

RESEARCH AND PRACTICE OF HEALTH PROMOTION

**EVALUATION AS A
QUALITY ASSURANCE
TOOL IN HEALTH
PROMOTION**

VOLUME 6, 1

Publisher: Federal Centre
for Health Education

The Federal Centre for Health Education (FCHE) is a government agency, based in Cologne, responsible to the Federal Ministry of Health. Its remit is to design and implement measures aimed at maintaining and promoting health.

It develops campaign concepts and strategies, produces summaries of media and methods, cooperates with a variety of workers and agencies in the health education field, and carries out education measures both for the population as a whole and covering selected topics for specific target groups.

The FCHE uses research results to plan and implement its work, as well as to evaluate its effectiveness and efficiency. This research includes projects on selected individual topics, evaluation studies, and the commissioning of representative repeat surveys. In order to promote an exchange of information and experience between theory and practice the FCHE holds national and international conferences.

These studies and assessments, along with the results of specialist meetings, are published by the FCHE in its specialist booklet series on “Research and Practice of Health Promotion”. This is to be seen as a forum for scientific discussion. The aim of the series – like the existing series on sex education and family planning – is to further extend the dialogue between theory and practice.

**RESEARCH AND PRACTICE OF HEALTH PROMOTION
VOLUME 6,1**

**EVALUATION AS A
QUALITY ASSURANCE TOOL
IN HEALTH PROMOTION**



An expert report by Gerhard Christiansen, Federal Centre for Health Education, on behalf of the European Commission, DG Health and Consumer Protection.

Federal Centre for Health Education (FCHE)
Cologne, 1999

Die Deutsche Bibliothek – CIP-Einheitsaufnahme

Christiansen, Gerhard:

Evaluation as a Quality Assurance Tool in Health Promotion : an expert report / by Gerhard Christiansen, Federal Centre for Health Education (FCHE), Cologne. On behalf of the European Commission, DG Health and Consumer Protection.

[Ed. Katharina Salice-Stephan ; Ingrid Schmitt, Transl.: Davis & Jungbluth, Köln]. - Köln : BZgA, 1999

(Research and Practice of Health Promotion ; Vol. 6,1)

Dt. Ausgabe u. d. T.: Christiansen, Gerhard: Evaluation – ein Instrument zur Qualitätssicherung in der Gesundheitsförderung

ISBN 3-933191-27-0

This volume forms part of the specialist booklet series "Research and Practice of Health Promotion", which is intended to be a forum for discussion. The opinions expressed in this series are those of the respective authors, which are not necessarily shared by the publisher.

Published by the
Bundeszentrale für gesundheitliche Aufklärung
(Federal Centre for Health Education - FCHE)
Ostmerheimer Str. 220, 51109 Köln, Germany
Tel.: +49(0)221/89 92-0
Fax: +49(0)221/89 92-3 00
E-Mail: Christiansen@bzga.de

All rights reserved.

Editors: Katharina Salice-Stephan, Ingrid Schmitt
Translation: Davis & Jungbluth, Köln
Composition: Salice-Stephan, Cologne
Printed by: Schiffmann, Bergisch Gladbach

Impression: 1.1.12.99

Printed on 100% recycled paper.

This publication can be obtained free of charge from:
BZgA, 51101 Köln, Germany
or on the Internet at <http://www.bzga.de>

Order No.: 60 806 070

PREFACE

In recent years the subject of quality assurance has become increasingly important in many sectors. In the field of health promotion, too, there is an intensive debate on which quality assurance tools and strategies can be used.

One quality assurance strategy which has already been applied and proved its value in health promotion is programme evaluation, which tests the efficacy of health promotion measures using previously established criteria. The Federal Centre for Health Education has been using this strategy for a long time in its national campaigns for the prevention of AIDS, drug prevention and sex education. For example, monitoring studies with a long-term design are carried out. These can be used, on the one hand, to examine the effectiveness of the campaigns and, on the other, to detect at an early stage whether individual elements of the campaigns need improvement. Another programme evaluation tool is used by the Federal Centre for Health Education when it subjects its TV and cinema advertising spots to pre-testing using a set procedure, for example. In this way, the communicative impact of the spots can be optimised before they are shown in cinemas or on television.

The Federal Centre for Health Education was commissioned and financially supported by the European Commission to investigate whether and in what way programme evaluation can be used systematically for quality assurance in health promotion. The results of this investigation are presented in this publication, Volume 6,1 of the specialist booklet series. The paper describes the most important elements of a quality-oriented evaluation concept which can be used to monitor and improve the development, implementation and success of health education measures.

By publishing this working paper, the Federal Centre for Health Education hopes to contribute at both a national and a European level to the discussion on the methods and strategies of quality assurance in health promotion.

Cologne, December 1999

Dr. Elisabeth Pott
Director of the Federal Centre
for Health Education

FOREWORD

The health promotion measures developed by the European Commission for the Member States of the European Union are geared to the maximum possible efficacy and verifiable evidence. Decision-making regarding the planning, promotion and continuation of programmes and projects is primarily based on drawing on the optimum proof of their efficacy and sustained effects. Reference to ambitious objectives or the professional recognition of specific methods are not, on their own, sufficient for the implementation of health promotion projects given the difficult budget situation.

Over the past years, the European Commission has provided financial sponsorship and organisational support for a number of scientific projects on various topics with differing perspectives. The aim of these presentations was generally to demonstrably reflect the efficacy of health promotion measures. The study by the Federal Centre for Health Education published here, on evaluation as a quality assurance tool, is among these projects.

This study investigates how programme evaluation can be used for quality assurance. It describes a quality-oriented evaluation concept geared to collecting and feeding back information on the efficiency of each project activity as it is carried out, working right from the start of every phase of a project and not only on its conclusion. This is intended to provide "preventive" safeguarding or improvement of the effectiveness of subsequent phases of the project. Projects evaluated in this way exhibit a high degree of quality but, in particular, a high degree of demonstrable quality.

This study also contains a number of proposals for action and planning principles as to the extent to which evaluation research and project development should be organised and integrated, in order to ensure and to demonstrate a high standard of quality. These instructions for action can also be seen as recommendations for guidelines for the planning of health promotion programmes and projects, to which the implementation of such measures should be geared.

Luxembourg, December 1999

Horst Kloppenburg
European Commission
Directorate General
Health and Consumer Protection
Unit F3

OUTLINE OF THE PROJECT

Project title: Evaluation as an instrument for quality assurance
in health promotion

Objectives:

- Description of the opportunities for utilising programme evaluation for the purpose of quality assurance of health promotion projects
- Proposal of an evaluation model for quality assurance in health promotion
- Recommendations for action for quality-oriented evaluation projects

Implementation period: November 1996 – August 1998

Project implementation: Federal Centre for Health Education
Ostmerheimer Str. 220
51109 Köln
Germany
Tel.: (+49) (0) 221 / 89 92-0
Fax: (+49) (0) 221 / 89 92-3 00

Project co-ordinator: Helene Reemann

Author of report: Gerhard Christiansen

Commissioned and financially supported by: European Commission
Directorate General Health and Consumer Protection
Unit F3
Jean Monnet Building
Rue Alcide de Gasperi
2920 Luxembourg
Luxembourg

CONTENTS

1.	THE TASK	13
1.1.	Initial situation	14
1.2.	Contents of the working paper	18
2.	KEYWORDS AND ELEMENTS OF A QUALITY-ORIENTED EVALUATION CONCEPT	19
2.1.	Evaluation	20
2.2.	Programme	22
2.3.	Quality	24
2.4.	Quality-oriented evaluation concept	27
2.5.	Summary	29
3.	PROGRAMME THEORIES	31
3.1.	What is a theory?	33
3.2.	Theory development: sources	34
3.3.	Theory development: working steps	35

3.4.	Programme theory and quality assurance	43
3.5.	Summary	44
4.	MONITORING OF IMPLEMENTATION PROCESSES	47
4.1.	Monitoring	48
4.2.	Analysis of the programme context	50
4.3.	Evaluation of programme components	52
4.4.	Monitoring of implementation quality	52
4.5.	Monitoring and programme evaluation	53
4.6.	Monitoring and quality assurance	54
4.7.	Summary	55
5.	IMPACT ANALYSIS	57
5.1.	Planning impact analyses	59
5.2.	Quality assurance in evaluation research	63
5.3.	Theory-driven impact analysis	67
5.4.	Summary	70

6.	SUMMARY AND RECOMMENDATIONS FOR ACTION	71
7.	REFERENCES	79

THE TASK

1

1. THE TASK

Evaluation plays a major role in the European Community action programme on health promotion, information, education and training. For example, Item 24 of the preamble of the action programme states: *“Whereas, in order to increase the value and impact of the action programme, a continuous assessment of the measures undertaken should be carried out, with particular regard to their effectiveness and the achievement of objectives at both national and Community level and, where appropriate, the necessary adjustments should be made.”*¹

Item 4 (7) of the 1997 Work programme of the Community action programme states: *“Evaluation and quality assurance will be developed as an integral part of the programme.”*² Evaluation of the project is the subject of the respective agreements when granting subsidies for health promotion projects.

This working paper was commissioned by the Service of the European Commission and makes a contribution to achieving this objective. It aims to examine how programme evaluation can be used for the purposes of quality assurance in health promotion projects of the European Commission and is designed to illustrate the demands which must be placed on quality-oriented evaluation projects.

1.1. INITIAL SITUATION

In recent years, various efforts have been made to spread the idea of quality assurance in European health promotion and to develop tools which can be used to improve the quality of concrete health promotion measures.

Various workshops and conferences have contributed to the opinion-forming process and the development of concepts at the European level. For instance, the International Union of Health Promotion and Education/Europe regularly holds conferences on the subject of effectiveness in health promotion and health education. Conferences directly addressing

1 European Commission: Resolution No. 645/96/EC of the European Parliament and Council dated 29.3.1996 on a Community action programme on health promotion, education and training within the framework for action in the field of public health (1996–2000), published in the Official Journal of the European Communities, L95/3 EN, dated 16.4.1996.

2 European Commission: Work programme 1997 of the Community action programme on health promotion, information, education and training (97/C18/04), published in the Official Journal of the European Communities, C18/11 EN, dated 17.01.1997.

the subject of quality were held in Turin in September 1996³ and, most recently, in Helsinki and Tallinn in May 1999, the latter entitled “Best Practices – Effectiveness and Quality of Health Promotion”. The WHO/EURO has set up a Working Group on quality and health promotion in the framework of the European Committee for Health Promotion Development (ECHPD).⁴ A European conference entitled “Quality Management in Workplace Health Promotion” was held in Erkner near Berlin in November 1996, specifically covering the field of health promotion at the workplace.⁵

Information on activities in the individual Member States can be found in the report on evaluation standards and quality assurance in health promotion interventions in the EU Member States, which was compiled on behalf of the European Commission.⁶

APPROACHES FOR QUALITY ASSURANCE ACTIVITIES IN EUROPE

Three approaches can be identified in quality assurance activities in Europe:

1. Criteria for action for the standardisation of quality in health promotion projects

The first approach attempts to get the quality of health promotion projects up to certain standards with the aid of criteria for action. This procedure is particularly suitable for spheres of work in which substantial empirical knowledge already exists. Examples of this include the criteria for health promotion at the workplace, proposed in the proceedings of the above-mentioned conference on “Quality Management in Workplace Health Promotion”⁷, or the “Adiposity quality grid for the assessment of measures aimed at weight reduction”⁸ published by the Federal Centre for Health Education (1994).

2. Continuing education measures for the improvement of planning quality in health promotion measures

The second approach tries to improve the quality of the planning of health promotion measures by way of continuing education programmes, handbooks or brochures. This is likewise aimed at standardisation, in that checklists or keywords are used to illustrate

3 MacDonald, 1997, contains conference proceedings.

4 The development of this discussion to date is illustrated in Parish (1995), Lethbridge/Parish/Hagard (1996), and in Springett (1998).

5 The proceedings were published by the Bundesanstalt für Arbeitsschutz und Arbeitsmedizin (Federal Institute for Occupational Safety and Health) (Quality management in workplace health promotion. Proceedings [Tb 81] of the workshop held in Erkner near Berlin in November 1996 [Dortmund/Berlin, 1997]).

6 International Union for Health Promotion and Education/Europe (1996): Presentation of existing standards for evaluation and quality assessment of health promotion interventions in the EU Member States (Commissioned by the Commission of the European Communities), Final Report. Woerden, NL. Netherlands Institute for Health Promotion and Disease Prevention. See also Bühnen-Armstrong & Bengel, 1997.

7 Bundesanstalt für Arbeitsschutz und Arbeitsmedizin (1997).

8 See also Töppich (1993).

those individual tasks and activities which are considered necessary for successful health promotion projects. Examples include the following:

- The handbook entitled “*Assuring Quality in Health Promotion*”, published by the British Health Education Authority (Evans/Head/Speller, 1994),
- The Swedish brochure entitled “*Succeeding with Health Promotion Projects – Quality Assurance*” (Berenson et al., 1996),
- The “*20-Point Questionnaire*” developed by the Department for Public Health Science of the Swedish Karolinska Institute for Health Promotion Training,⁹
- The Promotion Effectiveness Fostering Instrument (PREFFI) developed by the Netherlands Institute for Health Promotion and Disease Prevention.¹⁰

A study conducted by the Flemish Institute for Health Promotion (Van den Brouke/Lenders, 1997) shows that this quality assurance approach is indeed capable of improving the quality of planning: the quality of the planning of health promotion projects was measured with the aid of 26 questionnaire items in 9 groups, both before and after a continuing education measure based on practice-oriented planning handbooks. The results indicated a significant improvement in the quality of the planning, particularly in relation to the specification of target groups and objectives.

3. Programme evaluation as a means of quality assurance in health promotion

Programme evaluation is increasingly being discussed as a third approach. This is not only expressed in the Community action programme, but has also been examined by other organisations in recent years, such as the European Committee for Health Promotion Development (ECHPD) of the WHO/EURO.¹¹ Conferences on quality assurance in health promotion also often include the subject of evaluation on their agenda.¹²

One of the reasons for this occupation with the subject of evaluation and quality assurance is probably also the fact that, up to now, no independent procedures for quality measurement have been developed in health promotion, in contrast to health services, for example, where quality management procedures are currently being developed.¹³ It would now appear to be an obvious move to close this gap with the aid of evaluation research. Over the last 30 years, evaluation research has developed a number of analytical approaches and research techniques which can be used to examine the social processes

9 In: B. J. A. Haglund et al., *Creating supportive environments for health. Stories from the Third International Conference on Health Promotion*, Sundsvall, Sweden. ([Public Health Action Series, No. 3] Geneva: WHO, 1996).

10 See Van Driel & Keijsers (1997). The instrument is based on the results of a study on the effectiveness of health promotion and health education, sponsored by the European Commission and conducted by the International Union for Health Promotion and Education. See summary in C. A. Veen (1995).

11 See E. Ziglio, *Methods for the evaluation of health promotion programmes and policies*. Working paper. (Copenhagen: WHO/EURO, 1994). An evaluation working group was set up at the ECHPD on the basis of this planning paper.

12 A current example is given in Pelikan/Dietscher/Novak-Zezula (1998).

13 See Brook/MacGlynn/Cleary (1996) in this context.

and procedures taking place in social programmes, and which could also be put to use for health promotion projects.

Evaluation research gathered its experience in numerous programmes using a wide range of different intervention strategies and methods. It can therefore safely be assumed that evaluation research can also be applied to the various approaches and methods of disease prevention and health promotion, i.e. both for projects aimed at bringing about changes in knowledge, attitudes and behaviour, and also for projects designed to exert an influence on environmental and structural conditions or other health determinants.

Thus, the particular advantage of evaluation research lies in the fact that it can fall back on available resources. Its application represents a comparatively low-outlay strategy for quality assurance in health promotion.

Moreover, there are points of methodological contact between evaluation research and quality assurance. For instance, one of the goals of programme evaluation is to apply its results in order to improve the programmes examined, as formulated by the evaluation researcher Carol Weiss as long ago as 1972: *“Evaluation research means the immediate and direct improvement of social programmes”* (Weiss, 1972). The equivalent to this objective in quality assurance is the concept of “continuous improvement” of work processes as a fundamental principle of quality management – a quality philosophy which has also been discussed in health promotion in recent years (Parish, 1995; Oeretveit, 1996; Saan, 1997).

1.2. CONTENTS OF THE WORKING PAPER

Before going into a detailed description of the role that programme evaluation can play in quality assurance in health promotion, there is first a need to clarify the most important keywords – *evaluation*, *programme* and *quality* – in order to more precisely define and align the terminology arising from the different scientific traditions of evaluation research, health promotion and quality management. This is, however, primarily necessary in order to specify how the individual terms are used in the text that follows. The terms are defined in *Chapter 2*, at the end of which there is a brief overview of the most important elements of a quality-oriented evaluation concept.

The subsequent chapters elaborate the evaluation concept in more detail, indicating possible ways of realising it in the planning, implementation and result assessment of health promotion programmes.

Chapter 3 shows how the quality of health promotion projects can already be improved in the planning phase by elaborating a thoroughly formulated and differentiated intervention concept in the form of a programme theory. *Chapter 4* deals with measuring quality during the implementation phase, particularly with the possibilities offered by constant, systematic observation of the intervention process as regards the improvement of its effectiveness and quality. *Chapter 5* relates to impact analysis, i.e. that part of evaluation which is concerned with measuring the results and demonstrating the effectiveness of the programme. *Chapter 6* summarises the working paper in a brief recapitulation and reiterates the essential features of a quality-oriented evaluation concept in the form of recommendations for action.



**KEYWORDS AND ELEMENTS
OF A QUALITY-ORIENTED
EVALUATION CONCEPT**

2.1. EVALUATION

The starting point for the following clarification of terms is a brief definition: *evaluation means the systematic collection of information for the assessment of programmes*. The explanations given below of the four elements of this definition are intended to illustrate what evaluation is taken to mean in this working paper.

1. The systematic *collection of information* means that intervention programmes are examined with the help of empirical data. There are no restrictions as regards the tools to be used for obtaining the data. Psychological measuring instruments of the kind used in before/after control experiments, for example, are just as acceptable as the personal, face-to-face interviews used in survey research, or as qualitative intensive interviews. Other methods, such as the use of written questionnaires for participants or clients, the systematic observation of implementation processes, the analysis of time series or files, as well as the comparative meta-analysis of results of individual evaluation studies, are also regarded as suitable data collection techniques. This also applies to the typical information tools used in quality management, such as audits, quality circles or quality control based on standardised rating scales. Thus, evaluation is not limited to certain methods and procedures. Rather, it is “*pluralistic*” when it comes to data collection and analysis.¹

2. *Systematic* means that the information is collected in accordance with methodological rules. This includes, for example, public verifiability of how the data are obtained, as well as making every possible effort during collection, analysis and interpretation to ensure that the data obtained are free of bias.

However, a systematic approach also means that specifications must be drawn up as regards which data are to be collected and how a decision is to be reached as to whether or not the results expressed in the data can be interpreted as the effects of intervention. Evaluation encompasses a host of such systematising approaches: research designs, selection procedures, statistical analysis techniques, or also the theoretical analysis of intervention concepts. These rules and approaches are mainly those used in empirical social research – not least when it is a question of evaluating health promotion programmes or projects. This systematic approach to programme evaluation is taken to mean *evaluation research*.²

1 The term “pluralistic” was adopted from L. Sechrest (1993).

2 See Rossi/Freeman (1993).

3. The expression “*assessment of programmes*” is intended to indicate that the evaluation results are *fed back* to the programme directly: they are designed to assist in making concrete decisions in concrete programmes.

There is a host of possible decisions for which information can be provided, and they can be grouped under three general perspectives which are repeatedly emphasised in the literature on evaluation³:

- The perspective of accountability,
- The perspective of programme development,
- The perspective of expanding knowledge.

Among other things, the accountability aspect examines whether the anticipated success has been achieved and thus whether the use of the financial resources expended on the project was justified. The programme development perspective specifically looks for ways to improve the programme or individual programme elements. The perspective of expanding knowledge means that evaluation research also often helps to create general, empirically-based knowledge concerning the effectiveness of specific intervention strategies.

Consequently, the contents of evaluation research are not predefined from the outset: a separate decision as to which questions are to be examined must be made for each evaluation project. Two mutually complementary procedures are of particular importance in this context:

- On the one hand, the programme concept can be reformulated to produce a precise, realistic action theory, which contains predictions as to which interventions will induce which changes, and from which concrete research questions can be derived. This procedure is referred to as *theory-driven evaluation* (Chen, 1990; Chen/Rossi, 1989; Lipsey, 1993; Pawson/Tilley, 1997).
- On the other hand, the questions to be dealt with in the evaluation study are developed on the basis of close cooperation between the project developers, on the one side, and the evaluation researchers, on the other. One of the central tasks of this cooperation is to establish and precisely formulate the questions whose answers will be of the greatest benefit to the persons responsible for the project, the project developers and the project staff. The term *utilisation-focused evaluation* has been coined in the literature for this definition of evaluation contents based on the benefit of the information (Patton, 1997).

4. In order to emphasise the fact that the term programme is to count as one of the main keywords, it will be explained in detail in the next section.

³ For example E. Chelimsky (1997).

2.2. PROGRAMME

The language used in evaluation research does not completely correspond to that used in politics and administration, where more comprehensive and longer-term plans are often referred to as “programmes”, such as the EU’s Action programme on health promotion. Projects, on the other hand, are usually plans with a limited time frame, which are financed within a specific budget and often form part of a programme.

In the literature on evaluation *programme* is the term most commonly used to denote the *object of the evaluation research*. The word “project” is also used when less extensive programmes are involved. In evaluation research, sub-programmes or sub-projects are referred to as *programme components* (Bickman, 1985). This refers to a defined set of activities targeting a specific sub-goal of the programme. Components can also be evaluated separately, such as the individual TV spots in a TV campaign for health education, for example. In the present working paper, these terms are used in line with the terminology of evaluation research.

In the field of health promotion, the term *programme* can encompass completely different activities, such as a comprehensive health campaign for the general public running over several years, the reorganisation of workflows with the goal of developing a health-promoting hospital, or also a training course for promoting health-related skills.

A PROGRAMME AS A COOPERATIVE SYSTEM

Independently of programme goals and contents of this kind, a programme can be seen as a cooperative system. In detail, this means:

- A programme usually involves cooperation between a number of actors with occasionally different ideas, intentions and interests: the initiators and the beneficiaries of the programme from the fields of politics and administration, the financing agencies, the project developers and practitioners, and the participants and clients. This cooperative system also includes – last but not least – the evaluators. All these parties involved in a programme are often grouped under the term “*stakeholders*”, which is also used here.
- The term “cooperative system” also means that the cooperation between the stakeholders is seen as a process. This can be roughly broken down into three phases:
 - *Programme planning*: development of an action concept
 - *Programme implementation*: realisation of the concept
 - *Result assessment*: demonstration of effectiveness

- The programme process can be described as a strategy. The most important elements of programme strategies are *interventions*: methods, techniques and activities, such as motivating smokers to participate in a group discussion about addiction to smoking. Interventions are thus interactions with a specific group of people. In evaluation research, this group of people is generally referred to as the *participants* or *clients*. These people are usually recruited from a more or less precisely defined target group.

An intervention is intended to bring about specific changes in the participants, e.g. that the participants in a quit-smoking course give up smoking. The intended or actual effects of interventions are referred to as *intervention results* or *outcomes*.

- Every intervention is based on a more or less precisely formulated action concept. This concept can also be seen as an (*intervention*) *theory*, which describes the *mechanisms of action* via which the interventions are intended to bring about the intended outcomes and how the intervention works: who does what, with whom, and with what consequences.
- The practical realisation of a programme, its implementation, takes place in a special situation of circumstances and conditions: in a specific setting organised in a particular way, in a specific region with a specific culture; the intervention is aimed at a particular target group, which can be described in terms of gender, age, schooling and social stratum, and is carried out by a team of workers with specific training. All these framework conditions play a role in determining how the mechanisms of action work in reality and which intervention results or outcomes are actually achieved. The sum total of these framework conditions is referred to below as the context (Pawson/Tilley, 1997).

2.3. QUALITY

It is extremely difficult to find a general and, at the same time, wide-ranging definition of “quality”. For this reason, the discussion of quality in industry (Garvin, 1984) or in the field of health care (Donabedian, 1988; Blumenthal, 1996) led to the proposal that, instead of finding a global definition, a distinction should be made between different points of view and that these should be described. This proposal is to be taken up here.

A rough distinction can be made between three perspectives in the discussion of quality in health promotion:

- the accountability perspective,
- the stakeholder perspective, and
- the expert perspective.

THE ACCOUNTABILITY PERSPECTIVE

The *accountability perspective* is primarily concerned with the question of how the financial resources provided for intervention programmes can be used efficiently. Thus, quality assurance requires the coordination of available resources and targeted programme results with the aim of improving efficiency.

This quality perspective, in particular, is currently also to be found in many publications on the subject of quality and health promotion – usually in conjunction with a reference to the increasing scarcity of financial resources. These publications offer two versions of the argument that health promotion can be justified by more quality: in the first, quality improvement means creating additional effectiveness or quality on a given financial basis, while the second version argues that quality assurance means maintaining the previous level of quality on a lower budget.

THE STAKEHOLDER PERSPECTIVE

The *stakeholder perspective* considers the point of view of those persons who in some way expect a benefit from a product or a service, such as a health promotion measure. The stakeholder perspective means that the expectations, wishes and needs of the various stakeholders should be incorporated into the contextual definition of quality in each individual case. This particularly applies to customers, patients, clients and other recipients of products or services. The stakeholders whose quality expectations have to be taken into account also include external cooperation partners, financing agencies or interested

political institutions. Not least, they also include the staff and their expectations as regards the quality of the advance services to be rendered by other organisational units within the programme (“internal customers”). Not only the explicitly expressed wishes should be used in order to define quality in this context, but also the needs resulting from the lifestyles and problems of the respective stakeholders.

THE EXPERT PERSPECTIVE

The *expert perspective* is that of the programme developers, project managers and the staff. This so-called “technical quality” has two dimensions: firstly, it must provide suitable and effective instruments and, secondly, it must use these in the right way, so that the intended outcomes of interventions are achieved (“doing the right things right”).

In health promotion, technical quality is founded on knowledge concerning the relationships between the effects of interventions, based on experience and scientific insights.

QUALITY ASSURANCE AS A MANAGEMENT CONCEPT

Where quality assurance is to be implemented in concrete health promotion programmes, all three perspectives have a role to play, and none of them can generally be prioritised above the others. It is the task of the “programme management” responsible for planning and running the programme to achieve a balance between the various quality perspectives. Programmes can therefore not be regarded as a simple sequence of target-setting, implementation and outcome. Rather, they run in accordance with complex strategies which coordinate the (quality) expectations of the stakeholders, the technical demands of an intervention process which is as effective as possible, and the resources of the programme. Thus, quality assurance is often seen as being a management concept – in this case, a management concept for organising health promotion programmes. A whole series of specific principles of action have developed for quality management of this type.

Five of these principles, which appear to be of particular importance for the planning and implementation of health promotion programmes, will be explained briefly below:

Outcome orientation

Quality assurance in this sense requires that the programme be geared to clearly describable and essentially demonstrable intervention outcomes. These outcomes are intended for a closely defined client group, with the quality and benefits anticipated by this group also being taken into account. Another important factor is that the outcomes should be economically feasible. Programmes geared *exclusively* to highly rated, general goals are not regarded as being sufficiently justified.

Stakeholder orientation

The interests, expectations, lifestyles and working conditions of the stakeholders are all taken into account in programme management, when establishing the outcomes and during the programme planning and implementation phases. This applies, in particular, to the clients and staff, but also to cooperation partners, interested political bodies or the agencies responsible for the budget.

However, stakeholder orientation also means that it is impossible to simply assume a consensus between the various groups involved. The prerequisite for formulation of a clear quality assurance concept is a process of coordination and consensus formation with the goal of establishing a common view of the quality of a health promotion measure.

The principle of continuous quality improvement

The knowledge of the (intervention) experts on which technical quality is based is not regarded as constant. It changes in a cumulative process of *continuous quality improvement*. This principle can be summed up in just a few words as follows: working procedures and processes are subjected to constant empirical review (quality measurement). This review is primarily regarded as a possibility for identifying quality-impairing problems and discovering approaches towards solving them, in order to use these insights to improve working procedures during subsequent planning work. In keeping with the preference for simple schematic representations which prevails in the management sciences, this concept is represented in the form of a circle, known as the Shewhart or Deming cycle, which interconnects the elements Plan, Do, Check and Improve (Act) as a cyclic procedure (Deming, 1986).

Process orientation

The principle of continuous quality improvement is not primarily geared to the end-product, but to the entire *process* of services, interventions and other activities that are intended to bring about the intervention outcomes. Decisions and measures aimed at improving quality and increasing effectiveness should preferably be taken at the start of the process.

The principle of preventive action

The principle of constant improvement creates the opportunity of learning from experience. It becomes easier to identify possible mistakes. As a result, project planning activities can be geared to avoiding mistakes from the outset, rather than to retrospective correction of the end-product.

This brief attempt to clarify the term “quality” is geared to the possibilities of evaluation research as a quality control tool. Thus, no mention is made of the numerous quality instruments used in industry or in health care, and to some extent also in health promotion, in order to achieve, safeguard and improve quality.

In the following paper, evaluation research is treated as a quality assurance tool, and a description is given of an evaluation concept which is geared to the objectives of quality assurance as defined above.

2.4. QUALITY-ORIENTED EVALUATION CONCEPT

2

OUTCOME ORIENTATION

First of all, it is an obvious move to use evaluation research to measure quality: with its emphasis on demonstrable programme results, *outcome orientation* demands the use of valid, reliable research methods for measuring the intervention results and their change over time. Evaluation research can provide appropriate methods and research techniques.

CLIENT AND STAFF ORIENTATION

Being social research, evaluation research is also particularly suitable for obtaining reliable and unbiased information for the realisation of *client and staff orientation*: after all, one of the most important fields of application of empirical social research is the interviewing of voters, consumers, patients, participants in social programmes, and also of the staff in companies, hospitals and other organisations. The subjects covered by interviews in this context include, for example, the rating and acceptance of specific services or the consequences for the behaviour of the users of these services.

PROCESS ORIENTATION

However, the *principle of continuous improvement* demands that the evaluation project should not be restricted to measuring the ultimate results. It must also gather information on how the intervention outcomes came about. Only in this way can the research results be used to draw conclusions as to whether the implementation procedures or the action concept can be improved. Consequently, the evaluation study should provide data on the entire effect chain of the programme – starting with data on contacting the clients and continuing with activities and procedures which constitute the actual intervention, all the way to data on the intended (and also unintended) changes in the knowledge, attitudes and behaviour of the programme participants. The subject examined in quality-oriented evaluation research is thus the entire programme – entirely in the spirit of *process orientation*.

THE PRINCIPLE OF PREVENTIVE ACTION

The *preventive orientation* of quality assurance, geared to avoiding mistakes, demands that the systematic collection of information should start as early as possible and examine the extent to which the intervention concept is implemented faultlessly even during the implementation phase itself. However, avoiding mistakes at an early stage also means reviewing the actual concept before it is applied. This can be done by drawing up the most precise possible formulation of the relationships between the intervention activities, their mechanisms of action, the conditions of the context in which the programme is implemented, and the outcomes. This makes it possible to examine in advance whether the planned intervention activities are feasible – given the available resources – and it can be estimated whether or not they will have the intended effects at all. This prior review of feasibility is, above all, also a necessary precondition for the practicability of the evaluation study: only interventions and outcomes that can be realised can serve as measurable objects of research.

COMPONENTS OF A QUALITY-ORIENTED EVALUATION CONCEPT

Quality-oriented evaluation is thus interpreted as being a process which accompanies a programme from beginning to end – from elaboration of the programme concept and its implementation to measurement of the outcomes and final assessment of the efficacy of the programme, all the way to proposals for potential improvements. The organisational prerequisite for this evaluation concept is close cooperation between the project developers and the evaluation researchers, and this should likewise start during the planning and concept development phase.

A proposal for an evaluation concept geared to these quality principles is presented below. It is made up of three components:

1. The *programme theory* describes the intended mechanism of action: who does what, how, with whom and with what results? The aim is to review the achievability of the results, improve the effectiveness of the envisaged interventions and provide a theoretical basis for the evaluation study.
2. *Implementation monitoring* checks whether the mechanism of action has been triggered and is running as envisaged in the programme, thus also creating possibilities for improvement at an early stage, even while the programme is in progress.
3. The *impact analysis* examines which intervention outcomes have occurred and the extent to which they were induced by the interventions applied, the aim being to achieve continuous improvement of the programme, e.g. during the next round of implementation, or under different conditions.

These three sub-tasks are mutually dependent and cannot be carried out separately as an independent “formative” evaluation (of the implementation process) or “summative” evaluation (of the outcomes).

2.5. SUMMARY

In the debate on quality assurance in health promotion, evaluation research is regarded as one possibility for measuring, assessing and improving the quality of individual intervention programmes or projects for health promotion. It would be one of several quality tools (e.g. the quality circle), the applicability of which in the field of health promotion is also the subject of discussion or trials.

The application of evaluation research would have the advantage that use could be made of the organisational experience and effective methodological apparatus that it has developed over the past few decades in the context of evaluating social, health and other intervention programmes. However, evaluation research needs to be adapted to the objectives envisaged in quality assurance.

Quality assurance is interpreted as being a management concept which works on the basis of the following principles, among other things, when planning and realising (health promotion) projects:

- *Outcome orientation*: quality assurance presupposes that the programme pursues a clearly described, substantially, organisationally and economically feasible and essentially measurable intervention result.
- *Stakeholder orientation*: the services and outcomes offered in the programme give consideration to the quality expectations, needs and living conditions of the clients, staff, cooperation partners and other persons with an interest in the programme.
- *Principle of continuous improvement*: quality is safeguarded by a cumulative process of constant improvement.
- *Process orientation*: the improvements primarily relate to the activities and procedures taking place in the course of the implementation process.
- *Preventive behaviour*: improving the quality of planning makes it possible to identify potential mistakes at an early stage, avoid difficulties from the outset, and enhance the effectiveness of interventions as a result of decisions made as early as possible.

An evaluation concept oriented to these quality principles comprises the following three components:

1. Precise formulation of the action concept in the form of a programme theory,
2. Monitoring of the implementation process, and
3. Impact analysis.

Each of these components has its own task: the programme theory describes the envisaged mechanism of action, implementation monitoring checks whether the mechanism of action has been triggered and is running as intended by the programme, and impact analysis examines whether and how well it worked.

The prerequisite for practical applicability of this evaluation concept is continuous cooperation between the programme developers/project staff and the evaluation researchers, starting at the time of programme planning.

PROGRAMME THEORIES

3

3.

PROGRAMME THEORIES

*“There’s nothing so practical as a good theory.”*¹ This guiding principle was formulated at the very beginning of evaluation research in the social sciences. It is ascribed to Kurt Lewin, whose social psychology research work in the USA in the 1930s and 40s laid important foundations for social intervention programmes and research into them.² Later on, however, not very many evaluation studies took up this suggestion, resulting in occasional complaints about the atheoretical nature of evaluation research (Lipsey et al., 1985; Shadish/Epstein, 1987; Wortman, 1983). The responsibility for this was primarily attributed to the methodological doctrine according to which programme effects had to be demonstrated by (randomised) control-group experiments. As a general rule, this led to a situation where only the effects of the intervention were considered (by comparing the outcomes in the intervention group with those in a control group), the result being that the internal procedures of the actual intervention process were treated like a “black box” and thus largely ignored (Lipsey, 1993; Pawson/Tilley, 1997). And in many cases, social programmes were launched without giving much thought to how and under what circumstances the anticipated effects would come about (Sechrest/Figueroa, 1993).

Since the mid-1980s, more attention has been focused on the idea of taking theoretical considerations as a guide in programme development and research planning. The goal is theory-driven evaluation (cf. *Section 2.1.*). The evaluation goals and the evaluation plan are derived from a programme theory, which is elaborated during the concept phase of the programme and serves, as it were, as a guideline for defining the intervention measures, for implementing them in a specific social context, for selecting the research questions to be examined in the evaluation study and, not least, for improving the quality of the programme. Over the last ten years, a more theory-driven direction of this kind has increasingly become one of the maxims in health promotion practice.³

1 Cited after H. E. Lück (1996). See also Marrow (1969).

2 These are primarily the research work which Kurt Lewin directed in the USA in the 1930’s and 40’s after his emigration there from Germany. The most well-known are the field experiments on the effects of educational styles on the social behaviour of children in 1937 and 1938 (Lewin/Lippitt/White, 1939) and the experiments carried out in 1942 on the effects of different communication strategies in small groups on changes in eating habits (Lewin, 1951).

3 See for example Green/Kreuter (1991), or Bartholomew/Parcel/Kok (1998), and the commentary on the latter by C. A. Heaney (1998).

3.1. WHAT IS A THEORY?

Like empirically oriented social scientists in general, the advocates of theory-driven programme evaluation take a theory to be a set of assumptions, principles and statements, on the basis of which specific forms of social action (activities, attitudes) and changes in them can be explained. In some cases, theories of this kind contain assumptions formulated in relatively general and abstract terms, such as the assumption, often used as a basis when planning education campaigns, that the communication of information leads to a change in attitudes and that, in turn, a change in attitudes leads to a change in behaviour. Above all, however, a theory consists of specific, reality-based statements or hypotheses concerning the effects of particular *independent (causal) variables* on *certain dependent variables*. These research hypotheses can additionally contain *intermediate variables*, which explain how an independent variable influences a dependent variable.⁴

This formal description of what the word “theory” is intended to mean can be supplemented by a number of content-related criteria which are mainly of importance for programme theories (Chen, 1990:40ff.).

- Programme theories are *outcome-oriented*: The theory defines which programme results can fundamentally be achieved, thus enabling a detailed decision to be made as to which intervention outcomes are to be achieved in the prevailing context and with the given resources.
- Programme theories are *action-oriented*: The independent (causal) variables primarily describe highly specific activities (commitment of resources, intervention and implementation activities).
- Programme theories are *intervention-oriented*: They are not so much formulated in abstract concepts, but rather give an explicit description of the course of the effect mechanism (in a precisely defined context), i.e. of how the intervention activities lead step-by-step to the anticipated results.

⁴ In the language of evaluation research, the causal factor is described as an independent variable, the factor caused as a dependent variable. The variables occurring between the independent and dependent variables are known as intervening variables. However, in this working paper, the term intermediate variable is used instead of intervening variable. This is intended to make it clear that intervening variables cannot necessarily be equated with (health-promoting) interventions.

3.2. THEORY DEVELOPMENT: SOURCES

Existing theories from *social science* can be used as programme theories; however, this is only possible in rare cases because, when it comes to health promotion (and also the majority of other fields of application of social programmes), there are only few reality-based, empirically confirmed theories which can be applied directly to the subject of specific health promotion programmes.

Nonetheless, individual theories from the fundamental disciplines of health promotion can be used as a starting point or reference framework for elaborating intervention strategies and their action mechanisms. For instance, in the “Michigan Jobs Program” (which is frequently cited as an example below), a course based on the concepts of modern learning theory was developed for improving the practical and mental skills necessary when looking for a job (Caplan/Vinokur/Price, 1997).

The planners of education campaigns occasionally choose the model of attitude change formulated by McGuire (1989) as their theoretical starting point.⁵

Other sources used in theory development are research syntheses or meta-analyses of existing evaluation studies, or of other research on relevant topics. The results of reviews of this kind are particularly useful if they indicate whether and how the intervention strategies dealt with actually work. For example, a research synthesis of drug prevention studies was commissioned by the Federal Centre for Health Education and yielded a number of pointers towards the utility of the life skills approach (Künzel-Böhmer/Bühringer/Janik-Konecny, 1993).

Another example of research syntheses on the efficacy of health promotion interventions is the study on the effectiveness of health promotion and health education sponsored by the European Commission and conducted by the International Union for Health Promotion and Education.⁶

In the case of particularly extensive and expensive programmes, it is advisable to perform *preparatory studies* in order to check individual elements of the theory, such as hypotheses as to the efficacy of certain intervention components, in reality prior to starting implementation.

5 See also J. Töppich, Qualitätssicherung bezüglich Zielgruppenansprache und Akzeptanz (Quality assurance with respect to addressing target groups and acceptance). (In: Bundesvereinigung für Gesundheit e.V. und Spitzenverbände der gesetzlichen Krankenkassen [Hrsg.]: Qualitätsstrategien in Prävention und Gesundheitsförderung – Leitlinien, Praxisbeispiele, Potentiale. Bonn: proceedings, 1996.)

6 A list of 16 individual reports on individual areas of health promotion can be found in the summarising report on this project by C. A. Veen (1995).

However, the most important sources for theory development are the concepts and the *professional knowledge* of the programme developers, project staff and other important stakeholders. Their general objectives and fundamental conceptual ideas on the programme – and also their ideas concerning concrete procedures and routines – implicitly contain theoretical assumptions or hypotheses relating to the effects of certain measures or the causes of specific outcomes. Formulating a programme theory means turning these implicit action theories into explicit ones.

3.3. THEORY DEVELOPMENT: WORKING STEPS

3

A programme theory can be elaborated in several working steps:

- Problem analysis,
- Identification of the effective intervention components,
- Definition of the outcomes,
- Description of the causal mechanisms.

The working steps cannot always be handled in a fixed order during the elaboration of a theory in practice, and the framework conditions imposed by the other working steps must be taken into account in every step. However, problem analysis is generally the first move.

PROBLEM ANALYSIS

Problem analyses for health promotion programmes describe the non-medical causes of health problems or the processes from which they originate. Examples of these are problems which can be attributed to the living situation of specific groups of the population, such as a particularly high prevalence of certain diseases or particular social preconditions for the development and maintenance of good health. Or they may be problems arising as a consequence of general social problems, such as the impairment of mental health as a result of unemployment.

A problem analysis will describe these undesirable circumstances as precisely as possible and demonstrate the extent of the problem, where and in which problem groups it occurs, and what its other consequences are. As far as possible, these data will be presented on the basis of epidemiological data. Other avenues of investigation are the questions of why the problem will not be self-resolving or self-limiting, and how it will develop in the future in the absence of intervention.

It is of particular importance for the selection of an intervention strategy and the elaboration of a programme theory to ascertain the causes of the problem and to describe, as precisely as possible, how the problematic and undesired state (of health) normally arises (without the influence of an intervention). In this context, it is not solely a matter of determining the underlying social causes. Programme analysis should examine whether intermediate variables can be identified between the cause and its undesired effects, which could then be used to understand the way in which the social cause leads to the problem. This is particularly important because awareness of the way in which the problem arises facilitates a decision as to what form intervention must take in order to achieve a realisable change in the undesired state. It is desirable for the analysis of the problem to be based on empirical research results.

One example of a problem analysis of this type can be found in studies of unemployment in the Detroit region of the USA during the recession in the early 1980s. The results of these studies were subsequently used as a basis for the theory-driven planning and evaluation of the “Michigan Jobs Program”, which will be cited as an example several times in the course of this chapter, precisely because of this theory-driven approach (and its proven effectiveness).

The point of departure for the problem analysis of this programme was the observation that, among many unemployed people, the loss of their job resulted in impaired mental health, taking the form of anxiety or depression, for example. These observations were confirmed by surveys of employed and unemployed people, and people who had previously been unemployed but were now back in employment. These surveys also showed that the symptoms disappeared on returning to employment. A whole series of possible intermediate variables was examined to explain the way in which unemployment leads to depression. The following causal chain was identified: in the event of prolonged unemployment, loss of a job leads to a poor financial situation. The general stress burden induced by this then causes the impairment of mental health. Hence, the most important intermediate variable was identified as being the stress arising for financial reasons. Another intermediate variable with a slightly less significant effect was the reduced ability of the unemployed to cope with additional crises in their lives, which further impaired mental health. Other potential intervening factors frequently occurring as a consequence of unemployment, such as marital conflict or reduced social contacts, had no effect. However, a number of modifying factors were noted – such as social support, feelings of self-esteem, or the ability to cope with personal crises – which, when present, attenuated the psychological effects of job loss (Kessler/Turner/House, 1987; 1988. Cf. Fig. 1).

The simple but empirically based model, found in the problem analysis, regarding the relationship between job loss and mental health indicates the most effective form of intervention: a change in the most powerful intermediate variable. In other words: the mental situation of the unemployed is most likely to be improved by a change in their financial situation.

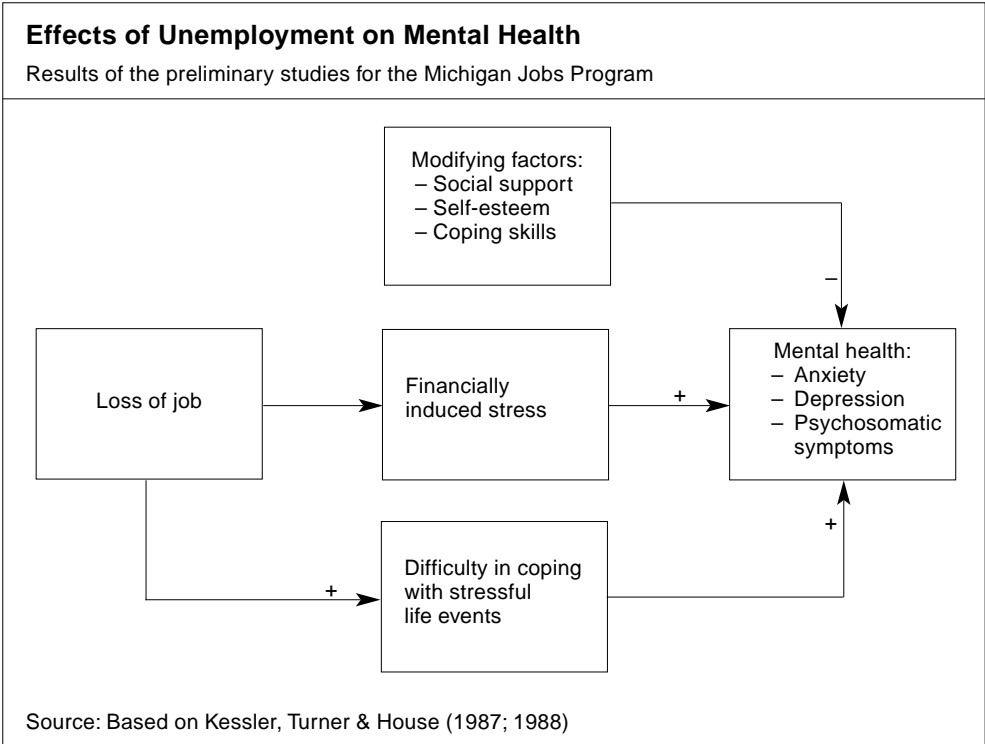


Fig. 1

This intervention option was not open to the project management, the Michigan Prevention Research Center of the Institute for Social Research at the University of Michigan. A programme was therefore selected which increased the likelihood of re-employment at the earliest possible stage, so that the stress situation would have no time to take effect and be able to impair mental health. The programme consisted of providing a course for the unemployed, aimed at improving their practical and mental skills in terms of job-seeking, with as little delay as possible after the loss of a job.

This example demonstrates the most important task of programme analysis: to develop a theory regarding the causes of the problem that throws light on the chain of effect of causal and intermediate variables, so that it is easy to see the point in this causal system at which intervention must take place in order to effectively improve the problem situation – always bearing in mind the available resources.

ESTABLISHING THE EFFECTIVE INTERVENTION COMPONENTS

Establishing the effective intervention components means specifying the intervention strategy selected. The individual intervention activities and services to be offered by the

programme are defined and described, taking financial and human resources into account. Determining those components and individual measures which would appear to be most effective with respect to the desired outcome is of particular importance in this context.

The “Michigan Jobs Program” (Price/Vinokur, 1995) can again be used as an example of the establishment of intervention components. The intervention strategy provided for a course for job-seekers who had been unemployed for no more than thirteen weeks. The course would teach them how to apply for jobs, i.e. the practical skills involved in job-seeking, writing applications and attending interviews. However, the decisive factor was building up and strengthening a repertoire of behaviours for coping mentally with job-seeking.

The overall intervention concept was divided into five individual learning goals for the course:

1. Successful recruitment of participants,
2. Creating trust in the trainers and the other course participants,
3. Increasing the motivation to apply effective action strategies in job-seeking,
4. Improving procedural knowledge and coping skills,
5. Transfer of what has been learned.

The didactic concept for the achievement of these learning goals was based on goal-oriented application of modern psychological learning theories, from which four general learning strategies were derived:

1. Active learning,
2. Social modelling,
3. Graded exposure to the learning content,
4. Reinforcement of self-efficacy and role-taking.

In turn, the learning goals, on the one hand, and the learning strategies, on the other, formed the basis for determining the actual intervention activities, i.e. for concrete behaviour on the part of the trainers towards the course participants. A matrix made up of the five learning goals and the four learning strategies was used to work out twenty of these effective intervention activities (Caplan/Vinokur/Price, 1997), which themselves formed the basis for precise instructions for action given to the trainers for their work.

The detailed establishment of concrete intervention components, as described above, should preferably be supplemented by an assessment of the strength of the intervention (treatment strength), in other words the intensity or “dosage” of the intervention. The majority of intervention tools and procedures achieve their intended effects only if they are applied sufficiently often and/or over an adequate period (Sechrest et al. 1979).

The following are possible dimensions and indicators of intervention strength:

- Contact intensity (duration of counselling session, length of an information text, duration of a group discussion),
- Frequency of application of a component per unit of time (e.g. number of course events during the intervention period, number of TV spot broadcasts per month),
- Duration (e.g. period over which a campaign or an offered course is run).

Establishing the intervention strength in the programme theory is also primarily of importance because it enables decisions of particular relevance in terms of cost to be derived, i.e. decisions concerning the employment of staff and other resources.

DEFINITION OF THE OUTCOMES

Definition of the outcomes is often the most difficult step in theory development. It cannot be achieved without a discussion process, which is often lengthy. At the end of the discussion process, politicians, financial sponsors, programme developers, evaluation researchers, and sometimes also other stakeholders, such as clients or staff, must reach a joint decision on a programme outcome, taking into account the interests of all those concerned, on the one hand, and formulated in clear, concrete and realistic terms, on the other.

This discussion is often problematic because it is frequently conducted on the basis of relatively abstract objectives or mere declarations of intent regarding the application of certain measures. These goals or measures are often described in very general terms, such as: “The development of a supportive, family-centred and skill-creating intervention system for families and children.” In this situation, the general concepts relating to measures need to be reformulated to yield a specific statement relating to the clients, from which it can be discerned what is intended to happen in whom. In the example above, a client-oriented formulation of this type might state: “As the outcome of the programme, parents will care for their children appropriately” (Patton, 1997:157). Such a statement would be a starting point for the further process of fine-tuning to specify the outcome precisely. Thus, there is also a need to decide what type of parents are involved, what activities are meant by “caring for children”, and what “appropriate” means in this context. The fine-tuning process extends up to the point at which the outcome is operationalised by defining the procedure used to record data relating to the respective outcome, the criterion used to decide whether the outcome has been achieved, and the means by which this decision is to be made. In some cases, minimum values can be specified that must be achieved at all events, or the maximum which can be achieved under favourable conditions. It is also important to define the period within which the outcome is intended to occur. In other words, the crux of the matter is to define the outcomes as precisely as possible.

An appropriate definition of the content of the outcomes is also sometimes difficult. It is logical in health promotion projects to select primarily *health outcomes* as the target variables for the programme, i.e. measurements of health attributes and conditions, such as are investigated in epidemiological studies. There are some projects where health outcomes are directly influenced by the intervention, meaning that they can then also be used as indicators for measuring effectiveness. But, in the majority of cases, using health outcomes to prove impact causes major practical problems. Even in the “Michigan Jobs Program”, in which the health outcome “mental health” was relatively closely linked to the intervention, the evaluation study had to use several measurement times extending over two-and-a-half years because restoration of mental health was the result of job-seeking which, in many cases, was successful only after a fairly long period of time. However, the chain of effect between intervention and health outcomes is often far longer still and the latter are usually also influenced by a host of other variables, so that the direct effects of the intervention are almost impossible to measure. It is therefore advisable to select a variable as the outcome which is, firstly, directly influenced by the health promotion programme and, secondly, itself has an effect on the health outcome (McKinnon/Dwyer, 1993; Tones, 1992). The programme theory should clearly illustrate the causal relationship between intervention, intervention outcome and health outcome, and this relationship should be capable of substantiation by existing research results if possible.

It is proposed that a further criterion for defining the content of outcomes be that benefits to the clients are described, in other words that the intervention brings about an improvement or changes in the attitudes, behaviour or living conditions of clients. An outcome is something that participants in a programme do, have or are as a reaction to the services provided by the programme. Simply taking part in a programme is not in itself an outcome and the number of participants is not an outcome variable (Plantz/Greenway/Hendricks, 1997).

Outcome definition is facilitated by problem analysis: outcome variables correspond to the undesired conditions which are intended to be changed by the intervention. The variables to be changed can be defined far more easily and unequivocally if the problem analysis reveals the social influencing factors and regularities which give rise to them. In the case of the “Michigan Jobs Program”, the problem analysis empirically confirmed that improvement of mental health impaired by the loss of a job would have to be the central result of the programme. In addition, this variable had already been precisely defined by the preliminary investigations. Firstly, a distinction had been made between the individual categories (anxiety, depression, psychosomatic symptoms) of the mental health impairment and, secondly, the indicators and research techniques for operationalising these abstract categories had already been selected.

The definition of the outcomes may, of course, also take the programme theory as its starting point. Once the intervention components have been established and the mechanisms of action described, possible outcomes which can in principle be achieved under these

preconditions will also become recognisable. The more precisely the programme theory is elaborated, the more likely this is to be the case. Another outcome of the programme can be derived from the programme theory of the “Michigan Jobs Program”, apart from mental health: re-employment (within a specified period), since this is the decisive mechanism of the intervention strategy that maintains or restores mental health.

The programme theory can also be used to discern the extent to which the intervention would cause unintentional effects. It is therefore important to specify undesirable outcomes of this type and their potential causes, because it may thus be possible to recognise ways of avoiding undesirable consequences of the intervention. If these cannot be avoided, it is often important to be able to incorporate them in evaluation of the programme. In the case of politically controversial programmes, unintentional (undesirable) effects are frequently actually the focal point of debate. It is then often the task of the evaluation study to clarify whether or not the programme actually does have such a negative impact as expected by some parties.

DESCRIPTION OF THE CAUSAL MECHANISMS

When it comes to describing the causal mechanisms, the aim is to illustrate the process which creates the relationship between the effective intervention components and the outcomes. This is achieved by working out and recording a causal flow between the independent variables of the intervention and the dependent outcome variables. As far as possible, this causal chain defines all the variables which act as connecting links between the components and services provided by the intervention, and the outcomes.

Figure 2 outlines a causal system of this type for the “Michigan Jobs Program”: the diagram first shows an overview of the content areas of the intervention, i.e. of the various learning strategies, on the one hand, and the targeted learning goals, on the other. (The individual components of the intervention, derived, as described above, from the respective combinations of learning goals and learning strategies, are not shown in detail.) In overall terms, the intervention brings about changes in a whole series of attitudinal and self-perception variables. If it is successful in generating these changes in attitude, this will have an impact on further intermediate variables which describe job-seeking behaviour. According to the programme theory, job-seeking is more effectively and more persistently pursued if the relevant attitudes are present. In turn, the behaviour of the course participants when seeking employment will most likely result in the desired outcome of re-employment and, indeed, re-employment in a position which is at least equivalent to the job which was lost. According to the preliminary problem-analysis studies, a further outcome expected to result from re-employment was an improvement in mental health.

Another mechanism of action is also conceivable: the changes in attitude generated by the course programme, such as the improvement of self-esteem, facilitate coping with

stress burdens if financial losses occur as a result of lengthy periods of job-seeking. This increased capacity for coping means that mental health is less severely impaired.

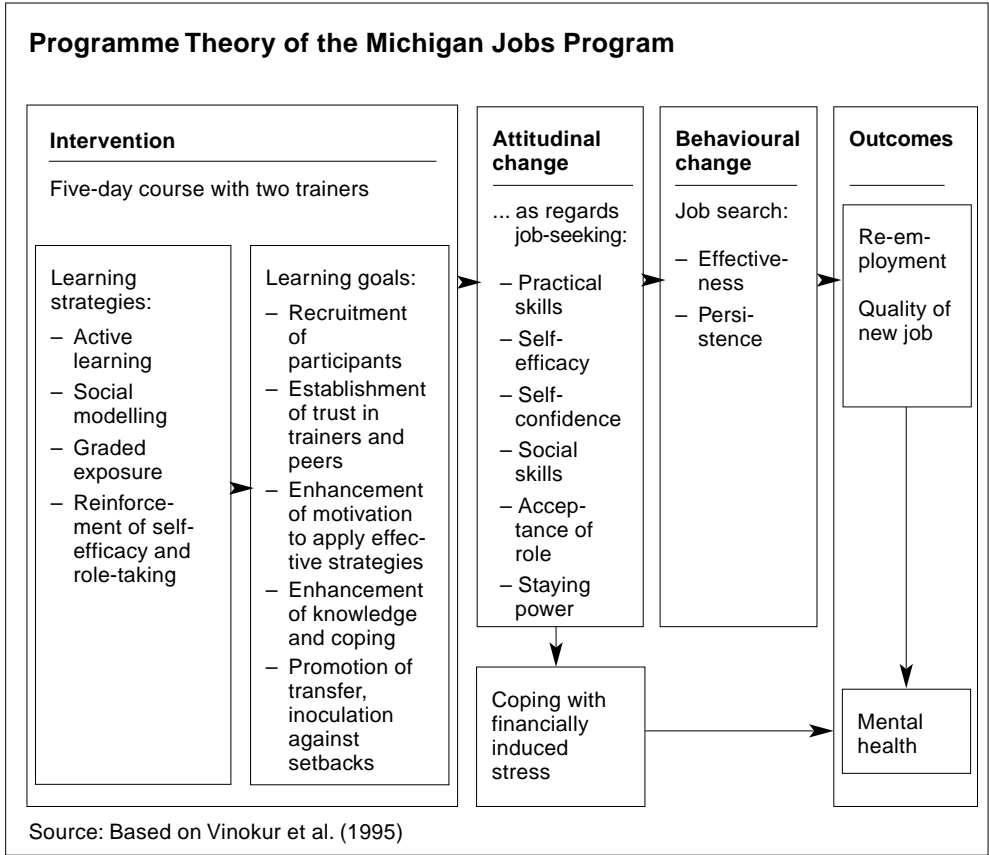


Fig. 2

In some health promotion programmes, some of the previously described working steps in the theory development process already form part of the concept development and programme planning stages. However, elaborating a programme theory goes beyond these: comprehensive consideration should be given to all the working steps, and the definitions, specified statements and relationships elaborated in this context should be systematised and organised in such a way that the end result is a fully formulated (written) model of the programme. This model can then be used as a guide for programme implementation, evaluation and, last but not least, quality assurance of the project concerned.

3.4. PROGRAMME THEORY AND QUALITY ASSURANCE

The practical benefit of a good theory, as promised in Lewin's guiding principle cited above, lies mainly in the fact that a theory which is based on reality and empirically oriented yields a model of the practical processes of the programme. This model can be used to simulate the course of the programme to a certain extent before it actually commences. In other words, realistic and specific predictions can be made regarding the precise circumstances under which highly specific effects can be anticipated. The earlier a model of this type is developed, the greater its practical benefit – particularly for quality assurance in terms of the five quality orientations listed in the previous chapter (see *Section 2.4.*):

- Theoretical “simulation” of the programme at an early stage is *preventive* in the sense of avoiding mistakes. It can be ascertained, before major investments are made, whether the programme can function at all in theory, whether it is feasible in the planned context or whether mistakes might perhaps develop which can be prevented before the start of the programme by altering the plans. The explicit formulation of the assumptions on which the intervention approach is based creates the preconditions for exposing logical inconsistencies and for considering whether there is, in fact, any causal relationship between the intended intervention and its effects.
- Elaborating, discussing and altering the programme theory means greater “planning quality”. As far as possible, quality assurance takes place at the start of the process ranging from planning to examining the results of the programme. This fulfils the quality management theory demand for *process orientation*, i.e. that the quality of products or services should primarily be achieved by means of high-quality development and elaboration processes.
- The cycle of *continuous improvement* also starts as early as the elaboration of the programme theory. First and foremost, a precisely formulated programme theory makes it possible to estimate the effectiveness of the programme in advance and, if necessary, to plan improvements resulting in greater efficacy. One possibility for improving efficacy lies in preparing implementation in such a way that the specifications of the programme theory are realised with as great a degree of “high-fidelity” as possible. Above all, however, the strength or “dosage” of the intervention activities can be increased if it is suspected that they will be insufficiently effective (Sechrest et al., 1979).
- With respect to *stakeholder orientation*, the greater degree of transparency alone is already an important step towards quality assurance. A thoroughly formulated pro-

gramme theory is a comprehensive and unequivocal model for negotiations regarding the outcomes and resources of the programme or for providing information to staff or clients.

- But, above all, a programme theory makes it possible to clarify at an early stage whether the programme is feasible at all in terms of *outcome orientation* and financial accountability. It is not only a tool which can be used to predict whether the anticipated outcomes can occur in theory, it can also be used as a basis for a cost estimate. Both of these items are prerequisites for setting up a cost/benefit budget which can be used to estimate whether the project is feasible with the given resources and under the given circumstances. And a decision can be made as to whether the relationship between the expected outcomes and the available resources is such that implementation of the project is justifiable.

The practical significance of programme theories for *evaluation research* lies primarily in the fact that they show evaluation researchers what are the right questions to ask in order to obtain answers which will allow assessment and quality improvement appropriate to the programme. A concrete description of the relationships between effective intervention components, mechanisms of action and outcomes is, at the same time, a compilation of the possible research hypotheses for the evaluation project. It determines the choice of the descriptive variables and indicators and allows unerring selection of those individuals who can provide the most reliable information on the course and effects of the programme. It is an important guideline in designing data recording tools and in making decisions regarding the evaluation methods.

3.5. SUMMARY

Evaluation research is geared increasingly to the concept of theory-driven evaluation. The evaluation objectives, plan and methods are derived from a *programme theory*. This describes the relationships between the effective intervention components, causal mechanisms and outcomes.

A programme theory is thus a systematically developed concept of the programme, formulated with reference to reality. The most important aspect in this context is to render explicit the theories of action of the programme planners, which are often only implicit. It can therefore be used as instructions for action, laying down the individual intervention measures and specifying how these should be implemented in the social context involved.

If the programme theory is elaborated in the initial concept phase and with the involvement of the major stakeholders, it can be used for “preventive” quality assurance at as early a stage as possible, particularly if questions regarding the feasibility of the programme are posed while it is being worked out:

- Is the programme logically possible, i.e. is there a causal relationship or a causal mechanism between the intervention and the outcome?
- Are the intervention measures sufficiently effective, do the interventions have a strong enough effect to bring about the desired outcomes?
- Is the programme feasible given the available resources?

The first quality improvements can be derived from the answers to these questions:

- Planning of “realistic” interventions, causal mechanisms and outcomes,
- Strengthening of the intervention activities,
- Coordination of intervention strategy and resources, cancellation of the project if necessary.

Thus, working out the programme theory at an early stage means that the cycle of continuous improvement can commence even at the planning stage.

The first step in theory development is usually analysis of the problem. Problem analysis for health promotion programmes describes the non-medical causes or processes of development of health problems. It specifies the undesirable outcomes, defines the (target) groups in which the problem occurs, and specifies the mechanisms and intermediate variables occurring between the fundamental social causes and the problem. If possible, this is done in so much detail that effective and practicable possibilities for intervention can be discerned.

The programme theory should describe the interventions of the programme in as much detail as possible and should, in particular, bring out the actual effective intervention components. This also includes an estimate of the intervention strength, i.e. of how strong the intervention activities need to be in order to bring about the intended outcomes.

Furthermore, the programme theory also states which outcomes are to be achieved, and also which are to be regarded as unintentional side-effects of the intervention. The outcomes should be defined as precisely as possible, stating which attitudes, modes of behaviour or living conditions of the clients are to be improved or altered, the period in which this can occur, and the way in which the outcomes can be operationalised and measured.

Above all, the programme theory describes the causal mechanisms – the process by which the impact of the effective intervention components is transmitted, step by step, to the outcomes. This is achieved by defining the intermediate variables which act as connecting links between the independent variables of the intervention and the dependent outcome variables.

**MONITORING
OF IMPLEMENTATION PROCESSES**



4.

MONITORING OF IMPLEMENTATION PROCESSES

There are two reasons for programmes failing:¹

1. The mechanism of action does not work as predicted by the programme theory (faulty concept).
2. The manner in which the concept was realised in a specific social context did not result in a (possibly effective) mechanism being set in motion in the first place (faulty implementation).

Two programme evaluation tasks can be derived from this. Firstly, programme evaluation must investigate the implementation process and collect data which can be used to document whether and how the programme concept was realised in a concrete social context. Demonstrating that the implementation process works means that the necessary precondition for the second part of the evaluation study, the impact analysis, has been met. The task of this analysis is to investigate whether the mechanism of action has led to the predicted outcomes. This chapter discusses evaluation of the implementation process.

Let us start by stressing once again that we are not saying that evaluators should deal only with the implementation process and dispense with evaluation of the results. Both of these programme evaluation tasks are regarded as being of equal importance.

4.1. MONITORING

Monitoring is a tool which is particularly suitable for evaluation of the implementation process. Monitoring means ongoing observation of the course of the programme using a systematically compiled set of data which are recorded regularly as a matter of routine. These include, for example, numbers of participants, participant data (socio-demographic attributes, perception and acceptance of intervention services), data relating to staff activities (contacts with participants, acceptance of and compliance with implementation specifications) or programme costs.

If the recorded data are continuously processed, evaluated and fed back into the system during the implementation process, this will lead to the development of a programme-related management information system. The data in this information system, which can

1 E. A. Suchman (1967) was the first to point out this distinction.

be called up regularly, can be used to reveal deviations and problems in the implementation process at an early stage, so as to be able to ensure that, wherever possible, quality improvement solutions can be found even while the programme is still in progress.

Monitoring of the implementation process primarily allows the following to be achieved:

- It provides time series data regarding the progress of the intervention with respect to inputs, services and participant responses.
- It checks the implementation quality, i.e. the accuracy with which the programme theory is applied to a real situation.
- It records changes in the original concept which become necessary in the course of the implementation process.
- Its results allow better interpretation and assessment of the measured programme effects.
- It increases the quality of data relating to the implementation process: on the one hand, recording data as soon as possible after the intervention activity achieves a high degree of measurement accuracy and, on the other hand, repeat measurements allow better estimates of random variations over time and hence an improvement in the reliability of data and results.

A host of requirements relating to content, and also to recording and processing methods, must be taken into account when planning a monitoring system, and some of these will have contradictory goals. For instance, on the one hand, the most comprehensive possible selection of indicators relevant to the content has to be made while, on the other hand, the time and effort expended on obtaining these must not make excessive demands either on the staff and participants or on the budget. It must be ensured, on the one hand, that changes during the intervention activities are documented as reliably as possible by way of repeated data recording, but, on the other hand, frequent data collection during the implementation process must not be allowed to alter the perception of the intervention by the participants or staff questioned in such a way as to have a distorting effect on subsequent outcome measurements. Planning a monitoring system is made easier by a programme theory which is comprehensive in content and has been elaborated in detail (see *Chapter 3*).

From the point of view of content, evaluation of the implementation process involves answering a few specific research questions which must be taken into account to a greater or lesser extent, depending on the programme concerned. These central research questions for implementation evaluation will each be described briefly in the sections below.

4.2. ANALYSIS OF THE PROGRAMME CONTEXT

The social context in which the programme is to be carried out should be examined thoroughly even before the start of actual implementation. This examination includes the extent to which the concept of the programme is consistent with the social reality of the context in which implementation is to take place. Many circumstances are conceivable that would make implementation more difficult: the qualifications of the project staff may be inadequate for specific intervention activities, it may not be possible to integrate the interventions into the lifestyle of the clients, or public opinion may be set against the project, to name but a few scenarios.

A checklist proposed by evaluation researchers as a basis for evaluating implementation processes may be useful for systematising potential contextual influences (Chen/Rossi, 1983; Chen, 1990):

Target groups (clients, participants, users):

- Social attributes (social situation, lifestyle),
- Values, attitudes,
- Compliance (willingness to cooperate),
- Receptiveness and receptive ability,
- Definition and selection rules for delimiting target groups.

Project team staff:

- Qualification,
- Specific qualification for the project,
- Motivation,
- Ability to make contact with the participants (rapport).

Circumstances of service provision:

- Technical framework conditions,
- Organisation of the setting,
- Acceptance of the intervention in the setting.

Organisation of implementation:

- Flexibility/rigidity of the organisational structure,
- The way in which working procedures are made routine,
- Quality control of working procedures,
- Capacity.

Relations with other organisations:

- Cooperating organisations,

- Competing organisations,
- Coordination processes and procedures.

Social environment of the participants:

- Social relations of the target groups or clients,
- Negative influences on the participants,
- Social support of the participants.

Macro-context (within the region or the Member State):

- National or regional framework conditions,
- Political influences,
- Social climate.

A similar checklist could conceivably be developed which is more precisely adapted to the problems and intervention strategies of health promotion programmes.

A checklist of this type can be used as a basis for elaborating the concrete forms which the items listed can take in the respective context. Hypotheses can be formulated as to which contextual conditions make the intervention more difficult or facilitate it, and in what way this occurs. These hypotheses can be checked in an independent research project. However, existing data or easily accessible sources of information will be sufficient for the majority of projects. Such sources of information might be as follows:

- Organisational knowledge and work experience of the project team regarding clients and settings,
- Expert knowledge from the field of social science regarding social structures, processes and regular features typical of the context,
- Analyses of official data and administrative documents,
- Interviews with typical clients, multipliers and other stakeholders with particularly good knowledge of specific areas of the social context,
- Reports of the results of evaluation studies of similar programmes in comparable contexts.

It is sometimes difficult to foresee the future development of specific contextual conditions. It is often necessary to monitor the influence of a particular context variable during the course of the programme. In such cases, suitable indicators will be developed and changes in these will be observed in the monitoring process.

4.3. EVALUATION OF PROGRAMME COMPONENTS

An evaluation of individual programme components should likewise be undertaken prior to the start of the actual implementation phase if possible. This may particularly be necessary if the programme comprises clearly delimited programme elements, such as TV spots, brochures or other media, but also if specific forms of counselling or other services are offered, which have been planned and developed more or less independently of the rest of the programme.

Component evaluations may be carried out in the planned context so that compatibility with the specific contextual conditions can again be checked here (e.g. the comprehensibility of individual media or acceptance among specific client groups). Or an investigation can be made, using laboratory experiments if appropriate, as to whether there are particular interactions with other intervention activities, e.g. whether the information effects of several different media are mutually reinforcing or cancel one another out. Component evaluations can also be conducted within the framework of a monitoring exercise, for example by considering a specific method of argumentation used by the trainers as an individual element in the context of a learning strategy and observing this element to see whether it changes over time.

An independent evaluation of individual programme components is primarily recommended because it is comparatively easy to generalise its results, so that standards for the construction and application of the relevant components can be worked out from a synthesis of these results (Bickmann, 1985). Starting points for this can be found in the design of TV spots for health promotion.

4.4. MONITORING OF IMPLEMENTATION QUALITY

Monitoring of implementation quality involves the question as to the extent to which implementation is faithful to the original concept, i.e. whether the concept is converted into those intervention activities with which the intended mechanisms of action can be triggered (Sechrest et al., 1979).

This conversion process is made considerably easier if there is an implementation manual describing the individual intervention activities and instructing the staff as to how they should behave in specific situations and with specific clients. A manual of this type also provides the potential for improving implementation quality simply by way of pro-

programme and context-related qualification of the project staff. However, it also facilitates the detection of deviations and mistakes which develop when the concept is realised in a specific social environment, and allows project-related criteria to be formulated for checking compliance with the concept. These can be drawn upon in the course of the monitoring process.

4.5. MONITORING AND PROGRAMME EVALUATION

Another primary application of monitoring is the systematic acquisition of information for analysing programme effects, i.e. for testing hypotheses regarding the relationships between effective intervention components, intermediate mechanisms and outcome variables. For example, the independent variables for measuring inputs into the intervention can be constructed from data relating to the frequency, coverage or duration of the intervention components applied. Survey or observation data recorded at the interfaces of the communication processes of the programme are particularly suitable for variables which map the mechanisms of action. These interfaces arise, for example, in the cooperation between individual organisational units of the programme, in situations in which staff and clients interact, in the use of programme media or as a result of contact between clients and individuals outside the programme. A wealth of preconditions for efficacy can be checked here, such as perception, understanding and acceptance of the content of communication.

The outcome variables can also be incorporated into the monitoring system. This presupposes that recording of the relevant data can be planned and organised as early as possible and can take place throughout the whole programme. In this way, the changes in the outcome variables can be observed in parallel to the intervention inputs in terms of time.

Measurement of these analytical variables does not necessarily mean that additional effort and expense is required for recording data. In many cases, data can be used for this purpose which are collected in any case for observing the implementation process or for project management.

4.6. MONITORING AND QUALITY ASSURANCE

A monitoring system is an important quality assurance tool. If they are commenced early, data collections and investigations relating to the social context, to individual programme components and to implementation quality provide an additional opportunity for improving a programme prior to its complete implementation. These improvements mean that the effectiveness of the programme can be increased at an early stage: inputs into the intervention which have initially proved to be too weak can be strengthened, the implementation quality can be described precisely and attempts can be made to control distorting and confusing influences relating to the context.

Continuous recording of data and observation of the course of the programme make programme monitoring comparable to certain central concepts of total quality management. The aim of the latter is to achieve quality primarily by means of having working processes constantly observed and, if appropriate, adjusted more accurately or changed with the objective of improving or maintaining the quality of the end-product in this way. Monitoring of health promotion programmes could also be similarly used for quality control by testing at regular intervals whether the course of the observed results of important outcome variables or intermediate variables corresponds to previously defined expectations.

In Total Quality Management, this type of quality control is backed by quality awareness on the part of the staff, which is geared to seeking out quality-impairing deviations and faults in the processes. A number of information tools are used to record and analyse any such deviations and faults, such as fault compilation lists, bar charts depicting fault frequencies, simple statistical analysis techniques or cause/effect diagrams. These tools can be used to convert the knowledge acquired from fault analysis into programme solutions and improvements. In turn, this quality-oriented observation of the process by the staff has similarities to the evaluation instrument of self-evaluation, where programme staff carry out recording, evaluation and processing of data themselves. They do this, for example, by observing and documenting their own working processes or interactions with clients and measuring and assessing these results against programme objectives. The particular advantage of self-evaluation is the fact that it arouses the attention of the project staff in terms of detecting implementation-related difficulties and problems, and increases motivation to think about problem solutions. The prerequisite is intensive training and supervision. In particular, it must be ensured that self-evaluation is compatible with the time-related, cognitive and emotional demands of the work.²

2 See M. Heiner (1996) regarding a project on the use of self-evaluation in child and youth welfare.

In addition, monitoring can be used as an information base for other quality management tools, such as for drawing up audits, i.e. for activity reports intended to list activities and services and allow an assessment of the extent to which quality demands are being met. Audits, in turn, can serve as a basis for quality certification.³

4.7. SUMMARY

Evaluation of the implementation process documents the course of the programme and proves that the programme could be realised under the conditions under which it was applied.

Monitoring is a tool particularly well-suited to this task. Monitoring is the constant observation of the course of the programme using a systematically compiled set of data recorded regularly and as a matter of routine. In addition to data relating to the administrative process, this includes information on the influence of contextual factors, on the efficacy and efficiency of individual programme components and on implementation quality faithful to the concept.

Monitoring is also suitable as a recording tool for the variables required for an impact analysis, i.e. for the strength of the effective intervention components, for the variables regarding the mode of action of the programme, and for outcome variables which can already be measured during the implementation process.

It is of decisive importance that the data collected be continuously processed and evaluated during the implementation process. These data, which can be called up regularly, can be used to detect any deviations and problems in the implementation process at an early stage, so that solutions for improving quality can be provided even while the programme is still running, if possible.

These improvements mean that the effectiveness of the programme can be increased early on. Inputs into the intervention which initially proved to be too weak can be reinforced, the implementation quality can be described more precisely and attempts can be made to control distorting and confusing influences arising from the context.

³ See the international standards on quality assurance systems, DIN ISO 9000 to 9004, in this context.

IMPACT ANALYSIS



5.

IMPACT ANALYSIS

The impact analysis studies the effects of the programme, ascertaining whether the anticipated outcomes have actually occurred, and investigating the extent to which they may be regarded as the result of the intervention and its causal mechanisms.

Measurement and analysis of the programme effects (“summative” evaluation) are just as important as the examination of implementation processes (“formative” evaluation). This needs to be pointed out once again in this context because the opinion has been repeatedly voiced in recent years that summative evaluation is not absolutely essential as long as a formative evaluation study is conducted (Sechrest/Figueredo, 1993). It is true that the importance of formative evaluation is stressed time and again in the evaluation literature, and the particular significance of measuring the quality of organisational procedures and processes is also referred to in the quality debate (Brook/McGlynn/Cleary, 1996). However, these references are primarily intended to make it clear that process quality plays a particular role in the quality of a product and that these production processes should therefore not be ignored in quality measurements or evaluations. Proof is nonetheless required as to whether these processes have yielded the expected results or not.

Impact analysis is of particular importance for quality assurance:

- It examines whether the programme was effective (outcome orientation),
- It also examines whether the mechanism of action worked in the way expected on the basis of the programme concept or whether the mode of action of the programme needs to be improved (continuous improvement),
- It attempts to answer the question as to whether quality has been achieved in the sense of client orientation,
- By answering these questions, it creates the basis for a decision as to whether the resources invested can be justified.

This working paper will have to restrict itself to a brief overview of the planning of impact analyses. A host of research techniques, control strategies, analytical and statistical procedures are available for detailed planning, which cannot be described in detail here. Extensive literature exists regarding practical evaluation research and adequate sources of advice on the methods and research techniques of empirical social and evaluation research can probably be found throughout the EU Member States.

Above all, it should once again be made clear at this point that evaluation planning, like the planning of other projects, is a complex decision-making process which needs to take a wide variety of conditions into account. On the one hand, this complexity offers the

advantage that “tailor-made” impact analyses are possible in principle, geared to the specific tasks and conditions and to the respective stakeholder interests. On the other hand, however, there are also enormous possibilities for making mistakes, entailing the risk of drawing incorrect conclusions regarding the quality of the programmes to be studied.

5.1. PLANNING IMPACT ANALYSES

A specific study concept, corresponding to the programme concerned and its contextual conditions, must be designed for the practical realisation of an impact analysis. This concept must lay down what is to be studied and how.

The general evaluation goals are determined in a process of coordination with the stakeholders. The outcome variables to be measured are selected and research hypotheses formulated regarding the relationships between effective intervention components, causal mechanisms and outcomes.

These study goals and research questions, in turn, form the specifications for deciding which methods and research techniques should be used to examine the individual questions. Two methodological problems generally have to be resolved here: procedures for measuring the outcomes and the other variables must be found or constructed, and a research design must be defined to control data collection and analysis.

5

MEASURING PROCEDURES

The following four methodological problems then have to be resolved when constructing measuring procedures:

1. Selection of the descriptive variables, i.e. the indicators which reproduce the terms used in the research hypotheses with the greatest possible substantial validity,
2. Selection or construction of recording instruments,
3. Selection of study units in which the data for measuring the variables can be obtained,
4. Definition of data analysis procedures.

RESEARCH DESIGN

It is the task of the research design to organise the data. Which data are collected when and from whom, which data are to be compared with which other data using which

statistical methods? The aim of the research design is to guarantee the validity of conclusions derived from the research results.

When evaluating social and health programmes, data analysis generally has to test the conclusion that the outcome variables have changed following exposure to the intervention. Changes of this type are best studied using a longitudinal design in which the outcome variables are measured at least once before the intervention and at least once at a later date at which the effect of the intervention is anticipated (before/after design).

Research designs based on the methodological principle of *control*, i.e. on the variation of specific variables being deliberately induced, limited or kept constant, are used to prove that the change in the outcome variables has been brought about causally by the intervention.

In order to create the clarity necessary in causal analyses regarding the fact that the independent causal variable precedes the dependent outcome variable, the former is controlled by producing it “experimentally”. In this (methodological) sense, programmes are regarded as “experiments”: the intervention is the independent variable which has been induced in a targeted and systematic manner.

MONITORING OF EXTERNAL INTERFERENCE FACTORS

In addition, in order to avoid incorrect conclusions being drawn, the influence of external interference factors must also be monitored. Changes in the outcome variables in before/after designs cannot be interpreted as proof of intervention effects without further ado, unless the possibility can be ruled out that other, external events between the before and after measurements have had an effect. Thus, for example, a public debate on smoking might possibly influence the “after” measurements of the results of a quit-smoking course, so that it is not clear whether the measured effects are attributable to the non-smoking training or to the public debate. External events or developments of this type, occurring between the before and after measurements, are described as time influences (*history*).

Other factors influencing the validity of longitudinal studies include personality changes (*maturation*) in the clients (independent of the intervention) or statistical regression influences.

External interference factors can also arise from the research methods used. These may be measurement effects in the respondents, if the “before” measurement influences the subsequent measurements (*testing*). Or the measuring instruments may change between the before and after measurements. This would be the case, for example, if staff conducting self-evaluation were to interpret their observations of client behaviour in a way which

was advantageous for the assessment of the programme and if this were to increase over time (*instrumentation*).¹

External interference factors can be monitored using control groups in whom no intervention takes place, but in whom before and after measurements are also taken. If there are external interference factors, their effects occur in both groups and the influence of the intervention can be demonstrated by comparing the results of the intervention group and the controls.

However, setting up control groups entails additional interpretation problems. For example, differences in the composition of the intervention and control groups make clear interpretation of the results difficult. However, this can, in turn, be controlled by random distribution of the test subjects over the two groups.

In the social sciences, randomised control group experiments with before and after measurements are frequently regarded as the instrument best suited to the detection of causal effects. It combines several monitoring possibilities in a single research design:

- Control of the independent variable by the intervention,
- Comparison of the experimental group and a control group with the same starting conditions at the time of the before measurement,
- Control of extraneous factors by comparing before and after measurements in the intervention group and the control group,
- Statistical control of other external influences by means of random distribution of the subjects between the intervention and control groups.

CONTAMINATING EFFECTS

However, this research design does not control all the possible interference factors. This applies, for example, to a group of interfering influences which can be summarised under the heading “contamination” (Mohr, 1995). This refers to distortions in the results of the “after” measurements, which arise from the intervention and/or from the experimental design of the evaluation project. Inadequate implementation, for example, may lead to such distortions, or mental processes of identification of the participants with the intervention team may lead to responses conforming to expectations in the “after” measurement. Specific participants in the control group may withdraw during the intervention because, unlike the participants in the intervention group, they are not particularly motivated to continue their participation, the result being that the two comparison groups

¹ The terminology comes from the classic classification of “threats” to validity of results of experimental and quasi-experimental study designs, devised by D. T. Campbell and J. C. Stanley (1966).

subsequently have different compositions. In addition, the measured difference between intervention group and control group may be contaminated if it is not possible to keep the members of the two groups strictly separate so as to prevent them exchanging information about their experiences with the intervention.

These contaminating effects are not controlled by control group designs. On the contrary, to some extent they only occur as a result of setting up control groups. They can be avoided by quality improvement in the intervention and by shielding against contextual influences.

STAKEHOLDER EXPECTATIONS

In addition to these methodological considerations concerning the measurement and control of the variables, expectations and criteria which derive from stakeholder interests must be taken into account when planning impact analyses. For instance, the study plan often has to take into account the fact that stakeholders not only expect proof of efficacy, but would also like to gain other benefits from evaluation studies, such as extending their knowledge of the subject of the investigation in general.

PROBLEMS OF RESEARCH ETHICS

Problems of research ethics, such as always occur when people are selected, surveyed or observed, are of particular importance. In the field of evaluation research, problems of research ethics are discussed primarily in connection with the use of control groups: is it acceptable to deny the members of the control group the improvement of life opportunities intended to be achieved by the intervention? Is it acceptable to monitor placebo effects by acting out an apparent social improvement for the members of the control group which, however, is bound to be virtually ineffective for methodological reasons? Control groups whose sole purpose is to monitor or measure the placebo problem are often rejected for the evaluation of social programmes because the pretence of a social improvement which does not actually occur is considered to be unethical (Young, 1966).

The fact that staff or clients might be overtaxed by the collection of data, or that organisational procedures might be impaired when implementing the programme, must also be taken into account.

FINANCIAL ASPECTS

Not least, financial accountability also plays a role. The general rule is that the costs of evaluation may never exceed a limited proportion of the total costs of the programme. An

approximate figure of ten percent is mentioned every now and again for health promotion projects. Lower figures occur where programme costs are very high, but where the programme can be investigated very satisfactorily with relatively economical evaluation projects. Higher figures are necessary if a ten percent share is insufficient to ensure an acceptable investigation of efficacy and/or if the decision as to continuing or stopping the programme would have major social consequences.

5.2. QUALITY ASSURANCE IN EVALUATION RESEARCH

The question of whether the quality of a programme can be proven depends on the quality of its evaluation study. Poor research increases the likelihood of making a programme appear effective, although it is in fact totally ineffective, and there is a risk that the clients of later applications of the programme will be deceived as to its efficacy. The consequences may be even more serious – particularly with respect to economic efficiency – if poor research results in an effect which is actually present not being discovered and proved, with the result that a health promotion programme which is actually effective is not used further.

A precondition for measuring efficacy is therefore that the study plan and recording methods of the evaluation study be designed so as to yield accurate, valid and reliable results. This means that quality assurance is also necessary for the evaluation research itself. Some possibilities for quality assurance of this type are listed below.

DISCUSSION OF PARADIGMS

In the past, various “paradigms” were proposed with the claim that these were the “correct” procedural approaches for the evaluation of various social programmes. There were essentially two paradigms (based on Van de Vall, 1993):

- The first paradigm required a study design involving randomised control group experiments, with quantitative selection and measurement techniques used for the recording procedures. The primary objective was to measure programme impacts and it thus focused on summative evaluation.
- The second paradigm contrasted this with a study design which provided for comparative methods using comparison of relevant sub-groups and required qualitative recording methods which were intended primarily to deal with the implementation process. It therefore focused on formative evaluation.

These two study strategies were considered to be opposites and were the subject of controversial debate for a long time. However, the debate was largely inconclusive and has subsided noticeably in recent times (Patton, 1997:290ff.).

There are major practical disadvantages inherent in the setting of such comprehensive research paradigms, since it is normally virtually impossible to adjust a programme to the methodological requirements of detailed study plans, such as those of an experimental study design, without compromising the intervention strategy. To do so would mean that a programme would be applied which was unable to fully exploit its own intervention potential because it had to take account of an allegedly “correct” research approach. Conversely, there is the risk of evaluation studies being performed which are not designed appropriately for the programme and thus yield incorrect evaluations. The rigidity of pre-defined research approaches contradicts the variety of practical programmes.

STANDARDISATION

Standardisation would be primarily expected to result in compliance with minimum research standards and the comparability of different programme evaluations. Here again, however, the question arises as to whether specifying particular research procedures for evaluation research in advance will not increase the likelihood of inappropriate evaluation approaches and hence of incorrect assessments. Given the current status of health promotion and evaluation research, setting standards for research procedures might sometimes result in less quality assurance in evaluation research, rather than more.

However, standardisation is not ruled out in principle and the first signs of standardisation can already be recognised. For example, recording instruments are standardised in that time-proven measuring tools, scales or questionnaire questions – e.g. for health behaviour indicators or regarding risks of disease – are documented, stating quality criteria, and recommended for use in evaluation studies, among other things.² This means that the quality of these measuring tools is relatively reliable, but that it is generally necessary to check whether they work in the relevant context (setting, target group).

It is a time-consuming and expensive process to ascertain the applicability of measuring tools under different conditions. Examples of this are the comprehensive research on the sense of coherence in the salutogenic model by Aaron Antonovsky³ or the documentation and research work by Ruut Veenhoven (1992) regarding satisfaction with life (happiness).

2 Some examples of relevance to health promotion are Bowling (1995; 1997) and Riemann/Gerber (1997).

3 A review in this context can be found in J. Bengel, R. Strittmatter and H. Willmann, What keeps people healthy? The current state of discussion and the relevance of Antonovsky's salutogenic model of health. (Research and Practice of Health Promotion, Vol. 4). Cologne: Bundeszentrale für gesundheitliche Aufklärung [Federal Centre for Health Education], 1998.

However, standardisation can contribute to increased quality in evaluation research only if it is founded on broadly-based research syntheses. The requirements for this are documentation of completed evaluation studies using descriptors which are as uniform as possible, collection and archiving of the studies and, above all, comparative analyses which investigate which measuring tools and research designs would be most appropriate for which health promotion approaches in which social contexts or settings. These results would then form the basis for a process of coordination between the programme developers and evaluation researchers with the task of proposing research standards on which agreement can be reached.

PROFESSIONAL QUALIFICATION

Keeping the qualifications of the researchers at as high and current a level as possible is another strategy for quality assurance. This can be achieved, on the one hand, by continuing training of the evaluation researchers, along the lines of survey research, for example, for which a number of special research and training institutions have been set up over the past decades. Another possibility would be to develop an evaluation research infrastructure which could offer expert evaluation research. The prerequisite here would be sufficient demand for evaluation research in general – not merely for health promotion – so that an efficient and effective market could develop.

ERROR CONTROL

The quality assurance procedure currently preferred in research in the social sciences is that of error control. It consists in anticipating potential sources of error and creating effective possibilities for controlling these in the research design.

Error theories and error systems are used to decide which errors should be controlled, and these have been developed in the various scientific fields on which evaluation research is based. There is not yet, however, any interdisciplinary, complete and clear standard system. Some of the error systems in existence tend to refer more to measurement errors, others to errors in analysis or design. There are major differences in terminology and the wide variety of potential errors listed makes it more difficult to decide which errors should be checked.⁴

⁴ An error system frequently used in evaluation research is the compilation of potential influences on the validity (threats to validity) of the results of experimental and quasi-experimental study designs, developed by Campbell and Stanley, (1966), and Cook and Campbell, (1979). This error system is largely aimed at controlling systematic distortions and has been expanded to a current figure of 33 possible sources of error.

The error control measures during the planning of the research design generally cover the measurement of additional variables, the introduction of additional comparison groups and the collecting of additional cases.

For example, sampling variance can be reduced, and the precision of data recording thus improved, by increasing the sample size. The variance of survey instruments can be reduced, and their reliability thus improved, by setting more individual questions or items for the variables to be measured, and combining these to form scales or tests if appropriate. Similarly, the “accuracy of content” or validity of measuring tools can be checked by comparisons with other measuring tools with different content and method. Additional attributes of the members of comparison groups can be recorded to check whether these groups are comparable in terms of their composition. Potential systematic errors in experimental study designs are often checked by using additional comparison groups.

The quality assurance strategy of error control generally consists of taking precautions to ensure that errors can be measured. This means that it can be proved retrospectively that the results have not been (or have been only slightly) distorted by these errors. If errors occur, they are isolated as independent variables when the data are analysed, and are taken into account in an appropriate manner in the substantial results. This is achieved by means of complex multivariate analysis methods.

Error control is undoubtedly an important procedure for safeguarding the results of social science research in general, but this procedure is usually not suitable for evaluating health promotion projects. The use of certain control techniques, such as large samples or experimental study designs, is often not permissible on grounds of costs or because this would create practical difficulties in the implementation of the programme. Nor is it in any way guaranteed that all truly relevant errors will be taken into account. And there is also the fact that the results of multivariate analyses are not very transparent and difficult to present.

These disadvantages can be avoided if attempts are made to avoid errors at the outset, instead of measuring them retrospectively. This is the precept of the quality assurance strategy known as theory-driven impact analysis⁵, the basic features of which will be explained in the next section.

5 This term originates from H. T. Chen (1990).

5.3. THEORY-DRIVEN IMPACT ANALYSIS

Theory-driven impact analysis uses two fundamental tools to avoid errors: firstly, the programme theory, which is used to some extent as a methodological procedure (Lipsey, 1993), and, secondly, monitoring of the implementation process. Thus, quality assurance of the impact analysis utilises the same two elements that are also employed for quality assurance of the programme.

PROGRAMME THEORY

The programme theory serves firstly as a guideline for reliable *measurement of the variables*, to be used in the impact analysis, in a manner valid in terms of content:

- The theory determines the selection of the descriptive variables. It describes the relationship between the variables and thus defines what the outcome variables are, which independent variables describe the intervention and which should be regarded as intermediate variables.
- The theory precisely specifies the meaning of the content. It defines the variables and thus establishes what they mean. The theoretical context of the variables also specifies what they do not mean.
- Precise specification of the meaning permits more precise operationalisation. It becomes easier to establish subject areas and to select reliable indicators. If the theory also makes some statement as to the period after which the effects of the programme can be anticipated, the times at which reliable measurements can be made can also be determined. The sensitivity required of the measuring tools in order to be able to measure the effects of the programme as accurately as possible can also be estimated on the basis of statements regarding the assumed strength of the effects of the intervention.
- All of this then improves the detailed planning of the recording procedures. It is possible to determine more precisely which information must be collected, who will be able to provide this information with the least difficulty and the greatest truthfulness, and which instruments will enable the information to be obtained, recorded and processed most reliably.

Overall, theory-driven measurement aims for an improvement of *construct validity*. An attempt is made to define the variables to be examined as accurately as possible, to delimit them from one another and to operationalise them in such a way that what is actually measured is what is intended to be measured, and that this is achieved as unequivocally as possible and without overlap of the individual variables in terms of content.

The construct validity of the independent (intervention) variables also depends on the quality of the intervention. The precautions taken to improve the effectiveness of the interventions also mean that the (independent) causal variable becomes more clearly effective, i.e. is distorted as little as possible by unclear, contradictory or otherwise ineffective intervention activities, and is largely unaffected by contextual conditions with contrary effects. An effective intervention results in contaminating factors being eliminated from the outset as far as possible.

All these methodological improvements not only enhance the measurement characteristics of the tools, they also improve the *proof of the programme effects*:

- The improved measurement characteristics and the minimisation of contaminating influences result in greater clarity in terms of content and greater precision of the individual outcome measurements. This means that the differences observed between before and after measurements can be interpreted with a large degree of certainty as being the result of the programme intervention.
- At the same time, the influence of extraneous factors is reduced: effective and accurately targeted implementation, on the one hand, and a high level of sensitivity in measuring the outcome variables, on the other, mean that the difference between before and after measurements is so clear that external interference factors can have only a comparatively small effect.

A well-formulated programme theory can also tell us what external interference factors might be involved in the first place. In particular, it is possible to look for specific influencing factors whose content can be described. This is achieved by using the programme theory and additional knowledge regarding the context to derive “rival hypotheses” which could also be used to explain differences between before and after measurements (apart from the effects of the intervention). Thus, even when planning the programme, opportunities can be sought for guarding against the influence of these specific interference factors during the intervention or for controlling this influence by other means. Or (economical) data sources may be sought which can be used to test the rival hypotheses. The awareness (founded in the programme theory and analysis of the context) that it is unlikely that any significant external interference factors will occur is even more important. In this case, there is much to support dispensing with a control group design, particularly if this would involve higher costs or practical difficulties for implementation of the programme.

MONITORING OF THE IMPLEMENTATION PROCESS

Monitoring of the implementation process offers further opportunities for improving the proof of the programme effects. The prerequisite for this is that the outcome variables be incorporated in the monitoring system as early as possible and measured continuously.

Initially, however, the continuous measurements of the monitoring process improve the quality of the data. The collection of data as soon as possible before, during and after the intervention permits a high level of measurement accuracy; at the same time, the repeated measurements allow better estimation of random variations over time and hence an improvement in the reliability of data and results.

In addition, the early stages of the curve for the outcome variables can be used to ascertain whether the preceding measurements are distorting the subsequent ones (testing, instrumentation). These interference factors will become visible in time-series measurements before they smother the intervention effects. Checks can be made as to whether these effects actually occur at all and, if they are recognisable, they can be separated from the effects of the intervention.

Above all, however, the time-series design offers the opportunity of checking specific content-related hypotheses with respect to the efficacy of the programme, and of proving that it is effective. If the curve for the outcome variables changes noticeably in the expected direction during the intervention, this suggests efficacy.

Interrelationship hypotheses regarding the causal mechanisms can also be tested. If it can be confirmed that the intervention activities alter specific variables of the intermediate causal mechanism and if these, in turn, are linked to the outcome variables, this may also be interpreted as proof of the efficacy of the intervention.

5.4. SUMMARY

The impact analysis checks the effects of the programme and investigates the extent to which these can be regarded as being a result of the intervention and its mechanisms of action. The results should be processed in such a way as to allow improvements to the intervention strategy to be derived from them.

Impact analysis planning can be geared to the model of theory-driven impact analysis, which is directed by the following research strategy targets:

- The study plan should be “tailor-made” to fit the programme and should thus preferably be drawn up on the basis of a programme theory. Ready-made research standards (“research paradigms”) increase the risk of drawing incorrect conclusions regarding the efficacy of the programme owing to an inappropriate impact analysis.
- The programme should be compromised as little as possible by the evaluation study. If complex error-control procedures, such as randomised experimental study designs, are taken too greatly into account in planning the programme, this may jeopardise the efficacy and efficiency of the programme.
- Avoiding errors from the outset takes priority over costly and complicated procedures for the retrospective control and measurement of errors.
- Above all, errors should be avoided when selecting or designing the measuring tools. A clearly defined and well-formulated programme theory, an intervention with quality assurance, and continuous measurements using monitoring are all of assistance here.
- The fundamental research design for demonstrating efficacy is a longitudinal study with before and after measurements. This design is enhanced by high-quality measuring tools, continuous time-series measurements in the context of the monitoring system and by empirical testing of the hypotheses derived from the programme theory in relation to the efficacy of the intervention.

This quality assurance strategy of a theory-driven impact analysis is generally based on the fact that there is a direct link between the quality assurance of the programme and the quality of the evaluation research: the better the (theory-driven) planning and (quality-assured) implementation of a social or health programme, the more likely it is possible to develop targeted and precise measuring tools and study plans for the analysis of efficacy.



**SUMMARY AND
RECOMMENDATIONS FOR ACTION**

6.

SUMMARY AND RECOMMENDATIONS FOR ACTION

This working paper examines the way in which programme evaluation can be used for the purpose of quality assurance. It describes a quality-oriented evaluation concept, geared primarily to collecting and feeding back data at all stages of a project, so that the quality of the subsequent phases of the project can be preventively safeguarded or improved.

A number of recommendations for action and planning principles can be derived from this evaluation concept, which can be regarded as guidelines for planning health promotion projects and their evaluation studies. These recommended actions are summarised once more in this final chapter, this compilation being geared to the three quality aspects principally used (and described in more detail in *Section 2.3*):

- Quality from the expert perspective (quality assurance of intervention development),
- Quality from the stakeholder perspective,
- Quality from the accountability perspective.

The intention of this description is to illustrate once again that programme evaluation is not understood here simply as a check of success or some kind of quality control which takes place after completion of the intervention. Rather, as an integral part of project management, it can contribute right from the start to quality assurance of the entire programme process. Quality assurance through programme evaluation is therefore not solely the task of evaluation experts, but the responsibility of all those involved in making decisions for the project or programme. This includes the experts in prevention and health promotion, those with political or financial responsibility, other stakeholders and, in particular, the project management team responsible for coordinating the various (quality) perspectives.

QUALITY ASSURANCE OF INTERVENTION DEVELOPMENT

Quality assurance of intervention development is geared to the quality strategy of a continuous improvement process (CIP), into which the programme evaluation procedure can be incorporated without any difficulty. They are both based on the same principle, i.e. the feedback of data relating to procedures, processes and results into the decision-making processes of the programme.

The most important potential feedback opportunities in the three fundamental phases of a programme will be described below:

- in the concept development phase,
- in the implementation phase, and
- in the impact analysis phase.

All three feedbacks together can be integrated into a closed system of quality assurance for (health promotion) programmes.

Concept development phase

The concept development phase provides the first opportunity for feedback, where the quality assurance of concept development can be supported by elaborating and thoroughly formulating a programme theory specific to the intervention, in the sense of the strategy of “theory-driven evaluation” which has gained increasing ground over the past decades.

As a systematically elaborated realistic programme concept the programme theory describes the precise links between the activities which make up the intervention (the effective programme components), the intermediate causal mechanisms set in motion by these, and the resultant concrete outcomes which can in principle be measured. This description also takes into account the context in which the programme is to be implemented. Thus, a programme theory is not limited to listing schematic targets, target groups and general intervention measures or methods. It is of far greater significance to turn the ideas of the concept, which are often initially only very general, into precise, specific statements on (causal) relationships, and to formulate the ideas of the programme developers, which are often only implicit, so as to arrive at explicit specifications for action in concrete intervention processes.

During the concept development phase, the programme theory is used for early “preventive evaluation” of the programme concept, confronting the statements of the theory with existing knowledge and relevant research results.

This makes it possible to ask and answer questions relating to the feasibility and probable efficacy of the concept, such as:

- Can the programme work at all, or do other or additional effective programme components need to be applied?
- Is the programme practicable in the envisaged context, or does it need to be better adjusted to the staff, clients or specific contextual influences?
- Is the strength of the intervention sufficient or does it need to be increased so that the intended outcomes really are achieved?

Answering these questions permits an initial fundamental improvement in quality:

- the avoidance of fundamental errors in concept development,
- realistic implementation of the concept, geared to the specific conditions of the context, and
- adequate strength of the intervention measures.

A quality-assured programme theory is, in turn, an important foundation for quality assurance of the implementation process, since precise specification of the actions to be implemented means that the implementation quality achieved can be as close as possible to the concept.

Implementation phase

The second opportunity for feedback arises in the implementation phase. Systematic data collection during the implementation process, based on monitoring, makes information available which can be used to detect difficulties in the course of the programme and thus to avoid mistakes at an early stage.

Monitoring of the implementation process means continuous observation of the course of the programme using a systematically compiled set of data which are measured regularly and as a matter of routine. An important prerequisite for selecting the indicators required is, again, a good programme theory. Apart from data relating to administrative handling, information is recorded regarding the influence of contextual factors, the efficacy and efficiency of individual programme components, and realisation of the implementation specifications in compliance with the concept.

In this context, it is of decisive importance that the data be processed and analysed constantly during the implementation process. These data, which can be called up regularly, can be used to detect deviations and problems during the ongoing programme, so that solutions can be sought and found even during the implementation phase. In particular, the effectiveness of the programme can be increased early on. Inputs into the intervention which initially proved to be too weak can be strengthened, the implementation quality can be made more precise, and attempts can be made to protect against and control interfering influences arising from the context.

Impact analysis phase

The third possibility for feedback makes use of the results of the impact analysis. This can contribute to improving the quality of (future) interventions primarily if it is possible to investigate and prove how the effects observed came about, or – if no effects or only minor effects have occurred – how the effects were inhibited or why they turned out to be so weak.

The impact analysis is based on a research plan which must be drawn up individually for the specific programme concerned. On the one hand, the research plan establishes how the programme outcomes are to be measured; on the other hand, it states the method to be used to examine whether the outcomes can be regarded as effects of the intervention and its causal mechanisms.

It is advisable to pay particular attention to the measurement of the variables when setting up the research plan. This includes measurement of both the outcome variables and also of those variables representing the interventions and the intermediate causal mechanisms.

High-quality measurement is of such importance because, the more accurate and reliable the measurement of the individual variables, the more accurate the investigations to demonstrate the efficacy will be. Measurement quality can be improved as follows:

- Consistent quality assurance of the implementation process specifies the interventions and causal mechanisms, and hence the corresponding variables, more precisely.
- A programme theory which has been elaborated in as much detail as possible allows precise definitions of what is to be measured and facilitates the design of measuring tools.
- Measurement of the variables while the programme is running, e.g. in the framework of a monitoring system, allows more accurate data to be obtained.
- The design of measuring tools which take into account the level of information and the willingness to provide information of the individuals being questioned in each case increases validity and reliability of measurements.

The study design for demonstrating efficacy should be geared to the realities of the programme and not to certain methodological specifications (such as the demand for the use of randomised control group experiments). It should primarily reveal the changes brought about by the intervention. A study design of this type, appropriate for the programme concerned, can be set up using the following measures:

- The measurement of changes in the outcomes and other important variables over time, ideally in the context of the monitoring process, with several measurement times before and after the intervention.
- Empirical testing of specific hypotheses, drawn from the programme theory, relating to the efficacy of the intervention.
- Testing only of those “rival hypotheses” for possibly explaining the occurrence of the outcomes which must be regarded as of genuine importance on the basis of the underlying programme theory, under the given contextual conditions, the quality of implementation and the measurement quality.

This quality assurance strategy is generally based on the fact that there is a direct link between the quality assurance of the programme and the quality of the evaluation research. The better the (theory-driven) planning and the (quality-assured) implementation of a social or health programme, the more likely it is possible to develop and to apply targeted and precise measuring tools and study plans for analysing its efficacy.

QUALITY ASSURANCE FROM THE STAKEHOLDER PERSPECTIVE

Close cooperation between programme developers and evaluation experts, starting with the preparation of the project and extending at least up to the concluding assessment of the results of the project, is of fundamental importance for the quality assurance of health promotion projects.

Cooperation of this type, accompanying the project and commencing as early as possible, is necessary simply on the basis of the preventive quality assurance approach. In addition,

synergistic effects can be realised as early as the concept development stage if evaluation researchers are involved at this stage of the project. They can contribute theoretical and empirical knowledge relating to the development of concrete, reality-based hypotheses when elaborating the programme theory, themselves acquiring an improved understanding of how the programme works and what it is intended to achieve.

Cooperation between all the important stakeholders should also be set up as early as possible for the purpose of consensual establishment and unequivocal definition of the intervention outcomes. This is particularly necessary in order to achieve agreement as early as possible on the criteria to be used to decide whether the intended effects have occurred or not.

Evaluation experts, project management and project staff should also cooperate in setting up a monitoring system for continuous observation of the implementation process.

The task of this cooperation is, working on the basis of a joint analysis of the working and interaction processes, to define a system of indicators which reflect the procedures and results of the intervention processes as reliably and expressively as possible and involve the lowest possible costs for collecting and processing the data. Cooperation with the project staff can help to avoid a situation where their work is so burdened by the collection of data for monitoring purposes that the quality of implementation might suffer as a result.

Quality will be improved if the monitoring data are available to the staff. This creates transparency and increases willingness to implement quality improvements. Similarly, in the case of self-evaluation of the implementation by the project staff, the advantages of greater transparency and greater awareness of quality must be weighed against the risk of an excessive workload and the resulting losses of quality.

Joint decisions must be made by the project developers and evaluation researchers regarding the research plan for the impact analysis. However, the programme must take priority. Thus, the programme concept must be compromised as little as possible by the evaluation study. Costly and complex procedures for controlling errors, such as randomised, experimental study designs, may put the efficacy and efficiency of the programme at risk.

QUALITY ASSURANCE FROM THE ACCOUNTABILITY PERSPECTIVE

The quality-oriented evaluation concept proposed here facilitates the planning and application of the financial resources for health promotion projects.

The programme theory permits the degree of transparency required for precise planning of the efficient investment of financial resources. With an accurate description of the in-

dividual effective intervention elements and of the staff required for this, it offers a particularly good basis for deciding on how scarce resources must be used or how the concept can be altered for reasons of economy. A feasibility study undertaken on the basis of the programme theory as early as the concept development stage makes it possible to ascertain, early on and not only once the project has been completed, whether it is worth while continuing with the project at all given the financial resources available.

The quality assurance measures proposed in the evaluation concept allow an increase in the productivity and efficiency of both the intervention and the evaluation.

For instance, the future costs of an intervention can be reduced if the evaluation results can be used to demonstrate that a certain impact can also be achieved with a lesser strength of effect. In turn, well-planned implementation that adheres to the concept permits a more thrifty evaluation design that is also able to dispense with extensive, costly control procedures (e.g. randomised control group designs) because precisely implemented interventions permits more unequivocal measurement of the impacts and more reliable demonstration of the mechanism of action.

All in all, quality assurance through evaluation research leads to greater efficiency in health promotion, primarily because it prevents incorrect conclusions being drawn as to the efficacy of programmes: on the one hand, it reduces the probability of health promotion projects that are in fact totally ineffective being regarded as effective, and thus receiving further financing, while, on the other hand, there is a better chance of truly effective interventions being recognised as such and thus put to use in order to improve the health situation of the population.

REFERENCES



7.

REFERENCES

- Bartholomew, L. K. / Parcel, G. S. / Kok, G. (1998): Intervention Mapping: A process for developing theory and evidence-based health education programs. *Health Education and Behavior*, 25 (5), 545–563.
- Bengel, J. / Strittmatter, R. / Willmann, H. (1999): What keeps people healthy? The current state of discussion and the relevance of Antonovsky's salutogenic model of health. (Research and Practice of Health Promotion, Vol. 4). Cologne: Bundeszentrale für gesundheitliche Aufklärung (Ed.).
- Berenson, K. / Granath, M. / Urwitz, V. / Björklund, E. G. (1996): Succeeding with Health Promotion Projects – Quality Assurance. Stockholm: Federation of Swedish County Councils.
- Bickman, L. (1985): Improving established statewide programs: a component theory of evaluation. *Evaluation Review*, 9, 189–208.
- Blumenthal, D. (1996): Quality of care: what is it? *New England Journal of Medicine*, 335, 891–93.
- Bowling, A. (1995): Measuring disease. Buckingham: Open University Press.
- Bowling, A. (1997): Measuring health. (2nd edition) Buckingham: Open University Press.
- Brook, R. H. / McGlynn, E. A. / Cleary P. D. (1996): Measuring quality of care. *New England Journal of Medicine*, 335, 966–970.
- Bühren-Armstrong, B. / Bengel, J. (1997): Qualitätsstandards in Prävention und Gesundheitsförderung – Nationale und internationale Erfahrungen. *Prävention*, 20 (2), 42–46.
- Bundesanstalt für Arbeitsschutz und Arbeitsmedizin (Eds.) (1997): Quality management in workplace health promotion. Proceedings (Tb 81) of the workshop held in Erkner near Berlin in November 1996. Dortmund/Berlin.
- Bundeszentrale für gesundheitliche Aufklärung (Ed.) (1994): Qualitätsraster Adipositas zur Beurteilung von Maßnahmen gegen Übergewicht. Cologne.
- Campbell, D. T. / Stanley, J. C. (1966): Experimental and quasi-experimental designs for research. Chicago: Rand McNally. (Reprint from Gage, N. L. (Ed.) [1963]: Handbook of research on teaching. Chicago: Rand McNally.)
- Caplan, R. D. / Vinokur, A. D. / Price, R. H. (1997): From job loss to reemployment: field experiments in prevention-focused coping. In: Albee, G. W. / Gullotta, T. P. (Eds.): Primary prevention works. Newbury Park, CA: Sage.
- Chelimsky, E. (1997): The coming transformations in evaluation. In: Chelimsky, E. / Shadish, W. R. (Eds.): Evaluation for the 21st century. Thousand Oaks: Sage.
- Chen, H. T. (1990): Theory-driven evaluations. Newbury Park, CA: Sage.
- Chen, H. T. / Rossi, P. H. (1983): Evaluating with sense: the theory-driven approach. *Evaluation Review*, 7, 283–302.
- Chen, H. T. / Rossi, P. H. (1989): Issues in the theory-driven perspective. *Evaluation and Program Planning*, 12, 4, 299–306.
- Cook, T. D. / Campbell, D. T. (1979): Quasi-experimentation: design and analysis issues for field settings. Chicago: Rand McNally.
- Deming, W. E. (1986): Out of the crisis. Cambridge: Massachusetts Institute of Technology.
- Donabedian, A. (1988): The quality of care: how can it be assessed? *Journal of the American Medical Association*, 260, 1743–48.
- Evans, D. / Head, M. J. / Speller, V. (1994): Assuring Quality in Health Promotion. London: Health Education Authority.
- Garvin, D. A. (1984): What does product quality really mean? *Sloane Management Review*, 25–43.
- Green, L. W. / Kreuter, M. W. (1991): Health promotion planning: an educational and environmental approach. Mountain View, CA: Mayfield.

- Haglund, B. J. A. / Pettersson, B. / Finer, D. / Tillgren, P. (1996): Creating supportive environments for health. Stories from the Third International Conference on Health Promotion, Sundsvall, Sweden (Public Health Action Series, No. 3). Geneva: WHO.
- Heaney, C. A. (1998): Intervention Mapping and the new health promotion. *Health Education and Behavior*, 25 (5), 564–568.
- Heiner, M. (Ed.) (1996): *Qualitätsentwicklung durch Evaluation*. Freiburg: Lambertus-Verlag.
- International Union for Health Promotion and Education/Europe (1996): Presentation of existing standards for evaluation and quality assessment of health promotion interventions in the EU Member States (Commissioned by the Commission of the European Communities), Final Report. Woerden: Netherlands Institute for Health Promotion and Disease Prevention.
- Kessler, R. C. / Turner, B. J. / House, J. S. (1987): Intervening processes in the relationship between unemployment and health. *Psychological Medicine*, 17, 949–961.
- Kessler, R. C. / Turner, B. J. / House, J. S. (1988): Effects of unemployment on health in a community survey: main, modifying, and mediating effects. *Journal of Social Issues*, 44, 69–85.
- Künzel-Böhmer, J. / Bühringer, G. / Janik-Konecny, T. (1993): *Expertise zur Primärprävention des Substanzmißbrauchs*. (Schriftenreihe des Bundesministeriums für Gesundheit, Vol. 20). Baden-Baden: Nomos.
- Lethbridge J. / Parish, R. / Hagard, S. (1996): Quality and health promotion – developing concepts and principles. Discussion document for the meeting in Ormoz, Slovenia, October 1996.
- Lewin, K. (1951): *Field theory in the social sciences*. New York: Harper and Brothers.
- Lewin, K. (1963): *Feldtheorie in den Sozialwissenschaften*. Ausgewählte theoretische Schriften. Bern: Huber.
- Lewin, K. / Lippitt, R. / White, R. K. (1939): Patterns of aggressive behavior in experimentally created “social climates”. *Journal of social psychology*, 10, 271–299.
- Lipsey, M. W. (1993): Theory as method: small theories of treatments. *New Directions for Program Evaluation*, 57, 5–38.
- Lipsey, M. W. / Crosse, S. / Dunkle, J. / Pollard J. / Stobart, G. (1985): Evaluation: The state of the art and the sorry state of the science. In: Cordray, D. S. (Ed.): *Utilizing prior research in evaluation planning*. New directions for program evaluation. San Francisco: Jossey-Bass.
- Lück, H. E. (1996): *Die Feldtheorie und Kurt Lewin*. Weinheim: Psychologie Verlags Union.
- MacDonald, G. (Ed.) (1997): Development and Measurement of Quality in Health Promotion. *International Journal of Health Promotion and Education*, 4 (2).
- Marrow, A. J. (1969): *The practical theorist*. New York: Basic Books.
- McGuire, W. J. (1989): The theoretical foundations of campaigns. In: Rice, R. E. / Atkin, C. K. (Eds.): *Public communication campaigns*. (2nd edition). Newbury Park, CA: Sage.
- McKinnon, D. P. / Dwyer J. H. (1993): Estimating mediated effects in prevention studies. *Evaluation Review*, 17, 144–158.
- Mohr, L. B. (1995): *Impact analysis for program evaluation*. (2nd edition). Thousand Oaks: Sage.
- Oeretveit, J. (1996): Quality in health promotion. *Health Promotion International*, 11, 1, 55–62.
- Parish, R. (1995): Quality issues in health promotion. Background paper for the meeting in Dublin, March 1995.
- Patton, M. Q. (1997): *Utilization-focused evaluation*. (3rd edition). Thousand Oaks: Sage.
- Pawson, R. / Tilley, N. (1997): *Realistic evaluation*. London: Sage.
- Pelikan, J. M. / Dietscher, C. / Novak-Zezula, S. (1998): Evaluation als Strategie der Qualitätssicherung in der Gesundheitsförderung. In: Dür, W. / Pelikan, J. M. (Eds.): *Qualität in der Gesundheitsförderung*. Vienna: Facultas-Universitätsverlag.
- Plantz, M. C. / Greenway, M. T. / Hendricks, M. (1997): Outcome measurement: showing results in the nonprofit sector. In: Newcomer, K. E. (Ed.): *Using performance measurement to improve public and nonprofit programs*. *New Directions for Evaluation*, 75, 15–30.

- Price, R. H. / Vinokur A. M. (1995): Supporting career transitions in a time of organizational downsizing. In: London, M. (Ed.): *Employees, Careers and Job Creation*. San Francisco: Jossey-Bass.
- Riemann, K. / Gerber U. (1997): *Standardisierung von Fragestellungen zum Rauchen. (Forschung und Praxis der Gesundheitsförderung, Vol. 1)*. Cologne: Bundeszentrale für gesundheitliche Aufklärung (Ed.).
- Rossi, P. H. / Freeman, H. E. (1993): *Evaluation: a systematic approach. (3rd edition)*. Beverly Hills, CA: Sage.
- Saan, H. (1997): Quality revisited. In: Macdonald, G. (Ed.): *Development and Measurement of Quality in Health Promotion. International Journal of Health Promotion and Education*, 4 (2).
- Sechrest, L. (Ed.) (1993): *Program evaluation: a pluralistic enterprise. New Directions for Program Evaluation*, 60. San Francisco: Jossey-Bass.
- Sechrest, L. / Figueredo, A. J. (1993): Program evaluation. *Annual review of psychology*, 44, 645–674.
- Sechrest, L. / West, S. G. / Phillips, M. A. / Redner, R. / Yeaton, W. (1979): Some neglected problems in evaluation research: strength and integrity of treatment. *Evaluation Studies Review Annual*, 4, 15–35.
- Shadish, W. R. Jr. / Epstein, R. (1987): Patterns of program evaluation practice among members of the evaluation research society and evaluation network. *Evaluation Review*, 11 (5), 555–590.
- Springett, J. (1998): What sort of quality for what sort of health promotion. Discussion document for the meeting in Brighton, April 1998.
- Suchman, E. A. (1967): *Evaluative research: principles and practice in public service and social action programs*. New York: Sage.
- Tones, K. (1992): Measuring success in health promotion. *Hygie (International Journal of Health Education*, 11, 10–14.
- Töppich, J. (1993): Der “Qualitätsraster Adipositas” der BZgA – als Beitrag zur Qualitätssicherung in der Gesundheitsförderung. *Prävention*, 16 (1), 18–20.
- Töppich, J. (1996): Qualitätssicherung bezüglich Zielgruppenansprache und Akzeptanz. In: Bundesvereinigung für Gesundheit e.V. und Spitzenverbände der gesetzlichen Krankenkassen (Hrsg.): *Qualitätsstrategien in Prävention und Gesundheitsförderung – Leitlinien, Praxisbeispiele, Potentiale*. Bonn: proceedings.
- Van den Brouke, S. / Lenders, F. (1997): Monitoring the planning quality of health promotion projects in Flanders. *Promotion and Education*.
- Van de Vall, M. (1993): *Angewandte Sozialforschung*. Weinheim/Munich: Juventa Verlag.
- Van Driel, W. G. / Keijsers, J. F. E. M. (1997): An instrument for reviewing the effectiveness of health education and health promotion. *Patient Education and Counseling*, 30, 7–17.
- Veen, C. A. (1995): Evaluation of the IUHPE project on the effectiveness of health promotion and health education. Utrecht: International Union for Health Promotion and Education, Regional Office for Europe.
- Veenhoven, R. (1992): *Bibliography of happiness; 2472 studies on subjective appreciation of life*. Rotterdam: University of Rotterdam experimentation: design and analysis issues for field settings. Chicago: Rand McNally.
- Vinokur, A. D. / Price, R. H. / Caplan, R. D. / van Ryn, M. / Curran, J. (1995): The jobs preventive intervention for unemployed individuals: short and long-term effects on reemployment and mental health. In: Murphy, L. R. / Hurrell, J. J. / Sauter, S. L. / Keita, G. P. (Eds.): *Job stress interventions (pp. 125–138)*. Washington, D.C.: American Psychological Association.
- Weiss, C. H. (1972): Evaluating educational and social action programs. In: Weiss, C. H. (Ed.): *Evaluating Action Programs*. Boston: Allyn and Bacon.
- Wortman, P. M. (1983): Evaluation research: a methodological perspective. *Annual Review of Psychology*, 34, 193–214.
- Young, M. (1966): *Innovation and research in education*. London: Routledge and Kegan Paul.
- Ziglio, E. (1994): *Methods for the evaluation of health promotion programmes and policies*. Working paper, Copenhagen: WHO/EURO.

In the specialist booklet series "Research and Practice of Health Promotion" has been published previously:

Volume 1 – *Gender-related Drug Prevention for Youths*

Practical Approaches and Theory Development.

Final report of a research project by Peter Franzkowiak, Cornelia Helfferich and Eva Weise commissioned by the FCHE.

Order No. 60 802 070

Volume 2 – *Ecstasy: Use and Prevention*

Empirical Research Results and Guidelines.

Documentation of a FCHE status seminar held in Bad Honnef from 15 to 17 September 1997

Order No. 60 801 070

Volume 3 – *Quality Assurance in AIDS Prevention*

Report of the Expert Conference in cooperation with the World Health Organization/Regional Office for Europe (WHO/EURO) from 13 to 15 November 1995.

Order No. 60 803 070

Volume 4 – *What Keeps People Healthy?*

The Current State of Discussion and the Relevance of Antonovsky's Salutogenic Model of Health.

An experts' report by Jürgen Bengel, Regine Strittmatter and Hildegard Willmann commissioned by the FCHE.

Order No. 60 804 070

Volume 5 – *Child Health*

Epidemiological Foundations.

Documentation of an expert seminar held by the Federal Centre for Health Education.

Order No. 60 805 070



BZgA

**Bundeszentrale
für
gesundheitliche
Aufklärung**

ISBN 3-933191-27-0