European Monitoring Centre
for Drugs and Drug Addiction

EMCDDA
MANUALS

Guidelines for the evaluation of drug prevention

A manual for programme planners and evaluators
Second edition

8

European Monitoring Centre
for Drugs and Drug Addiction

Guidelines for the evaluation of drug prevention

A manual for programme planners and evaluators
Second edition

8

Legal notice

Cataloguing data can be found at the end of this publication.

Luxembourg: Publications Office of the European Union, 2012

ISBN 978-92-9168-499-1

doi: 10.2810/51268

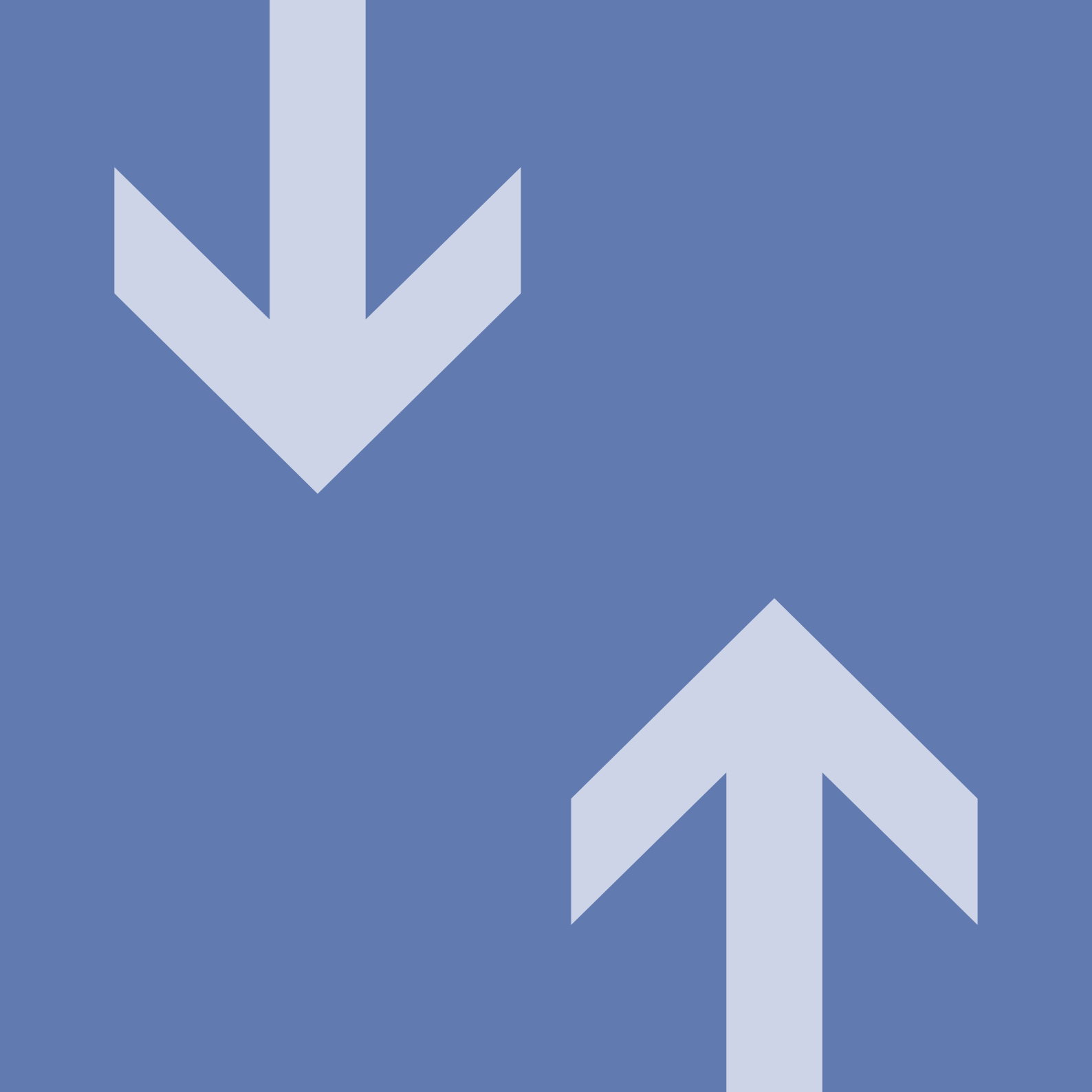
Printed in Luxembourg

PRINTED ON WHITE CHLORINE-FREE PAPER

European Monitoring Centre
for Drugs and Drug Addiction

Cais do Sodré, 1249-289 Lisbon, Portugal
Tel. (351) 211210200 • Fax (351) 218131711
info@emcdda.europa.eu • www.emcdda.europa.eu

Contents



Acknowledgements

Project leader

Gregor Burkhart, European Monitoring Centre for Drugs and Drug Addiction (EMCDDA), Portugal.

Authors

For the updated version and Chapters 2 and 7 concerning theory-driven evaluation: Liliana Leone, CEVAS, Rome.

For the original version: Christoph Kröger, Heike Winter, Rose Shaw, IFT Munich.

Contributors to the original 1998 edition

Wim Buismann, Jellinek Centre, Amsterdam.

Gerhard Christiansen, Federal Centre for Health Education, Cologne.

Mark Morgan, Education Research Centre, Dublin.

Alice Mostriou, Athens University Medical School, Athens.

Jorge Negreiros, Universidade do Porto, Porto.

Teresa Salvador, Centro de Estudios sobre la Promoción de la Salud, Madrid.

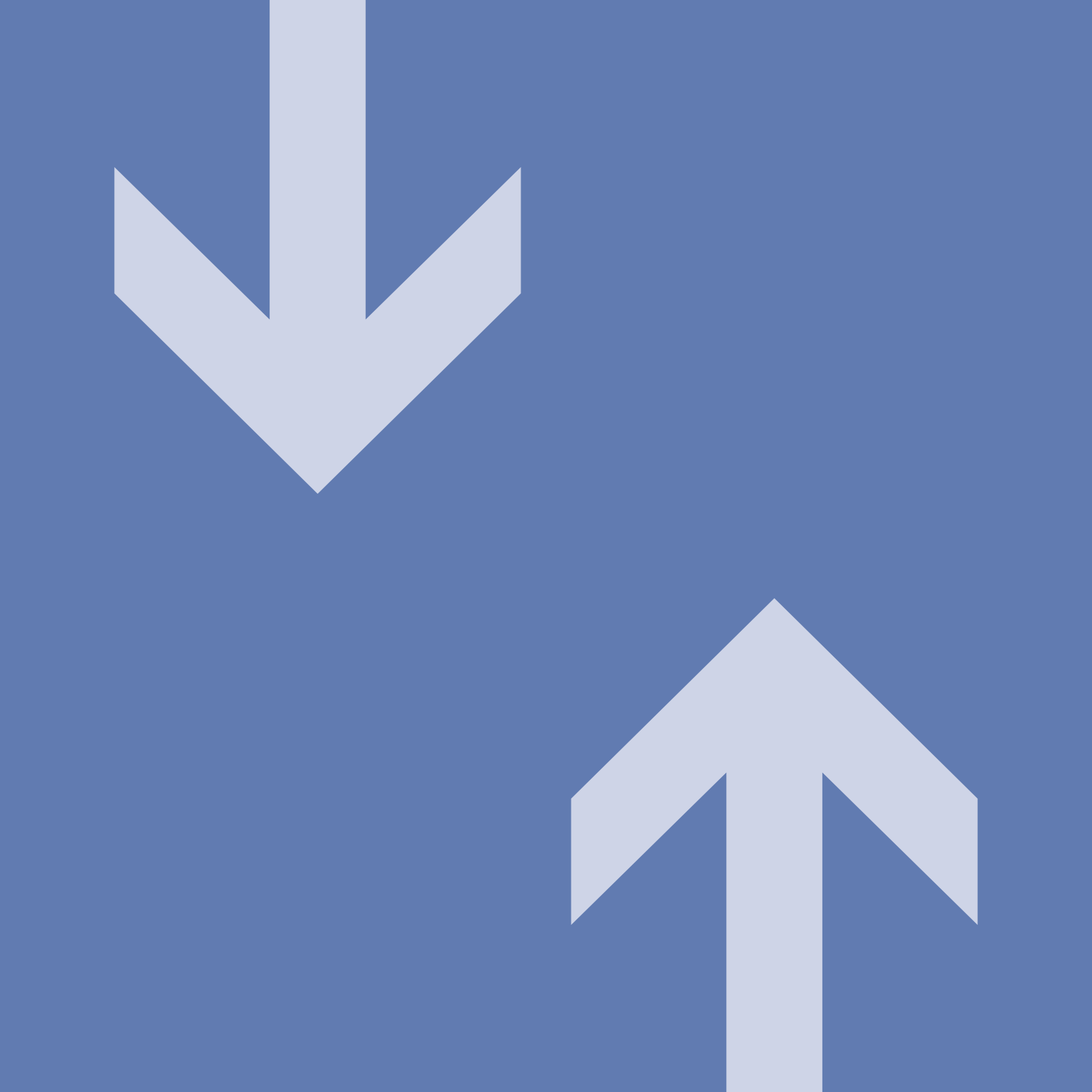
Anne-Marie Sindballe, Sundhedsstyrelsen, Copenhagen.

Zili Sloboda, National Institute on Drug Abuse, Rockville, MA, United States.

Alfred Springer, Ludwig Boltzmann Institute, Vienna.

Jürgen Töppich, Federal Centre for Health Education, Cologne.

Alfred Uhl, Ludwig Boltzmann Institute, Vienna.



Preface

Prevention is a key consideration when discussing substance use and related problems. Drug prevention is any activity that delays or reduces the use of substances. Preventing substance use potentially makes other interventions, especially targeting treatment and harm reduction, unnecessary. Evaluation of the various interventions is essential to develop and improve existing approaches and to work towards creating a best practice in drug prevention.

The EMCDDA first published the *Guidelines for the evaluation of drug prevention* in 1998 to respond to a need expressed by professionals and policymakers to have a tool and stimulus to evaluate prevention interventions in Europe. These guidelines aimed to facilitate and encourage project leaders to evaluate their own interventions and to design them accordingly.

This second edition is an updated version of the original EMCDDA guidelines, which contains new methods, concepts and examples of currently available drug prevention interventions and which provides a framework for carrying out an evaluation. The content has also been updated with developments in online EMCDDA drug prevention and evaluation resources.

Since publication of the first edition, prevention professionals who continue to apply these guidelines have contributed to the increased attention given to evaluation and prevention. However, although guidance on drug prevention is available in some Member States, only a minority of interventions in the European Union have an outcome evaluation. Even large, well-funded projects are still not fully evaluated and the availability of empirical evidence concerning prevention is weak.

These guidelines are intended for professionals working in the drug prevention and evaluation field and can provide valuable guidance when planning and evaluating prevention activities. The intention to develop best practice in drug prevention, as expressed in the 2005–8 and 2009–12 EU drugs action plans, and the fact that evaluation is now widely used in Member States' drug policy, testify to the importance placed on drug prevention and evaluation in the European Union.

The aim of this publication is to raise the profile of both drug prevention and evaluation methodologies and to provide prevention practitioners with an up-to-date framework, to enable their interventions and evaluations to reach their full potential.

Wolfgang Götz
Director, EMCDDA

Introduction

The discipline of evaluation has rapidly developed in the last decade. Public authorities, funding institutions, professional and social–health practitioner associations, healthy schools networks, international bodies and local health promotion coalitions call for evidence-based practice and effective interventions. They have been recommending, commissioning or carrying out thousands of evaluations and monitoring reports. However, there remain two weaknesses: while there is a high level of consensus about the value and the need for evaluation, there is still a disagreement about more suitable → evaluation approaches ([1]) of drug prevention programmes; and there is also confusion about the different meanings and uses of 'theory' in evaluation and planning practices.

The aim of this second edition of the EMCDDA guidelines is not to enter into debate, suggesting 'the best way' or to explain in depth an approach; it is rather to open a window of knowledge to the newcomers of evaluation, thereby widening the possibility to appreciate the 'logic of evaluation', to develop some reference points to avoid common misunderstandings and discover methods that are more tailored to the objectives of a specific situation. Professionals who want to evaluate drug prevention initiatives should consider employing a broad range of information-gathering procedures using multiple methods. It is important to recognise the value of methodological pluralism and the role of theory in the design and evaluation of programmes.

The main differences of this second edition of the evaluation guidelines are:

(1) giving more attention to the role of theory in evaluation and clarifying some terminological and conceptual distinctions among approaches or similar terms;

(2) updating many practice examples on the basis of more recent and well-conducted evaluations;

(3) providing some practical tips about how to conduct the first steps of an appropriate evaluation – how to develop evaluation questions and how to choose appropriate evaluation designs;

(4) sustaining the use of multiple methods, since the purposes of an evaluation and the level of complexity of an intervention are different: from multi-placed community initiatives with long-term expected impacts to single, standardised micro-interventions;

([1]) '→' indicates terms explained in the Glossary (see Part Three of this publication).

(5) considering, in evaluation processes, other issues such as: relations with stakeholders, ethical issues, different uses of evaluation, development of evaluative questions, gender and ethnic target differences, costs and administrative instruments, differences between evaluation and monitoring, integration with other policies and sectors;

(6) offering a framework to evaluate drug prevention interventions more coherently with the latest advances of the EMCDDA Prevention and Evaluation Resources Kit (PERK) ([2]), with regards to the role of theory in drug prevention planning.

It is proven that the evaluation of drug prevention interventions has become pervasive and is now part of everyday activity, either in public, non-profit or voluntary sectors. Private and public funding often involves the obligation, for even small-scale projects, to undergo or carry out an evaluation. Notwithstanding the increased pressure to evaluate, there is a deep dissatisfaction about the low level of methodological quality and the usefulness of knowledge generated by evaluations. In fact, the accumulation of empirical evidence about effectiveness is of limited value to the practitioner and stakeholders, unless accompanied by general theoretical principles which might inform wider application. The issue, therefore, becomes how to integrate theory into the evaluation process.

Theory-driven approaches of evaluation have developed significantly in the last fifteen years, being mostly used in the areas of health promotion, risk prevention, complex socio-economic interventions and comprehensive community-based initiatives. The purpose of theory is intended not to offer universal explanations or predictions, but rather to enhance comprehension of complex situations; knowledge is seen as contextual and contingent. A well developed and fully articulated 'programme theory' can be very useful for framing key evaluation questions, designing evaluation, identifying the most critical dimensions to the programme's success, selecting outcome variables and identifying timing of measures and instruments. A broader use of theory in evaluation design can:

- enhance understanding of complex situations, build common understanding amongst programme staff and modify a planned intervention to avoid negative consequences;
- help people to better design evaluation plans considering the influence of contextual factors;
- include in the process also the experience of practitioners and communities.

([2]) EMCDDA (2010), *Prevention and Evaluation Resources Kit (PERK)*, EMCDDA Manuals 4, European Monitoring Centre for Drugs and Drug Addiction, Lisbon.

A further basic idea stressed by these guidelines is that different circumstances may yield different reactions to the same intervention, and putting in place alternative → mechanisms may produce different results. Interventions and programmes cannot be intended as simple treatments to be repeated, not taking into account the differences, backgrounds and contexts of individuals and without regard to implementation processes. Interventions evolve, while trying to combat social and health problems and their success depends upon the trust and the credibility they have gained in the past, in practice.

What is evaluation?

Evaluating a → prevention intervention ([3]) means systematically collecting, analysing and interpreting information about how the intervention works, with which effects and for whom, by the rigorous use of scientific social research methods, to judge the merit and value of an intervention. The knowledge that arises from an evaluation can be used to decide how to improve an intervention, whether to expand it or to abandon it altogether and to draw lessons to design future programmes or to improve existing programmes. Each evaluation should be tailored to fit programmes and social contexts. There is no such thing as the best method, approach or evaluation design. Essentially, evaluation should answer the following questions:

- Is the project or programme adequately planned and designed?
- Are the interventions' strategies based on scientific evidence and recommendation?
- How were different community-level stakeholders and agencies involved?
- What is the → programme theory? Is it consistent?
- How is the intervention being implemented?
- Does the intervention actually reach the → target group?
- What are the → results and the outcomes? How do they change in relation to different contexts and circumstances?

([3]) Throughout these guidelines, the term 'prevention intervention' is preferred to 'project' or 'programme', as these other terms can mean different things to different people. We maintain the term in the case of 'programme theory' to avoid confusion with 'intervention theory', which in literature refers to other meanings.

- Are there any positive or negative unintended impacts? Does it take into account the problem of equity and risks of health inequalities?
- Is the programme efficient (value for money) and relevant?
- Are we able to distinguish intervention failures due to poor designs and wrong theoretical assumptions from failures in achieving successful results due to inaccurate management and implementation?
- What lessons can we draw from the evaluation results? Is it possible to generalise them and transfer some understanding to broader contexts?

Answers to these questions help to distinguish useful prevention interventions from ineffective, inefficient ones. This is not only important in terms of improving the level of knowledge about prevention, but it can also be the basis upon which policymakers and those financing projects decide which projects to support and which to ignore.

However, despite the widely accepted need for and use of evaluation in theory, very few prevention interventions have actually been evaluated in Europe. One of the reasons for this lack of evaluation may be that there is insufficient knowledge and a lack of confidence about how to tackle evaluating prevention interventions in the field of substance use.

Theoretical background of the guidelines

These guidelines take a → mixed method approach. While some users may prefer a structured, → quantitative approach of research, other users may prefer more in-depth information based on → qualitative methodology. Over the course of previous years, many evaluations of drug prevention interventions have been successfully carried out combining different approaches. For more information on the theory and methodology of evaluation procedures, see the 'Further reading' section in the References. The companion volume to this manual is 'Evaluating drug prevention in the European Union' ([4]). This monograph provides in-depth information on many different aspects of planning and evaluation.

Who can use the guidelines?

These guidelines are intended to help people from a variety of backgrounds plan and evaluate drug-prevention interventions. They are designed especially for those with more practical experience and less knowledge of evaluation, but they can also be helpful for more practised evaluators. In order to address the needs of both these target groups, the manual is divided into three parts. This division is based on the assumption that the more experience the reader has of evaluation, the less he or she will need to read, and vice versa.

Applying the guidelines

First and foremost, the guidelines are intended to help people working in the drug prevention field evaluate their intervention. But the guidelines can also be used in other ways, for example, as a teaching aid in evaluation training or in drawing up and evaluating funding proposals, reports and even other guidelines.

When can the guidelines be used?

The guidelines focus on evaluating prevention interventions. They do not provide information on how to → design a specific prevention intervention. Instead, they can be applied as soon as the concept of a prevention intervention begins to be discussed. At this stage, they can help reflect how the intervention was planned while later on, they can be used to evaluate its implementation

([4]) 'Evaluating drug prevention in the European Union', *EMCDDA Scientific Monograph Series No 2* (please refer to the References section).

and results. They are suitable not only for untested interventions, but also for those that have already been conducted and that are routinely applied.

Structure of the manual

There are three parts to this manual: the first concise section contains the basic guidelines themselves; the second provides more detailed background information and practical examples; and the third is a glossary of the terms used.

Part 1: The guidelines

This section provides the framework for the manual, covering all the steps and issues that should be considered during an evaluation. It covers four main areas:

- the planning phase;
- the quality and process evaluation;
- the outcome evaluation; and
- communicating the evaluation's results.

Part 2: Examples

This section contains background information on the issues covered by the guidelines, as well as practical examples suggested by the feasibility study. Part 2 will be particularly helpful for those less experienced in project planning and evaluation.

Part 3: Glossary

The Glossary contains more detailed descriptions, definitions and explanations of the technical and methodological terms used in the guidelines. As with Part 2, this will help those who are less familiar with methodological issues to understand and use the manual.

How to use the manual

These guidelines attempt to incorporate the most important elements for consideration when planning and evaluating prevention activities. While many other issues could have been included, for practical reasons the manual has been limited to the bare essentials. Many of those working in the prevention field may not have access to the financial and personnel resources necessary to evaluate a prevention intervention fully. The reader is nevertheless strongly encouraged to follow systematically the basic steps laid out in these guidelines, to ensure the quality of the prevention intervention. It is recommended that all the questions in Part 1 be read, even though in many cases not all of them will need to be answered. In general, practitioners and evaluators should try to answer all the questions about the planning phase (Chapter 1) and – depending on their particular evaluation plan – proceed to the next chapters. Outcome evaluation is very important, although it is not always feasible for smaller prevention interventions. Planning an outcome evaluation requires at least basic statistical knowledge which cannot be conveyed in a manual such as this. Chapter 5, on the use of the results, is relevant for all kinds of evaluation.

This manual will guide the reader through the complex, costly and time-consuming procedure of evaluating drug-prevention interventions. Evaluation allows services to be planned more effectively, and these guidelines will, in turn, enable evaluations to be planned more effectively.

Users

As evaluations are undertaken at different levels, by practitioners and by expert evaluators with different degrees of expertise, complexity of evaluation designs and different budgets, the guidelines have two main target groups: (a) prevention professionals and social-health workers involved in self-evaluation practices and (b) professional audience of evaluators.

In order to facilitate the reading of the text and avoiding an excess of information, we have used the symbol ☞ at the beginning of paragraphs which are addressed to more experienced prevention practitioners.

PART ONE

Guidelines for planning and realising evaluation

CHAPTER ONE

Chapter 1
Assessing intervention planning

Assessing a project proposal/plan and an intervention planning is undertaken after the → prevention intervention is planned and designed and after its goals, means, methods, → target group and work-plan are chosen. Usually, the intervention plan conforms to a predefined format, according to the requests of sponsors or funding institutions. Assessing this phase involves judging:

- the adequacy of the intervention's → strategy;
- the knowledge of context, → needs and resources assessment;
- the quality and feasibility of the proposal;
- the potential → impacts;
- the clarity and solidity of assumptions at the core of the intervention;
- the use of → scientific evidence as the basis of the → prevention approach;
- the coherence of different steps of the proposal and the adequacy and strength of partnership arrangements.

The evaluation can be carried out either by officers responsible for prevention funding (i.e. a call for tender to select and co-finance some drug prevention proposals) or an external → evaluator, or by the person in charge of the intervention. Information should be gathered from the planners and the team that will actually carry out the intervention and from the project proposal document.

The phenomenon

The starting point for the planning phase should be a description of the nature, extent and location of the phenomenon to be addressed by the intervention. The term 'phenomenon' is preferred in these guidelines, because the term 'drug-related problem', which could also be used, has a more negative connotation and because a drug prevention programme can be addressed to tackle many kinds of phenomena. This phenomenon should always be defined as a → substance-use behaviour which the planned intervention aims to prevent, or as an → attitude toward psychoactive substance use and lifestyle. The characteristics of those affected by the

phenomenon should also be set out. The following questions need to be answered during the assessment.

- What phenomenon is the intervention designed to prevent?
- Are the sources of information of good quality: updated, reliable, trustworthy?
- What are the socio-demographic characteristics of those affected by the phenomenon compared with those who are not?
- Where does the phenomenon occur, and where does it not?
- How long has the phenomenon been known? Have its size, impact and relevance changed over time?

The conceptual background and the theoretical framework

After defining the phenomenon, the theory behind any impressions the practitioners (or the planners) may have of the cause, modification and control of the phenomenon should be outlined. This should clarify why they chose a specific strategy and approach to drug prevention intervention and why those particular → objectives and methods for the intervention were identified (see below). When designing an intervention, they should explain what their theoretical background is and why they anticipate that such an intervention will result in a positive outcome within a specific target group.

Usually, the practitioners refer to both implicit (not explicit) theories of action and scientific theories or results of literature (i.e. evidence and results of systematic reviews). To assess the quality of a project proposal, the evaluator should consider the scientific and practical reliability of the → 'programme theory' underlying the intervention plan. 'Programme theory' is a group of interrelated assumptions, principles or propositions that enable us to explain or guide a social action. It identifies the necessary actions that should be taken and the conditions required to achieve the planned outcomes of the intervention.

The use of → 'programme theory' in planning and in evaluation is now commonplace and many project funders require proposals for projects to include a logic model or programme theory in recognition of its value for planning, as well as for evaluation.

The following questions need to be answered, as well as examples of strategies and activities, which could change (or which have changed) the phenomenon:

☞ 'Theory of the programme' is not to be confused with an intervention based only on scientific → explanatory theories (such as 'social cognitive', 'theory of planned action', etc.). It is necessary to reconstruct the underlying assumptions of the intervention, the stakeholders' tacit theories and conjectures guiding the choices during the different stages of actual implementation. It usually starts with an assumption about what will work – sometimes called a → theory of change and the evaluator helps the programme stakeholders (policymakers, practitioners and participant(s)) to bring their latent or tacit theories to the surface ([1]); he/she can discover that different members of staff have different assumptions and, therefore, it can be useful to work with multiple theories which may even compete with each other. On the basis of scientific literature, the evaluator can choose, using professional judgement or by working with key stakeholders, the more plausible and more relevant hypotheses. Programme stakeholders have an insider understanding of the interventions in which they are involved and, therefore, constitute key informants in the evaluation process.

([1]) See *Prevention and Evaluation Resources Kit (PERK)*, EMCDDA Manuals No 4, 2010, pp. 27–34.

- Which explanation best describes the origin of the phenomenon?
- What factors are responsible for the continuation of the phenomenon?
- As an initial step of evaluation, has the intervention been described in a detailed way?
- How accurate is the reconstruction of the 'programme theory'?
- Does the reconstruction of the theory of the intervention include either accumulated scientific theories and conjectures or explicit hypotheses of planners? How does the evaluator select intervention theories?

The need for a prevention intervention

That the phenomenon exists to a degree that warrants an intervention must also be verified. This analysis of the need for a particular intervention must calculate the number of people affected by the phenomenon, make the case for this particular intervention as opposed to any other, and describe how (and if) it fits in with other activities. The following questions need to be answered:

- Do the proposals explain how many people are affected by the phenomenon? How many new cases are there and how often do they appear (→ prevalence, → incidence)?
- How is the phenomenon likely to develop if nothing is done? On what grounds?
- How should the need for the intervention be described?
- Are there different opinions as to the need for an intervention (varying perspectives on need)?
- How was the need for the intervention assessed (needs assessment)?
- Are any related interventions being carried out or planned? Will the current intervention cooperate with these activities?

The target group

Next, the group to whom the specific intervention is addressed (target group) should be considered. Two kinds of target group usually can be differentiated: an ultimate target group, which is most at risk from the drug phenomenon; and an → intermediate target group, or intermediate beneficiaries, such as parents, teachers, gym trainers, social and health workers, stakeholders of the community, peer-groups, non-profit-making organisations, youth leaders, and the general population. If the intervention only addresses an intermediate target group, this group should be described as the target group. How the target group is to be approached and motivated to participate in the intervention should also be outlined. The following questions need to be answered:

- Is the target group the ultimate target group or an intermediate target group?
- What are the socio-demographic characteristics of the target group, the scale of the phenomenon and the size of the group?
- Why was this target group chosen?
- How many people will the intervention reach?
- Where and how will the target group be contacted, recruited and motivated (→ selection effects, → coverage, → bias)?
- How can it be guaranteed that the target group will stick with the intervention (→ attrition)?
- Even if the planned intervention exclusively addresses an intermediate target group, what are the characteristics of the ultimate target group?

Objectives

The intervention's objectives need to be totally clear, with its expected effects on both substance-use behaviour and on → mediating variables defined. What effects it is hoped the intervention will achieve for any intermediate target group should also be described.

The following questions need to be answered:

- How will the intervention affect substance-use behaviour in the ultimate target group?
- How will the intervention affect substance-use behaviour and related risks in the ultimate target group?
- What objectives are considered for other mediating variables (→ life skills, → risk factors, → protective factors, → problem behaviour, → structural changes, changes in → lifestyle and → cultural habits)?
- What is the relationship between these mediating variables and substance-use behaviour?
- What are the objectives concerning the intermediate target group?
- How are the objectives for the intermediate target group and the ultimate target group related?

Methods

It is also essential to be clear about the methods and strategies that will be employed to meet the objectives. The empirical evidence for these strategies should be described, as should the timetable and overall timescale required for the intervention. The following questions need to be answered:

- What strategies, components and methods will be used in the intervention?
- Who will be involved in the intervention?
- Is there any empirical evidence for the success of the methods chosen (e.g. scientific literature, research papers)?
- How long will the intervention last?
- What is the planned timetable of the intervention (number of activities, duration and frequency of each activity, etc.)?
- Will the feasibility of the intervention be tested?

Resources

Having clarified the objectives and methods, the resources available (including staff time) must be examined to ensure there are no → barriers that might affect implementation or evaluation. The following questions need to be answered:

- Which staff members will carry out the intervention, and what qualifications are required?
- How time-consuming will the intervention be for each of these staff members?
- What is the budget and who is providing it?
- What additional resources are available (e.g. people, organisations, rooms, materials, etc.)?
- What could hinder the implementation or evaluation (barriers)?

Planning the process evaluation

When evaluating the planning phase, the subsequent stages must also be borne in mind. At this point, it should be decided whether a process evaluation will be undertaken and if so, who will carry it out. Chapters 2 and 3 deal with process evaluation more fully, but if planning to carry out such an evaluation, the following questions should be answered:

- Is a process evaluation planned?
- What resources are available to undertake such a process evaluation?
- Who will carry out the process evaluation?

Reflecting on the planning phase

When the planning phase has been evaluated, the whole process of information gathering and the lines of communication involved should be examined. The following questions need to be answered:

- Who was involved in the planning phase?
- What is the overall assessment of the planning phase process?

Planning checklist

The evaluation of the intervention's planning phase will now be complete and the following points clarified:

- what is to be tackled;
- how to explain the phenomenon;
- why the intervention is necessary;
- whom the intervention is intended to help;
- what the objectives are;
- how the intervention will be carried out; and
- what resources the intervention will use.

The review of the whole planning process will illustrate how decisions were made. The following chapters will help in monitoring the intervention's process and outcomes.

CHAPTER TWO

Chapter 2
Designing and carrying out a theory-driven evaluation

There is a list of terms and labels referred to 'theory' that have been used in evaluations: theory-driven, theory-based, theory-oriented, theory-of-change and implementation theory. For both the major approaches, theory-based evaluation (Weiss, 1997a, 1997b) and realist evaluation (Pawson and Tilley, 1997), the programme theory consists of a series of assumptions, mostly implicit, concerning how the programme actions are supposed to achieve the outcomes. It is expressed in the form: 'If we do this ... then it will happen that ...'. Programme theory can be defined as a set of propositions regarding what goes on in the black box of the programme/intervention during the transformation of input to output, that is, how a bad situation is transformed into a better one through treatment inputs (Lipesy, 1993).

Programme theory concerns the process through which programme components are presumed to affect outcomes and the conditions under which these processes are believed to operate (Donaldson, 2001). It offers explanations of how or why something happens, by identifying key variables and establishing the relationship between them, helping people to make predictions. To get an in-depth understanding of an intervention or a programme, evaluators should use: 'scientific theories'; main stakeholders' implicit theories and assumptions; research and evaluation results referring to the same phenomena to which the intervention is addressed; or other sectors indirectly involved. With regards to scientific drug prevention theories, refer to the online version of the Prevention and Evaluation Resources Kit (PERK), Step 2b 'Theory' ([5]).

☞ According to Carol Weiss, there are two kinds of theories (Weiss, 1997a, 1997b): the programme theory and implementation theory. We call the combination of both theories the programme's theories of change. While the implementation theory focuses on the delivery of programme services, the programme theory deals with the relationship that occurs between interventions and the achievement of goals. The main difference is that 'programme theory' focuses on → mechanisms of change and places the emphasis on the responses of the target groups to the interventions. The second, the implementation theory, refers to the implementation logic box with

([5]) EMCDDA (2010), http://www.emcdda.europa.eu/publications/perk/resources/step2b (last accessed on 10 May 2012).

linear connections between inputs, activities, outputs, outcomes and impacts (sometimes, additional concepts are used such as context, assumptions and external factors) and assumes that interventions are conducted as planned, with reliability, intensity and quality, so that anticipated outcomes can be reached. This version takes the shape of a diagram with arrows linking different components of the box. It represents a good start to focusing on outcomes rather than activities; however, this simple diagram, in contrast with 'programme theory', does not examine the causal → mechanisms involved in programmes and does not explore alternative causal strands and possible explanations (the so-called 'why questions') of the achieved outcomes.

Objectives and use of the evaluation

Evaluations are undertaken for a variety of purposes; it is possible to distinguish two main purposes: → accountability and learning. Evaluation objectives should be determined by careful consideration of the possible decisions to which the evaluation's results should contribute. Possible decisions include: (a) refining or redesigning the intervention; (b) identifying whether modifications in the delivery of some interventions could improve the efficiency or efficacy; (c) whether and how the programme should be modified in relation to different contexts; (d) whether the intervention is 'relevant' and contributes to satisfy needs considered a priority by the community; (e) whether the implementation network adopts an intervention strategy based on scientific evidence. Evaluations also provide information for high-level decision-making concerning how different interventions and policies, delivered at community level, work together as integrated strategies, or are incoherent and achieve opposite goals. Depending on the objectives and number of resources available, it is possible to perform more than one type of evaluation in the same timeframe.

☞ The motivations and the scope of evaluation are connected to its use, so it is necessary from the beginning to understand what type of use it is expected to have and to choose if it is opportune to undertake an evaluation. There are three main types of → use of evaluation: (1) instrumental; (2) conceptual; and (3) symbolic use.

(1) Instrumental use involves providing information for decision-making and improving programmes and applying evaluation results in direct ways;

(2) Conceptual use involves using the results of evaluation for general enlightenment;

(3) Symbolic use involves using evaluation results to legitimatise and sustain predetermined positions. If the primary purpose consists of hidden agendas, then an evaluator should reconsider whether he/she should be doing an evaluation at all.

The following questions need to be answered:

- Have the evaluation objectives been specified and are the expectations realistic? Have the purposes and different audiences for the evaluation been clearly specified?
- How are the evaluation results expected to contribute to the decision process?
- What decisions concerning the programme will be made as a result of the evaluation, who will make them and when?
- Are the conditions good enough to carry out an evaluation (i.e. a substantial interest, sufficient resources) or would it be better to avoid it? Are there any conflicts of interest which may compromise the evaluation?

How to develop evaluative questions

After the first step of an evaluation – the description of both the intervention and the 'programme theory' – the second step concerns the identification of specific evaluative questions. Whilst we can monitor every aspect of an intervention, we cannot evaluate everything: we need to focus on specific evaluative questions ranked in order of priority and relevance.

☞ An evaluator will need to elicit from stakeholders involved in the programme their hypotheses, how they intend to achieve their aims and to verify whether the connections between their actions and the intended results are plausible. Programme theory can be used to better understand the real nature of the programme/intervention and develop 'good questions' and 'testable hypotheses', which can be investigated using the variations among different contexts, circumstances and levels of implementation. As successful drug-prevention interventions occur in some circumstances and not in others, the realist evaluation (Pawson and Tilley, 1997) aims to verify 'what works, for whom, in what circumstances'. Possible variations (between interventions and within the same programme) are not considered threats to validity due to confounding variables, as a 'noise' to be screened out, but rather as opportunities to test out and focus on different intervention hypotheses.

The following questions need to be answered:

- Are the evaluative questions adequately formulated and specific? Are they relevant and responsive to the priority and the objective of the evaluation?
- Are the evaluative questions shared with those who commissioned the evaluation and main intended users (e.g. project coordinators, practitioners of drug prevention interventions, head of schools)?
- Are too-generic questions avoided, such as 'Is the project effective and efficient? How many young students have been reached?'

How to decide on the evaluation design and choose the method: a tailored evaluation

A main task of an evaluator is creating a study → design that fits the questions to be asked. An evaluation of an intervention should be tailored to specific questions, to the nature of the intervention, to time, economic or other practical constraints, as well as organisational issues. The best model of evaluation, good for every use and situation, does not exist and there are no general recipes for all the situations. Since a drug prevention intervention is not a simple treatment replicable in a broad social context, it is necessary to carefully understand the real-life condition of an intervention. Generally, mixed-method designs are preferred, as qualitative evaluation has particular strengths that often complement those of quantitative approaches. A further topic concerns the role of the evaluator and the question of ethics. One of the weaknesses of an evaluation is the potential conflict of interests of the evaluator, for example, being a member of the same institution and team responsible for the concept and planning of the intervention. People involved in an evaluation need to consider the possible negative side-effects of intervention or evaluation techniques, as well as the participants' right to privacy, respect, informed consent and knowledge about the results and how the results will be used.

☞ It is suggested to 'triangulate' several sources of information: administrative data, micro-data coming from other researches or by enforcement and health systems (i.e. illicit drug seizures data, market trends of performance-enhancing drugs and nutritional supplements, health-system records) with data coming from internal monitoring systems, socio-economic and demographic indicators and primary data collected during the evaluation. If there are standardised and continued interventions, it could be useful to analyse the performance and standard of services. For studying a shifting intervention, the evaluator should take frequent periodic measures of the programme's effect, rather than collecting data to one point in time ([1]). Between the quantitative formal designs there are one-group designs: 'After only', 'Before and after'. It is possible to extend them in two directions: more data on what happens *during* the intervention or more data on what happens *before* and *after* the intervention (i.e. → Dose-response designs, → Time-series designs). The first group of designs can be extended by adding other groups to the study (group of people, work teams, classes, communities, etc.). In this case, the designs are: Before and after with the comparison group. In qualitative designs, a range of different traditions and methods of investigation are encompassed (i.e. ethnography, participant-observation, observation, informal interviewing, case study). Experimental designs provide an important contribution to understanding the effectiveness of programmes and single interventions. However, in practice, it is often difficult to carry out experimental studies of good quality, as certain constraints emerge concerning appropriateness, ethical issues, practical feasibility or over-simplification of the intervention structure (i.e. not mentioning bias in enrolment procedures, in the selection of samples, or failing to consider relevant intermediate variables and the influence of implementation → mechanisms). It could be dangerous to give more attention, credibility and funds to an intervention only because they have been evaluated with randomised controlled trials (RCTs), regardless of their merit. The main problem in evaluation, as well as with scientific research, is the attribution of causality – interpreting the observed changes of a variable (i.e. a lower rate of last-year prevalence in smoking) as a result of an intervention. It is necessary to declare and control different types of systematic distortion (→ bias) that can threaten the validity of an intervention evaluation research. Distortions due to non-response, for example, are very common since participation and enrolment of the organisations in the intervention are not by chance, but depend upon voluntary participation and upon students' well-being undertaken by the organisation. For example, a researcher has demonstrated that among cigarette smokers, the return rate of questionnaires in a survey is usually lower than the non-smoker's rate.

([1]) Please refer to the Resources section of the online version of PERK: http://www.emcdda.europa.eu/publications/perk/resources/step1 (last accessed on 10 May 2012).

The following questions need to be answered:

- Does the study design fit with the objectives of the evaluation and resources?
- Does the evaluation employ technically adequate designs and analyses that are appropriate for evaluation purposes? Does it employ systematic information collection, review, verification and storage methods?
- Did the evaluator ascertain that accurate and valid information will be used? Will the evaluation yield sufficiently dependable and consistent information for the intended uses?
- Does the evaluation design consider human and legal rights and maintain the dignity of participants?
- Are there any serious threats to validity that may change the evaluation design?
- Were the evaluative designs and the objectives modified during its implementation?

How to interpret and use results of an evaluation

Programmes have long implementation chains – often many stakeholders and many levels of government are involved (i.e. municipalities, provinces, regions, states). Interventions do not work by themselves; they only take effect through the reasoning and reactions of their recipients and practitioners.

Recipients are many and varied, they are rooted in different localities, institutions, cultures, histories; reactions to interventions thus differ and produce mixed outcomes. Interventions are limited activities that are embedded in complex social systems and occur in wider settings. They are not hermetically sealed but occur in open systems characterised by a multitude of interventions, each one interfering with the reception of another. Also, the schemes are not stable: practitioners learn and work by improving the delivery of interventions rather than preserving uniformity to meet evaluation and trial requirements.

The following questions generally needed to be answered:

- What ideas, plans, expectations and assumptions have gone into the making of the intervention? How was it supposed to work? Have different perspectives been taken into account on how the intervention is supposed to work?

- Which assumptions have been proved correct and which have failed? How has the intervention worked?
- How did the reconstruction of the programme theory and initial evaluative results contribute to the redesign of the ongoing intervention?
- Have alternative causal explanations been analysed? Have different perspectives on how the intervention could work been taken into account?
- Have the costs of the same intervention implemented in different places been compared?
- Is there a connection between → mechanisms of access to financing, rules to allocate resources, the duration and objectivity of the intervention?
- Have the contexts and global prevention strategies and policies at local/national level been considered? Have negative influences of other sectors or potential positive impacts of more integrated local policies been identified?
- Has an evaluation considered the history of an intervention over time, the broader strategies of the organisation, the relationship between the implementation partnership and with other services and policies?

CHAPTER THREE

Chapter 3

Process evaluation and monitoring system

Process evaluation has to be designed, whether it uses a qualitative or quantitative method of investigation. It assesses the implementation of an intervention and the reactions of practitioners, different stakeholders and the participants. It describes:

- how and if the → prevention intervention took place;
- how the different resources (time, money, structure ...) were spent;
- to which extent and how the designated → target groups were reached;
- what the participation level of different stakeholders is and how different partners collaborated.

It is also concerned with the → quality of the intervention. As process evaluation collects all the relevant data about an intervention's implementation or otherwise, it provides useful information that can be used to improve future interventions and to explain why some outcomes are or are not achieved. The description of the intervention implementation and development should cover all the activities that were actually undertaken. This allows the intervention to be assessed, whether or not it was implemented as originally planned (→ adherence, → fidelity, → reinvention, → unexpected changes, → intentional changes). There is always confusion between monitoring and evaluation. A → monitoring system can be defined as an observation system aimed to verify whether the project activities are happening according to intervention plan and whether means (money, time, instrumental resources) are used in a correct and efficient manner. The monitoring system is managed by internal staff and must supply the project managers with a continuous flow of information throughout the course of the intervention to make it possible to take the right decisions. During an evaluation, it is possible and useful, if not indispensable, to use monitoring data, particularly administrative data. While evaluation is aimed to express specific judgements on the basis of criteria, and sometimes standards, a monitoring system guarantees a flow of data gathering; it is a continuous management function that aims primarily at providing regular feedback and early indications of progress or lack in the achievement of intended intermediate results.

Planning the process evaluation

In planning the process evaluation, it needs to be decided which variables and indicators should be measured. In answering the following questions, it should be noted what has been measured, how and when.

- What variables and indicators will provide useful information on how the intervention was accomplished?
- What kind of information (qualitative or → quantitative) will the process evaluation assess?
- What methods and instruments will be used (interviews, questionnaires, observation instruments)?
- Where, when and how often will the process data be collected (→ design)?
- Who will provide the information needed for the process evaluation?
- How will the data be analysed?

Monitoring the prevention intervention

This is perhaps the most important point in the life of the intervention. The description of its implementation and development should cover all the activities that were actually undertaken. This allows the intervention to be assessed, whether or not it was implemented as originally planned (→ adherence, → fidelity, → reinvention, → unexpected changes, → intentional changes). The following questions need to be answered:

- What strategies, components and methods were actually implemented? Compare the answers with the original plan (see Chapter 1, 'Methods').
- What data sources and instruments were used to measure the intervention's implementation? Compare the answers with the original plan (see 'Planning the process evaluation', above).
- What resources were actually used? Compare the answers with the original plan (see Chapter 1, under the 'Resources' section).

The target group revisited

Whether or not the desired target group was reached should also be examined. The following questions need to be answered, and information on the number of participants, their age, gender, education and other relevant variables should also be given. As with the previous subsection, the answers should be compared with the original plans, in this case the target group as envisaged in Chapter 1:

- How many people did the intervention actually reach?
- What were their socio-demographic characteristics?
- How was this information collected?

Next, how much of the prevention intervention which actually reached the target group should be identified (→ exposure). The following questions need to be answered:

- How was exposure measured? Which data sources, instruments or indicators were used?
- How long did the prevention intervention actually last and how many prevention activities took place? Compare these answers with the original plan and evaluative questions (see Chapter 1, under the 'Methods' section and Chapter 2, under the 'How to develop evaluative questions' section).
- To what extent was the target group actually reached? Compare these answers with the original plan (see Chapter 1, under 'The target group' section).

Quality of the prevention intervention

As well as examining how the intervention was carried out, how well it was carried out must also be assessed. The quality of the intervention can be identified in terms of the reactions and attitudes of the target group towards it (e.g. acceptance, degree of identification, involvement, personal benefit, etc.). The following questions need to be answered. The answers should be compared with the original views on process evaluation envisaged at the start of this chapter:

- Who provided the information on the quality of the intervention?
- What indicators and instruments were actually used to evaluate the quality of the intervention?
- What are the results of the quality measurements?

Discussing the results of the process evaluation

When evaluating the planning of an intervention, the results of the process evaluation should be analysed and interpreted. These results should be compared with those obtained by other evaluations and relevant studies, and the analysis should also include suggestions for the future. The following questions need to be answered:

- How do the plans for the intervention compare with its actual implementation and the current evaluation? Are there any discrepancies and what are the possible reasons for them?
- What is the impact of any discrepancies on the intervention?
- What are the strengths and weaknesses of the way the intervention has been implemented? Compare these with results from other interventions.
- What suggestions can be made for the future implementation of a similar prevention intervention?
- What suggestions can be made for future process evaluations of this kind of prevention intervention?

Process checklist

The intervention's process evaluation is now complete. By now, the following should be clear:

- how the 'process' will be measured;
- what actually happened during the intervention;
- how many people were actually reached;
- how much of the target group was reached; and
- how 'good' the intervention was.

The actual implementation of the intervention should also have been reviewed alongside the original plans to see how much changed in practice. The next chapter will help in evaluating the intervention's outcomes; in other words, whether it really did what was required.

CHAPTER FOUR

Chapter 4

Outcome evaluation

Outcome evaluation looks at the effects of the intervention. It deals with whether it actually achieved its intended goals and is thus an essential tool for judging whether a particular intervention is worth continuing, adapting or discarding. The → design of an outcome evaluation has a very strong influence on the quality of the results. This chapter therefore begins by describing outcome-evaluation planning before asking how the results of the outcome evaluation can be presented.

Planning the outcome evaluation

Planning any outcome evaluation must begin before the intervention starts, as the decisions made in this phase may influence the intervention's timetable and data collection. To ensure that an effective outcome evaluation is planned, the following questions need to be answered:

- What are the data and/or indicators for outcome and how will they be measured?
- Will the information on outcome be collected following a → quantitative or a → qualitative approach?
- What indicators and → instruments will be used to collect information?

The following classification may prove useful:

(1) indicators and instruments to measure → substance-use behaviour for the → ultimate target group;

(2) indicators and instruments to measure → mediating variables related to substance-use behaviour for the ultimate target group;

(3) indicators and instruments to measure other mediating variables for the ultimate target group;

(4) indicators and instruments to measure → objectives for the → intermediate target group.

Guidelines for planning and realising evaluation

- What is known about the quality of the instruments (→ objectivity, → reliability, → validity)? Will the feasibility of the instruments be tested?

- From whom, when and how often will information on outcome be collected (design)?
- How will the information gathered be analysed? Which → statistical methods are adequate for the quality of the data and the design?

Achieving the outcome evaluation

Armed with this plan, the actual outcome evaluation should then be described. This should focus on changes or adaptations not only in the sample, but also in the design and the use of instruments. Unexpected and → intentional changes should also be distinguished. When answering the following questions, the previous subsection should be considered at all times.

- What was the design of the outcome evaluation?
- What instruments were applied?
- How were data collected, by whom, when and under what circumstances?
- How were data processed and what statistical analyses were performed?

The sample

Information on the sample used to provide the data for the outcome evaluation should be produced. If the sample corresponds to everyone reached by the intervention, or even to the → target group, then one should simply refer to the description in subsections as 'The target group' (Chapter 1) and 'The target group revisited' (Chapter 3). If not, then the sample's characteristics need to be outlined, as well as the recruitment process and the → attrition level.

The following questions need to be answered:

- How was the sample recruited?
- What were the sample's socio-demographic characteristics, size and so on?
- How do these characteristics compare with those of the whole target group?
- Was it possible to identify drop-outs? If so, what were their characteristics?

The outcomes

At some point in outcome evaluation, the outcome of the intervention must be examined. The results can be presented in tabular form, as complex statistical analyses or as a written statement. The following questions need to be answered:

- How did the intervention affect the ultimate target group's substance-use behaviour?
- How did the intervention affect mediating variables related to substance use in the ultimate target group?
- How did the intervention affect other mediating variables in the ultimate target group?
- How did the intervention affect objectives in the intermediate target group?
- Are different subgroups (e.g. men/women, age groups, risk groups, etc.) affected differently by the intervention?

Discussing the results of the outcome evaluation

Finally, the outcome evaluation's results need to be analysed and interpreted. As with process evaluation, these results should be compared with those obtained from other evaluations and relevant studies, and suggestions for the future should be made.

Guidelines for planning and realising evaluation

The following questions need to be answered:

- Did the intervention achieve the expected outcomes?
- What are the most relevant and significant results? Compare these with results from other studies.
- Is it certain that the intervention caused the results? Are there any alternative explanations for them? Are the study limitations declared?
- What explanation is there for negative results?
- Are there any other positive or negative unforeseen impacts? Are more general recommendations provided to policymakers?

- What suggestions can be made for the future use of similar interventions?
- What suggestions can be made for future outcome evaluations of this kind of prevention intervention?

Outcome checklist

The outcome evaluation is now complete and the entire evaluative process is nearly finished. By now, the following should be clear:

- how it was planned to measure 'outcome';
- how the outcome evaluation was actually conducted;
- from whom the outcome information was gathered;
- whether the intervention had any effect on target-group behaviour; and
- whether the intervention actually achieved its purpose.

The intervention has now been evaluated. The next stage is to publicise that fact and share the experiences gained. The best solution, when possible, is giving ongoing feedback about evaluation results to the programme-intervention staff, since they can better use recommendations and insights offered by evaluation and when necessary, decide on timely modifications of the intervention in order to address intended effects.

PART **TWO**

Examples

Part 2

Examples

The following section provides background information on all the questions asked by the guidelines, as well as examples of how to answer these questions. The quotations in some of the examples are taken from prevention interventions in which the guidelines were applied during the feasibility phase, and most have been summarised or shortened because of space considerations. Given the diversity of possible interventions, these examples cannot be truly representative. Instead, they are used here to highlight the wide range of possibilities that may be encountered when evaluating a project.

CHAPTER FIVE

Chapter 5

Communicating the results

The evaluation is now complete, but the work is not yet over. The use to which the conclusions will be put must now be considered.

Developing a communication plan

Evaluations can be conducted for many different reasons, but one of them should always be to provide a basis for future decision-making. Certain steps should be considered to ensure maximum use of the evaluation. The following questions must therefore be answered:

- Who should be 'in the know'?
- When do they need the information?
- What information will interest different people?
- Which forms of written communication should be used?
- Which forms of oral communication should be used?

CHAPTER SIX

Chapter 6

Assessing intervention planning

The phenomenon

What phenomenon is the intervention designed to prevent?

The phenomenon that should ultimately be addressed by a drug-prevention intervention is that of → substance-use behaviour in the → ultimate target group. This can refer to legal as well as illegal drugs, and holds true even if the phenomenon is only addressed indirectly, as in life-skills training or mass-media campaigns.

> *'We want to prevent the use of drugs with a major focus on alcohol as a legal drug and heroin as an illegal drug among people aged between 10 and 18.'*
>
> *'We want to prevent alcohol consumption at the workplace in a company with 750 employees.'*

What are the socio-demographic characteristics of those affected by the phenomenon compared to those who are not?

The most important socio-demographic characteristics are gender, age, race, socio-economic status and area of residence. Depending on the planned intervention, other relevant characteristics could be → risk factors, → protective factors, personality traits and so on, and characteristics may vary for different drugs or different localities. It is crucial to know what these characteristics are, as only then can the intervention be adequately focused on the → target group. However, as there could be dozens of group characteristics, it is best to focus on those that appear most relevant and are the best documented.

> *'According to a regional survey, there are four male drug users to every one female. The average age of all drug users is 27.6: 77 % are unmarried, 21 % have a high-school diploma and 45 % have a stable job.'*
>
> *'Clubbers who regularly use ecstasy are between 18 and 23 years' old. They either tend to be in regular employment or they go to school or university (over 80 % of them are students).'*

Where does the phenomenon occur, and where does it not?

Exactly where the substance-use behaviour, which the intervention aims to address, takes place must be described. Does it happen, for example, at home alone, with friends, at nightclubs, on the street or in the countryside? It is equally important to know where it does not happen, so that the intervention can be correctly targeted. The answers to these questions are as essential as knowing who is affected by the phenomenon as, without them, the intervention will be unfocused.

> *'Alcohol consumption in our community occurs in private as well as public places (bars, discos, streets). Heroin is consumed mainly in private, but also in public parks. Both seldom occur in school.'*
>
> *'In general, non-urban areas have lower levels of drug use.'*

How long has the phenomenon been known? Have its size, impact and relevance changed over time?

It is important to be able to make predictions about the future development of the phenomenon. Furthermore, an increase in the phenomenon justifies a more rigorous prevention intervention.

> *'After increasing rapidly in the last decade, heroin consumption seems to be stabilising. There is also a considerable increase in the use of ecstasy at clubs and "raves". Drug use is no longer an essentially urban phenomenon and, according to recent field research, poverty and insecurity are more concentrated among drug addicts.'*

Finally, throughout, sources of information should be described and the quality both of the information and of the information sources noted. National surveys, literature reviews, other journal articles, personal surveys and so on can all be used: The 'UK national survey of schoolchildren's drug use, and local surveys in the UK. These findings are broadly supported by research from other similar cultures, especially the USA and Australia.'

The conceptual background

Which explanation best describes the origin of the phenomenon?

There might be more than one possible theory or explanation for a phenomenon, so the one that appears most relevant to the planned intervention should be outlined. Why this is the preferred theory should also be explained.

'There are many theories to explain teenage substance use. Although all of them have some merit, the most widely accepted are the theories of social learning and interaction. The main assumption is that teenage behaviour is heavily influenced by parents and peer groups, and from this viewpoint, the first experience with drugs is the result both of → exposure to consumption-favouring models and of a vulnerability to social pressure.'

'... Alcohol and other drugs (AOD)-related harms do not occur equally in society, societal factors contribute to particular categories of people disproportionately engaging in risky AOD behaviours and experiencing AOD-related harms. These factors include culture, social capital, employment policies, welfare policies, education policies and urban planning policies.'

What factors are responsible for the continuation of the phenomenon?

In some cases, the factors that cause a phenomenon are different from those that maintain and exacerbate it. It can sometimes, therefore, be more important to focus on these maintaining factors, especially as it will often be too late to affect the causal factors.

'The attitudes, habits and → norms among adults towards alcohol and drug consumption are influencing factors for young people. They are also affected by the media.'

'There is an attitude of "no future" among young people in our area, which results partly from high unemployment. This negative self-image as a community is also affected by society's permissive social norms towards drug use.'

Examples of strategies and activities which could change (or which have changed) the phenomenon should also be described. Any specific successful interventions that have been conducted should be described in as much detail as possible. Even without knowledge of a similar successful intervention, the elements and features considered necessary for a particular type of intervention can still be mapped out.

'The implementation of primary prevention activities on a regional or national level is expected to influence the development of the phenomenon. Activities which could slow down the increase in drug use include: public-information campaigns; school-based prevention programmes; the provision of information to pre-existing and especially established parent groups; and programmes aimed at school leavers.'

As the initial step of evaluation, has the intervention been described in a detailed way?

An initial step of the evaluation process is to describe the programme in detail. This collaborative activity can create a mutual understanding of the programme, the evaluation process, and programme and evaluation terminology. Developing a programme description also helps ensure that programme activities and objectives are clearly defined and that the objectives can be measured.

> *'The EmPeCemos programme is made up of three components. The family component is a parent-training programme composed of twelve sessions. (...) Basic contents are: promoting positive behaviours and the bonding between parents and children, promoting school adjustment, limit-setting and managing disruptive behaviours. Additionally, some sessions are devoted to communication, self-control and problem-solving skills. Besides promoting a good family climate, these sessions support children's cognitive and emotional development. The children's programme also consists of twelve sessions to be applied along with the parents' component.'*

Does the reconstruction of the 'programme theory' include either accumulated scientific theories and conjectures or explicit hypotheses of planners? How does the evaluator select intervention theories?

The first step is to reconstruct the assumptions at the base of the architecture of the intervention and the conjectures guiding the choices during the implementation. It usually starts with an assumption about what will work – sometimes called a → theory of change – but those assumptions are no guarantee. Sometime we can commence by reading an official project document and 'declared' theory of reference.

> *'The cannabis Show intervention is based on a reasoned selection of several behavioural theoretical models. The theories or model used are: the Stage of Change Model, the Health Belief Model, Attribution Theory, Social Learning Theory and theories of planned or reasoned behaviour.'*

> *'The interactive, bi-directional approach to programme theory development described above strives to produce parsimonious programme theories that can be used to develop a common understanding of the programme amongst stakeholders, and to help identify and prioritise evaluation questions (Donaldson, 2003, p. 361). While we believe it is important to acknowledge some programmes and their effects may be very complex (...), at least a version of a programme theory should focus on the main pathways between a programme and its presumed outcomes.'*

How accurate is the reconstruction of the 'programme theory'?

> *'Concept mapping is a technique that combines qualitative data with quantitative analysis. (...) The overall question drafted for the concept mapping process was, "In what way can/does the Youth Centre contribute to prevention?" This question was thought to be general enough to encompass the Youth Centre's overall global approach to prevention. At the initial meeting, respondents were asked to brainstorm and express their thoughts and ideas on the theme question in the form of short statements.'*

The reconstruction of the 'programme theory' should not be limited only to the 'logical frame' and implementation aspects and it needs to consider also why, and thanks to what → mechanisms, some behaviours and attitudes should change.

> *'Conceptual theories linking youth tobacco use etiology to proposed mediators:*
>
> - *Refusal skills = youth smoke because they do not have adequate skills to resist interpersonal and media pressures to smoke;*
> - *Social norms = youth who perceive their immediate social environment as tolerant of smoking are more likely to smoke;*
> - *Perceived prevalence = youth smoke because they believe "everybody does it";*
> - *Tobacco attitudes = smoking is determined by a motivation to experience the benefits and avoid the consequences of smoking;*
> - *Availability of tobacco = with less access to tobacco, youth will be less likely to smoke.'*
>
> *'The analysis of the impact of drug testing focused on the change in substance use. The analysis of the impact of arrest referral was organised around the causal → mechanism through which arrest referral was expected to affect the behaviour of young people. The causal → mechanism is:*
>
> – arrest referral improves referral to and access of services;
>
> – access to services, reduces the risk of offending and substance use; and
>
> – improving the factors that influence a young person's risk of offending will reduce the chances that a young person would reoffend.

Restrictions on the research design and the availability of data meant that several different research designs and combinations of data were used to undertake different elements of the analysis.'

The need for a prevention intervention

How many people are affected by the phenomenon? How many new cases are there and how often do they appear (→ prevalence, → incidence)?

In order to estimate the relevance of the phenomenon, it is necessary to gauge its size and the number of people affected by it. This estimate can either be a general, national one, or on the more specific community level, depending on the scale of the proposed intervention. See → needs assessment for more on how to obtain the necessary information.

'In Spain, (...) the majority of drivers (70 %) are regular drinkers (survey of 12 000 drivers by Alvarez et al., 1995), 51 % of them recognise to have driven after drinking in the past year, 7.3 % drive always, no matter if they have drunk and 14.3 % has done it in a "drunken state". (DGT 2000, 10).'

How is the phenomenon likely to develop if nothing is done? On what grounds?

It is important to have a well-founded idea of the future development of the phenomenon if the intervention were not to take place. Essentially, this means ascertaining whether the phenomenon would get better or worse and whether consumption patterns or the user group would change if nothing were done. Without such answers, the relevance of the planned intervention remains unproved.

'Young people's substance use is likely to continue to rise. Alcoholic "soft" drinks are now widely available and it is generally believed that under-age drinkers are the target of this marketing exercise. Research has indicated that the use of a substance from a young age predicts the future use of other drugs.'

How should the need for the intervention be described?

Apart from the number of people affected by the phenomenon, the reasons for undertaking the planned prevention intervention must be identified.

'..In the past decade, our understanding of alcohol and other drug (AOD) issues has grown. It is better recognised that AOD use is not always harmful and can, in fact, be beneficial (for example, in terms of social bonding). As such, the focus has shifted from AOD use per

se to risky or harmful use behaviours (such as sharing injecting equipment or driving while intoxicated) and to specific AOD harms (such as HIV or alcohol-related road accidents).'

Are there different opinions as to the need for an intervention (→ varying perspectives on need)?

Different sources may have differing views on whether or not a prevention intervention is needed. These differing perspectives can cause problems. In each case, the target group, policymakers and professionals must all agree on the need for the prevention activity. Without full agreement, a particular group with a particular agenda – a local politician, for instance, coming up for re-election – could 'hijack' the intervention for their own purposes.

> '*Local project workers, other prevention and treatment experts and policymakers* all agree on *the need for this prevention intervention.*'

> '*During the planning phase, the project planner identified varying perspectives on the need for a prevention intervention. These differences are due to the particular field (alcohol use) and the national profile of some of the experts.* → *Cultural habits differ with regard to the use of alcohol in social situations and the amount and frequency of consumption.*'

How was the need for the intervention assessed (needs assessment)?

As different techniques can yield different results, the methods must be specified carefully.

> '*We consulted key people with many years' experience in drug-prevention work, education and teacher training. Other than themed discussions, no formal needs analysis was carried out.*'

Are any related interventions being carried out or planned? Will the current intervention cooperate with these activities?

In order to avoid unnecessary overlap and duplication, it is useful to find out about existing or planned prevention activities in the same area. Of course, this could also be helpful for exchanging ideas and experiences.

> '*There have been some attempts to use the Internet for prevention, but our planned intervention will be the first systematic effort in the country.*'

'There are two other mass-media drug campaigns, one targeting the general population and the other, teachers. As our prevention campaign addresses young people, there is no overlap.'

The target group

Is the target group the ultimate target group or an → intermediate target group?

This needs to be clarified or the intervention will not be clear.

'We are focusing on an intermediate target group (teachers).'

'The project target groups are pupils (ultimate target group) as well as teachers and parents (intermediate target groups).'

What are the socio-demographic characteristics of the target group, the scale of the phenomenon and the size of the group?

These characteristics include age, gender and race, as well as socio-economic status and living environment. Any other target group characteristics relevant to the intervention should also be identified, the phenomenon located within the group itself and the size of the group indicated. If there is more than one target group, this exercise should be undertaken separately for each.

'The target groups live in a socially and economically deprived area – A community with almost no infrastructure, on the outskirts of a small town, and separated from green and open space by a motorway. Some 150 families live in this so-called "social residential quarter" and drug use is rife.'

Why was this target group chosen?

The choice of a particular target group might be influenced by theoretical considerations (it could be a high-risk group), more practical considerations (it was an easy group to reach or it was highly motivated) or a combination of the two.

'This target group was chosen because the incidence of drug use is high and because the presence of a village hall gave the opportunity of reaching the families via the community channel.'

'We chose fourth and eighth graders as the ultimate target group because of their high levels of heroin use. The intermediate target group (the local media) was chosen because it wanted to participate and because of its role as opinion-former.'

How many people will the intervention reach?

The answer to this question is crucial. Goals for the intervention must be set, and they need to be realistic. If the potential target group is very large, as for example in a regional or national intervention, it is better to give a rough estimate of the number of people that will probably be reached, and then attempt to fulfil that estimate, than simply to say 'we want to reach as many as possible'.

> *'All 150 families in the community, about 500 people.'*
>
> *'We expect to reach between 500 and 5 000 people in the first six months. Due to the nature of the Internet however, predicting the exact number of people visiting our site is difficult.'*

Where and how will the target group be contacted, recruited and motivated (→ selection effects, → coverage, → bias)?

An intervention which no one notices is a waste of time and money. It must be recognised that a target group does not necessarily have any more motivation to participate in an intervention than a non-target group, and so – above all else – one should be concerned with how to motivate potential targets to seek out the intervention and participate in it. Possible selection effects, bias and under-coverage (or even over-coverage) should also be taken into account.

> *'In our workplace programme, the director will contact all employees directly by letter as well as by placing a notice on the bulletin board. They will be entitled to time off work for the duration of the programme.'*
>
> *'Children and young people (the ultimate target group) will be approached in school, leisure centres, youth clubs, police stations and voluntary organisations. Parents (the intermediate target group) will be approached through the school, at work and through information in the media.'*

How can it be guaranteed that the target group will stick with the intervention (→ attrition)?

As with initial motivation, continued participation cannot be guaranteed – it is unlikely that contracts of obligation will have been signed with all participants, and so there is nothing to stop people walking out of the intervention. The risk of a high attrition rate has to be borne in mind and some attempt should be made to address it or to encourage people to stay with the intervention.

'Teachers will be offered an accredited certificate in drug-prevention education upon satisfactory completion of the course.'

'The school programme is mandatory for all girls in ninth grade.'

'The design and management of Searching Family Treasure is based on principles of efficacy for family-based prevention programmes (...), namely: focus on family; length of the programme; the existence of strategies to keep the families on the programme (giving incentives); include all the family, intense support (...). The programme has evidence-based content, and uses strategies indicated in the literature to ensure high rates of recruitment and retention.'

Even if the planned intervention exclusively addresses an intermediate target group, what are the characteristics of the ultimate target group?

This is important if substance-use behaviour is not being addressed directly, as it can help keep the final aim of all prevention in sight.

'The prevention intervention addresses secondary-school teachers as an intermediate target group. Their target group is pupils in sixth grade (aged 11 and 12), most of whom do not yet consume illegal drugs, although they are at risk from doing so.'

Objectives

How will the intervention affect substance-use behaviour and related risks in the ultimate target group?

An idea of how the intervention will affect substance-use behaviour is necessary. Even if it does not directly target such behaviour, it should ultimately affect it. In this case, it is more important to describe the → objectives that are expected to affect the → mediating variables or the intermediate target group.

Two kinds of mediating variables are usually distinguished:

- mediating variables that are directly related to substance use and demand, such as attitudes towards drugs, intention to use drugs;
- mediating variables that are indirectly related to substance use referred to both the supply and demand side, such as life skills, risk and protective factors, lifestyle, cultural habits and problem behaviour, peers' drug consumption rates, drug availability and drug prices.

Even if an outcome evaluation of the intervention is not planned, there should still be some objectives. These objectives might include preventing children from using drugs, postponing the age of onset of drug use, reducing the amount and/or frequency of substance use, and so on.

> *'The intervention's objectives are: to decrease regular/excessive consumption of alcohol; to decrease the number of young people regularly smoking tobacco; and to postpone first experimentation with drugs.'*

> *'The indirect objective of the project (which is aimed at improving family functioning) is to reduce substance use and related risks.'*

> *'The expected outcomes are: 1. measurable and sustained reduction in the initiation of drug and alcohol use among children and adolescents; 2. A reduction in drug- and alcohol-related deaths and injuries, especially among children, adolescents, and young adults; 3. A decline in the prevalence of health problems related to or exacerbated by alcohol and drug abuse; 4. A reduction in on-the-job problems and accidents related to alcohol and drugs; 5. A reduction in drug-related crime.'*

How will the intervention affect mediating variables related to substance-use behaviour in the ultimate target group (→ knowledge about substance use, → attitudes towards drugs, → intention to use drugs, normative beliefs)?

Apart from the direct influence on substance-use behaviour itself, the intervention might also aim to change mediating variables related to substance-use behaviour. Primary prevention interventions will often focus more on these mediating variables than on substance-use behaviour itself.

> *'The intervention aims to correct normative beliefs about cocaine and performance-enhancing substances' consumption among students of 15–18 years old in the District of ...'.*

What objectives are considered for other mediating variables (→ life skills, risk factors, protective factors, → problem behaviour, → structural changes, changes in → lifestyle and cultural habits)?

As well as mediating variables that are directly related to substance use (knowledge, attitudes, etc.), any intervention can also have profound effects on more general variables. This should be remembered at the start of the planning phase.

> *'The intervention will improve problem-solving, enhance communication skills, strengthen self-esteem and encourage creativity.'*

What is the relationship between these mediating variables and substance-use behaviour?

In order to affect mediating variables, the relationship between these variables and substance-use behaviour must be explained. The question everyone will be asking is how can giving someone a leaflet about drugs change substance-use behaviour or make an individual more self-confident?

> *'Autonomy, good affective relationships and social support from family and friends are considered to be protective factors against substance use.'*

What are the objectives concerning the intermediate target group?

If the intervention covers an intermediate target group, this section should effectively recapitulate the last four questions, but for this group, not for the ultimate target group.

> *'For teachers, the intervention's objectives are as follows: to increase awareness and understanding of substance use; to develop understanding of effective prevention and educational strategies; and to develop appropriate skills for working in this field (e.g. teaching, communication and counselling skills).'*

> *'Improve the relationship between parents and their children.'*

How are the objectives for the intermediate target group and the ultimate target group related?

The relationship between the objectives for all the target groups must be identified. It is essential to understand how changes in the intermediate target group will affect the ultimate target group. Without a clear rationale for the choice of these objectives, their implications for drug prevention will remain unclear.

> *'Parents play an important role as models for the behaviour of adolescents. To affect smoking behaviour in this ultimate target group, parents' behaviour must be influenced as well.'*

Methods

What strategies, components and methods will be used in the intervention? How were they adapted to organisation and local contexts and needs?

This is the core of the plan – the detailed description of the proposed prevention activities. It is better to avoid a mechanic transposition of an intervention model from one context to another and

considering the issue of feasibility and time constraints together, by the practitioners responsible for the implementation.

> *'The major strategy for facilitating adoption of Life-Skills Training by the Adoption of Drug Abuse Training (ADAPT) Project has been to assists teachers with the infusion of LST into their normal teaching.'*

> *'The educational method used by the health education programme is that of "active learning" which aims to support and enable the development of a student's basic skills. It will be implemented in the three grades of elementary school and will cover three topics (tobacco, alcohol, illegal drugs). Methods include class discussions, group discussions, → interviews, research/surveys, writing, dramatisation, role play, films, etc.'*

> *'The intervention is designed for school teachers and drug educators to increase their knowledge of the most commonly used substances, to improve their communication skills and to enhance small group teaching and counselling skills. It is a four-module programme resulting in a certificate in drug prevention and education. The modules are based on the relevant literature and research.'*

Who will be involved in the intervention?

This refers not only to the people who will carry out the intervention, but also to those who will communicate the intervention's message to the ultimate target group. If the intervention is aimed specifically at the ultimate target group (schoolchildren, for example) then perhaps – although by no means always – the people who are conducting the intervention will be the only ones involved in it. However, if an intermediate group is targeted, it is likely that they will in turn become involved in communicating the intervention's lessons (local radio stations, for instance, running features on substance use, teachers giving drug-education lessons). All those expected to be involved in the intervention should therefore be covered.

> *'Teachers and parents will act as the main agents for change/intermediate target group.'*

Is there any empirical evidence for the success of the methods chosen (e.g. scientific literature, research papers)?

If similar interventions (or elements of them) have already been tested in another country or another setting, this should be pointed out and the relevant results summarised.

'There is no direct empirical support for the success of prevention via the Internet, but in some related areas, especially "telemedicine", studies on client satisfaction have been conducted. These studies show that in some cases, the patients even prefer communicating with a computer than with a practitioner.'

How long will the intervention last?

As well as knowing the who, how and where, it is equally important to know how long an intervention is expected to last.

'The programme director and policymakers have agreed on five years with the option of extending it for another five.'

What is the planned timetable of the intervention (number of activities, duration and frequency of each activity, etc.)?

The duration of each of the intervention's elements should also be described in more detail.

'Three community group meetings for two hours each; 12 working-group meetings for two hours each; five to eight additional seminars for parents. For the ultimate target group: fourth-graders – eight one-hour lessons each year; eighth-graders – four one-hour lessons each year.'

'Weekly classes (one hour each) for one year, making a total of about 40 weekly classes.'

Will the feasibility of the intervention be tested?

Before starting a new intervention or a tried-and-tested old one in a new field, whether the intervention will be accepted by the target group and whether it is likely to achieve the expected results should be examined. In a mass-media campaign, for example, the materials that are to be used may be 'test-driven' within a 'focus group' to discover their reactions. Alternatively, some teachers might be asked for their views on a manual for a school project, or a full 'dry-run' of the intervention may even be undertaken.

'We plan to use a "drug-information suitcase" filled with material related in some way to drugs or drug-use behaviour, such as brochures, texts about drug use, and so on. We plan to test out this suitcase with several people without any prompting and to ask them for their opinions.'

'We want to ask 25 people about a poster on risk-taking behaviour which we have designed to be distributed to schools. People will be asked to describe what they see

on the poster, what they think the most important message is, and whether they think it suitable for this purpose.'

Resources

Which staff members will carry out the intervention, and what qualifications are required?

Only the people directly involved in the prevention project should be listed. All additional resources should be described in the following sections.

'Major project implementer – formal characteristics: psychotherapeutic training, professional experience in working with adolescents. Informal characteristics: credibility, grew up in community where prevention intervention takes place, able to identify with young people.'

How time-consuming will the intervention be for each of these staff members?

It is important to plan realistically to avoid overloading those involved in the intervention. One should always be aware of the pitfalls of underestimating the time needed to carry out the intervention.

'Project leader: two hours per day over a period of two years. Project implementer: 100 %, i.e. 40 hours per week.'

What is the budget and who is providing it?

The overall budget, as well as the source(s) of funding, should be given. For example:

'Between EUR 200 000 and 300 000. Funded by the National Institute of Health.'

What additional resources are available (e.g. people, organisations, rooms, materials, etc.)?

These resources should be identified, as they can provide invaluable support.

'An office is available for drug prevention and counselling, fully equipped with personal computer, telephone and fax machine.'

'Audiovisual and projection tools (television displays), computer.'

What could hinder the implementation or evaluation (→ barriers)?

It is important to anticipate potential obstacles in order to find ways to avoid them if and when they occur.

'Potential barriers: language difficulties, time-consuming travel, lack of financial support and the difficulties professionals may have in finding enough time to participate fully.'

Planning the process evaluation

Is a process evaluation planned?

A simple 'yes' or 'no' answer will suffice. If the answer is 'yes', more in-depth answers will be needed at a later date (see Chapter 3).

What resources are available to undertake such a process evaluation?

Whether the necessary financial and personnel resources are available to perform the process evaluation adequately must be assessed. The costs in terms of time and money are often underestimated. It is advisable to reserve 10 to 30 % of the financial resources available for a prevention intervention for process and outcome evaluation.

'The evaluation will be performed by our unit's research team.'

Who will carry out the process evaluation?

The name of the organisation or individuals who are likely to carry out the evaluation should be given. Their role (internal or external to the intervention) and their formal and informal characteristics (member of staff, qualifications, etc.) should be described.

'Psychologist – a member of our staff will assist him. An external → evaluator will carry out the process evaluation. Contacts already exist at the Mental Health Research Institute of the local university.'

Planning the outcome evaluation

Is an outcome evaluation planned?

A simple 'yes' or 'no' answer will suffice here too. If the answer is 'yes', more in-depth answers will be needed at a later date (see Chapters 2 and 4).

What resources are available to undertake the outcome evaluation?

Whether the necessary financial and personnel resources are available to perform the process evaluation adequately must be assessed. The costs in terms of time and money are often

underestimated. It is advisable to reserve 5 to 8 % of the financial resources available for a prevention intervention for process and outcome evaluation.

> *'The outcome evaluation will be performed by the research team of the cooperating local university.'*

Who will carry out the outcome evaluation?

The name of the organisation or individuals who are likely to carry out the evaluation should be given. Their role (internal or external to the intervention) and their formal and informal characteristics (member of staff, qualifications, etc.) should be described.

> *'There will be an external evaluator from the university research team.'*

Reflecting on the planning phase

Who was involved in the planning phase?

Whose ideas and wishes were taken into account when planning the intervention should be considered. This could include the team implementing the prevention intervention, the target group (intermediate/ultimate), policymakers and researchers.

> *'The project was set up as an "action research model". The first step was to assess the need for a prevention intervention among the target group. To this end, we interviewed families in the community, asking them where they expected to see interventions (in the school, family, community, etc.), how effective they expected these interventions to be, and what they felt the key areas of an intervention should be. Thus, the target group was directly involved in the planning phase, along with the research team itself, consisting of two psychologists and two social workers.'*

What is the overall assessment of the planning phase process?

If any difficulties arose which should be addressed differently in the future, they should be noted now.

'Some aspects of the planned intervention were not sufficiently thought through. During planning, it became apparent that we had to have more concrete ideas on how to transfer the benefit of the intervention from our intermediate target group (teachers and social workers) to the ultimate target group (schoolchildren).'

CHAPTER SEVEN

Chapter 7

Designing and carrying out a theory-driven evaluation

The evaluation plan is similar to a research project and demonstrates how to collect and analyse information generally, while the intervention is operating and sometimes after its end. The information helps the individual to know what to change during the intervention, to help people to reach the intervention goals (formative evaluation). One can also use the information to answer evaluation questions at the end of the intervention (summative evaluation).

☞ The theoretical framework has to be explicit, since it allows us to formulate assumptions with regards to the relationship between an intervention and its effects, identifying possible → mediating variables, otherwise the selecting process and outcome indicators would be entirely random.

Objectives and use of the evaluation

Have the evaluation objectives been specified and are the expectations realistic? Have the purposes and different audience for the evaluation been clearly specified?

Sometimes, particularly when the objectives are not clear and negotiated, evaluation results can become questioned by commissioners or other stakeholders.

> *'Much about the evaluation is questioned by some of the stakeholders: (...) whether telephone surveys are a reasonable method for collecting data on illegal personal behaviour; whether there was adequate implementation research to explain why no impacts were found.*
>
> *Equally significant questions are raised – usually by other stakeholders – about the intervention itself: Was the intervention a model that could be evaluated in the ways that demonstration programmes have been evaluated in the past? Was it a powerful enough intervention to expect measurable reductions in alcohol and drug use?'*

How the evaluation results were expected to contribute to the decision-making process?

The evaluator needs to understand what kind of decisions about the intervention will be made as a result of the evaluation, who will make them, and when.

'Having demonstrated its appeal to youth in the various locations, the purpose of the evaluation was twofold: (a) to identify and describe those components or factors that have a significant impact on the preventive nature of the Youth Centre and (b) to generate pertinent information to be used in planning the new Youth Centre. The transitional nature of the Youth Centre had to be taken into consideration in selecting the specific approach for this evaluation. In essence, this evaluation involved a programme that would no longer exist in its present format when the results would be available. (...) These, in turn, could be used in the planning and organising of the new Youth Centre. It also could contribute to developing a theory of global prevention in community settings.'

'Project Freedom carefully evaluated the effects of an intervention designed to reduce sales of alcohol and tobacco to youth. In-depth analysis of all changes brought about by a coalition is time and cost intensive, however. Accordingly, care is needed in selecting potentially powerful changes in programme, policy, or practice for intense study.'

Are there good conditions to carry out an evaluation (i.e. a substantial interest, enough resources) or would it be opportune avoiding it? Are there any conflicts of interests that may compromise the evaluation?

'This manuscript describes a case study using a comprehensive system for evaluating community coalitions. This research was part of a collaborative relationship between a university research team and staff of Project Freedom, a substance abuse coalition in Wichita, Kansas (...).'

'Initial financial support for the coalition came from the Wichita Public Schools (...), and planning and implementation grants from the Kansas Health Foundation. Project Freedom was composed of nearly 100 organisations and over 750 individuals with an interest in reducing use and abuse of alcohol and other drugs.'

How to develop evaluative questions

Are the evaluative questions adequately formulated and specific? Are they relevant and responsive to the priority and the objective of the evaluation?

'Four research questions were identified by the second evaluation team: How successfully was "Fighting Back" implemented? Was the demand for alcohol and drugs reduced?

Was harm due to alcohol and drug abuse reduced? To what extent did the project generate fundamental and sustainable system change?'

'The evaluation system yields a variety of information, including data for five key questions of importance to the coalition and its funding agents: (a) Was the community mobilised to address substance abuse (process)?; (b) What changes in the community resulted from the coalition (outcome)?; (c) Is there a change in reported use of substances by youth (outcome)?; (d) Does the coalition have a community-level impact on substance abuse (impact)?; and (e) Is community-level impact related to changes facilitated by the coalition (impact)?'

'Threats of sanction do not act in the same way on all populations and under all conditions (i.e. at all times). (...) Can the fact that the first interview takes place in the Local Health Unit (and not in the Territorial Government Office, as foreseen by the legislative measure) influence the effectiveness of the intervention? Is there a different way in which the young drug consumer elaborates and reinterprets his/her experience depending on the context of the interview?'

Are the evaluative questions shared with the commissioner and main intended users (i.e. project coordinators, practitioners of drug prevention interventions, heads of schools)?

'The evaluation team engaged stakeholders in a participatory process that involved constructing models of how their programmes work, and then used those models to guide question formation, data gathering, and evaluation. To achieve this, the evaluation team facilitated numerous meetings and discussions of the programme models and theories of change, evaluation design, data collection methods, feedback loops, and evaluation reports.'

'Specific attention was given to identifying and measuring realistic outcomes and indicators that would result from specific programme efforts. In addition, each evaluation was tailored to answer questions deemed of most importance by each project team.'

Are too-generic questions avoided such as 'Is the project effective and efficient? How many young students have been reached?'

The evaluation plan should also include evaluative questions about implementation but it does not consist only of answering the simple question 'Did this happen?', gathering evidence about each of the components in the logic model.

'More specifically, the study seeks to answer the following two questions: (1) What types of organisational characteristics are related to successful use of each of the programming processes (i.e. planning, implementation, evaluation and sustainability) that are part of comprehensive programming frameworks?; and (2) What are the similarities and differences in the organisational patterns correlated with successful use of each of the programming processes?'

How to decide on the evaluation design and choose the method: a tailored evaluation

If evaluators intend to carry out a tailored evaluation whose results are to be 'useful', it is necessary to comprehend the real nature of an 'intervention'; it is useful to know not only 'what' the intervention is expected to achieve but also 'how' it is expected to achieve it.

Does the study design fit with the objectives of evaluation and the resources?

'... the management of the institution decided to undertake an evaluation that would be carried out and managed regionally. The research proposed was inspired by D. T. Campbell and consisted of using case studies to test the hypothesis on which the institution's policy towards companies was founded. Given that this form of inquiry is quite demanding, it was only possible to work on a limited number of case studies.'

Generally, a programme adopts an intervention model, for example Life-Skills Training (also LST) without any variation, but consistent variations sometimes can occur aimed to improve the original scheme or adapting it to new contexts. In these cases, the evaluation design can compare the two schemes: the original one and the modified scheme (intervention 1 versus intervention 2).

'... teachers were trained in the Life-Skills Training (LST) approach by an approved LST trainer. In the I-LST, schoolteachers were trained by members of the ADAPT staff in LST principles. The comparison schools did minimal drug prevention education during the period of this study. Implementation fidelity and student dosage was monitored using teacher self-rating forms that were filled out every lesson delivered. These forms indicated the lesson implemented or the topic covered (...), as well as the list of students in attendance.'

Does the evaluation employ technically adequate designs and analyses that are appropriate for the evaluation purposes? Does it employ systematic information collection, review, verification, and storage methods?

'The following methods were used to address the evaluation questions:

1. *individual interviews with different responsible members of the services and four group interviews with the social workers and the staff of practitioners from four different locations: three Local Health Units and one Territorial Government Office (also TGO);*
2. *analysis of the monitoring data provided by the TGO itself (database with micro-data of 5 000 subjects) and by the local health units with reference to the two-year pilot programme;*
3. *survey through a semi-structured telephone interview with a sample of 100 young people coming from the whole province of Milan who had been interviewed in the previous two months (to compare the pilot to the old programme);*
4. *three focus groups with the beneficiaries of the treatment provided by the different Local Health Units (while noting that the treatment follows interviews in only 10–14 percent of cases).'*

'Schools were selected on the basis of inclusion criteria and of willingness to cooperate, and randomly assigned to either intervention (102 schools) or control group (68 schools). Of these schools, 77 intervention schools and 64 control schools continued participation throughout the study.'

Did the evaluator ascertain that accurate and valid information will be used? Will the evaluation yield sufficiently dependable and consistent information for the intended uses?

'Observers were trained to complete ratings following the same guidelines as the programme specialists. Unlike programme specialists, however, observers made ratings for each major sub-activity of each session taught.'

Does the evaluation design consider human and legal rights and maintain the dignity of participants?

Some of the most well studied side-effects are the labelling → mechanism related to selecting procedures of most at risk pupils and negative effects of knowledge oriented and of scaring tactics of drug-prevention interventions.

'For YOT (youth offending team) workers, treatment providers and court staff, the key issue when assessing suitability for an intervention was the level of motivation a young person was assessed to have for addressing substance use issues and engaging with the requirement. (...) Legal advisers in some sites described how there were different consent requirements depending on the age of the young person, with younger individuals requiring parental consent in addition to their own.'

'The lack of suitable accommodation for assessments reported during early implementation was not found during mainstream implementation. Locations used for assessments included arrest referral premises, agency rooms, community settings, custody suites and young people's homes.'

Are there any serious threats to validity that suggested to change evaluation design?

It is necessary to declare potential threat to reliability and validity of data and design of research and indicate corrections that have been introduced.

'Life-Skills Training, in particular, may have selected a biased sample in the six year follow up in which only those subjects who had received 60 % or more of the programme (high fidelity sample) were included.'

'Finally, the lack of scientific control in the field-based evaluations could pose a threat to the reliability and validity of the data. Given these potential threats, (...) the investigators employed several procedures to enhance data quality, including (a) use of standard evaluation tools in the NOT curriculum; (b) provision of facilitator training on evaluation and the use of programme evaluation tools; (c) provision of ongoing and readily available technical assistance to facilitators (via local and national offices); (d) manual participant survey reviews for elimination of errors, inaccuracies, and unreliable data (by researchers); and (e) ongoing evaluation consultation meetings between the ALA and researchers.'

Were the evaluative design and the objectives modified during its implementation?

'One of the main challenges we experienced in the current evaluation was making the transition from formative to summative evaluation. In short, for obvious reasons, some stakeholders seem to resist the reallocation of resources from programme improvement to summative evaluation activities.'

How to interpret and use results of evaluation

What ideas, plans, expectations, assumptions have gone into the making of the intervention? How was it supposed to work? Have different perspectives on how the intervention should work been taken into account?

It is important to gather information on the theories that underpin the intervention, examining existing evaluations and research to find evidence on the success of the 'programme theory'. We can consider similar theories of interventions coming from health promotion literature.

> *'The way in which the actors understood the programme is crucial. We observed significant differences among the Local Health Units (LHUs): the use of sanction-based behaviour and of administrative measures rather than mandatory treatment programmes (both measures pertaining to the individual rights) proved to be quite variable. The percentage of individuals who had an administrative sanction imposed on them in the 3 Local Health Units, varies from a minimum of 4.6 % (LHU B) to a maximum of 13 % (LHU C); similarly, the percentage of individuals who were sent for treatment varies from 10.7 % to 16 %. The differences related to the drop-out from treatment process (from 4 % in the LHU B to 24 % in the LHU C) are noteworthy, i.e. those who did not attend the group treatment were about to receive a sanction.'*

Which assumptions have proved being correct and which have failed? How has the intervention worked?

The evaluator will verify which parts of the plan have come to fruition and how. He/she will seek to understand which programme theories work for whom, in what circumstances and in what respects it is unlikely that interventions are ever implemented in the same way. In doing so, he/she will try to improve the implementation and targeting of the intervention.

> *'The institutional and organisational context of the interview affected significantly the way the young consumers perceived the measure "Article 75" in all its aspects: from the police detention to the subsequent interview. The hypothesis was confirmed. The use of illicit substances was more likely to be considered as a "health problem" if the interview took place in the LHU context. (...) The interviewees in the TGO systematically reported "the experience of fear"), while those interviewed in the LHUs reported "the experience of ease". The connection between the experience of feeling at ease and undergoing the*

*interview in the LHUs (67.6 % in the LHU3 against 8.5- Chi squared = 6.791 gdl = 1 ** p<.01) is statistically significant.'*

Have alternative causal explanations been analysed? Have different perspectives on how the intervention could work been taken into account?

'To ensure that the negative advertising coefficients imply that recall of advertising leads to lower marijuana and cocaine/crack use and are not due to omitted-variable bias (e.g. omission of variables such as exposure to other anti-drug programmes), we examined the correlation between the advertising-recall variable and the estimated equation error.'

How did the reconstruction of the programme theory and initial evaluative results contribute to the redesign of ongoing intervention?

The process of developing programme theory often reveals that an intervention is not ready for full-scale evaluation and needs to be redesigned.

'In this case, substantial time and resources are saved by redirecting efforts toward further programme development and/or implementation activities, as opposed to summative evaluation certain to reveal null effects. Secondly, evaluation questions are prioritised in this approach, which helps stakeholders decide how to allocate evaluation resources in a cost-effective manner. We have found that developing programme theory usually enables stakeholders and evaluators to make informed decisions about evaluation design and methods, often leading to cost-effective evaluation.'

'A central goal of the theory-driven evaluation of the Work and Health Initiative was to facilitate continuous programme learning, problem solving, and programme improvement throughout the life of the initiative.'

Have the costs of the same intervention implemented in different places been compared?

In order to provide valuable recommendations to policymakers and intervention managers, it is necessary to consider also the evaluative criterion of efficiency, analysing and comparing the costs of similar output or unit cost of different interventions.

'The following key findings emerged from the analysis of the economic costs of drug testing.

- *Set-up ranged from about GBP 7 000 in Site 2 to around GBP 35 000 in Site 3.*

- *Running costs ranged from around GBP 10 000 in Site 2 to about GBP 44 000 in Site 3.*
- *The unit cost of a drug test ranged from GBP 57 in Site 3 to GBP 121 in Site 4.*
- *The unit cost of positive tests ranged from GBP 1 219 in Site (...).'*

'Given that there was little activity on drug testing and the fact that the nature of drug testing varied from case to case, it is difficult to derive any meaningful conclusions from the analysis undertaken. (...) In terms of the different agency contributions to the delivery of the drug testing in most sites, the Yot project provided the majority of these resources.'

Have the contexts and global prevention strategies and policies at local/national level been considered? Have negative influences of other sectors or potential positive impacts of more integrated local policies been identified?

All the issues related to 'health promotion' and 'health inequality' need more coordination between different sectors of policies. The evaluation should provide a bigger contribution to the development of prevention policies, considering the global context with more attention, such as: in relation to enforcement and treatment drug policies, the connection with illicit drug market and licit drug consumptions, the influence of alcohol selling rules.

'However, the limited available evidence suggests that little money is spent on primary prevention activities (...). Estimates of drug control expenditures for the Netherlands (Rigter, 2006) and Sweden (Ramsted, 2006) show that prevention programmes account for a very small part of the total, 2 % in the case of the Netherlands and 1 % in the case of Sweden. Australia is an outlier, with an estimated 23 percent of the drug control budget going to prevention (Moore, 2005).'

Have the relations between administrative rules and instruments been investigated? Is there a connection between → mechanisms of access to financing, rules to allocate resources, the intervention's duration, equity issues and effects?

It is impossible to control all and to measure every variation. However, it is possible to understand the connection between → mechanisms and impacts related to other policies. Change is gradual but some factors or constraints re-emerge in different contexts. For example, the characteristic of the bids, the norms and restrictions that rule the access to funding, the length of a Local Health public agency strategic plan can influence the choices of prevention strategies adopted by funded

organisations of a third sector. Also, the different phases of the interventions need to be analysed. For instance, in the recruitment phase, many interventions have to create waiting lists: the length of waiting time can influence the motivation or the condition of the target groups and have negative or positive consequences.

> *'... Some restrictions relating to those who have access to resources can have an unexpected influence on the type of measures supported. (...) The fact that lawmakers have indicated some subjects as beneficiaries of financing ensures that in planning and negotiating groups and, finally, in the prevention projects themselves, other organisations, which could have an important role, do not appear at all as partners.'*

> *'If we analyse over 3–5 years the ability to attract economic resources made available under different specific laws and from European Union funds (...), we notice that particular districts, provinces or even more specifically particular councils, are systematically excluded: young people from those areas are therefore heavily disadvantaged in terms of opportunity of access to services. This phenomenon occurs systematically and depends on the following two factors. Local administrations and non-profit organisations tend to develop specific know-how in relation to the means of planning and the capacity to access public funds, which leads them, year after year, to increase their competitive advantage.'*

> *'... the results of projects are not connected negatively or positively to the mere presence of more partners. The broad partnership may at the same time be a key to success but also a disadvantage, since it brings with it an increase in complexity in the management stage, in conflict and in coordination costs. Therefore, it is necessary to reduce in public competitions – and in particular at the assessment stage – the value attributed to the criterion of "partnership" and avoid the partnership being created with the sole purpose of more easily accessing financing.'*

The interventions are generally part of a broader strategy to manage drug prevention funds planned by a local health authority, a city council or a foundation. Practices of public competitions (bids, grants) with a 1-year duration are highly at risk in a sector with long-term results and should be eliminated for drug prevention projects.

> *'... there is a positive correlation between the reduced duration of projects (less than 1 year), weakness of strategic drug prevention plan at the local health authority level and the prevalence of initiatives based on information strategies with low or doubtful evidence*

of effectiveness. These are projects which are easy to manage and which do not require complex coordination processes, but which translate into barely effective initiatives or which risk creating perverse effects (see increase in the interest of young people in relation to some little known substances). We must therefore enhance the planning processes which are, vice versa, connected to forms of initiatives which are considered more effective.'

Has the evaluation considered the history of an intervention across time, the broader strategies of the organisation, the relation inside the implementation coalition and with other services?

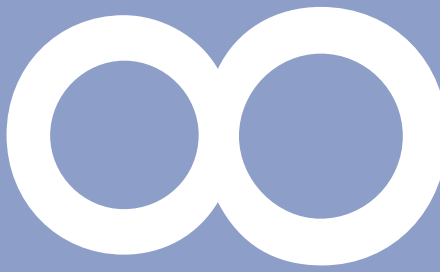
Durable behavioural change requires the coordination of a range of interventions and services delivered for a long time by different partners and public agencies. Accordingly, one pressing problem for evaluation is to investigate the success of coordination, and the overlapping and synergy of different interventions. Since an intervention is embedded in a specific context, previous histories and experiences, relations, long-term aims of each partner, count and influence expectations.

'The initiative was conceived as a major foundation effort and investment, but it was extended in time, entailing even greater financial commitments. One result was that the starting vision and cast of project stakeholders differed in important ways from the intervention and stakeholders twelve years later.'

'(...) the degree of overlap between local social policies and social plans and drug addiction prevention financed under Law 45/99 varies significantly from region to region: where the fund has been managed by the local authority health division, the local health authorities have had a dominant role and governed the planning process, whereas, vice versa, the Law 45/99 fund, together with other funds for laws in the sector, has fallen under the competence of the regional social policies division, the governance role of local administrations has prevailed and consequently the tool of Area plans under Law 328/00 has also prevailed.'

'To avoid risks of the proliferation of "coordination groups" and the dispersion of energy and resources, (...) both local managers and technical operators have highlighted the need to enhance moments of dialogue among coordination groups created in different areas and Planning offices: "there are too many groups around tables, only carpenters are getting anything out of it!", "We don't have the time and resources to take part in all the groups!"'

CHAPTER EIGHT



Chapter 8

Process evaluation and monitoring system

Planning the process evaluation

What variables and indicators will provide useful information on how the intervention was accomplished? What kind of information (→ qualitative or → quantitative) will the process evaluation assess?

There are four ways of conceptualising helpful variables and indicators:

(1) Will the intervention be implemented as designed? This could be backed up by, for example, teachers' reports on the use of an intervention manual.

(2) How much of the intervention will the target group receive? This can be answered with information on the number and the socio-economic characteristics of the beneficiary groups linked to information about duration and frequency of the specific activity.

(3) Will the prevention intervention be changed during implementation? Self-reports from trainers about any adaptation during the intervention's implementation are helpful here.

(4) What is the quality of the intervention? This could be assessed by stakeholders' satisfaction with the intervention or methodological adequacy of an intervention.

The distinction between quantitative and qualitative information refers to whether the information can be expressed in numbers (quantitative) or in verbal descriptions (qualitative). An example of a quantitative technique used to gather data is the use of a → questionnaire; an example of a qualitative technique is an unstructured → interview about an individual's experience of the intervention. Remember that data gathered thorough a focus group, a qualitative technique, can be successively analysed also by means of quantitative techniques of content analysis.

> *'Variables that will be used to assess the realisation of the prevention intervention: observation and report on whether the intervention has been delivered; the number and duration of the training sessions; teacher and pupil satisfaction. As well as these quantitative measures, group interviews on the implementation process will be conducted with all teachers.'*

What methods and → instruments will be used (interviews, questionnaires, → observation instruments)?

Questionnaires, interviews, reports, checklists and written records can all be used to measure the process. Another possibility might be the use of observation during the intervention's implementation.

> *'Qualitative instrument used: Final Evaluation of Teacher's perception of skills achieved and development of the programme.'*

> *'Questionnaires, semi-structured interviews and informal telephone interviews with the leader of the team will all be undertaken. Semi-structured interviews will be carried out with the course tutors and participants, as well as observation of a number of training sessions and focus groups with evaluation personnel.'*

> *'Programme-specific instruments used (e.g. self-constructed collection instrument): Annual Coaching Teachers "Growing Up Playing" Instrument; qualitative instruments of teacher's perception of the development of the programme (NGO Prevenir) Annual qualitative instrument of teacher's perception of the training (Centro de Formacao de Professores, Ministry of Education).'*

Where, when and how often will the process data be collected (→ design)?

> *'During dance events, and after each intervention.'*

> *'Classroom questionnaires will be distributed after each training session, and observations will take place during the third, sixth and ninth sessions. Teacher interviews will also be conducted after the fifth training session.'*

Who will provide the information needed for the process evaluation?

It must also be decided who will be asked or examined (e.g. intervention participants, trainers, teachers, independent → observers, etc.).

> *'We will obtain information from the schoolteachers who will implement the programme, the head teachers in the schools where the programme will be implemented and students who will attend the programme.'*

How will the data be analysed?

At this point, whether to apply any special statistical procedures or merely to describe the data should be decided. For the former, the necessary preconditions for complex data analysis (statistical knowledge and technical equipment) must be available.

> *'The questionnaire data will be analysed by the statistics programme entitled "Statistical package for the social sciences" (SPSS). The statistical procedure that will be used is Manova.'*
>
> *'The statistical analysis carried out was the Variance analysis (General Linear Model, repeated measures) showing tests of Within-Subjects Effects/Contrast; tests of Between-Subjects Effects, Mauchly's Test of Sphericity).'*
>
> *'Student interviews (satisfaction with the intervention, interest in the training, personal benefit, increased knowledge, etc.) will be analysed and the results will be proportionally described.'*

Monitoring the prevention intervention

What strategies, components and methods were actually implemented?

Compare the answers with the original plan (see Chapter 1, under the 'Methods' section). What actually happened must now be described in detail.

Components

The educational materials used by the health education programme are divided into three parts. The first deals with smoking ('counter-attack on smoking', 'the language of persuasion', 'the first offer', 'giving up the habit'). The second deals with alcohol ('the general picture', 'in-between', 'one among many', 'ask yourself'). The third deals with drugs ('pictures', 'a pill for every patient', 'why not try'). The Appendix has three sections ('drugs – general information', 'self-awareness', 'how we breathe'). All components have been implemented in the three secondary-school grades. For example, parts one and two of the first section dealing with smoking are implemented in the first and second grades, part three in the second grade and the remainder in the third grade. In other words, the appropriate materials have been chosen according to the students' age.

Methods

The educational method used by the intervention is that of 'active learning', which aims to support and enable the students' basic skills to develop. The activities include: class discussion; group discussion; interviews; research/survey; attitudes; classification; debate; writing; role play; the media; posters; community involvement; reflection; etc.

→ Intermediate target groups

> *'The secondary-school teachers were trained by the scientific team in a three-day seminar on the intervention's principles and methods. This training was approved by the Department of Education and the teachers' participation was voluntary. Alongside the permission required from the Department of Education, the consent of the head teacher and the Association of Teachers has been indispensable. The training was carried out according to the original plan. Parents were also told about the goals of the health education programme. They could join a special "parent group" and, throughout, the programme has tried to promote close cooperation between schools and the local community.'*

What data sources and instruments were used to measure the intervention's implementation? Compare the answers with the original plan (see Chapter 3, 'Planning the process evaluation')

These sources could include participants, trainers, the → evaluator, an independent observer, or anyone else involved in the implementation process. Instruments can include questionnaires, reports, checklists, interviews or discussion groups.

> *'The people involved in this measurement exercise were: the teachers who took part in the programme; the head teachers of the schools that implemented the programme; and the students who attended the programme.'*
>
> *'Municipal teams for community prevention of drug addiction were entrusted with organising the implementation of the programme in their regions, and thus gave invaluable information. Teachers, parents and schoolchildren were also sources of data.'*
>
> *'The "Notebook for teachers", as well as allowing for personal comments from every teacher who implemented the programme, included questions such as: Which element of*

the intervention has been implemented? How helpful have the manual's instructions been? Can these instructions be improved? Were there any difficulties during implementation?'

What resources were actually used? Compare the answers with the original plan (see Chapter 1, 'Resources' section)

It is vital to determine whether the resources were used as planned or whether changes occurred. Knowing this will be extremely helpful in any future application of the intervention.

> *'The resources were used as planned with the following exception. The fact that 20 % of the pupils had moved away from the area meant that the administrator needed to spend more time chasing up the new addresses in order to send the follow-up questionnaire. Instead of 20 hours a week, she needed to work 30 hours for a period of three months.'*

The target group revisited

How many people did the intervention actually reach?

This is an obvious question, but a crucial one for making any headway in understanding the overlap between the actual and the planned interventions.

> *'Some 450 pupils were contacted.'*

What were their socio-demographic characteristics?

This is just as important as the number of people reached. The answer will allow the planned target group to be matched with the actual target group. It will also help to detect → selection effects.

> *'Gender: 45 female and 75 male participants. This is a gender ratio of 1:1.7. Education: 10 junior high-school students, 35 high-school students, 75 college students.'*

> *'Primary school: 10 % of participants. Secondary school: 90 % of participants. The annual income of the families reached: < USD 15 000 – 10 %; USD 15 001 to USD 35 000 – 60 %; > USD 35 000 – 30 %.'*

> *'Other relevant information: the participants belong to an ethnic minority (Moroccan women, Muslims).'*

How was this information collected?

As different methods yield different results, it is important to know how the information was collected. Different methods of data collection can lead to possible distortions in the answers, for example, participants may over-estimate their income in group discussions.

> *'The socio-demographic data were collected using an anonymous questionnaire. This questionnaire was handed out after the first session and collected at the next session.'*

Throughout, the answers should be compared with the original plans, in this case the target group as envisaged in Chapter 1, 'The target group'. Deviations concerning the target group when setting up the intervention may lead to differences in its implementation and in the final results.

> *'The target group was reached as planned.'*
>
> *'We planned to reach young people aged between 15 and 17 visiting "raves". The actual participants were older (mean age: 18.3 years).'*
>
> *'The gender ratio and the ratio of German to other nationalities was supposed to be the same in all three experimental groups. In fact, in the → control group there were 10 % more Turkish pupils than in the other two groups.'*

Exposure

How was → exposure measured? Which data sources, instruments or indicators were used?

A crucial element for understanding whether the intervention reached the proposed target group is to estimate how prominent the intervention actually was. For instance, in a public-information campaign, how widely information leaflets have been distributed will need to be known. Who was asked must also be clarified in order to gather the relevant data. These data sources could be all those involved in the intervention or just a select few. Once again, what data-gathering instruments were used must be spelled out.

> *'The teachers in the schools implementing the programme were the main data source. The "Notebook for teachers" asked for comments on the programme's implementation, as well as for information on the number of teaching hours actually carried out in each class.'*

How long did the prevention intervention actually last and how many prevention activities took place?

Compare the answers with the original plan (see Chapter 1, under the 'Methods' section). This description of the duration and the number of activities carried out will help in assessing the level of exposure.

> *'In the first two years of implementation, each class had 12 lessons of 45 minutes each.'*

> *'The number of sessions ranged from three to 10, with a mean of 8.76 sessions. The average length of each session ranged from less than 30 to over 50 minutes, with a mean of 40 minutes. Time differences were mostly due to variations in the length of the classes, which ranged from 45 to 55 minutes. The whole programme consisted of 10 sessions.'*

To what extent was the target group actually reached?

Compare the answers with the original plan (see Chapter 1, under 'The target group' section). This is a key question, because even if a prevention intervention is fully delivered to the target group, some of that group may not receive the intervention because of absence, illness or truancy.

> *'In total, 1 500 students took part in the programme. Some 85 % took part in all 12 lessons, 93 % took part in 10.'*

Quality of the prevention intervention

Who provided the information on the → quality of the intervention?

These sources could include participants, trainers, the evaluator, an independent observer, or anyone else involved in the implementation process.

> *'Teachers and students were the source of information on the quality of the actions developed.'*

What indicators and instruments were actually used to evaluate the quality of the intervention?

Indicators can include active participation, attitudes towards the intervention, personal benefit or degree of identification. Instruments could include questionnaires, reports, checklists, interviews or discussion groups.

'Teacher indicators used: perceived efficiency, degree of satisfaction, participation experience, versatility of the programme, curricular integration, attractiveness of supporting material. Pupil indicators used: recognition of the programme, linguistic adaptation, personal experience of the time dedicated to the programme, perceived efficiency.'

'Teachers received a questionnaire that included questions about class participation in the programme, class disruption and achievement. A global rating of how well the programme was implemented was also assessed (1 = very poorly to 4 = very well). In addition, the research staff observer conducted random observations of implementations in each school, as well as periodic telephone calls and meetings with teachers and head teachers. Criteria for rating implementation were discussed among research and programme staff prior to actual observation, until a consensus (using hypothetical examples) was reached. Items included observed class participation, interest and teachers' completion of session activities, as well as an overall rating of implementation (1 = very poorly to 4 = very well). The observer's overall rating was compared to the teachers' report of overall intervention implementation.'

'Knowledge of the participating nurses and their fidelity to the protocol were checked by regular supervision. There was no other instrument to measure this, than the review of the filled files within the regular interviews with the collaborating nurses. Their performance was mostly satisfactory.'

What are the results of the quality measurements?

This is one of the most crucial questions if the results of the particular intervention are to be used to inform future interventions.

'The assessment of teaching methods, attainment of → objectives and involvement of students revealed the following: for 8.6 %, teaching methods were inefficient; for 10.4 %, objectives were not achieved; and for 13.4 %, student involvement was low. However, overall quality ratings show 52.7 % of the ratings to be excellent, 46.7 % to be good and only 0.6 % to be poor.'

Discussing the results of the process evaluation

How do the plans for the intervention compare to its actual implementation and the current evaluation? Are there any discrepancies and what are the possible reasons for them?

Any deviations and discrepancies in the intervention's implementation, its target group and its exposure should be summarised and explained. Any discrepancies cannot merely be presented as a measure of the distance between expected versus obtained output indicators or results. It is necessary to explain and to interpret the meaning of observed processes and any deviations which occur. Sometimes, practitioners have 'good reasons' for changing planned standard interventions. Moreover, thanks to these discrepancies, we can discover unexpected results; for example, that many evidence-based prevention programmes did not work well with female target groups and, therefore, advancing and refining our knowledge about the 'theory of the programme'.

> *'The evaluation found that the programme had been implemented in all 17 intervention classes, but the quality of the programme delivery was rather low. It was evidenced by low rates of family evenings' participation, with 30 % of students attending with at least one parent, low rates of parental participation (50 % of parents completed at least half of booklet activities), and low rates of students' satisfaction (43 % of students were satisfied). Evaluation also identified some reasons of poor programme delivery. Implementers of the programme were challenged by lack of sufficient funds from local authorities to support a two-year programme delivery. Most participating teachers were forced by their school principals to run this programme. Their initial motivation to run the programme was rather low, and most of them had no previous experience in prevention activities.'*

What is the impact of any discrepancies in the intervention?

If there were any discrepancies, their implications for the intervention should be discussed. This will help in understanding their significance for the intervention's implementation.

> *'Due to the lowered exposure rate, it is difficult to draw conclusions as to the intervention's effects. The negative results may be related to this fact.'*

> *'The patients' recruitment started on 1 July 2001 and had to be interrupted early in March because of the mandatory removal of the saliva screening-reagent due to some detected defects.'*

What are the strengths and weaknesses of the way the intervention has been implemented? Compare these with results from other interventions

Answering this question will help avoid similar problems and improve implementation in the future. If possible, compare the results of the intervention with similar interventions reported in the literature. This is very important, as it forms the basis for any judgement concerning the value of the intervention.

> *'Results of this study indicated a high quality of programme realisation, as measured by the level of exposure and overall judgement of the quality of implementation. One problem concerning generalisation of the results is the assistance the teachers received from the project staff. Teachers were contacted by telephone and could talk over any problems. This assistance probably contributed to the high motivation of the teachers. Whether similar results can be gained when teachers do not receive special treatment is still to be ascertained. The findings of process evaluation are consistent with the results of other school-based life-skills programmes.'*

What suggestions can be made for the future implementation of a similar prevention intervention?

Having completed the prevention intervention, the implementer is well placed to make suggestions for other planners.

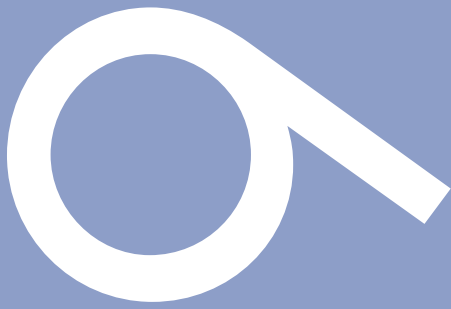
> *'In the future, we envisage increased involvement from teachers, a further analysis of attitudes and → lifestyle, and a modification of the videotape to include a specific section on new drugs.'*

What suggestions can be made for future process evaluations of this kind of prevention intervention?

The implementer is also in a good position to make recommendations about process evaluation.

> *'After each session, we asked the pupils to fill out a questionnaire on the session. We used the same questionnaire after each session and noticed that pupils became less and less careful when answering. We would therefore suggest the inclusion of specific questions relating to each session in order to make answering them more interesting for the pupils.'*

CHAPTER NINE



Chapter 9

Outcome evaluation

Planning the outcome evaluation

What are the → indicators for outcome and how will they be measured?

Deciding how to measure outcome is not always easy, but it is a crucial decision. Given the constraints of money and time, a good outcome measure is one that is not only feasible to use, but also more or less directly related to the goals of the intervention. In order to know whether the intervention has achieved its aims, the criterion for those goals must be clearly defined. In other words, it must be 'operationalised' and defined in measurable terms. This is especially important with → mediating variables, which are often not easily observable.

> *'The analysis of the impact of drug testing focused on the change in substance use. The analysis of the impact of arrest referral was organised around the causal → mechanism through which arrest referral was expected to affect the behaviour of* young people. *The causal → mechanism is: arrest referral improves referral to and access of services; access to services reduces the risk of offending and substance use; and improving the factors that influence a young person's risk of offending will reduce the chances that a young person would re-offend.'*

> *'The main outcome indicator will be the self-report WHO → questionnaire, "Survey on the student population". Questions concern tobacco and alcohol use, medicine and illegal drugs, knowledge about substances, attitudes towards substances, intentions to use substances, and antisocial behaviour.'*

Will the information on outcome be collected following a → quantitative or a → qualitative approach? What indicators and → instruments will be used to collect information?

> *'The intervention is planned to increase the knowledge and skills of drug educators. Knowledge will be measured by a written test after the intervention with 50 questions on drug-related topics. Education skills will be measured by asking the participants to produce an outline drug-education programme for a named client group within a specified*

educational setting; to create a lesson plan; and to present the lesson plan. The participants will evaluate the content and presentation of each other's lesson plans.'

'We will apply a sub-scale ("locus-of-control") of the "Life-skills training student Questionnaire."'

What is known about the quality of the instruments (→ objectivity, → reliability, → validity)? Will the feasibility of the instruments be tested?

In order to assess the results of any evaluation correctly, it is important to provide information on the quality of the instruments used. For instruments that are already well established, it is usually sufficient to refer to the handbook or manual. If instruments have been specially constructed for the intervention, however, it may be useful to study their reliability and validity. Testing the feasibility of an instrument is obviously advisable in such a case, but this should also be carried out when using an instrument constructed for another language, cultural background, area, and so on. In such situations, it might be more suitable to have the instrument rated externally for comprehensibility and logic.

'We plan to use the "Children's self-concept attitudinal inventory". As a paper-and pencil test, it is objective. As evidence for reliability, the scales in the inventory have an average Cronbach alpha of 0.80. There is also evidence for validity of the scales based on the "know-group-method" of comparing high- and low-performance students in school. We had to translate the questionnaire into our own language, and so we plan to conduct a small feasibility study. We will ask 10 pupils to fill out the inventory and to indicate for each question whether they understand what we mean by it.'

'The Spanish version of the AUDIT (cut-offs of 6 for females and 8 for males) and the AUDIT-C questionnaire (cut-offs of 4 for females and 5 for males); the Attribution of Injury Scale and the Readiness-to-Change Ruler.'

From whom, when and how often will information on outcome be collected (→ design)?

It is also important to understand how recent and how widespread the outcome information is.

'Using a quasi-experimental design, data will be collected from an experimental group and not a randomly assigned → control group. The data collection will take place twice (one week before and one week after the intervention).'

'(The) Self-report questionnaire was administered to students four times: (1) before the programme (base-line); (2) after the Program Domowych Detektywow (PDD) and prior to

the Fantastyczne Możliwości (FM) implementation; (3) just after the FM was completed; and (4) one year after the FM was completed.'

How will the information gathered be analysed? Which → statistical methods are adequate for the quality of the data and the design?

If those implementing the intervention will be analysing the data, the statistical procedures and technical equipment to be used should be described. It may, however, be advisable to hand data analysis to an experienced researcher. Who will process the data should also be indicated.

> *'Questionnaire data will be analysed by the computer package SPSS PC+. Processing entails frequencies, multiple variable tables, mean values and comparisons between means per group (→ Chi-square and students T).'*

The sample

How was the sample recruited?

How the sample was approached (i.e. via a notice board, a newspaper advert, word of mouth, etc.) is as important as how the whole → target group was recruited. Whether the sample participated voluntarily or compulsorily is also significant, as is the issue of reimbursement.

> *'The questionnaire was distributed to classes in the first grade of secondary schools before the implementation of the health education programme, as well as to the classes of two control schools. The completion of the questionnaire was compulsory and took place over two consecutive teaching hours.'*

> *'The experimental group was made up of all those who took part in the intervention. The control group was recruited from students of the same age, gender and socio-demographic characteristics, attending the same school levels in the same geographical area. Both groups chose to participate in the outcome evaluation.'*

What were the sample's socio-demographic characteristics, size and so on?

The characteristics of the sample and – if applicable – the control group should be described separately and compared for any significant differences. Any such differences lessen the degree of comparability between the two groups.

'Experimental group (n = 720): 60 (50 %) female, 60 (50 %) male. Average age 10.8 years; standard deviation 2.4; range 5.5 to 17.3. Control group (n = 110): 80 (73 %) female, 30 (27 %) male. Average age 11.0 years; standard deviation 2.3; range 6.5 to 17.6.'

'→ Intermediate target group: five schoolteachers, two youth and community police officers, four drug educators. The participants ranged in age from 21 to 55 years. Ethnic groups: three black, eight white. Six female, five male.'

How do these characteristics compare with those of the whole target group?

It is, of course, vital to know whether the sample is representative of the whole target group.

'As planned, we gathered results on outcome from a subgroup (20 %) of the target group. Since we had administered the first questionnaire on → substance-use behaviour to all participants in the intervention, we were able to compare socio-demographic characteristics between the sample and the whole target group. There were no significant differences in age, gender or other characteristics.'

Was it possible to identify drop-outs? If so, what were their characteristics?

If those that do not stay in the sample until the end of the data collection differ significantly from the others, for example in age or gender, it is likely that the sample is no longer representative.

'We made a statistical comparison of the differences between the drop-outs and the remaining sample (age, gender, socioeconomic status, etc.).'

'Due to technical reasons, only half of the sample (N=372) participated in the one-year follow-up. This group consisted of students who at the base-line were fourth graders, older students (base-line fifth graders moved to middle schools and were not practically available for one-year follow-up).'

The outcomes

This crucial section should follow the same format as the earlier description of the intervention's → objectives (see Chapter 6, 'Objectives'). In light of the considerations that we have shown up to this point in Chapter 7 ('How to analyse the programme theory'), some theoretical hypotheses are necessary in order to identify significant outcome indicators and mediating variables. The first

three questions refer to outcome variables in the → ultimate target group, the rest to outcomes in the intermediate target group.

How did the intervention affect the ultimate target group's substance-use behaviour?

> *'Results of multiple analyses of variance (Manova) with an individual subject as the unit of analysis indicated no significant differences between intervention and reference group in the alcohol use scale both at the end of the programme Fanstastyczne Możliwości and in the one-year follow-up. The results of bi-variate comparisons suggested a significant reduction of incidences of drunkenness (Chi square = 4.62; $p<0.05$) among students in the intervention group. These effects appeared just after the programme was completed and in the follow-up one year later.'*

> *'At 6-month follow-up, between-group effects were found on measures of 30-day alcohol use, marijuana use, polydrug use, and total substance use. Relative to girls in the control group, girls exposed to the Internet-based intervention reported lower rates of use for these substances.'*

How did the intervention affect mediating variables related to substance use in the ultimate target group?

> *'Moreover, girls receiving the intervention achieved gains over girls in the control group on normative beliefs and self-efficacy at post-test and 6-month follow-up, respectively.'*

How did the intervention affect other mediating variables in the ultimate target group?

The discussion of the effects are to be connected to the initial theoretical model or assumptions that guided planners or to theoretical constructs chosen by evaluator and to results of other researches, otherwise it could not be possible to 'recognise' and study mediating variables and their relation to demographic, cultural, socio-economic characteristics of the target group.

> *'So far, significant differences were found only for one mediating variable. Mean values for social competence were significantly higher in the experimental group than in the control group ($p<0.05$).'*

> *'Theoretical models suggest that girls are more influenced by family protective factors, whereas boys are more influenced by school or community environments. Among girls, self-esteem is strongly dependent on a positive relationship with parents.'*

How did the intervention affect objectives in the intermediate target group?

> *'Our mass-media campaign had a positive influence on the intermediate target groups (parents and teachers): 80 % said they now think more about their responsibility for children and adolescents; 6–7 % said they had learned a lot about substance use; and 45 % said they now think more about their own consumption of alcohol and tobacco.'*

Are different subgroups affected differently by the intervention (e.g. men/women, age groups, risk groups, etc.)?

This is one of the most neglected issues in the last twenty years of drug prevention studies that led to inequality of health and hindered new insights and improvement of prevention programmes.

> *'Results indicated that protective factors increased for high Socio-economic Status (SES) classrooms receiving the PDP (Prevention Dimension Programme) and decreased for the low classrooms receiving the PDP. Results for risks indicated an increase during the PDP for the sixth grade classrooms, and the increase in risk factors was greatest in the low SES classrooms.'*

> *'Overall, prevention effects were strongest for cigarette smoking, but were also evident for alcohol. Significant gender by programme interactions, showing differential programme effects for males and females, were found for cigarettes and marijuana, but not for alcohol. There was a non-significant trend suggesting that the programmes were most effective for Asians and least effective for whites.'*

Discussing the results of the outcome evaluation

Did the intervention achieve the expected outcomes?

Discuss any discrepancies between expectations and results, addressing possible reasons for them and their impact on the study. At this point, any arguments against how the outcome results were interpreted should be anticipated.

> *'One of the objectives of the intervention was the enhancement of → life skills. At present, only one mediating variable was significantly different at post-test – a scale measuring social competence. (...) Together with a lower prevalence of smoking in this group, it is also an indirect confirmation of the theoretical assumption that improvement of life skills is an important primary-prevention intervention.'*

What are the most relevant and significant results? Compare these with results from other studies

> *'A similar primary-prevention effect on smoking (namely, the delay of initiation into smoking) has frequently been reported as a result of life-skills training.'*

Is it certain that the intervention caused the results? Are there any alternative explanations for them? Are the study limitations declared?

Without an answer to these questions, doubt can always be cast on an intervention. It is extremely important, therefore, to address the issue of cause and effect.

> *'Since we did not have a control group, we cannot say for certain that the effects are really caused by our intervention. An important influence to be considered is the maturing of the participants over the course of the year.'*

> *'The study's limitations must be considered as well. Those students whose parents failed to return the consent form or denied consent cannot be assumed to be the same as those students with more compliant parents. Due primarily to school districts' reluctance to implement individualised drug use detection procedures (e.g. CO testing), usage data probably reflect some level of self-report bias. Attrition from sixth to eighth grade was not random, but was associated with reported sixth grade marijuana use. Even though assignment of schools to an intervention was random in each district, schools self-selected into the study. Each of these factors limits the generalisability of the present findings.'*

What explanation is there for negative results?

It can be very discouraging if an outcome evaluation does not show the expected result. However, there could well be a number of explanations for negative results, such as deficiencies in the intervention itself, its implementation (the process evaluation can be useful here to highlight relevant aspects), flaws in the design, inadequate measurement instruments (some could be too insensitive), or perhaps external problems out of the intervention's control (a fire in a school, for instance). Whatever the reason for unexpected results, an explanation for them should always be found to enable conclusions to be drawn for the future.

> *'Contrary to our expectations, there were no effects on alcohol consumption. This might be due to the fact that the prevalence of alcohol consumption is very low in this age group, and so it is possible that effects will occur later. A second follow-up will therefore*

be conducted after one year. It can also be argued that the topic of alcohol consumption is irrelevant for this age group.'

'There is insufficient evidence in this evaluation report to support wider rollout of drug testing. While no single stream of evidence is sufficient to draw a conclusive verdict about drug testing, the combination of different evidence strands is insufficient to support the roll-out of drug testing. (...) There has been a very low proportion of positive tests. (...) Limitations in the design of the impact analysis have prevented any clear conclusions regarding impact.'

Are there any other positive or negative unforeseen impacts? Are more general recommendations provided to policymakers?

An added value of evaluation consists also in the learning process developed during the process that can be returned to policymakers in the form of direct or indirect (more general issues) recommendations. General recommendations are addressed to improve the policy and planning processes, while specific suggestions regard a single programme or intervention.

'The intermediate effects of tobacco control policies are currently not well understood and documented, (...) (they) include, for example: (1) ease of youth access to tobacco products; (2) number of spontaneous tobacco purchases; (3) percentage of tobacco companies reporting data accurately and on time; (4) consumer awareness of harmful health effects of tobacco products; and (5) types of tobacco products consumed.'

'There is a need to understand better the economic context of tobacco product regulation. Very little information is currently available to DG SANCO on key indicators of the market. Such indicators include reliable data on the number of manufacturing businesses, number of retailers or employment in the sector. This impact assessment has had to rely substantially on self-reported costs from industry, given the lack of such cost data elsewhere. This has implications because self-reported data on costs from industry are more often than not overestimated and difficult to verify independently. Therefore, developing independent expertise on the regulated industry sector will be essential to allow for an impartial assessment of the effects of current and future regulation.'

What suggestions can be made for the future use of similar interventions?

It is now time to look to the future. Should other people copy this intervention, and if so, what modifications should be made?

'The positive outcome results justify an implementation on a much broader basis. However, the results of process evaluation should also be taken into account and changes in the structure of the training should be made according to the teacher comments.'

'Effective programmes were equally successful with marijuana, alcohol and tobacco. Interactive programmes were significantly more effective than Non-interactive programmes for reduced marijuana use. But of greater importance is the finding that these very same programmes had similar results for alcohol and tobacco (... 5). For prevention planners, this means marijuana use does not need to be addressed in a singularly focused programme. The more generic approaches that also incorporate alcohol and tobacco can be used with adolescents in 6th through 8th grades.'

What suggestions can be made for future outcome evaluations of this kind of prevention intervention?

The experiences and suggestions of the intervention's implementers can be of crucial importance for other → evaluators and practitioners. They can profit from positive results, but also from negative ones, by avoiding making the same mistakes. As a result, complete honesty is essential.

'One reason for the lack of positive results may be the lack of high-quality instruments for pupils aged 10 to 12. Future research should therefore focus on the development and use of sensitive instruments for this age group.'

'These findings provided some important implications. First, the validity of long-term effects might be threatened by maturity and history. (...) Second, more rigorous statistical methods that control for confounding variables could provide smaller, but more accurate estimates of the effect size.'

'In our study, positive expectations towards substances did not significantly modify the programme's effect in either gender. This was also the case for an indicator of self-esteem, although the data in this case suggested a differential modifying effect of self-esteem on the programme's effectiveness in the two genders, indicating that girls with low self-esteem had the least benefit from the programme. A differential gender effect linked to self-esteem would not be surprising, and should be thoroughly investigated in larger studies. In fact, there is some evidence that lack of self-esteem can be a stronger risk factor for drug use among girls than among boys.'

CHAPTER TEN

10

Chapter 10

Communicating the results

Developing a communication plan

Who should be 'in the know'?

Potential audiences include funding agencies, policymakers, administrators, the intervention's → target group, service providers, community groups and the media.

> *'We have made our evaluation available to the partners in the project, the authorities and the public.'*

> *'First, we will inform the planning committee, as it is a primary user of information. We will give special attention to the committee chairperson, who has vast experience and is a respected opinion leader. We will inform her individually before the general committee meeting and discuss potentially controversial findings.'*

When do they need the information?

If the results of the evaluation are to be useful, the timing of an evaluation report can be critical. The results should be reported before too long, as changes are more likely to occur soon after an evaluation and while the intervention is still vivid in people's memory.

> *'To support continuous programme improvement throughout the life of the Initiative, the evaluation team created three primary vehicles for providing continuous programme improvement feedback over time. These included: (1) mid-year evaluation reports, (2) year-end evaluation reports, and (3) annual 360 degree feedback from grantees (...) that is, grantees were given the opportunity annually to evaluate the evaluation team and foundation staff. In addition, these were supplemented with several interim evaluation reports and frequent telephone and e-mail communications designed to provide timely feedback throughout the year.'*

On the other hand, the results should not be reported too early. → Evaluators are often pressed to give a 'first impression' of effectiveness, but such impressions have limited use, as insufficient data will have been collected to allow for any valid conclusions. First impressions also tend to mean that 'last impressions' go unnoticed.

'There will be two intermediate reports and one final report to the funding agency. Teachers will be informed about results continuously in the meetings taking place every two months.'

What information will interest different people?

The message should be tailored to the specific audience. Different audiences will need different information. Staff, for instance, will probably be more interested than funders in the details of the implementation. The latter may also be too busy to read a full report, and so an executive summary is more important for them. Such a summary should contain a brief overview of the evaluation, an explanation as to why it was conducted, as well as the major findings, conclusions and recommendations.

> *'We will carry out team discussions with teachers as the main → intermediate target group throughout the period of process evaluation and continually feed back the information we receive. The primary user of the outcome evaluation is the funding organisation, which expects a comprehensive written report on the findings at the end of the evaluation.'*

Which forms of written communication should be used?

The target audience must be identified and the report written to address their needs, including the use of non-technical language and a user-friendly format. Evaluation results can be communicated in a variety of forms and through the use of different media (i.e. social media), which again have to be chosen according to the relevant audience. Examples of written communication include the evaluation report itself, an executive summary, memos, press releases, articles in scientific journals or newspapers, posters and leaflets.

> *'We will produce a research report to be distributed to schools and other public and private agencies working in the field of youth problems and their prevention.'*

Which forms of oral communication should be used?

Paper is not everything. In fact, many findings can best be disseminated through personal discussions, short films, presentations at conferences or public hearings, programme or institutional websites and media appearances.

> *'Results of the programme will be presented at a National Association for Addiction conference, as well as during European Drug Prevention Week.'*

PART **THREE**

Glossary

Part 3
Glossary

Accountability is an obligation for the actors participating in the introduction or implementation of a public intervention, to provide political authorities and the general public with information and explanations on the expected and actual results of an intervention, with regard to the sound use of public resources (taken from the Glossary of European Commission DG Regional Policy Inforegio 'Evalsed').

Adherence describes one aspect of an intervention's implementation, along with fidelity and reinvention. It usually measures whether a programme was implemented in the experimental group rather than in the control group and whether both groups adhered to their respective experimental conditions. Indicators of adherence in the experimental group can include whether the programme was implemented sufficiently rigorously to conclude that it was delivered, or whether the programme lasted long enough for the target group to notice it.

Attitudes towards drugs comprise all opinions, beliefs and norms that people have about drugs. Examples include 'drinking alcohol makes people act stupidly', 'people who use cannabis have more fun', and so on. Attitudes towards drugs are mediating variables and are often used as indicators in outcome evaluation, although a causal relationship to drug-use behaviour cannot always be found. It is, however, believed that positive attitudes towards drugs will lead to an increase in drug-use behaviour, whereas negative attitudes will decrease such behaviour.

Attrition is the process of reducing the effectiveness of a prevention intervention. The investigation of drop-outs is crucial for medium and long-term prevention interventions. A drop-out is an intervention participant who took part in the initial data collection(s), but who did not remain in the sample for the full duration of the intervention or the data collection. A large number of drop-outs can threaten the validity of the outcome results by building in bias. Drop-outs may also create problems for statistical analysis by decreasing the size of the sample.

Barriers The financial situation, staffing, politics, administrative difficulties and many other conditions can hinder the evaluation of any prevention intervention. Barriers can include a loss of funding, negative staff attitudes, a decision not to publish the evaluation report or a school's refusal to allow a survey of its students.

Bias refers to all kinds of unplanned and often unnoticed variations which occur during the data-collection process and which can prejudice the results of the evaluation. An example of bias is the extent to which only specific subgroups of the designated target group participate in the intervention (only highly motivated children, for instance). Such a sample is 'skewed' and the results could be invalid. Bias can also be introduced via attrition and logical mistakes in the evaluation design.

Control group is a group of people who do not participate in the prevention intervention that is being evaluated. Instead, this group receives either no intervention, or one that is not related to the prevention activity.

Control-group design (or 'experimental design') compares the data obtained from the control group with the results from the experimental group to identify any differences in the variables which the intervention is supposed to change. There are two types of control-group design: the 'true' experimental design; and the quasi-experimental design. Under the first, participants are randomly assigned to groups. To ensure random allocation, each person or each group of people in a target population must have the same chance of being selected for either the experimental or the control groups. This is considered to be the optimal approach because it avoids systematic differences (e.g. selection effect, bias) between the groups and it increases the validity of outcome results. However, it is expensive and in the evaluation of prevention interventions, randomisation is rarely accomplished. Instead, quasi-experimental designs tend to be used, under which people are assigned to groups by other procedures, such as matching (when a control group is selected that most closely resembles the experimental group).

Coverage means the extent to which an intervention reaches its intended target group. It can be defined as the ratio of the number of actual participants to the number of intended participants, and low coverage may increase the bias.

Cultural habits In this context, cultural habits are substance-use behaviours influenced by the cultural or social norms of a specific society. Examples include the consumption of wine with a meal in southern Europe or the consumption of beer after work in Germany. Cultural habits determine the use or non-use of drugs in social situations, the amount and frequency of consumption, and can influence the perception of drug-use behaviour. One objective of a prevention intervention might be to change these cultural habits to make drug-free alternatives more acceptable.

Data quality The quality of the data determines how an evaluation will proceed and what statistical procedures can be used. These statistical procedures are grouped as 'scales'. The most primitive scale is the 'nominal'. In a nominal scale, objects and events are merely classified, such as male/female, smoker/non-smoker, and so on. Nominal data can be analysed by techniques such as the Chi-square test. The next level is the Ordinal scale, which ranks objects and events (e.g. 1 = very poor, 2 = poor, 3 = good, 4 = very good). It is not, however, possible to add, subtract, multiply or divide numbers in an ordinal scale. Hence, the measurement for central tendency in ordinal data is, for example, the mode rather than the mean. Ordinal data can be analysed by techniques such as the Mann-Whitney U-test or the Wilcoxon test. The next level is the 'interval scale'. An interval scale identifies how far apart the scores are because each unit on the scale is of a fixed size (such as degrees Celsius). Interval data can be analysed by the T-test or by analysis of variance. The highest level is the 'ratio scale'. This incorporates the concept of a 'true zero', which means that relationships between two scores are the same, such as in height or weight. Data in the social sciences are most frequently on the nominal or ordinal level, sometimes on an interval level and almost never on a ratio level.

Design is a plan which indicates how often, when and from whom information will be gathered during the course of an evaluation. A good design is essential if the results of an evaluation are to have any future use. A design with at least one experimental group and one control group is known as a control-group design; a time-series design uses only one experimental group, but at least three data collections; and designs that do not use a control group or time-series analysis are the pre- and post-test designs.

Dose-response designs In some cases with a one-group design, an internal comparison can be designed. Evaluation can compare participants who received a large quantity of service ('a high dose') with those who received a lower quantity. The assumption is that if intervention is doing good, then more intervention will do better. The evaluator assesses if high-dose recipients obtain better outcomes (Weiss, 1997a, 1997b).

Evaluation is the systematic application of scientific methods to assess the design, implementation, improvement or outcomes of an intervention. The term 'intervention' may include any organised action such as media campaigns, service provision, educational services, public policies, research projects, etc.

Evaluator The person who acts as the evaluator should be familiar with evaluation planning, social-research methodology, statistical designs and related problems. They should have sufficient

statistical knowledge to analyse the data gathered during the intervention's implementation, and should be familiar with specific computer packages (such as SPSS, SAS). Since systematic evaluation is based on social science research, many evaluation specialists also have a basic social science training. There are two fundamental models for the evaluator's relationship with the organisation carrying out the intervention – they can perform an 'internal evaluation' (when they are part of the implementing organisation) or an 'external evaluation' (when they work for a research institute, consultancy or university). There are pros and cons to both options. Since internal evaluators have better access to the programme staff and administration, they may be more familiar with the intervention than an external evaluator. A lot of information that is gained informally will not be available to an external evaluator. Being well known and trusted, an internal evaluator usually finds the staff more willing to devote time to the evaluation, to admit problems and to share confidences. The results of the evaluation may also be fed back to the project staff more easily and more informally. The clear disadvantage to an internal evaluator may be a lack of objectivity imposed, first, by their dependence on the organisation and, second, by their personal relationships with the programme staff. Internal and external evaluators can perform all three types of evaluation – planning, process and outcome. However, certain situations are better suited to one or other type of evaluator. It may be more appropriate for planning and process evaluation to be undertaken by an internal evaluator, whereas outcome evaluation may be better served by an external evaluator.

Exposure measures how much of an intervention the target group actually experienced. This includes the number and length of intervention sessions and the materials used. The degree of exposure also relates to levels of participation and whether participants were actually reached by the intervention.

Fidelity Along with adherence and reinvention, fidelity is an aspect of programme implementation. Fidelity measures whether the programme was implemented as originally designed – that is, how true it stayed to plan. It can be measured either by an evaluator's subjective judgement or by more objectively documenting procedures, such as whether the required number of programme sessions were delivered.

Incidence is defined as the number of new cases displaying a particular phenomenon arising in a specific geographical area during a specific timescale.

Indicator is a measure that reflects a particular problem or condition. Indicators are used to substitute an objective or concept which cannot be measured directly or which will only be

observed in the future. The selection of appropriate indicators has to be founded in the literature, theories or previous research. Indicators are used to measure the quality of implementation or the outcome of an intervention.

Instruments refer to all the methods used to collect information on the target group, the evaluation, and so on. The most widely used instruments in evaluation are self-report questionnaires, tests, ratings, interviews and observation instruments. It is advisable to use instruments with well-founded objectivity, validity and reliability. The feasibility of the instruments should be pre-tested before using them on a wider scale.

Intentional changes are initiated deliberately to improve the intervention or evaluation.

Intention to use drugs is a mediating variable and is often used as an indicator in outcome evaluation. It deals with whether an individual believes that he or she will use drugs in the future. It is especially useful in primary-prevention programmes which target young children, when it does not make sense to ask if they have already used drugs.

Intermediate target group is a group of people targeted by an intervention who play a mediating role. It is hoped that they will be able to influence the future substance-use behaviour of the ultimate target group by passing on the contents of the intervention. Multiplier-centred approaches, peer-group approaches and family-oriented approaches all address an intermediate target group.

Interview In evaluation research, interviews are used to assess data on the implementation process and outcome. Interviews can differ in their degree of standardisation (whether structured, semi-structured or unstructured interviews), the type of contact (face-to face, telephone or written) or the number of people interviewed at the same time (individual or groups).

Knowledge about substance use This refers to the factual knowledge participants have about substance use. It is sometimes assumed that a lack of knowledge leaves people exposed to substance use risks, but this is a moot point since an awareness of the relevant facts and a better knowledge about substance effects sometimes can result in an increase of experimentation of new illicit drugs. Curiosity, experimentation or wanting to take risks are the first factors that draw many people into trying illicit drugs. Knowledge about substance use cannot be considered as a mediating variable or a proxy indicator of outcome evaluation.

Life skills enable people to deal effectively with the demands and challenges of everyday life through teaching adaptive and positive behaviour. These skills enhance individual

competence, reduce susceptibility to drug use and promote the health and well-being of children and adolescents. The following life skills are often targeted in prevention interventions: decision-making; problem-solving; creative thinking; critical thinking; effective communication; interpersonal-relationship skills; self-awareness; empathy; coping with emotions; coping with stress; and resilience.

Lifestyle relates to specific attitudes towards drugs among certain groups and in specific social or environmental conditions. The dance scene is an example of a lifestyle, where clubbing is associated with the use of synthetic drugs. A prevention intervention can have lifestyle change as one of its goals.

Mechanisms Realist evaluation seeks to understand why a programme works through an understanding of the action of mechanisms. The concept of mechanism refers to the choices and capacities which lead to regular patterns of social behaviour. They generate those patterns of behaviour which are deemed social problems and which are the rational of 'social programmes'. (Pawson and Tilley, 1997).

Mediating variables are supposed to be linked to substance-use behaviour by encouraging the changes in substance-use behaviour brought about by the prevention intervention. Two kinds of mediating variables can be distinguished: (1) mediating variables that are directly related to substance use and demand, such as attitudes towards drugs, intention to use drugs; (2) mediating variables that are indirectly related to substance use (on the supply and demand side), such as life skills, risk factors, protective factors, structural changes, lifestyle, cultural habits and problem behaviour, drug availability and drug prices.

Monitoring A continuing function that uses systematic collection of data on specified indicators to provide management and the main stakeholders of an ongoing development intervention with indications of the extent of progress and achievement of objectives and progress in the use of allocated funds (OECD, Glossary 2002).

Needs assessment (or needs analysis) is the systematic appraisal of the perceived phenomenon and the appropriateness of the proposed intervention. It is essential in order to avoid misjudging the size and character of the specific problem and therefore the need for the specific intervention. Different techniques can be used in needs assessment. The 'key-informant approach' involves identifying, selecting and consulting experts working in the field. The value of this technique is that a broad impression can be gained of the needs and services required for the target group. The limitation is that this impression could be based upon the experts' pre-existing prejudices or

biased information. It is therefore a good strategy to draw up a question structure to be used with all the experts. This will allow the answers given by different experts to be compared. The questions should address specific and concrete information (who, where, what and how) as another check against biased information. The 'community-forum approach' is based on an open meeting of community members. It can be used to gather information about the prevalence and incidence of a problem and about the characteristics of the target population. As in the key-informant approach, there is still the chance of biased information, either because of an under-representation or an over-representation of individuals affected by the problem within the community forum. Best results can therefore be obtained if the evaluator draws on a cross-section of the community. Again, it is helpful to ask specific questions to ensure the validity of the information. The 'rate-under-treatment approach' estimates the target population by analysing the experience of a similar prevention activity in another community. The assumption underlying this approach is that the characteristics and size of the two groups will be similar. The 'indicators approach' estimates the target population on the basis of epidemiological data from national statistical sources. These statistics are normally generated by various government agencies. Often, data are available on the populations of cities, towns and counties. Finally, specially commissioned surveys to assess the nature and extent of the problem are among the most direct and often most accurate way of estimating the need for an intervention.

Norms are unwritten rules of behaviour that are cognitively represented as beliefs or opinions. In the context of substance use, normative beliefs have been identified in empirical research as strong mediating variables for initiating substance-use behaviour. One objective of a prevention intervention can be to influence or to change these normative beliefs.

Objectives are specific and measurable statements regarding the desired outcome of the prevention intervention. For evaluation purposes, the formulation of objectives must specify the variables to be changed and establish measurable success criteria. A plausible, testable assumption must link programme activities to objectives, and objectives to intended outcomes. Unless vague goals are formulated into specific objectives, it will not be possible to implement an intervention or to assess the effectiveness of the intervention.

Objectivity is, along with reliability and validity, an important indicator for the quality of an instrument. It refers to the fact that the results yielded by the instrument must be independent of the person measuring the data – different people using the same instrument should achieve the same results.

Observation instruments are used to assess a specific situation or condition. In the case of prevention evaluation, observation is usually used to measure the implementation of an intervention, especially its fidelity. It is used either as the sole source of information or as an additional source to validate other measures of implementation. Observation by research staff or independent observers is considered the most objective measurement. It can be carried out either in narrative detail or by using standardised question structures or rating schemes.

Observer An observer attends a prevention-intervention activity to listen to and see the verbal and non-verbal behaviour and interaction of the participants and project organisers. In contrast to everyday observation, this empirical observation requires a plan which specifies the behaviours to be viewed, whether an interpretation of the observation is allowed, and where, when and how the observation takes place and is recorded. The observer should be introduced to the concept of the specific prevention intervention in order better to understand the observation's meaning, without explaining the specific evaluation hypothesis to avoid observer bias. Additionally, observers should be trained in the use of observation instruments.

Pre- and post-test The pre- and post-test design is a simple way to plan an outcome evaluation without the benefits of a control group. In this design, the only people measured are those who receive the intervention. They are tested (for example, on their knowledge, attitudes or intentions) before and after the intervention. The differences between the two measurements are then checked for statistical significance. The advantage of this design is its simplicity and the fact that it is not very time-consuming. The major drawback is that without a control group, it is not clear whether the results are really due to the intervention or to some other confounding factors.

Prevalence is the number of cases with a given condition or characteristic in the population of a particular geographic area at a certain time (e.g. the number of people who have taken cannabis in the last year).

Prevention intervention describes an activity carried out to prevent a substance-use behaviour. Prevention interventions can be realised in different settings and with different methods and contents. The duration can vary between one-off activities and long-term projects running for several months or more.

Problem behaviour Certain problem behaviours can be considered risk factors for drug use. These can include inappropriately shy or aggressive behaviour, lying, stealing, truancy, anxiety, and so on.

Programme A joint number of coordinated activities or interventions. The programme aims to achieve general objectives related to drugs.

Programme theory The theory of the programme explains how planners and other stakeholders expect the intervention to reach its objective. Describing this often implicit set of assumptions allows us to understand what is being implemented and why (or why not). The programme theory represents a hypothesis that can be tested and further developed (Chen, Rossi, 1987). It is the construction of a plausible and sensible model of how a programme is supposed to work (Bickman, 1987). It is a set of propositions regarding what goes on in the black box during the transformation of input to output, i.e. how a bad situation is transformed into a better one through treatment inputs (Lipsey, 1993). It is also the process through which programme components are presumed to affect outcomes and the conditions under which these processes are believed to operate (Donaldson, 2001).

Protective factors are a personal or social condition assumed to decrease the probability – in this case – of substance use. They are therefore the mirror image of risk factors. They may alter, or even reverse, predictors of negative developments and enable individuals to cope with negative life events. Individual protective factors that are consistently identified in the literature include social competence, problem-solving skills, autonomy and self-efficacy, sense of purpose and managing the future. Examples of protective factors within the family are care and support, boundary setting, high consideration of children and encouraging children's participation and involvement. More general protective factors include success in school and strong bonds with pro-social institutions.

Qualitative approach Qualitative approaches to evaluation aim to understand a programme or particular aspects of it as a whole. Instead of entering the study with a pre-existing set of expectations for examining or measuring processes and outcomes (quantitative approach), the emphasis is on detailed description and in-depth understanding as it emerges from direct contact and experience with the programme and its participants. Qualitative techniques rely on observation, interviews, case studies and other means of fieldwork. This approach can be appropriate alone or in combination with quantitative approaches, for example when a programme emphasises individualised outcomes, when there is concern for programme quality, or when the goals of a programme are rather vague. Qualitative data cannot easily be summarised in numerical terms, but they may be transformed into quantitative data.

Quality of intervention implementation The quality of implementation refers to how the participants or the practitioners rate the intervention and its quality. Indicators relating to perceptions of the intervention could include the acceptance of its contents, the degree of identification with or credibility of those contents, satisfaction with the intervention, personal benefit and relevance to the problem. Indicators which relate to perceptions of the intervention's quality could include the persuasiveness of the practitioners, their motivation and the interaction between practitioners and participants.

Quantitative approach Quantitative data are observations that can easily be represented numerically, such as answers to structured questionnaires. Quantitative approaches to evaluation are concerned primarily with measuring a finite number of specified outcomes. The emphasis is on measuring, summarising, aggregating and comparing measurements and on deriving meaning from quantitative analyses. Techniques often used in quantitative approaches are experimental designs and employment of control groups. They are particularly important when the primary aim of the evaluation is to establish programme effectiveness (qualitative approach).

Questionnaire A questionnaire is a list of questions, the answers to which can be systematically assessed. Depending on the answering mode, questionnaires may have open answers (where the respondents have to formulate the answers themselves) or closed answers (where they have to choose between several possible answers).

Realism An approach to evaluation and research based on a philosophy of science that is concerned with 'real world' problems and phenomena but believes these cannot simply be observed. It seeks to open the black-box within programmes or policies to uncover the mechanisms that account for what brings about change. It does so by situating such mechanisms in contexts and attributing to contexts the key to what makes mechanisms work or not work. Different mechanisms come into play in different contexts which is why some programmes or policy instruments work in some but not all situations.

Reinvention is an aspect of implementation along with adherence and fidelity. It refers to alterations in programme contents and implementation from the originally developed standard. It differs from lack of adherence by involving intentional and planned changes which were initiated in order to enhance the programme's effectiveness (as opposed to lack of acceptance, non-cooperation or unplanned change). Reinvention is especially important when programme effects could be boosted by tailoring its content to certain environments or populations.

Reliability, validity and objectivity are important indicators for the quality of an instrument. The reliability of an instrument is defined as the degree to which identical scores or values would be obtained in repeated data collections with the same subjects. Reliability ratings show whether the instrument yields consistent results. The effect of unreliability is to dilute or obscure real differences when they exist. If, for example, the outcome of an effective intervention is measured with an unreliable instrument, it may appear to be less effective than it actually is and vice versa.

Risk factors are personal or social conditions that are assumed to increase the probability of – in this case – substance use. They therefore mirror protective factors. Evidence suggests that children who are socially incompetent and aggressive are at greater risk of developing drug problems than those who are better integrated. A number of studies have shown that antisocial behaviour such as aggression can predict – as early as the first grade – an early initiation into substance use and later substance use disorder. The literature differentiates between early childhood risk factors (such as lack of social competence, lack of social support in the family), late childhood risk factors (lack of problem-solving skills, detrimental family norms, lack of self-esteem) and adolescent risk factors (negative influence of peers, reduced self-esteem due to adolescence).

Selection effects reduce the representative nature of a sample. They can refer to the fact that the participants that are easiest to reach are also the most likely to change. Projects relying on voluntary cooperation are therefore most often affected by selection effects. (See also 'Bias').

Statistical methods Depending on the kind of data, a number of different statistical procedures have to be followed when making group comparisons. Examples of procedures for group comparison of two variables include the Chi-square test, the T-test and analysis of variance (ANOVA). Procedures for more than two dependent variables are called 'multivariate comparisons'. An example of this is the multivariate analysis of variance. For the analysis of repeated measures (e.g. pre- and post-test), statistical procedures have to be used that are suitable for dependent samples. This rules out Chi-square tests.

In order to compare outcome data with process data, 'regression analysis' has to be used. This procedure determines, for example, whether and to what extent the effects of an intervention are due to the way it was implemented.

- **Analysis of variance** ANOVA is used to examine differences in the means of two or more independent groups, analysing how unlikely any observed difference would be in a set of random samples from a single population.

- **Chi-square test** The Chi-square test is used to compare data from two or more different groups. It can be used for categorical variables such as gender (as opposed to continuous variables such as age). The test converts a variable into categories and computes a Chi-square statistic. The statistic thus calculated provides information on whether the groups are comparable or whether they are significantly different.
- **T-test** The T-test checks whether the mean of a variable for subjects in one group differs significantly from that in a control group. It can be used on samples which are independent from, or dependent on, one another.

Structural changes The structural approach aims to change the environment, including the social environment so that individuals are more likely to behave in a desired way. Structural approaches include projects that address drug-related social problems, counselling centres for partner problems, and drug-free alternatives like youth centres and sports facilities.

Stakeholder Stakeholders include but are not limited to programme staff, programme clients, decision-makers and evaluators.

Substance-use behaviour refers to the consumption of a substance. Substance-use behaviour can be described in terms of the substances used (alcohol, heroin, cocaine, cannabis, etc.), the patterns of use (occasional use, regular use, recreational use, substance use disorder, addictive use, etc.) and the frequency of use.

Target group The target group is the group of people, households, organisations, communities or any other identifiable unit towards which a prevention intervention is directed. Two kinds of target group can be identified: ultimate target groups; and intermediate target groups. A careful analysis and estimation of the size and nature of the target group are essential preconditions when documenting the need for a prevention activity. It will also increase the quality and effectiveness of the project.

Theory-based evaluation examines conditions of programme implementation and mechanisms that mediate between processes and outcomes as a means to understand when and how programmes work (Weiss, 1997a, 1997b).

Theory of action All the assumptions made by funding authorities, planners and managers to explain how a public intervention will produce its effects and achieve its aim. The theory of action consists of relations of cause and effect linking outputs, results and impacts. It is often implicit, or at least partly so. Evaluation helps to clarify and to test those assumptions and theories.

Theory of change We can distinguish two types of theory of relevance to the health promotion and drug prevention intervention: explanatory theory and change theory. Explanatory theory sheds light on the nature of the *problem* and helps to identify the range of factors that interventions might seek to modify. In contrast, change theory informs the development and implementation of intervention strategies. Without a full, rational appraisal of the problem and possible solutions, interventions might easily address wrong or inappropriate variables or tackle only a proportion of the combination of variables required to have the desired effect. Theory provides the basis for judging whether all the necessary elements of a programme are in place. For example, a programme designed to encourage a particular behaviour, such as condom use, could not be expected to succeed unless it addressed the known determinants of that behaviour. Providing young people with information about condoms will have little effect unless they also have the skills to obtain and use condoms, they are able to be assertive in negotiating condom use with their partner, condoms are available, and so on (National Cancer Institute, 1997).

Time-series design is another extension of one-group designs extending the evaluation both backward and forward in time. This design involves a series of measurements on key outcomes at periodic intervals before and after the programme duration. It allows the evaluator to answer questions such as: 'Was the condition stable over the time period before the programme began, was it getting worse, or was it getting better even without the intervention?' (Weiss, 2007).

Ultimate target group is the group of people who will finally profit from the intervention. These individuals can be addressed directly by the intervention or indirectly via the intermediate target group. Two concepts may be used to identify the ultimate target group: population at risk; and population in need. The first covers a segment of the population with a high probability of developing the substance-use behaviour (e.g. children from broken homes, children with drug-dependent parents or siblings). In contrast, the term 'population in need' defines the target population as a unit with specific characteristics (e.g. all fifth graders).

Unexpected changes in the programme implementation or evaluation are changes or deviations from the programme plan that were unplanned and unforeseen. They can be negative in their consequences (lack of programme acceptance, school-district changes, budget cuts, etc.). But they can also be positive, such as unexpectedly high participation rates and additional sponsors.

Use of evaluation Instrumental use involves providing information for decision making and improving programmes and applying evaluation results in direct ways. Evaluation findings lead to immediate and specific actions such as modifying the delivery of a service, expansion of

a programme. *Conceptual use* involves using the results of evaluation for general enlightenment; more general learning takes place as a result of evaluation and the knowledge can be transferred to other interventions. Lastly, the *symbolic use* involves using evaluation results to legitimatise and sustain predetermined positions. It represents a political tool to legitimatise some choices or practices.

Validity, reliability and objectivity are important indicators of the quality of an instrument. Judgements of validity answer the question of whether an instrument really measures what you want to know and whether it is appropriate.

Varying perspectives on need Potentially, professionals, policymakers and target groups all have different perspectives on what constitutes a problem. What appears as a problem in one group may not be perceived as such by another. Research obviously cannot settle the issue of which perspective is the 'right' one, but it can eliminate conflicts that might arise from approaching drug phenomena from different perspectives. Part of the planning evaluation may include needs assessment from the several perspectives that may be involved in the intervention.

Bibliography

The publications listed below do not represent a comprehensive overview of evaluation literature. Rather, these are the articles and books which were helpful sources when designing the guidelines. Those marked with an asterisk (*) are deemed to be especially useful when conducting an evaluation to improve methodological bases.

*Chen, H.T., Rossi, P. (1987), 'The theory-driven approach to validity', *Evaluation and Programme Planning*, Volume 10, pp. 95–103.

Center for Substance Abuse Prevention (CSAP) (2002), 'Making prevention effective for adolescent boys and girls: Gender differences in substance use and prevention', *Monograph Series No 4,* Substance Abuse and Mental Health Services Administration (SAMHSA) www.samhsa.gov

Davidson, E.J. (2000), 'Ascertain causality in theory based evaluation', *New Direction for Evaluation*, Volume 87, pp. 17–26.

Donaldson, S.I. (2001), 'Mediator and moderator analysis in program development', in S. Sussman (Ed.), *Handbook of program development for health behavior research and practice*, pp. 470–496, Newbury Park, CA: Sage.

Donaldson, S.I., Gooler, L.E. (2003), 'Theory-driven evaluation in action: lessons from a $20 million statewide Work and Health Initiative', *Evaluation and Programme Planning*, Volume 26, pp. 355–366.

European Monitoring Centre for Drugs and Drug Addiction (2006), *A gender perspective on drug use and responding to drug problems*, Annual report 2006: Selected issues, EMCDDA, Lisbon: http://www.emcdda.europa.eu/html.cfm/index34880EN.html

European Monitoring Centre for Drugs and Drug Addiction (2010), *Prevention and evaluation resources kit (PERK)*, EMCDDA Manuals No 4, EMCDDA, Lisbon.

European Monitoring Centre for Drugs and Drug Addiction (1998), *Evaluating drug prevention in the European Union,* Scientific Monograph Series No 2, EMCDDA, Lisbon.

Green, J. (2000), 'The role of theory in evidence-based health promotion practice', *Health Education Research*, Volume 15(2), pp. 125–129.

*National Cancer Institute (1997), *Theory at a glance: a guide for health promotion practice* http://rex.nci.nih.gov/NCI_Pub_Interface/Theory_at_glance/HOME.html.

*National Institute for Health and Clinical Excellence (NICE) (2007), 'Behaviour change at population, community and individual levels', *Public Health Guidance No 6*, London.

*Patton, M. (1989), *How to use qualitative methods in evaluation,* (Third edition), Beverly Hills, CA, Sage.

*Pawson, R., Tilley, N. (1997), *Realistic Evaluation*, Sage publications, London http://www.communitymatters.com.au/RE_chapter.pdf (last accessed on 15 August 2011)

Rogers, P.J., Weiss, C. (2007), 'Theory-based evaluation: Reflections ten years on: theory-based evaluation, past, present, and future', *New Directions for Evaluation*, Issue 114, pp. 63–81.

*Rossi, P.H., Freeman, H.E., Lipsey, M.W, (1999), *Evaluation: a systematic approach* (Sixth edition), Sage Publications, CA.

*Weiss, C.H. (1997a), *Evaluation*, Second edition, Prentice Hall.

*Weiss, C.H. (1997b), 'How can theory-based evaluation make greater headway?', *Evaluation Review,* Volume 21, No 4.

*Yarbrough, D. B., Shulha, L. M., Hopson, R. K., and Caruthers, F. A. (2011), *The programme evaluation standards: guide for evaluators and evaluation users* (Third edition), Sage publications, CA.

Further reading

Botvin, G., Griffin, K., Diaz, T., Miller, N. and Ifill-Williams, M. (1999), 'Smoking initiation and escalation in early adolescent girls: One-year follow-up of a school-based prevention intervention for minority youth', *Journal of American Medical Women's Association*, Volume 54(3), pp. 139–143 http://www.med.cornell.edu/ipr/PDF/Botvin-et-al-1999-JAMWA.pdf

Connell, James P. and Kubisch, Anne C. (1998), 'Applying a theory of change approach to the evaluation of comprehensive community initiatives: progress, prospects, and problems', The Aspen Institute, USA
https://communities.usaidallnet.gov/fa/system/files/Applying+Theory+of+Change+Approach.pdf

Conrad, K. et al. (1991), 'Threats to internal validity in work site health promotion programme research: common problems and possible solutions', *American Journal of Health Promotion*, Volume 6(2), pp. 112–122.

Copple, B. (1997), 'Evaluating community coalitions for prevention of substance abuse: the case of Project Freedom', *Health Education Behaviour*, Volume 24(6), pp. 812–828.

Donaldson, S.I., Graham, J.W., Piccinin, A.M. and Hansen, W.B. (1995), 'Resistance-skills training and onset of alcohol use: Evidence for beneficial and potentially harmful effects in public schools and in private Catholic schools', *Health Psychology, Volume 14*, pp. 291–300.

Eisen, M., Zellmanb, G.L., David, M. Murray, D.M. (2003), 'Evaluating the Lions-Quest "Skills for Adolescence" drug education program. Second-year behavior outcomes', *Addictive Behaviors*, Volume 28, pp. 883–897.

Fawcett, S.B., Lewis, R.K., Paine-Andrews, A., Francisco, V.T., Richter, K.P., Williams, E.L., Copple, B. (1997), 'Evaluating community coalitions for prevention of substance abuse: the case of Project Freedom', *Health Education Behaviour*, Volume 24(6), pp. 812–828.

Gorman, D.M., Huber, J.C. (2009), 'The social construction of "evidence-based" drug prevention programs. A reanalysis of data from the Drug Abuse Resistance Education (DARE) Program', *Evaluation Review*, Volume 33(4), pp. 396–414.

Graham, J.W., Johnson, C.A., Hansen, W.B., Flay, B.R., Gee, M.S. (1990), 'Drug use prevention programs, gender, and ethnicity: Evaluation of three seventh-grade project SMART', *Preventive Medicine,* Volume 19, pp. 305–313.

Hansen, W. (1996), 'Pilot test results comparing the All stars programme with seventh grade DARE: programme integrity and mediating variable analysis', *Substance use and misuse*, Volume 31(10), pp. 1359–1377.

Hansen, W., et al. (1991), 'Programme integrity as a moderator of prevention programme effectiveness: results for fifth-grade students in the adolescent alcohol prevention trial', *Journal of Studies on Alcohol*, Volume 52(6), pp. 568–579.

Homel, R. (1997), 'Preventing drunkenness and violence around nightclubs in a tourist resort', *Situational Crime Prevention: Successful Case Studies,* ed. R. Clarke, Harrow and Heston, Guilderland, New York, pp. 263–82.

Homel, R., Hauritz, M., Wortley, R., McIlwain, G. and Carvolth, R. (1997), 'Preventing alcohol-related crime through community action: The Surfers Paradise Safety Action Project', in Holel, R., 'Policing for prevention: reducing crime, public intoxication, and injury', *Crime Prevention Studies,* Volume 7, pp. 35–90.

Hoorn, K., Dino, G., Kalsekar I., Mody, R. (2005), 'The impact of not on tobacco on teen smoking cessation: end-of-programme evaluation results 1998 to 2003', *Journal of Adolescent Research*, Volume 20, pp. 640–661.

Klepp, K., Tell, G. and Vellar, O. (1993), 'Ten year follow up of the Oslo youth study smoking prevention programme', *Preventive Medicine*, 22, pp. 453–462.

Home Office (2007), 'Evaluation of drug interventions. Programme pilots for children and young people: arrest referral, drug testing and drug treatment and testing requirements', Home Office Online Report 07, London http://ahfh.co.uk/ldngov/Substance%20Misuse%20%5BDrugs%5D/Evaluation%20of%20DIP%20Pilots%20-%200707%20%5BDrugs%5D.pdf (last accessed 28 March 2012).

Leone, L. (2008), 'Realistic evaluation of an illicit drug deterrence programme: analysis of a case study', *Evaluation*, Volume 14, pp. 9–28. Online Evaluation report available at: http://www.cevas.it/wordpress/wp-content/uploads/2010/10/Sperimentazione-Prefettura-Regione-Lombardia-2004.pdf (last accessed on 30 September 2011).

Leone, L., Celata, C. (a cura di) (2006), 'Per una prevenzione efficace. Evidenze di efficacia, strategie di intervento e reti locali nell'area delle dipendenze', Il Sole 24 Ore, Milan.

Leone, L., Scatigna, M., Pesce, C. (2011), 'Bridging supply and demand: relationship between cocaine seizures and prevalence estimates of consumption', submitted to *Public Health*. http://dx.doi.org/10.1016/j.puhe.2012.01.009 (last accessed on 5 July 2012).

Lincoln, R., Mustchin M. (2000), 'Clubs and violence: a follow-up evaluation of the Surfers Paradise Safety Action Plan', Faculty of Humanities and Social Sciences, Humanities and Social Sciences papers, Bond University, Australia.

Livet, M.L., Courser, E.M., Wandersman, A. (2008), 'The prevention delivery system: organizational context and use of comprehensive programming frameworks', *American Journal of Community Psychology*, Volume 41, pp. 361–378.

Mackenzie, M., Koshy, P., Leslie, W., Lean, M., Hankey, C. (2009), 'Getting beyond outcomes: A realist approach to help understand the impact of a nutritional intervention during smoking cessation', *European Journal of Clinical Nutrition*, Volume 63(9), pp. 1136–1142.

MacKinnon, D.P., Taborga, M.P., Morgan-Lopez, A.A. (2002), Mediation designs for tobacco prevention research, *Drug and Alcohol Dependence* Volume 68, pp. 69–83.

Mercier, C., Piat, M., Normand Peladeau, N., Dagenais C. (2000), 'An application of theory-driven evaluation to a drop-in youth center', *Evaluation Review*, Volume 24(1), pp. 73–91.

National Centre on Addiction and Substance Abuse (2003), 'CASA's new report calls for nationwide overhaul in prevention and treatment programs, big differences in why girls and boys use cigarettes, alcohol and drugs'
http://www.casacolumbia.org

National Institute on Drug Abuse (NIDA) (1997), *Preventing drug use among children and adolescents: a research-based guide*, (Second edition), Rockville, MD.

Nilsson, S., Spak, F., Marklund, B., Baigi, A., Allebeck, P. (2005), 'Attitudes and behaviours with regards to androgenic anabolic steroids among male adolescents in a county of Sweden', *Substance Use and Misuse*, Volume 40(1), pp. 1–12.

Pawson, R. (2006), *Evidence-based policy: A realist perspective,* Sage publications, London.

Pawson, R., Greenhalgh, T., Harvey, G., Walshe, K. (2004), 'Realist synthesis: an introduction, ESRC Research Methods Programme', RMP Methods Working Paper Series, 2/2004, London http://www.communitymatters.com.au/RE_chapter.pdf (last accessed on 15 August 2011).

Pesce, C., Donati, A., Magrì, L., Cereatti, L., Monacelli, C., Giampietro, M., and Zelli, A. (2004), 'Behavioral and psychological factors related to the use of nutritional ergogenic aids among preadolescents', *Pediatric Exercise Science*, Volume 16, pp. 231–249.

Petróczi, A., Aidman E. (2008), 'Psychological drivers in doping: The life-cycle model of performance enhancement, *Substance Abuse Treatment, Prevention and Policy*, Volume 3(7).

Petticrew, M., Cummins, S., Ferrell, C., Finday, A., Higgins, C., Hoy, C., Kearns, A., Sparks, L. (2005), 'Natural experiments: an underused tool for public health?' *Public Health*, Volume 119(9), pp. 751–757
http://www.ncbi.nlm.nih.gov/sites/entrez?cmd=Retrieve&db=PubMed&list_uids=15913681&dopt=AbstractPlus (last accessed on 15 August 2011).

Pentz, M., Trebow, E., 'Implementation issues in drug abuse prevention research', In Leukefeld, D., Bukoski, W., *Drug abuse prevention intervention research: methodological issues*, Rockville, M D: NIDA, 1991.

Reuter, P.H., Trautmann, F., Pacula, R.L., Kilmer, B., Gageldonk, A., van der Gouwe, D. (2009), 'Assessing changes in global drug problems, 1998–2007, main report', RAND/TR-704-EC, http://www.rand.org/pubs/technical_reports/TR704.html (last accessed on 28 March 2012).

Rogers, P. (2002), 'Programme theory: not whether programs work but how they work', *Evaluation Models*, Volume 49, II, pp. 209–232.

Saint-Jean, G., Carlos A. Martinez, C.A., Crandall, L.A. (2008), 'Psychosocial mediators of the impact of acculturation on adolescent substance abuse', *Journal of Immigrant and Minority Health*, Volume 10, pp. 187–195.

Schwinn, T.M., Schinke, S.P., Di Noia, J. (2010), 'Preventing drug abuse among adolescent girls: outcome data from an internet-based intervention', *Prevention Science*, 11, pp. 24–32.

Sherwood, K.E., 'Evaluation of the Fighting Back Initiative' (2005), *New Directions for Evaluation*, Issue 105, pp. 15–38.

Sloboda, Z., and David, S. (1997), 'Preventing drug use among children and adolescents. A research based guide' (Second edition), National Institute on Drug Abuse (NIDA), Rockville, MD http://m.drugabuse.gov/sites/default/files/preventingdruguse.pdf

Spooner, C. (2009), 'Social determinants of drug use – barriers to translating research into policy', *Health Promotion Journal of Australia*, Volume 20(3), 2009, pp. 180–185.

Stame, N. (2004), 'Theory-based evaluation and types of complexity', *Evaluation* Volume 10(1), pp. 58–76.

Tatchell, T.W., Waite, P.J., Tatchell, R.H., Durrant, L.H., Bond, D.S. (2004), 'Substance abuse prevention sixth grade: The effect of a prevention programme on adolescent's risk protective factors', *American Journal of Health Studies*, Volume 19(1), pp. 54–61.

Tiessen, J., Hunt, P., Celia, C., Fazekas, M., de Vries et al. (2011), 'Assessing the impacts of revising the Tobacco Products Directive. Study to support a DG SANCO Impact Assessment', prepared for the European Commission Directorate-General for Health and Consumers, RAND Corporation.

Tilley, N. (2004), 'Applying theory-driven evaluation to the British Crime Reduction Programme', *Criminal Justice*, Volume 4(3), pp. 255–276.

Tobler, N.S., Lessard, T., Marshall, D., Ochshorn P., Roona, M. (1999), 'Effectiveness of school-based drug prevention programs for marijuana use', *School Psychology International,* Volume 20(1), pp. 105–137.

Vaeth, P. et al. (1995), 'Examining the link between provider roles and programme development: findings from a process evaluation of a community based prevention program', *Journal of Primary Prevention*, Volume 16(1), p. 5573.

Vigna-Taglianti F., Vadrucci S., Faggiano F., Burkhart, G., Siliquini R., Galanti M.R. (2009), 'Is universal prevention against youths' substance misuse really universal? Gender-specific effects in the EU-Dap school-based prevention trial', *Journal of Epidemiology and Community Health*, Volume 63, pp. 722-728.

Weiss, C.H. (1997), *Evaluation*, Second edition, Prentice Hall.

Weiss, C.H. (1995), 'Nothing as practical as good theory: exploring theory-based evaluation for comprehensive community initiatives for children and families', in *New approaches to evaluating community initiatives: concepts, methods, and contexts*, The Aspen Institute, (ed.) James Connell et al., Washington, DC.

European Monitoring Centre for Drugs and Drug Addiction

EMCDDA Manuals No 8

Guidelines for the evaluation of drug prevention
A manual for programme planners and evaluators
Second edition

Luxembourg: The Publications Office of the European Union

2010 – 135 pp. – 21 x 21 cm

ISBN: 978-92-9168-499-1
doi: 10.2810/51268

HOW TO OBTAIN EU PUBLICATIONS

Free publications:

- via EU Bookshop (http://bookshop.europa.eu);
- at the European Union's representations or delegations. You can obtain their contact details on the Internet (http://ec.europa.eu) or by sending a fax to +352 2929-42758.

Priced publications:

- via EU Bookshop (http://bookshop.europa.eu).

Priced subscriptions (e.g. annual series of the *Official Journal of the European Union* and reports of cases before the Court of Justice of the European Union):

- via one of the sales agents of the Publications Office of the European Union (http://publications.europa.eu/others/agents/index_en.htm).

About the EMCDDA

The European Monitoring Centre for Drugs and Drug Addiction (EMCDDA) is one of the European Union's decentralised agencies. Established in 1993 and based in Lisbon, it is the central source of comprehensive information on drugs and drug addiction in Europe.

The EMCDDA collects, analyses and disseminates factual, objective, reliable and comparable information on drugs and drug addiction. In doing so, it provides its audiences with an evidence-based picture of the drug phenomenon at European level.

The Centre's publications are the prime source of information for a wide range of audiences including policymakers and their advisors; professionals and researchers working in the drugs field; and, more broadly, the media and general public.

The EMCDDA's Manuals are practical handbooks aimed at professionals and grassroot practitioners working in the drugs field.

TD-31-12-508-EN-C

Publications Office

ISBN 978-92-9168-499-1