA.: Sage Publications.
- Lewis, V., Dunbar, G., and Hill, R., (1997) Children's knowledge of danger, attentional skills and child/parent communication: relationships with behaviour on the road. *Road Safety Research Report No. 10*. London: Department for the Environment, Transport and the Regions.
- MacGregor, A.S.T., Currie, C.E. and Wetton, N. (1998) 'Eliciting the views of children about health in schools through the use of the draw and write technique', *Health Promotion International*, Vol. 13, No. 4., pp 307-318.
- McNeish, D. and Downie, A. (2002) *Evaluation Toolkit: A practical guide to project evaluation*. Available from Banardos and Leeds Health Action Zone.
- McWhirter, J. with Pegg, C. and Charles, K. (2002) 'It makes you feel how it is'. *An evaluation study of the 'S'Cool to be Safe: On the Road' – a Theatre-in-Health Education project*. Available from the RoSPA website <http://www.rospace.com>.
- Patton, M.Q. (1986) *Utilization-focused evaluation*, 2nd edn, Newbury Park, CA: Sage.

Pawson, R. and Myhill, A (2001) Learning lessons: Enhancing evaluation through research review, TRL report 507. Crowthorne: TRL Limited.

Savill, T., Bryan-Brown, K., and Harland, G., (1996) The effectiveness of child cycle training schemes. *TRL Report 214*. Crowthorne: Transport Research Laboratory.

Thompson, N.J. and McClintock, H.O., (1998) *Demonstrating your program's worth: A primer on evaluation for programs to prevent unintentional injury*. Atlanta, Georgia: National Center for Injury Prevention and Control.

de Vaus, D.A. (1996) *Surveys in Social Research*, 4th Edition, London: UCL Press.

Warden, D., Moran, E., Gillies, J., Mayes, G., and Macleod, L., (1997) 'An evaluation of a children's safety training programme.' *Educational Psychology*, 17, 4, pp 433-448.

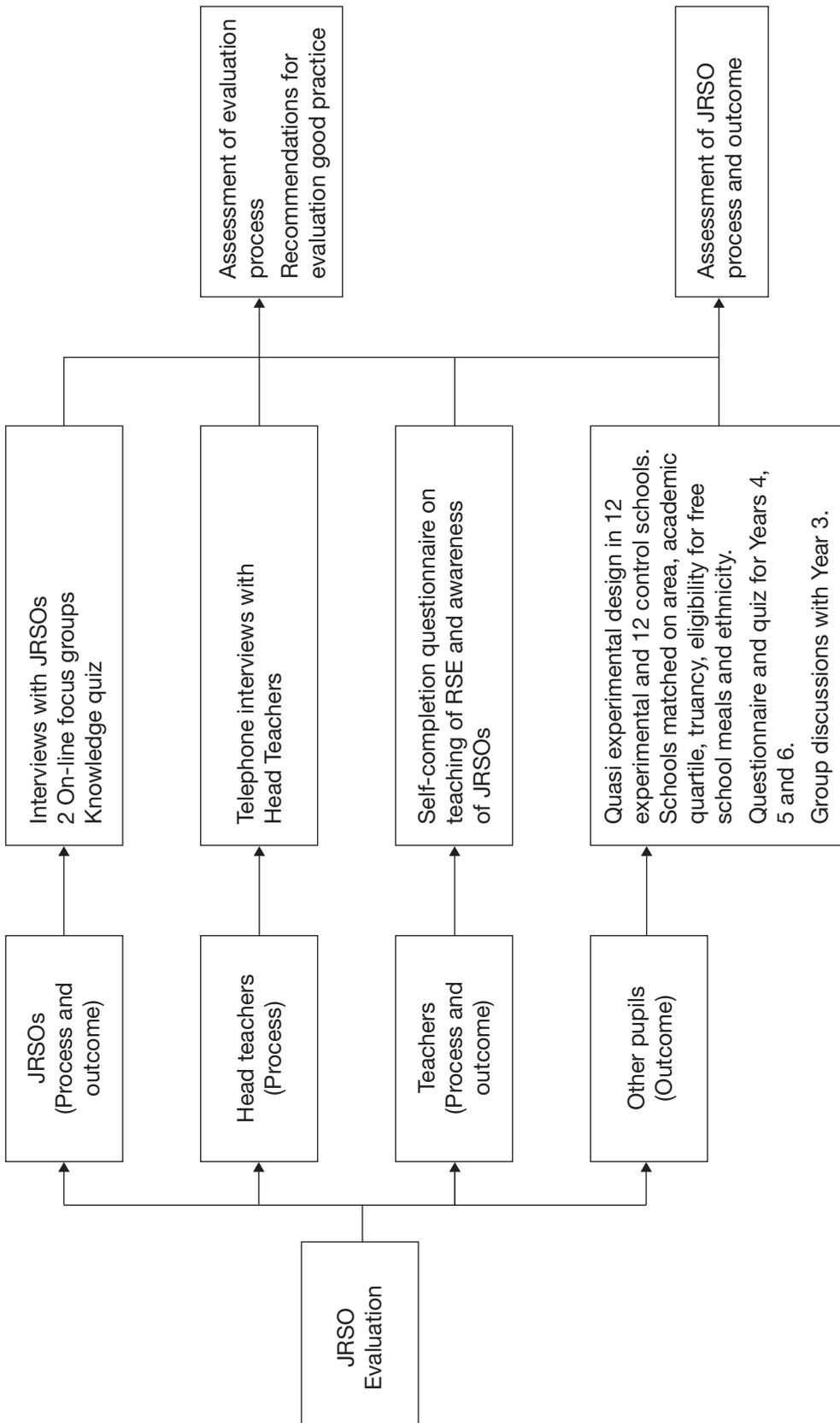
Waterman, A., Blades, M., and Spencer, C., (2001) 'Is a jumper angrier than a tree?', *The Psychologist*, Vol 14 No 9, pp 474-477.

Whitebread, D. and Neilson, K. (1998). *Cognitive and metacognitive processes underlying the development of child pedestrian skills*. London: Department for the Environment, Transport and the Regions.

## Appendix A: Case study profiles

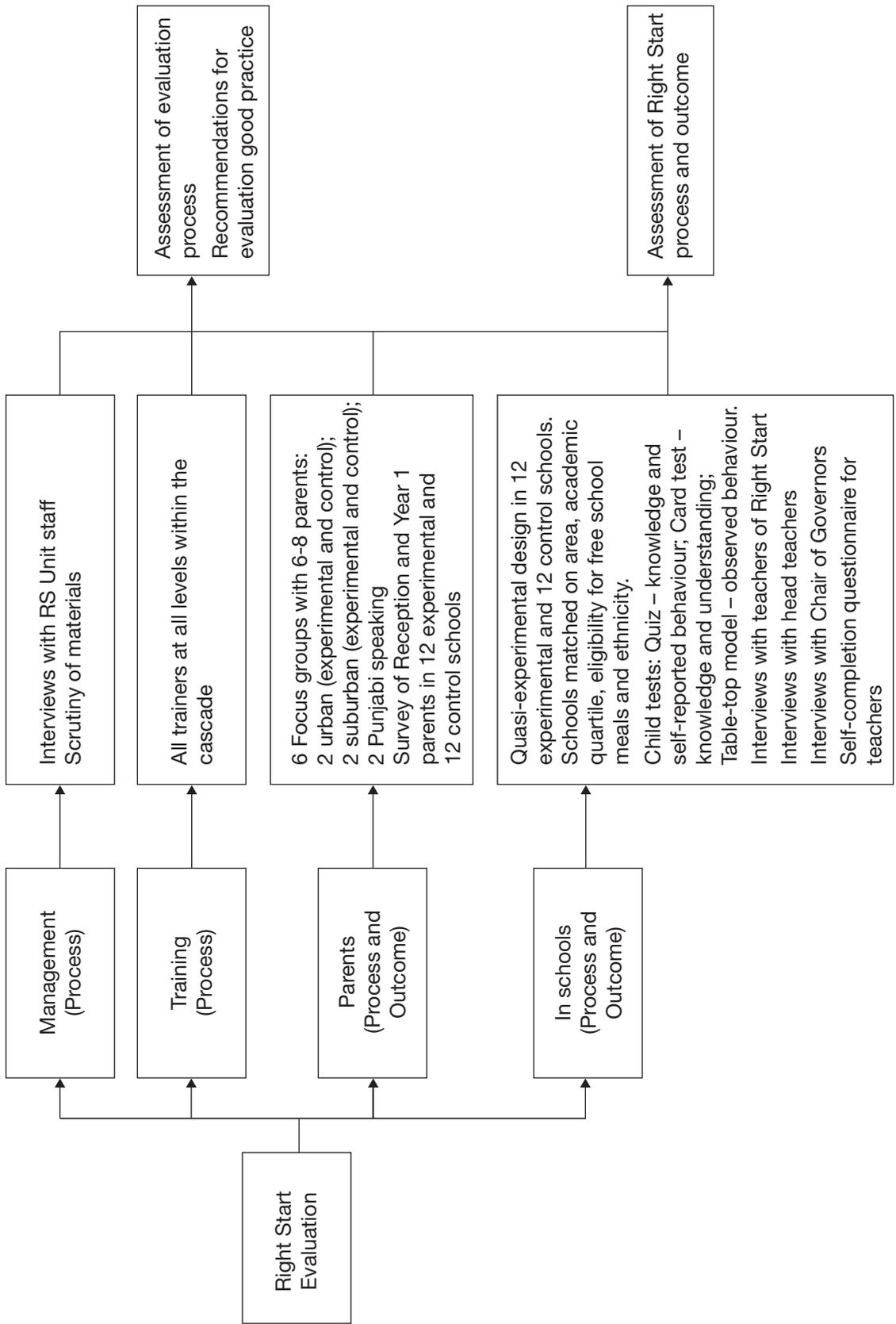
<b>Name of scheme</b>	Leicestershire Junior Road Safety Officers.
<b>Type of scheme</b>	Peer delivery.
<b>Target groups</b>	Intended to reach all Junior and Infant pupils in participating schools. No specific adaptations for pupils with special needs in special or mainstream schools.
<b>Participation rate</b>	Offered to all schools in Leicestershire and the City of Leicester. Taken up by 60% of schools in any one year, in the full range of environments from rural to inner-city.
<b>Teaching activities</b>	Scheme co-ordinator sends monthly newsletters and promotional materials on road safety themes. These are used by JRSOs as the basis for talks at school assemblies, competitions, noticeboards and sometimes other activities.
<b>Delivered by</b>	Scheme co-ordinator is an RSO employed by Leicestershire County Council. JRSOs are Year 6 (sometimes Year 5) pupils who volunteer or are selected for the job. Generally two or three JRSOs are chosen from each school.
<b>Aims of scheme</b>	<ul style="list-style-type: none"> <li>• To encourage responsibility among primary school pupils;</li> <li>• To enhance road safety awareness;</li> <li>• To facilitate activity and encourage the involvement of children in road safety activities.</li> </ul>
<b>Brief history</b>	Scheme has been running for approximately 15 years (in 2002). It evolved from a scheme originally developed in Cheshire. JRSO schemes have been run or are running in a number of UK local authorities.
<b>Aims of evaluation</b>	<p>The evaluation research was designed to assess:</p> <ul style="list-style-type: none"> <li>• whether the scheme met its stated aims;</li> <li>• how and why schools used the scheme or why they chose not to participate;</li> <li>• for whom and in what circumstances the scheme met its aims.</li> </ul>

Figure 2: Outline of the Leicestershire JRSO evaluation



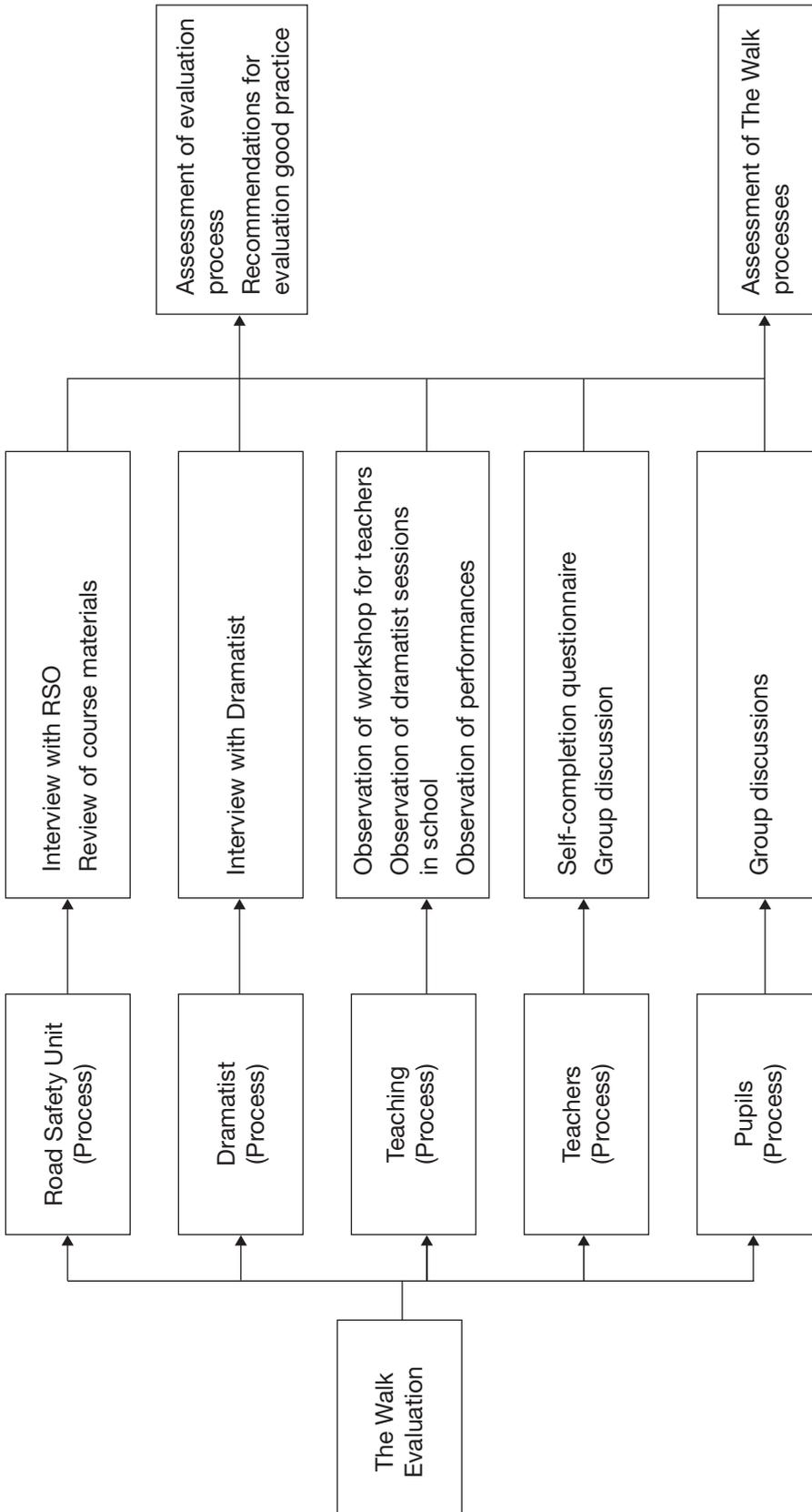
<b>Name of programme</b>	Right Start Practical Pedestrian Training Programme.
<b>Type of programme</b>	Roadside pedestrian training, with classroom preparation for and review of each practical training session.
<b>Target groups</b>	Pupils in Reception, Year 1 and Year 2 at participating schools. No specific adaptations for pupils with special needs in special or mainstream schools.
<b>Participation rate</b>	Offered to all schools in Lancashire. Approximately 10% have implemented Stage 1.
<b>Teaching activities</b>	Taught in three Stages, each organised in units of training: Other People, Safer Places, Stop, Look, Listen and Cross. Stage 1 (Reception) covers all six units; Stage 2 (Year 1) covers the first three Units and Stage 3 (Year 2) the latter three Units. Each Unit consists of three components: Let's Talk, Let's Do, and Let's Remember.
<b>Delivered by</b>	The programme is managed by a Road Safety Officer (RSO) with the assistance of Area Co-ordinators (ACs). Parents at participating schools volunteer to provide roadside training. Classroom components of each unit are usually delivered by the class teacher.
<b>Aims of programme</b>	To reduce child pedestrian casualties through a structured training programme designed to: <ul style="list-style-type: none"> <li>● encourage and develop the skills required to cope in the traffic environment as a pedestrian;</li> <li>● reinforce appropriate pedestrian behaviour;</li> <li>● allow children to develop strategies to make decisions in on-road situations;</li> <li>● develop awareness amongst parents, teachers and carers regarding child development and the ability of children of different ages to comprehend and deal with the traffic environment.</li> </ul>
<b>Brief history</b>	Programme was developed in 1997 for Lancashire County Council. After piloting in several schools, a modified Stage 1 programme was launched in June 1999 and Stage 2 followed in January 2001. Stage 3 is under development.
<b>Aims of evaluation</b>	The evaluation research was designed to assess: <ul style="list-style-type: none"> <li>● whether the programme met its stated aims;</li> <li>● how and why schools used the programme or why they chose not to participate;</li> <li>● for whom and in what circumstances the programme met its aims.</li> </ul>

Figure 3: Outline of the Right Start Evaluation



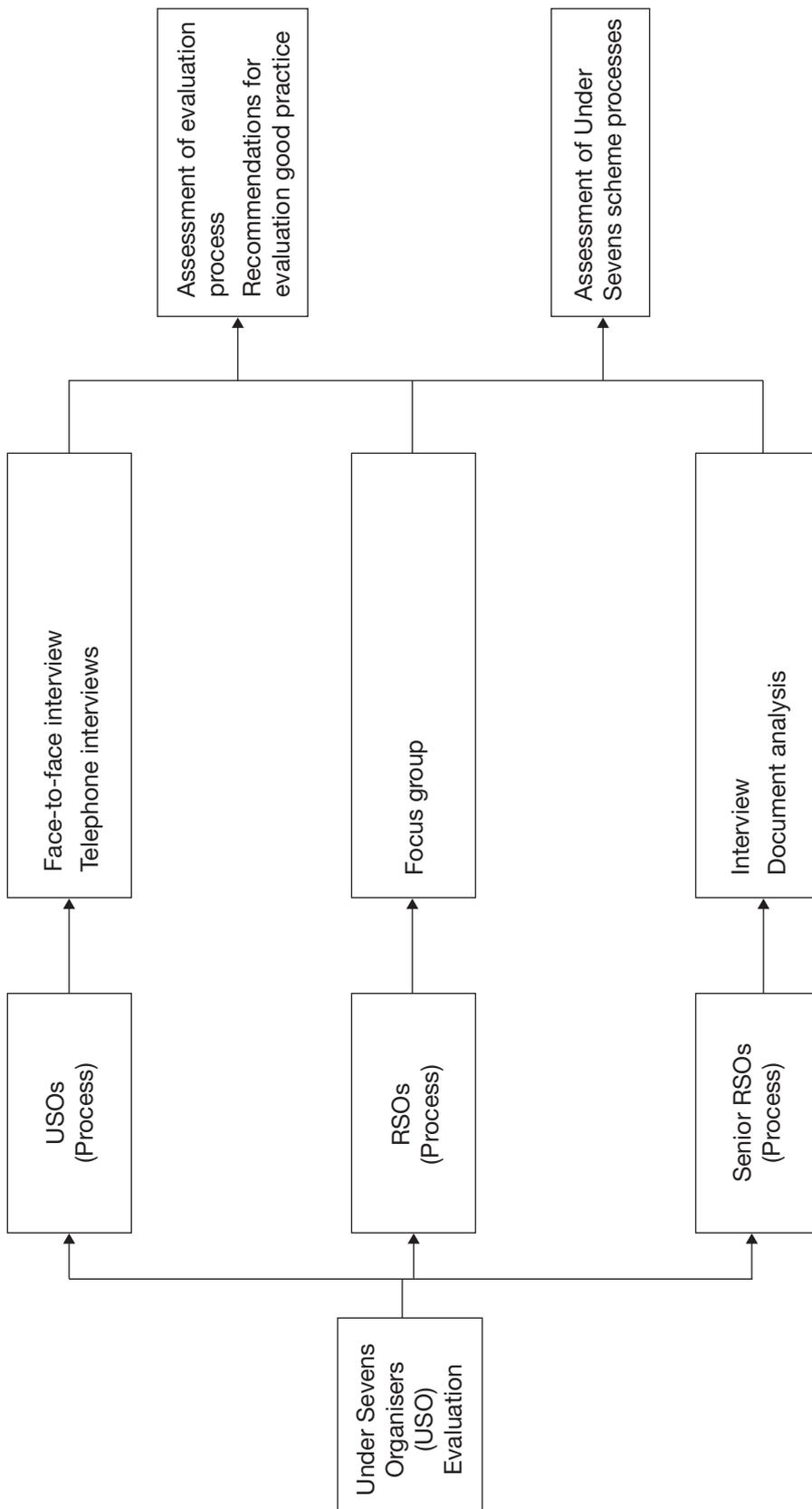
<b>Name of scheme</b>	The Walk.
<b>Type of scheme</b>	Drama-based education resource.
<b>Target groups</b>	Pupils in Year 5 at participating schools; teachers and parents.
<b>Participation rate</b>	Offered to all schools in the London Borough of Barnet.
<b>Teaching activities</b>	One teacher from each school attends a one-day training course incorporating drama workshops on the theme of walking to school. The teachers use their training and the resource pack provided to develop a performance with their pupils based on local issues about walking to school. This is presented to parents and other pupils.
<b>Delivered by</b>	A professional dramatist conducts the training course and provides in-class support to each teacher as he or she works with pupils to develop the workshops into a performance.
<b>Aims of scheme</b>	<ul style="list-style-type: none"> <li>• To influence parents and raise their awareness of the benefits of allowing their children to walk to school.</li> <li>• To provide an activity in which all Year 5 pupils could participate.</li> <li>• To provide a resource for Year 5, linking with existing pedestrian skills programmes available in Barnet for pupils in Years 4 and 6.</li> <li>• There were specific objectives for teachers, children and parents relating to school travel, safety and drama skills.</li> </ul>
<b>Brief history</b>	Scheme was developed by the London Borough of Barnet Accident Prevention Centre in 1996 and has been implemented annually since then.
<b>Aims of evaluation</b>	The scheme was evaluated primarily as a road safety intervention. The evaluation research was designed to assess scheme processes including the role of the dramatist, training and support available to teachers, content of the resource pack and effectiveness of teaching. Outcomes of the scheme were not considered.

Figure 3: Outline of The Walk Evaluation



<b>Name of scheme</b>	Under Sevens Scheme.
<b>Type of scheme</b>	Classroom activities.
<b>Target groups</b>	Pupils aged under seven years, at participating nurseries and infant schools. Children with special needs who attend special or mainstream schools are included in the scheme although the materials used may vary and the range of ages taught may be wider. Hearing-impaired children are taught with the help of a signer.
<b>Participation rate</b>	Offered to all nurseries and infant schools in the five local authorities within the Gwent region.
<b>Teaching activities</b>	Storytelling, including stories based on the original RoSPA Tufty materials. Road safety is the core of the curriculum but other safety issues are also addressed. A wider range of materials are now used and activities include role-play with roll-out zebra crossings, crossing patrol uniforms etc.
<b>Delivered by</b>	Under Sevens Organisers (USOs) each have a list of schools and are contracted for a fixed number of hours per term. They visit each school once per term and the length of visit depends on the size of the school.
<b>Aims of scheme</b>	<ul style="list-style-type: none"> <li>• To teach basic road safety concepts to children who are under seven years of age.</li> <li>• To make available Gwent Consultancy's Tufty Club membership and items to those schools/children wishing to join the club.</li> </ul>
<b>Brief history</b>	The scheme began with a parent volunteer who told the Tufty stories at her child's nursery. The Road Safety Officer (RSO) later offered to pay her expenses to visit other schools. When she left, the person who now co-ordinates the programme was asked to continue the school visits, supported by the RSO. Until recently the USOs worked on a volunteer basis; now they are paid for time spent in schools plus travel expenses.
<b>Aims of evaluation</b>	To examine scheme processes, specifically the management of the Under Sevens Organisers. Processes within schools and outcomes of the scheme were not considered.

Figure 4: Outline of the evaluation of Gwent's Under Sevens Organiser Scheme



## APPENDIX B: Pawson and Myhill's Evaluation Lessons

**Lesson 1.** Evaluation research should always begin with the recognition that programme efficacy depends on the 'ideas', 'individuals', 'institutions', and 'infrastructure' which make up the intervention. In any particular investigation, it is unlikely that all such features may be controlled and measured but efforts should always be made to observe, monitor and record basic information on all these features. It is particularly important that collective wisdom about the 'four I's' should accumulate across investigations.

**Lesson 2.** Programmes are not things but conjectures, and the evaluator needs to test them as conjectures. The typical programme under review takes the form of a theory that speculates, 'if we do X, we may persuade children to do Y'. The ideas embodied in a programme, however, may be met with acceptance, incomprehension, indifference or resistance on the part of children. A 'theory-driven approach' to evaluation is thus recommended, which considers the congruence of the conceptions of all concerned.

**Lesson 3.** Programme outcomes are the result of programme processes and evaluations are always enhanced to the extent that the study of one supports the understanding of the other. Programmes dependent on elaborate social process will always generate a complex footprint of outcome. Anticipating and understanding these patterns demands a partnership of 'process' and 'outcome' evaluation, and requires the use of both quantitative and qualitative methods.

**Lesson 4.** Randomised control trials are often considered the gold-standard method in evaluation since they offer great certainty about whether a particular programme has worked in a particular instance. But they throw up inconsistent results because they provide no understanding of the exchange of ideas within a programme nor of the context in which it takes place. They are not an absolute priority for the policy-maker.

**Lesson 5.** Policy-makers should desist from charging evaluation research with discovering 'what works' for no initiative will work for all subjects in all circumstances. The really instructive evaluation question is 'what is it about a programme which works for whom in what circumstances and in what respects?'

**Lesson 6.** Programmes which aim to teach children sets of behavioural codes/rules/guidelines have to appreciate that children do not just 'follow' rules but adapt them to their existing understanding. These processes of assimilation are quite uneven because certain rules will clash with other expectations created during their upbringing. The researcher's task is to uncover and explain the patterns of acceptance and resistance to rules.

**Lesson 7.** Programmes which aim to inform children via ‘peer-education’ and ‘club-activity’ have to anticipate wide variation in the ‘credibility of’ and a sense of ‘belonging to’ such manufactured social formations. Recipients will differ in the extent to which they find support and shared-identity. Researchers should expect considerable variability in the success of such networks and concentrate on discovering the successful combinations of message, messenger and subject.

**Lesson 8.** Programmes which aim to establish children as ‘partners’ in the development of the daily activity which makes up an initiative have to establish the boundaries of the child’s expertise. These programmes tend to be researched using a ‘formative approach’, which focuses on the day-to-day interpretations of those involved. This method, however, often lacks range in terms of coverage of viewpoints and the ability to trace outcomes. It frequently ends in advocacy rather than evaluation.

**Lesson 9.** Programmes should be developed with the aim of identifying and counteracting ‘risk factors’. But the preceding factors of social problems are always multiple. Evaluation thus has the task of understanding what risks a programme can and cannot confront, and to ensure that the tractable problems are pinpointed and attacked. The key priority is to ‘get the grease to the squeak’.

**Lesson 10.** Before any intervention or programme commences, there should be intelligence on the persons and situations at risk. And after any evaluation, such understanding should be refined considerably. Such discernment of risk and response is best gathered by a close analysis of the behaviour of particular sub-groups under the initiative.

**Lesson 11.** Having children as programme subjects and research subjects provides evaluation with a special concern with ‘age’. The aptness of an intervention never depends simply on age *per se*, but on the child’s developmental stage, their exposure to the contributory problem and their potential control over prevention strategies. The key strategy for promoting children’s participation in the research process lies in the adaptation of standard instruments, presenting questions in frameworks and formats with which particular age-groups are already familiar.

**Lesson 12.** Safety problems congregate in hazardous environments. Many of the risks children face are specific to local environments. Programmes aimed at children’s safety thus need to tap into local knowledge of the conditions that generate the problems they are set to tackle. Evaluation research needs to map differences in peoples’ everyday wisdom about their communities and measure success in terms of growth in understanding of, and responses to, the local hazards.

**Lesson 13.** Programme effects are always diverse and evaluators should never rely on singular measures to monitor programme success (or failure). Evaluators should always use multiple measures, thereby anticipating an uneven footprint of outcomes associated with the many processes involved in constructing, implementing and experiencing initiatives.

**Lesson 14.** A hierarchy of outcomes should be established for each family of programmes, following through the changes in understanding, attitudes and behaviour that interventions seek to affect. Outcome measurement is more secure the further it travels along the understanding ➔ attitude ➔ behaviour chain. It is often infeasible to measure long-term behavioural outcomes of a programme and actual outcome measures should be chosen to optimise the balance between the 'feasibility' and 'strength of proof' of potential measures.

**Lesson 15.** Programmes cannot be considered properly tested if they are assessed solely through the usage of 'approval ratings' (satisfaction scales). Such measurement methods are often seriously contaminated with error due to 'subject effects'. Sentiments of 'approval' for a programme are so distantly related from the question of whether it has changed behaviour that it is probably a waste of research effort to attempt to perfect such measures.

**Lesson 16.** Programmes cannot be considered properly tested if outcomes are measured solely through the usage of self-reported behaviour. The accuracy of self-report varies from topic to topic and from child to child in any programme area. Each field of programming needs to build up its own intelligence on the validity of self-report by considering the power of recall, level of stigma, degree of anonymity and capacity for deception associated with the task.

**Lesson 17.** Any education programme worth its salt will produce educational gains and *less* evaluation time and effort is required to show that this is the rule. The more 'proximate' to reality the assessment of knowledge-gain measure the better. Evaluations without the resources to carry through investigation to behavioural outcomes should concentrate on assessing the variability in knowledge gains across different subjects and different aspects of the programme curriculum.

**Lesson 18.** Measuring programme outcomes by simulating the situations in which they are intended to have effects is an important, if limited, method for the evaluator. Provided ethical and practical issues can be dealt with, the more true-to-life the outcome measures the better.

**Lesson 19.** Retrospective surveys can provide broad-brush information on the coverage of programmes and some evidence on their (otherwise obscure) longer-term impact. Caution must be observed in making causal inferences about programme impacts from a survey that has not closely monitored the subject's involvement with the programme.

**Lesson 20.** Evaluation research should always seek a multiplicity of outcomes, including those that fall outside the intentions and expectations of programme-architects. Unanticipated outcomes of programmes flow from programme complexity, subject over-reaction and bloody-mindedness, technological failure, legal loopholes and, occasionally, from nowhere.

**Lesson 21.** Researchers should never evaluate a programme *ab ovo*, they should always inspect previous broken eggs. Policy-makers attempt to produce new responses to old problems but the chances are some very similar initiative will have been put into place sometime, somewhere. Periodic reviews of existing research are thus vital. The process of review itself, however, takes on many forms from the catalogue-of-abstracts to quantitative meta-analysis. The most beneficial ways to conduct a review is to study each piece of research in the light of its contribution on ‘what works for whom in what circumstances.’

Source: Pawson and Myhill (2001)

## Appendix C: Useful websites

The following websites include copies of evaluation reports or provide information helpful to conducting or commissioning an evaluation.

AA Foundation for Road Safety Research:  
<http://www.aanewsroom.com/aafoundation/introduction.cfm>

British Medical Journal: <http://www.bmj.com>

Department for Transport: <http://www.dft.gov.uk/>

Injury Prevention Journal: <http://ip.bmjournals.com>

Kerbcraft Pedestrian Training Scheme: [www.kerbcraft.org.uk](http://www.kerbcraft.org.uk)

Royal Society for the Prevention of Accidents (RoSPA):  
<http://www.rospace.com/CMS/index.asp>

Scottish Executive: <http://www.scotland.gov.uk>

Scottish Road Safety Campaign: <http://www.srsc.org.uk>

Social Research Association: <http://www.the-sra.org.uk>

Sustrans Safe Routes to School: <http://www.saferoutestoschools.org.uk>

The Cochrane Collaboration: <http://www.cochrane.org/>

Transport Research Laboratory: <http://www.trl.co.uk/>

UK Evaluation Society: <http://www.evaluation.org.uk>

Young TransNet: <http://www.youngtransnet.org.uk>

## Appendix D: Feedback Form

We are very interested in finding out what you think of this guide and would like to have your feedback. If you would like to tell us, please answer these questions and post them to us. You can also contact us via email or phone.

### 1. In what context have you used the publication?

---

### 2. Were you working with a specific RSE programme? If so, please describe:

---

### 3. Which sections of the guidelines did you read? (please tick all that apply)

- |                         |                          |  |                          |
|-------------------------|--------------------------|--|--------------------------|
| Introduction            | <input type="checkbox"/> | General guidance                                   | <input type="checkbox"/> |
| Terminology             | <input type="checkbox"/> | Cost effective evaluations                         | <input type="checkbox"/> |
| Review past evaluations | <input type="checkbox"/> | Appendix A: Case study profiles                    | <input type="checkbox"/> |
| Choosing an evaluator   | <input type="checkbox"/> | Appendix B: Pawson and Myhill's Evaluation lessons | <input type="checkbox"/> |
| Doing an evaluation     | <input type="checkbox"/> |  |                          |

### 4. What is your opinion of the guidelines? (please tick one box on each line)

	Useful	Somewhat useful	Not useful	Did not read
Introduction	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Terminology	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Review past evaluations	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Choosing an evaluator	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Doing the evaluation –	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Develop an evaluation objective(s)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Identify your target group	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Develop evaluation measures	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Select methods appropriate evaluation measures	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Design and test the materials	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Administering the tests	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Interpretation of the results	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Writing an evaluation report	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
General guidance	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Cost effective evaluations	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Appendix A: Case study profiles	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Appendix B: Pawson and Myhill's Evaluation lessons	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

**5. How could these guidelines be improved?**

**6. Any other comments?**

**7. What is your occupation?**

Road Safety Officer

Other practitioner

Executive/manager

Academic/researcher

Other (please specify) \_\_\_\_\_

**Please return the feedback form to:**

Safety Group  
TRL  
Crowthorne House  
Nine Mile Ride  
Wokingham  
Berkshire  
RG40 3GA

Tel: 01344 773131  
Fax: 01344 770356  
E-mail: enquiries@trl.co.uk