

**FORSCHUNG UND PRAXIS DER GESUNDHEITSFÖRDERUNG**

**EVALUATION –  
EIN INSTRUMENT ZUR  
QUALITÄTSSICHERUNG  
IN DER GESUNDHEITS-  
FÖRDERUNG**

**BAND 8**

Herausgeberin: Bundeszentrale  
für gesundheitliche Aufklärung

Die Bundeszentrale für gesundheitliche Aufklärung ist eine Behörde im Geschäftsbereich des Gesundheitsministeriums mit Sitz in Köln. Sie hat zur Aufgabe, geeignete Maßnahmen zur Erhaltung und Förderung der Gesundheit zu erarbeiten und umzusetzen.

Sie entwickelt Konzepte und Strategien für Kampagnen, erstellt Übersichten über Medien und Methoden, kooperiert mit unterschiedlichen Akteuren der gesundheitlichen Aufklärung und führt Aufklärungsmaßnahmen für die gesamte Bevölkerung sowie für spezielle Zielgruppen in ausgewählten Themenfeldern durch.

Bei der Planung, Durchführung und Bewertung von Effektivität und Effizienz ihrer Arbeit stützt sich die BZgA auf Forschungsergebnisse.

Dazu werden in ihrem Auftrag repräsentative Wiederholungsbefragungen, Forschungsprojekte zu ausgewählten Einzelthemen und Evaluationsstudien durchgeführt. Zur Förderung des Informations- und Erfahrungsaustauschs zwischen Wissenschaft und Praxis veranstaltet sie nationale und internationale Konferenzen.

In ihrer Fachheftreihe zur Forschung und Praxis der Gesundheitsförderung veröffentlicht die Bundeszentrale für gesundheitliche Aufklärung diese Studien, Expertisen und Ergebnisse von Fachtagungen. Sie versteht sich als Forum zur wissenschaftlichen Diskussion. Ziel dieser Reihe, wie der bereits vorliegenden Fachheftreihe zur Sexualaufklärung und Familienplanung, ist es, den Dialog zwischen Wissenschaft und Praxis weiter auszubauen.

**FORSCHUNG UND PRAXIS DER GESUNDHEITSFÖRDERUNG  
BAND 8**

# **EVALUATION – EIN INSTRUMENT ZUR QUALITÄTSSICHERUNG IN DER GESUNDHEITSFÖRDERUNG**



Eine Expertise von Gerhard Christiansen, Bundeszentrale für gesundheitliche Aufklärung, im Auftrag der Europäischen Kommission, GD Gesundheit und Verbraucherschutz.

Bundeszentrale für gesundheitliche Aufklärung (BZgA)  
Köln 1999

Die Deutsche Bibliothek – CIP-Einheitsaufnahme

**Christiansen, Gerhard:**

Evaluation – ein Instrument zur Qualitätssicherung in der Gesundheitsförderung; eine Expertise / von Gerhard Christiansen.  
Bundeszentrale für gesundheitliche Aufklärung. Im Auftr. der Europäischen Kommission, GD Gesundheit und Verbraucherschutz  
[Red.: Katharina Salice-Stephan ; Ingrid Schmitt] – 2. Aufl. – Köln : BZgA, 2000  
(Forschung und Praxis der Gesundheitsförderung ; Bd. 8)

**ISBN 3-933191-26-2**

Die Beiträge geben die Meinung der Autorinnen und Autoren wieder, die von der Herausgeberin nicht in jedem Fall geteilt werden muß. Die Fachheftreihe ist als Diskussionsforum gedacht.

Herausgeberin:  
Bundeszentrale für gesundheitliche Aufklärung (BZgA)  
Ostmerheimer Str. 220, 51109 Köln  
Tel.: 02 21/89 92-0  
Fax: 02 21/89 92-3 00  
E-Mail: christiansen@bzga.de

Alle Rechte vorbehalten.

Redaktion: Katharina Salice-Stephan, Ingrid Schmitt

Satz: Salice-Stephan, Köln  
Druck: Schiffmann, Bergisch Gladbach

Auflage: 2.3.06.00

Gedruckt auf Recyclingpapier.

Band 8 der Fachheftreihe ist kostenlos erhältlich  
unter der Bestelladresse BZgA, 51101 Köln,  
und über Internet unter der Adresse <http://www.bzga.de>

Bestellnummer: 60 608 000

## **VORWORT**

In den letzten Jahren hat das Thema Qualitätssicherung in vielen Bereichen an Bedeutung gewonnen. Auch in der Gesundheitsförderung wird intensiv darüber diskutiert, welche Instrumente und Strategien der Qualitätssicherung Anwendung finden können.

Eine in der Gesundheitsförderung bereits angewandte und bewährte Strategie der Qualitätssicherung ist die Programmevaluation, die gesundheitsfördernde Maßnahmen nach zuvor festgelegten Kriterien auf ihre Wirksamkeit hin überprüft. Die Bundeszentrale für gesundheitliche Aufklärung wendet diese Strategie bereits seit langem in ihren bundesweiten Kampagnen zur Aids-Prävention, Suchtvorbeugung und Sexualaufklärung an. So führt sie beispielsweise langfristig angelegte Monitoring-Studien durch, mit deren Hilfe zum einen die Wirksamkeit der Kampagnen untersucht werden kann, zum anderen auch frühzeitig erkennbar ist, ob einzelne Kampagnenelemente verbessert werden müssen. Ein weiteres Instrument der Programmevaluation nutzt die Bundeszentrale für gesundheitliche Aufklärung, wenn sie z.B. ihre TV- und Kino-Spots mit einem festgelegten Vorgehen pretestet. Auf diese Weise kann die Kommunikationswirkung der Spots optimiert werden, bevor sie im Kino oder Fernsehen eingesetzt werden.

Die Bundeszentrale für gesundheitliche Aufklärung hat – im Auftrag und mit finanzieller Unterstützung der Europäischen Kommission – untersucht, ob und auf welche Weise die Programmevaluation systematisch für die Qualitätssicherung in der Gesundheitsförderung genutzt werden kann. Die Ergebnisse dieser Untersuchung werden in dem vorliegenden Band 8 der Fachheftreihe dargestellt. Es werden die wichtigsten Elemente eines qualitätsorientierten Evaluationskonzepts beschrieben, durch das sich Entwicklung, Implementation und Erfolg von Maßnahmen zur gesundheitlichen Aufklärung kontrollieren und verbessern lassen.

Mit der Veröffentlichung dieses Arbeitspapiers möchte die Bundeszentrale für gesundheitliche Aufklärung auf nationaler wie auch auf europäischer Ebene einen Beitrag zur Diskussion über die Methoden und Strategien der Qualitätssicherung in der Gesundheitsförderung leisten.

Köln, Dezember 1999

Dr. Elisabeth Pott  
Direktorin der Bundeszentrale  
für gesundheitliche Aufklärung



## VORBEMERKUNG

Die von der Europäischen Kommission für die Mitgliedstaaten der Europäischen Union entwickelten Maßnahmen zur Gesundheitsförderung sind auf eine größtmögliche Wirksamkeit und nachprüfbare Evidenz gerichtet. Die Entscheidung über die Planung, Förderung und Weiterführung von Programmen und Projekten basiert in erster Linie darauf, daß der beste Nachweis ihrer Wirksamkeit und Nachhaltigkeit herangezogen wird. Der Hinweis auf hochgesteckte Zielvorstellungen oder auf die fachliche Anerkennung bestimmter Methoden allein reicht für die Durchführung von Projekten der Gesundheitsförderung angesichts schwieriger Haushaltsbedingungen nicht aus.

In den letzten Jahren hat die Europäische Kommission eine Reihe von wissenschaftlichen Arbeiten zu unterschiedlichen Themen mit verschiedenen Perspektiven finanziell gefördert und organisatorisch unterstützt. Ziel dieser Darstellungen war es zumeist, die Wirksamkeit von Maßnahmen zur Gesundheitsförderung nachweislich widerzuspiegeln. Zu diesen Arbeiten gehört auch die hier veröffentlichte Studie der Bundeszentrale für gesundheitliche Aufklärung zur Evaluation als Mittel der Qualitätssicherung.

Diese Studie untersucht, wie eine Programmevaluation für die Qualitätssicherung genutzt werden kann. Sie beschreibt ein qualitätsorientiertes Evaluationskonzept, das darauf gerichtet ist, in allen Phasen eines Projekts – von Beginn an und nicht erst nach seinem Abschluß – Informationen über die Leistungsfähigkeit der jeweils abgelaufenen Projektaktivitäten zu sammeln und rückzukoppeln, um dadurch die Wirksamkeit der nachfolgenden Projektphasen „präventiv“ sicherzustellen oder zu verbessern. Auf diese Weise evaluierte Projekte besitzen ein hohes Maß an Qualität, vor allem aber ein hohes Maß an *nachweisbarer* Qualität.

Darüber hinaus enthält diese Studie einige Handlungsvorschläge und Planungsprinzipien, inwieweit die Evaluationsforschung und Projektentwicklung organisiert und integriert werden sollten, um einen hohen Standard an Qualität zu sichern und nachzuweisen. Diese Handlungsanweisungen lassen sich auch als Empfehlungen für Leitlinien zur Planung von Programmen und Projekten der Gesundheitsförderung verstehen, an denen sich die Durchführung derartiger Maßnahmen orientieren sollte.

Luxemburg, Dezember 1999

Horst Kloppenburg,  
Europäische Kommission,  
Generaldirektion Gesundheit  
und Verbraucherschutz, Referat F-3



## STECKBRIEF DES PROJEKTS

---

Projekttitel: Evaluation als Mittel der Qualitätssicherung  
in der Gesundheitsförderung

---

Ziele:

- Darstellung von Nutzungsmöglichkeiten der Programmevaluation für die Qualitätssicherung von Gesundheitsförderungsprojekten
- Vorschlag eines Evaluationsmodells für die Qualitätssicherung in der Gesundheitsförderung
- Handlungsempfehlungen für qualitätsorientierte Evaluationsprojekte

---

Durchführungszeitraum: November 1996 bis August 1998

---

Projektträgerin: Bundeszentrale für gesundheitliche Aufklärung  
Ostmerheimer Str. 220  
51109 Köln  
Tel.: 02 21/ 89 92-0  
Fax: 02 21/ 89 92-3 00

---

Projektkoordinierung: Helene Reemann, Gerhard Christiansen

---

Autor des Berichts: Gerhard Christiansen

---

Finanzielle Unterstützung: Europäische Kommission  
Generaldirektion Gesundheit und Verbraucherschutz  
Referat F-3  
Gebäude Jean Monnet  
Rue Alcide de Gasperi  
2920 Luxemburg  
Luxemburg

---



# INHALTSVERZEICHNIS

<b>1.</b>	<b>DIE AUFGABENSTELLUNG</b>	13
1.1.	Ausgangssituation	14
1.2.	Inhalt des Arbeitspapiers	18
<b>2.</b>	<b>SCHLÜSSELBEGRIFFE UND ELEMENTE EINES QUALITÄTSORIENTIERTEN EVALUATIONSKONZEPTS</b>	19
2.1.	Evaluation	20
2.2.	Programm	22
2.3.	Qualität	24
2.4.	Qualitätsorientiertes Evaluationskonzept	27
2.5.	Zusammenfassung	29
<b>3.</b>	<b>PROGRAMMTHEORIEN</b>	31
3.1.	Zum Theorieverständnis	33
3.2.	Theoriebildung: Quellen	34
3.3.	Theoriebildung: Arbeitsschritte	35

- 3.4. **Programmtheorie und Qualitätssicherung** 43
- 3.5. **Zusammenfassung** 45

## **4. MONITORING VON IMPLEMENTATIONSPROZESSEN** 47

- 4.1. **Monitoring** 48
- 4.2. **Analyse des Programmkontexts** 50
- 4.3. **Evaluation von Programmkomponenten** 52
- 4.4. **Kontrolle der Übertragungsqualität** 52
- 4.5. **Monitoring und Programmevaluation** 53
- 4.6. **Monitoring und Qualitätssicherung** 54
- 4.7. **Zusammenfassung** 55

## **5. WIRKUNGSANALYSE** 57

- 5.1. **Planung von Wirkungsanalysen** 59
- 5.2. **Qualitätssicherung der Evaluationsforschung** 63
- 5.3. **Theoriegeleitete Wirkungsanalyse** 67
- 5.4. **Zusammenfassung** 70

<b>6.</b>	<b>RESÜMEE UND HANDLUNGSEMPFEHLUNGEN</b>	71
<b>7.</b>	<b>BIBLIOGRAPHIE DER BENUTZTEN UND ZITIERTEN LITERATUR</b>	79



**DIE AUFGABENSTELLUNG**



# 1.

## DIE AUFGABENSTELLUNG

Im Aktionsprogramm der Europäischen Union zur Gesundheitsförderung, -aufklärung, -erziehung und -ausbildung spielt die Evaluation eine wesentliche Rolle. So heißt es unter Punkt 24 der Präambel des Aktionsprogramms: *„Um den Nutzen und die Wirksamkeit dieses Programms zu steigern, sollten die Aktionen insbesondere hinsichtlich ihrer Wirksamkeit und Verwirklichung der Ziele sowohl auf nationaler wie auf gemeinschaftlicher Ebene einer kontinuierlichen Bewertung unterzogen und gegebenenfalls entsprechend angepaßt werden.“*<sup>1</sup>

Im Arbeitsprogramm 1997 des Aktionsprogramms heißt es unter Punkt 4 (7): *„Evaluation und Qualitätssicherung sind integrierende Bestandteile des Programms und spielen als solche während des ganzen Programmverlaufs eine wichtige Rolle.“*<sup>2</sup> Bei der Vergabe von Beihilfen für Gesundheitsförderungsprojekte ist eine Evaluierung des Projekts jeweils Gegenstand des Vertrages.

Das vorliegende, vom Dienst der Europäischen Kommission in Auftrag gegebene Arbeitspapier ist ein Beitrag zur Umsetzung dieser Zielsetzung. Es soll untersuchen, wie die Programmevaluation für die Qualitätssicherung von Gesundheitsförderungsprojekten der Europäischen Kommission genutzt werden kann, und es soll aufzeigen, welche Anforderungen an qualitätsorientierte Evaluationsprojekte gestellt werden müssen.

# 1.1.

## AUSGANGSSITUATION

In den letzten Jahren hat es eine ganze Reihe von Bemühungen gegeben, die Idee der Qualitätssicherung in der europäischen Gesundheitsförderung zu verbreiten und Instrumente zu entwickeln, durch die sich die Qualität konkreter Maßnahmen der Gesundheitsförderung verbessern läßt.

Zur Meinungsbildung und Konzeptentwicklung auf europäischer Ebene haben verschiedene Workshops und Konferenzen beigetragen: So veranstaltet die International Union of Health Promotion and Education/Europe regelmäßig Konferenzen zur Effektivität in der

---

1 Europäische Kommission: Beschluß Nr. 645/96/EG des Europäischen Parlaments und Rates vom 29.3.1996 über ein Aktionsprogramm der Gemeinschaft zur Gesundheitsförderung, -aufklärung, -erziehung und -ausbildung innerhalb des Aktionsrahmens im Bereich der öffentlichen Gesundheit (1996–2000), veröffentlicht im Amtsblatt der Europäischen Gemeinschaften, L95/1 DE, 16.4.1996.

2 Europäische Kommission: Arbeitsprogramm 1997 für das Aktionsprogramm der Gemeinschaft zur Gesundheitsförderung, -aufklärung, -erziehung und -ausbildung (97/C18/04), veröffentlicht im Amtsblatt der Europäischen Gemeinschaften, C18/11 DE, 17.1.1997.

Gesundheitsförderung und Gesundheitserziehung. Konferenzen direkt zum Thema Qualität fanden im September 1996 in Turin<sup>3</sup> und zuletzt im Mai 1999 in Helsinki und Tallinn unter dem Titel „Best Practices – Effectiveness and Quality of Health Promotion“ statt. Im Rahmen des European Committee for Health Promotion Development (ECHPD) hat die WHO/EURO eine Arbeitsgruppe zu Qualität und Gesundheitsförderung eingerichtet<sup>4</sup>, und für den spezifischen Bereich der betrieblichen Gesundheitsförderung wurde beispielsweise im November 1996 in Erkner bei Berlin eine europäische Konferenz mit dem Titel „Quality Management in Workplace Health Promotion“ veranstaltet.<sup>5</sup>

Auskunft über die Aktivitäten in den einzelnen Mitgliedstaaten gibt der im Auftrag der Europäischen Kommission erstellte Bericht über Evaluationsstandards und Qualitätssicherung bei Interventionen der Gesundheitsförderung in den EU-Mitgliedstaaten.<sup>6</sup>

## ANSÄTZE VON QUALITÄTSSICHERUNGSAKTIVITÄTEN IN EUROPA

Bei den Qualitätssicherungsaktivitäten in Europa lassen sich drei Ansätze erkennen:

### 1. Handlungskriterien zur Standardisierung der Qualität von Gesundheitsförderungsprojekten

Der erste Ansatz versucht, mit Hilfe von Handlungskriterien die Qualität von Gesundheitsförderungsprojekten auf bestimmte Standards zu bringen. Dieses Vorgehen ist besonders für solche Arbeitsfelder geeignet, für die bereits ein erhebliches Erfahrungswissen existiert. Beispiele hierfür sind die im Tagungsbericht der oben erwähnten Konferenz „Quality Management in Workplace Health Promotion“ vorgeschlagenen Kriterien für betriebliche Gesundheitsförderungsmaßnahmen<sup>7</sup> oder das von der Bundeszentrale für gesundheitliche Aufklärung (1994) herausgegebene „Qualitätsraster Adipositas zur Beurteilung von Maßnahmen gegen Übergewicht“.<sup>8</sup>

### 2. Fortbildungsmaßnahmen zur Verbesserung der Planungsqualität von Gesundheitsförderungsmaßnahmen

Der zweite Ansatz versucht, über Fortbildungsmaßnahmen, Handbücher oder Broschüren die Planungsqualität von Gesundheitsförderungsmaßnahmen zu verbessern. Er zielt ebenfalls auf eine Standardisierung ab, indem in Form von Checklisten oder Merksätzen

3 Einen Konferenzbericht enthält MacDonald (1997).

4 Die bisherige Entwicklung dieser Diskussion verdeutlichen Parish (1995), Lethbridge/Parish/Hagard (1996) sowie Springett (1998).

5 Der Tagungsbericht wurde von der Bundesanstalt für Arbeitsschutz und Arbeitsmedizin herausgegeben (Quality management in workplace health promotion. Tagungsbericht [Tb 81] zum Workshop November 1996 in Erkner bei Berlin. [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. Siehe auch Bührlen-Armstrong/Bengel (1997).

7 Bundesanstalt für Arbeitsschutz und Arbeitsmedizin (1997).

8 Siehe auch Töppich (1993).

diejenigen Einzelaufgaben und -tätigkeiten aufgezeigt werden, die für erfolgreiche Gesundheitsförderungsprojekte für notwendig gehalten werden. Beispiele hierfür sind:

- das von der englischen Health Education Authority herausgegebene Handbuch „*Assuring Quality in Health Promotion*“ (Evans/Head/Speller, 1994),
- die schwedische Broschüre „*Succeeding with Health Promotion Projects – Quality Assurance*“ (Berenson et al., 1996),
- der vom schwedischen Karolinska Institute for Health Promotion Training, Department for Public Health Science, entwickelte „*20-Punkte-Fragebogen*“<sup>9</sup>,
- das vom Niederländischen Institut für Gesundheitsförderung und Krankheitsprävention entwickelte Promotion Effectiveness Fostering Instrument (PREFFI).<sup>10</sup>

Daß mit diesem Ansatz von Qualitätssicherung auch tatsächlich die Planungsqualität verbessert werden kann, zeigt eine Studie des Flemish Institute for Health Promotion (Van den Brouke/Lenders, 1997): Vor und nach einer Fortbildungsmaßnahme auf der Grundlage praxisnaher Planungshandbücher wurde die Planungsqualität von Gesundheitsförderungsprojekten mit Hilfe von 26 nach 9 Dimensionen gruppierten Fragebogen-Items gemessen. Es ergab sich eine signifikante Verbesserung der Planungsqualität, vor allem aber im Hinblick auf die Spezifizierung von Zielgruppen und Zielen.

### **3. Programmevaluation zur Qualitätssicherung in der Gesundheitsförderung**

Als dritter Ansatz kommt zunehmend die Programmevaluation ins Gespräch. Dies findet nicht nur im Aktionsprogramm der Europäischen Gemeinschaft seinen Ausdruck, sondern auch andere Organisationen, wie das European Committee for Health Promotion Development (ECHPD) der WHO/EURO haben sich in den letzten Jahren damit beschäftigt.<sup>11</sup> Konferenzen zur Qualitätssicherung in der Gesundheitsförderung setzen das Thema Evaluation ebenfalls häufig auf ihre Tagesordnung.<sup>12</sup>

Die Beschäftigung mit dem Thema Evaluation und Qualitätssicherung ist wohl auch dadurch begründet, daß in der Gesundheitsförderung bisher keine eigenständigen Verfahren der Qualitätsmessung entwickelt wurden, wie es z.B. zur Zeit im Qualitätsmanagement der Gesundheitsversorgung der Fall ist.<sup>13</sup> Es liegt nun nahe, diese Lücke mit Hilfe der Evaluationsforschung zu schließen. Diese hat in den letzten 30 Jahren eine Reihe von analytischen Ansätzen und Forschungstechniken entwickelt, mit denen sich die in Sozial-

---

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] Genf: WHO, 1996).

10 Siehe Van Driel & Keijsers (1997). Das Instrument beruht auf den Ergebnissen einer von der Europäischen Kommission geförderten und von der International Union for Health Promotion and Education durchgeführten Studie zur Effektivität von Gesundheitsförderung und Gesundheitsförderung. Siehe hierzu zusammenfassend C. A. Vein (1995).

11 Siehe E. Ziglio, Methods for the evaluation of health promotion programmes and policies. Working paper. (Kopenhagen: WHO/EURO, 1994). Auf der Grundlage dieses Planungspapiers wurde beim ECHPD eine Arbeitsgruppe zur Evaluation eingerichtet.

12 Ein aktuelles Beispiel bieten Pelikan/Dietscher/Novak-Zezula (1998).

13 Siehe hierzu Brook/MacGlynn/Cleary (1996).

programmen stattfindenden sozialen Prozesse und Vorgänge untersuchen lassen und die auch für Projekte der Gesundheitsförderung nutzbar gemacht werden können.

Da die Evaluationsforschung bei einer Vielzahl von Programmen mit einem breiten Spektrum unterschiedlicher Interventionsstrategien und -methoden ihre Erfahrungen gesammelt hat, kann davon ausgegangen werden, daß sie für die unterschiedlichen Ansätze und Methoden der Krankheitsprävention und Gesundheitsförderung angewendet werden kann – also sowohl für Projekte, die Änderungen von Wissen, Einstellungen und Verhalten zum Ziel haben, als auch für Projekte, die auf die Umwelt- und Strukturbedingungen oder andere Bestimmungsgründe der Gesundheit Einfluß nehmen wollen.

Der besondere Vorteil der Evaluationsforschung besteht also darin, daß man auf Vorhandenes zurückgreifen kann. Dies wiederum bedeutet eine eher wenig aufwendige Strategie zur Qualitätssicherung in der Gesundheitsförderung.

Darüber hinaus bestehen Berührungspunkte zwischen den Denkansätzen der Evaluationsforschung und denen der Qualitätssicherung. So ist es ein Ziel der Programmevaluation, ihre Ergebnisse für eine Verbesserung der untersuchten Programme zu verwenden, wie dies die Evaluationsforscherin Carol Weiss bereits 1972 folgendermaßen beschrieben hat: „*Evaluationsforschung bedeutet die unmittelbare und direkte Verbesserung von Sozialprogrammen*“ (Weiss, 1972). Dieser Zielsetzung entspricht in der Qualitätssicherung das Konzept der „kontinuierlichen Verbesserung“ von Arbeitsprozessen als Grundprinzip des Qualitätsmanagements – eine Qualitätsphilosophie, die in den letzten Jahren auch in der Gesundheitsförderung diskutiert wird (Parish, 1995; Oeretveit, 1996; Saan, 1997).

## 1.2. INHALT DES ARBEITSPAPIERS

Bevor im einzelnen beschrieben wird, welche Rolle die Programmevaluation für die Qualitätssicherung in der Gesundheitsförderung übernehmen kann, müssen zunächst die wichtigsten Schlüsselbegriffe – *Evaluation*, *Programm* und *Qualität* – geklärt werden. Dies ist notwendig, um die aus den unterschiedlichen Wissenschaftstraditionen von Evaluationsforschung, Gesundheitsförderung und Qualitätsmanagement stammenden Terminologien zu präzisieren und anzugleichen, vor allem aber, um festzulegen, wie die Begriffe im folgenden Text verwendet werden. Diese Begriffsklärungen erfolgen in *Kapitel 2*, welches am Schluß noch einen kurzen Überblick über die wichtigsten Elemente eines qualitätsorientierten Evaluationskonzepts gibt.

In den anschließenden Kapiteln wird das Evaluationskonzept weiter ausgearbeitet: Es werden Möglichkeiten aufgezeigt, wie es bei der Planung, Implementation und Erfolgskontrolle von Gesundheitsförderungsprogrammen realisiert werden kann.

*Kapitel 3* zeigt, wie sich durch die Erarbeitung eines ausformulierten und differenzierten Interventionskonzepts in Form einer Programmtheorie bereits in der Planungsphase die Qualität von Gesundheitsförderungsprojekten verbessern läßt. In *Kapitel 4* geht es um die Qualitätsmessung bei der Implementation, vor allem um die Möglichkeiten, die die laufende systematische Beobachtung des Interventionsprozesses für dessen Effektivitäts- und Qualitätsverbesserung bietet. *Kapitel 5* schließlich behandelt die Wirkungsanalyse, also den Teil einer Evaluation, der sich mit der Messung der Ergebnisse und dem Nachweis der Wirksamkeit des Programms befaßt. *Kapitel 6* faßt das Arbeitspapier in einem kurzen Resümee zusammen und gibt als Handlungsempfehlungen die wesentlichen Merkmale eines qualitätsorientierten Evaluationskonzepts wieder.



**SCHLÜSSELBEGRIFFE UND ELEMENTE  
EINES QUALITÄTSORIENTIERTEN  
EVALUATIONSKONZEPTS**

## 2.1. EVALUATION

Ausgangspunkt für die folgenden Begriffsklärungen ist eine kurze Definition: *Evaluation ist die systematische Informationssammlung für die Bewertung von Programmen.* Anhand von Erläuterungen zu den vier Komponenten dieser Definition soll verdeutlicht werden, wie Evaluation in diesem Arbeitspapier verstanden wird.

1. Systematische *Informationssammlung* besagt, daß Interventionsprogramme mit Hilfe von empirischen Daten untersucht werden. Bei der Erhebung der Daten existieren keine Beschränkungen im Hinblick auf die Instrumente, die dabei verwendet werden. Psychologische Meßinstrumente, die z.B. in Vorher-Nachher-Kontrollexperimenten angewendet werden, sind ebenso akzeptabel wie persönlich-mündliche Interviews der Umfrageforschung oder qualitative Intensivinterviews. Andere Verfahren, wie z.B. die schriftliche Befragung von Teilnehmern oder Klienten, die systematische Beobachtung von Implementationsabläufen, die Auswertung von Verlaufsstatistiken oder Akten sowie die vergleichende Meta-Analyse von Ergebnissen einzelner Evaluationsstudien, gelten ebenfalls als adäquate Techniken der Informationssammlung. Das gilt auch für die typischen Informationsinstrumente des Qualitätsmanagements, wie Audits, Qualitätszirkel oder Qualitätsprüfungen mit Hilfe standardisierter Bewertungsskalen. Evaluation ist somit nicht auf bestimmte Methoden und Verfahren festgelegt; sie ist, was die Sammlung und Auswertung von Daten betrifft, „pluralistisch“.<sup>1</sup>

2. *Systematisch* bedeutet, daß die Informationssammlung nach methodischen Regeln vorgeht. Dazu gehört z.B., daß – mit dem Ziel der prinzipiellen Überprüfbarkeit – nachvollziehbar festgelegt ist, wie die Daten erhoben werden, und daß bei der Erhebung, Auswertung und Interpretation alles getan wird, um verzerrungsfreie Daten zu erhalten.

Systematisches Vorgehen bedeutet aber auch, daß festgelegt werden muß, welche Daten gesammelt werden und wie entschieden werden soll, ob sich die in den Daten ausdrückenden Ergebnisse als Interventionswirkungen interpretieren lassen oder nicht. Evaluation umfaßt eine Vielzahl solcher systematisierenden Vorgehensweisen: Forschungsdesigns, Auswahlverfahren, statistische Analysetechniken oder auch die theoretische Analyse von Interventionskonzeptionen. Diese Regeln und Vorgehensweisen sind überwiegend die der empirischen Sozialforschung – nicht zuletzt, wenn es um die Evaluation von Gesundheitsförderungsprogrammen oder -projekten geht. Dieses systematische Vorgehen der Programmevaluation wird als *Evaluationsforschung* verstanden.<sup>2</sup>

---

1 Der Ausdruck „pluralistisch“ wurde von L. Sechrest (1993) übernommen.

2 Siehe Rossi/Freeman (1993).

3. Der Ausdruck „*Bewertung* von Programmen“ soll darauf hinweisen, daß die Evaluationsergebnisse unmittelbar zum Programm *rückgekoppelt* werden: Sie sollen helfen, konkrete Entscheidungen in konkreten Programmen zu treffen.

Die Vielzahl möglicher Entscheidungen, für die Informationen bereitgestellt werden können, lassen sich unter drei generellen, in der Evaluationsliteratur immer wieder hervorgehobenen Perspektiven zusammenfassen<sup>3</sup>:

- die Perspektive der wirtschaftlichen Verantwortlichkeit,
- die Perspektive der Programmentwicklung,
- die Perspektive der Wissenserweiterung.

Unter dem Aspekt der wirtschaftlichen Verantwortlichkeit wird beispielsweise untersucht, ob die erwarteten Erfolge eingetreten sind und sich somit die für das Projekt eingesetzten Finanzmittel rechtfertigen. Unter der Perspektive der Programmentwicklung wird gezielt nach Möglichkeiten geforscht, das Programm oder einzelne Programmelemente zu verbessern. Die Perspektive der Wissenserweiterung schließlich bedeutet, daß Evaluationsforschung häufig auch dazu beiträgt, generelles, empirisch begründetes Wissen über die Wirksamkeit bestimmter Interventionsstrategien zu schaffen.

Inhaltlich ist Evaluationsforschung also nicht von vornherein festgelegt: Bei jeder Evaluation muß neu entschieden werden, welche Fragestellungen untersucht werden sollen. Dafür sind zwei sich gegenseitig ergänzende Vorgehensweisen von besonderer Bedeutung:

- Zum einen kann das Programmkonzept zu einer präzisierten und realitätsnahen Handlungstheorie umformuliert werden, die Voraussagen darüber enthält, welche Interventionen welche Veränderungen herbeiführen, und aus der sich konkrete Forschungsfragen ableiten lassen. Diese Vorgehensweise wird als *theory-driven Evaluation* (theoriegeleitete Evaluation) bezeichnet (Chen, 1990; Chen/Rossi, 1989; Lipsey, 1993 und Pawson/Tilley, 1997).
- Zum anderen werden die Fragestellungen der Evaluationsstudie durch einen engen Kooperationsprozeß zwischen Projektentwicklern einerseits und Evaluationsforschern andererseits entwickelt. Eine zentrale Aufgabe dieser Kooperation ist es, die Fragen herauszufinden und zu präzisieren, deren Beantwortung für Projektverantwortliche, Projektentwickler und Projektmitarbeiter den größten Nutzen haben. Für diese am Informationsnutzen orientierte Eingrenzung von Evaluationsinhalten wurde in der Literatur der Begriff *utilization-focused Evaluation* (anwenderorientierte Evaluation) geprägt (Patton, 1997).

4. Um deutlich zu machen, daß der Begriff *Programm* zu den zentralen Schlüsselbegriffen gezählt werden soll, wird er im folgenden Abschnitt eingehender erläutert.

<sup>3</sup> So z.B. bei E. Chelimsky, 1997.

## 2.2. PROGRAMM

Der Sprachgebrauch der Evaluationsforschung deckt sich nicht vollständig mit dem von Politik und Verwaltung. Dort werden häufig umfassendere und längerfristige Vorhaben als Programme bezeichnet, wie z.B. das Aktionsprogramm Gesundheitsförderung der EU, während Projekte meist zeitlich begrenzte Vorhaben sind, die innerhalb eines bestimmten Budgets finanziell gefördert werden und oft Teil eines Programms sind.

In der Evaluationsliteratur ist der Begriff *Programm* die überwiegend verwendete Bezeichnung für den *Gegenstand der Evaluationsforschung*. Für nicht besonders umfangreiche Programme wird auch das Wort Projekt benutzt. Teilprogramme oder Teilprojekte bezeichnet die Evaluationsforschung als *Programmkomponenten* (Bickman, 1985). Darunter versteht man eine abgrenzbare Menge von Aktivitäten, die auf ein bestimmtes Unterziel des Programms gerichtet sind. Komponenten können auch getrennt evaluiert werden, wie z.B. die einzelnen TV-Spots einer Fernsehkampagne zur Gesundheitsaufklärung. Im vorliegenden Arbeitspapier werden diese Begriffe entsprechend der Terminologie der Evaluationsforschung verwendet.

In der Gesundheitsförderung können unter dem Begriff *Programm* völlig unterschiedliche Aktivitäten subsumiert werden. So kann es sich dabei z.B. um eine umfassende, sich über mehrere Jahre erstreckende, auf die Allgemeinbevölkerung abzielende Gesundheitskampagne oder um die Umorganisation von Arbeitsabläufen zur Entwicklung eines gesundheitsfördernden Krankenhauses handeln; es kann aber auch ein Training zur gesundheitsbezogenen Kompetenzförderung damit gemeint sein.

### PROGRAMM ALS KOOPERATIONSSYSTEM

Unabhängig von solchen Programmzielen und -inhalten läßt sich ein Programm als Kooperationssystem verstehen. Dies heißt im einzelnen:

- In einem Programm arbeiten in der Regel eine ganze Reihe von Akteuren mit zum Teil unterschiedlichen Vorstellungen, Absichten und Interessen zusammen: die Initiatoren und Nutznießer des Programms aus Politik und Verwaltung, die Finanziere, die Projektentwickler und Praktiker sowie die Teilnehmer und Klienten. Und nicht zuletzt gehören zu diesem Kooperationssystem auch die Evaluatoren. Alle diese Beteiligten an einem Programm werden häufig mit dem Begriff „*Stakeholder*“ zusammengefaßt, der auch hier verwendet wird.
- Kooperationssystem bedeutet weiterhin, daß die Kooperation der Stakeholder als Prozeß verstanden wird, der sich grob in drei Phasen einteilen läßt:

- *Programmplanung*: Entwicklung eines Handlungskonzepts,
  - *Programmimplementation*: Realisierung des Konzepts,
  - *Überprüfung der Ergebnisse*: Nachweis der Wirksamkeit.
- Der Programmprozeß läßt sich als Strategie beschreiben. Wichtigste Bestandteile von Programmstrategien sind *Interventionen*: Methoden, Techniken und Aktivitäten, wie z.B. die Motivierung von Rauchern, an einer Gruppendiskussion über die Abhängigkeit vom Rauchen teilzunehmen. Interventionen sind also Interaktionen mit einem bestimmten Personenkreis. Dieser Personenkreis wird in der Evaluationsforschung in der Regel als *Teilnehmer* oder *Klienten* bezeichnet, die sich meistens aus einer mehr oder weniger präzise definierten Zielgruppe rekrutieren.

Mit der Intervention ist beabsichtigt, bei den Teilnehmern bestimmte Veränderungen zu bewirken, z.B., daß die Teilnehmer eines Raucherentwöhnungskurses das Rauchen aufgeben. Die beabsichtigten oder tatsächlich eingetretenen Wirkungen von Interventionen werden als *Interventionsergebnisse* oder *Outcomes* bezeichnet.

- Jeder Intervention liegt ein mehr oder weniger ausformuliertes Handlungskonzept zugrunde, das sich auch als (*Interventions-*)*Theorie* verstehen läßt. Diese beschreibt, durch welche *Wirkungsmechanismen* die beabsichtigten Outcomes erzielt werden sollen, und wie die Intervention funktioniert: Wer tut was mit wem und mit welchen Konsequenzen.
- Die praktische Realisierung, die Implementation eines Programms, findet in einer besonderen Situation von Umständen und Bedingungen statt: in einem bestimmten Setting, das auf eine besondere Weise organisiert ist, in einer bestimmten Region mit einer spezifischen Kultur; die Intervention richtet sich an eine bestimmte nach Geschlecht, Alter, Schulbildung und sozialer Schicht beschreibbaren Zielgruppe und wird von einem Mitarbeiterstab mit einer spezifischen Ausbildung ausgeführt. Alle diese Rahmenbedingungen entscheiden mit darüber, wie die Wirkungsmechanismen in der Realität funktionieren und welche Interventionsergebnisse tatsächlich erreicht werden. Die Summe dieser Rahmenbedingungen wird im folgenden als Kontext (Pawson/Tilley, 1997) bezeichnet.

## 2.3. QUALITÄT

Eine generelle und gleichzeitig inhaltsreiche Definition des Begriffs „Qualität“ zu finden, ist äußerst schwierig. In der Qualitätsdiskussion der Industrie (Garvin, 1984) oder auch der Gesundheitsversorgung (Donabedian, 1988; Blumenthal, 1996) wurde deshalb vorgeschlagen, statt einer globalen Definition verschiedene Sichtweisen bzw. Perspektiven zu unterscheiden und zu beschreiben. Dieser Vorschlag soll hier aufgenommen werden.

In der Qualitätsdiskussion der Gesundheitsförderung lassen sich grob drei Perspektiven unterscheiden:

- die Perspektive der wirtschaftlichen Verantwortlichkeit,
- die Stakeholder-Perspektive und
- die Experten-Perspektive.

### **PERSPEKTIVE DER WIRTSCHAFTLICHEN VERANTWORTLICHKEIT**

Bei der Perspektive der *wirtschaftlichen Verantwortlichkeit* geht es vor allem um die Frage, wie sich die für Interventionsprogramme bereitgestellten Finanzmittel effizient einsetzen lassen. Qualitätssicherung erfordert also die Abstimmung von verfügbaren Ressourcen und angestrebten Programmergebnissen im Sinne von Effizienzsteigerung.

Gerade diese Perspektive der Qualitätsdiskussion findet sich gegenwärtig auch in vielen Veröffentlichungen zum Thema Qualität und Gesundheitsförderung – meist im Zusammenhang mit dem Hinweis auf knapper werdende Finanzmittel. Zwei Varianten des Arguments, Gesundheitsförderung ließe sich durch mehr Qualität rechtfertigen, lassen sich dort erkennen: Bei der einen Variante bedeutet Qualitätsverbesserung die Schaffung von zusätzlicher Effektivität oder Qualität bei einem gegebenen finanziellen Ansatz, während die zweite Variante die Aufrechterhaltung des bisherigen Qualitätsniveaus bei einer Reduzierung des Finanzbudgets meint.

### **DIE STAKEHOLDER-PERSPEKTIVE**

Die *Stakeholder-Perspektive* berücksichtigt die Sicht derjenigen, die in irgendeiner Weise einen Nutzen von einem Produkt oder einer Dienstleistung, wie z.B. einer Gesundheitsförderungsmaßnahme, erwarten. Sie besagt, daß in die inhaltliche Definition dessen, was im Einzelfall unter Qualität verstanden werden soll, die Vorstellungen, Wünsche und Bedürfnisse der verschiedenen Stakeholder einbezogen werden sollen. Ganz besonders gilt dies für Kunden, Patienten, Klienten und andere Adressaten von Produkten oder Dienstleistungen. Zu den Stakeholdern, deren Qualitätserwartungen berücksichtigt werden

müssen, gehören auch externe Kooperationspartner, Finanziere oder interessierte politische Institutionen. Nicht zuletzt gehören dazu die Mitarbeiter und deren Erwartungen an die Qualität der Vorleistungen, die von anderen Organisationseinheiten des Programms jeweils zu erbringen sind („interne Kunden“). Es geht auch nicht allein darum, zur Bestimmung der Qualität die ausdrücklich geäußerten Wünsche heranzuziehen, sondern auch die Bedürfnisse, die sich aus Lebensweisen und Problemlagen der jeweiligen Stakeholder ergeben.

## **DIE EXPERTEN-PERSPEKTIVE**

Die *Experten-Perspektive* ist die der Programmentwickler, Projektmanager und Mitarbeiter. Diese sogenannte „technische Qualität“ hat zwei Dimensionen. Sie muß zum einen angemessene und wirkungsvolle Instrumente vorsehen und zum anderen diese auf die richtige Art und Weise einsetzen, so daß die beabsichtigten Interventionsergebnisse erreicht werden („Doing the right things right“).

In der Gesundheitsförderung beruht die technische Qualität auf einem durch Erfahrung und wissenschaftliche Erkenntnisse begründeten Wissen über die Wirkungszusammenhänge von Interventionen.

## **QUALITÄTSSICHERUNG ALS MANAGEMENTKONZEPT**

Bei der Qualitätssicherung konkreter Gesundheitsförderungsprogramme spielen alle drei Perspektiven eine Rolle, wobei jedoch normalerweise keine eine eindeutige Priorität besitzen wird. Es ist Aufgabe des für Planung und Ablauf des Programms verantwortlichen „Programm-Managements“, die verschiedenen Qualitätsperspektiven miteinander ins Gleichgewicht zu bringen. Programme sind also nicht als eine einfache Abfolge von Zielbestimmung, Implementation und Ergebnis zu verstehen; vielmehr verlaufen sie nach komplexen Strategien, welche die (Qualitäts-)Erwartungen der Stakeholder, die technischen Anforderungen eines möglichst effektiven Interventionsprozesses sowie die Ressourcen des Programms aufeinander abstimmen. Qualitätssicherung wird dementsprechend vielfach als Managementkonzept verstanden – im vorliegenden Fall als Managementkonzept für die Organisation von Gesundheitsförderungsprogrammen. Für ein solches Qualitätsmanagement haben sich eine ganze Reihe spezifischer Handlungsprinzipien herausgebildet. Fünf dieser Prinzipien, die für die Planung und Realisierung von Gesundheitsförderungsprogrammen besonders wichtig erscheinen, werden im folgenden kurz erläutert:

### **Outcome-Orientierung**

Qualitätssicherung in diesem Sinne erfordert die Ausrichtung des Programms an eindeutig beschreibbaren und prinzipiell nachweisbaren Interventionsergebnissen. Diese Out-

comes sind für eine klar abgrenzbare Gruppe von Klienten bestimmt, wobei auch deren Qualitäts- und Nutzenerwartungen berücksichtigt werden. Nicht zuletzt sollen die Ergebnisse ökonomisch realisierbar sein. Es reicht nicht aus, wenn Programme *ausschließlich* durch hoch bewertete generelle Zielvorstellungen legitimiert werden.

### **Stakeholder-Orientierung**

Beim Programm-Management, bei der Festlegung der Outcomes, bei der Programmplanung und der Implementation werden die Interessen, Erwartungen sowie die Lebens- und Arbeitsverhältnisse der Stakeholder berücksichtigt – vor allem die der Klienten und Mitarbeiter, aber auch die der Kooperationspartner, der interessierten Politiker oder der Haushaltsverantwortlichen.

Stakeholder-Orientierung bedeutet jedoch auch, daß nicht ohne weiteres von einem Konsens zwischen den verschiedenen Beteiligungsgruppen ausgegangen werden kann. Voraussetzung für die Formulierung eines klaren Qualitätssicherungskonzepts ist ein Abstimmungs- und Konsensbildungsprozeß, dessen Ziel gemeinsame Vorstellungen von der Qualität einer Gesundheitsförderungsmaßnahme sind.

### **Prinzip der kontinuierlichen Qualitätsverbesserung**

Das der technischen Qualität zugrundeliegende Wissen der (Interventions-)Experten wird nicht als konstant angesehen: Es ändert sich in einem kumulativ ablaufenden Prozeß der *kontinuierlichen Qualitätsverbesserung*. Dieses Prinzip läßt sich in wenigen Worten folgendermaßen zusammenfassen: Arbeitsvorgänge und -prozesse werden laufend empirisch überprüft (Qualitätsmessung). In dieser Überprüfung wird vor allem die Möglichkeit gesehen, qualitätsmindernde Probleme zu identifizieren und Ansätze zu deren Lösung zu entdecken, um diese Erkenntnisse bei der anschließenden Planung zur Verbesserung von Arbeitsvorgängen einzusetzen. Entsprechend der Vorliebe der Management-Wissenschaften für einfache schematische Darstellungen, wird dieses Konzept als Kreis dargestellt, bekannt als Shewhart- oder Deming-Zyklus, der Planen, Ausführen, Überprüfen und Verbessern als zyklischen Vorgang miteinander verbindet (Deming, 1986).

### **Prozeßorientierung**

Das Prinzip der kontinuierlichen Qualitätsverbesserung ist nicht in erster Linie auf das Endprodukt ausgerichtet, sondern auf den *gesamten Prozeß* von Leistungen, Interventionen und anderen Aktivitäten, die zu den Interventionsergebnissen führen sollen. Entscheidungen und Maßnahmen zur Qualitätsverbesserung und Effektivitätssteigerung sollten möglichst am Anfang dieses Prozesses stattfinden.

### **Prinzip der Fehlervermeidung**

Das Prinzip der ständigen Verbesserung macht es möglich, aus Erfahrungen zu lernen und Fehlermöglichkeiten leichter zu identifizieren. So kann die Projektplanung darauf ausgerichtet sein, Fehler von vornherein zu vermeiden statt das Endprodukt nachträglich nachzubessern.

Dieser kurze Versuch, den Begriff Qualität zu klären, orientiert sich an den Möglichkeiten der Evaluationsforschung als Instrument zur Qualitätssicherung. Unerwähnt bleiben somit die vielen Qualitätsinstrumente, die in der Industrie oder in der Gesundheitsversorgung, zum Teil auch in der Gesundheitsförderung, angewendet werden, um Qualität zu erreichen, sicherzustellen und zu verbessern.

Im folgenden wird die Evaluationsforschung als Qualitätssicherungsinstrument behandelt, und es wird ein Evaluationskonzept beschrieben, das sich an den oben genannten Zielvorstellungen der Qualitätssicherung orientiert.

## 2.4. QUALITÄTSORIENTIERTES EVALUATIONSKONZEPT

### OUTCOME-ORIENTIERUNG

Zunächst einmal liegt es nahe, Evaluationsforschung zur Messung von Qualität zu verwenden: Die *Outcome-Orientierung* mit ihrer Betonung nachweisbarer Programmsergebnisse verlangt die Anwendung gültiger und zuverlässiger Forschungsverfahren zur Messung der Interventionsergebnisse und deren Veränderungen im Zeitverlauf. Die entsprechenden Methoden und Forschungstechniken hierfür kann die Evaluationsforschung zur Verfügung stellen.

### KLIENTEN- UND MITARBEITERORIENTIERUNG

Als Sozialforschung ist die Evaluationsforschung auch besonders dazu geeignet, zuverlässige und unverzerrte Informationen für die Realisierung der *Klienten- und Mitarbeiterorientierung* zu beschaffen; denn die Befragung von Wählern, Konsumenten, Patienten, von Teilnehmern an Sozialprogrammen, aber auch von Mitarbeitern in Betrieben, Krankenhäusern und anderen Organisationen ist eines der wichtigsten Anwendungsgebiete der empirischen Sozialforschung. Befragungsthemen in diesem Zusammenhang sind beispielsweise die Beurteilung und Akzeptanz spezifischer Leistungen oder die Verhaltenskonsequenzen bei den Nutzern dieser Leistungen.

### PROZESSORIENTIERUNG

Das *Prinzip der ständigen Verbesserung* verlangt jedoch, daß sich das Evaluationsprojekt nicht darauf beschränkt, die endgültigen Ergebnisse zu messen. Es muß auch Informationen darüber sammeln, wie die Interventionsergebnisse zustande gekommen sind.

Nur so läßt sich aus den Forschungsergebnissen schlußfolgern, ob Implementationsvorgänge oder das Handlungskonzept verbessert werden können. Die Evaluationsstudie sollte deshalb über die gesamte Wirkungskette des Programms Daten liefern – angefangen bei Daten über die Kontaktaufnahme mit den Klienten, dann zu Aktivitäten und Vorgängen, die die eigentliche Intervention ausmachen, bis hin zu Daten zu den beabsichtigten (oder auch unbeabsichtigten) Wissens-, Einstellungs- und Verhaltensänderungen der Programmteilnehmer. Untersuchungsgegenstand einer qualitätsorientierten Evaluationsforschung ist demnach – ganz im Sinne der *Prozeßorientierung* – das gesamte Programm.

## **PRINZIP DER FEHLERVERMEIDUNG**

Die auf Fehlervermeidung ausgerichtete *präventive Orientierung* der Qualitätssicherung verlangt, daß die systematische Informationssammlung möglichst früh einsetzt und bereits während der Implementationsphase untersucht, inwieweit das Interventionskonzept fehlerfrei umgesetzt wird. Frühzeitige Fehlervermeidung bedeutet aber auch, das Konzept selbst zu überprüfen, bevor es angewendet wird. Dies kann geschehen, indem die Zusammenhänge zwischen den Interventionsaktivitäten, deren Wirkungsmechanismen, den Kontextbedingungen, unter denen das Programm stattfindet, und den Outcomes möglichst präzise formuliert werden. Auf diese Weise kann vorher geprüft werden, ob die geplanten Interventionsaktivitäten – bei gegebenen Ressourcen – machbar sind, und es kann abgeschätzt werden, ob sie überhaupt die beabsichtigten Wirkungen haben werden. Diese vorherige Prüfung der Machbarkeit ist vor allem auch eine notwendige Voraussetzung für die Durchführbarkeit der Evaluationsstudie: Nur realisierbare Interventionen und Outcomes geben meßbare Forschungsobjekte ab.

## **KOMPONENTEN EINES QUALITÄTSORIENTIERTEN EVALUATIONSKONZEPTS**

Qualitätsorientierte Evaluation wird somit als Prozeß verstanden, der ein Programm von Anfang bis zum Ende begleitet: von der Programmkonzeption über die Implementation und die Messung der Outcomes bis hin zur abschließenden Einschätzung der Wirksamkeit und zu Vorschlägen von Verbesserungsmöglichkeiten. Organisatorische Voraussetzung für dieses Evaluationskonzept ist eine enge, ebenfalls bereits bei der Planung und Konzeptentwicklung beginnende Kooperation zwischen Projektentwicklern und Evaluationsforschern.

Im folgenden wird ein auf diese Qualitätsprinzipien ausgerichtetes Evaluationskonzept vorgeschlagen. Es setzt sich aus drei Komponenten zusammen:

1. Die *Programmtheorie* beschreibt den vorgesehenen Wirkungsmechanismus: Wer tut was auf welche Weise mit wem und mit welchen Ergebnissen? Ziel ist es, die Reali-

sierbarkeit der Ergebnisse zu überprüfen, die Effektivität der vorgesehenen Interventionen zu verbessern und eine theoretische Grundlage für die Evaluationsstudie verfügbar zu machen.

2. Ein *Implementations-Monitoring* überprüft, ob der Wirkungsmechanismus ausgelöst ist und programmgemäß abläuft, und schafft frühzeitig – noch während des Programmverlaufs – Verbesserungsmöglichkeiten.
3. Die *Wirkungsanalyse* untersucht, welche Interventionsergebnisse eingetreten sind und inwieweit diese durch die eingesetzten Interventionen hervorgerufen wurden. Sie zielt auf eine kontinuierliche Verbesserung des Programms ab, z.B. im Hinblick auf die nächste Einsatzrunde oder andere Einsatzbedingungen.

Diese drei Teilaufgaben sind voneinander abhängig und lassen sich nicht als eigenständige „formative“ Evaluation (des Implementationsprozesses) oder „summative“ Evaluation (der Outcomes) vornehmen.

## 2.5. ZUSAMMENFASSUNG

In der Diskussion über die Qualitätssicherung in der Gesundheitsförderung wird die Evaluationsforschung als eine Möglichkeit betrachtet, die Qualität einzelner Interventionsprogramme oder -projekte zur Gesundheitsförderung zu messen, zu bewerten und zu verbessern. Sie wäre eines unter mehreren Qualitätsinstrumenten (z.B. Qualitätszirkel), deren Anwendbarkeit in der Gesundheitsförderung ebenfalls diskutiert oder erprobt wird.

Die Anwendung der Evaluationsforschung hätte den Vorteil, daß man auf organisatorische Erfahrungen und einen effektiven methodischen Apparat zurückgreifen könnte, den sie in den letzten Jahrzehnten bei der Evaluation von Sozial-, Gesundheits- und anderen Interventionsprogrammen entwickelt hat. Es ist jedoch notwendig, die Evaluationsforschung an die Zielvorstellungen der Qualitätssicherung anzupassen.

Qualitätssicherung wird als Managementkonzept verstanden, das bei der Planung und Realisierung von (Gesundheitsförderungs-)Projekten unter anderem nach den folgenden Prinzipien vorgeht:

- *Outcome-Orientierung*: Qualitätssicherung setzt voraus, daß mit dem Programm ein eindeutig beschriebenes, inhaltlich, organisatorisch und ökonomisch realisierbares und prinzipiell meßbares Interventionsergebnis angestrebt wird.
- *Stakeholder-Orientierung*: Die in dem Programm angebotenen Dienstleistungen und Outcomes berücksichtigen die Qualitätserwartungen, Bedürfnisse und Lebensbedin-

gungen der Klienten, Mitarbeiter, Kooperationspartner und anderer Personen, die an dem Programm Interesse haben.

- *Prinzip der kontinuierlichen Verbesserung*: Qualität wird durch einen kumulativen Prozeß der ständigen Verbesserung sichergestellt.
- *Prozeßorientierung*: Die Verbesserungen richten sich vor allem auf die im Verlauf des Implementationsprozesses stattfindenden Aktivitäten und Vorgänge.
- *Prinzip der Fehlervermeidung*: Durch eine verbesserte Qualität der Planung lassen sich Fehlermöglichkeiten frühzeitig identifizieren, Schwierigkeiten von vornherein vermeiden und – dank möglichst frühzeitiger Entscheidungen – die Effektivität der Interventionen steigern.

Ein Evaluationskonzept, das sich an diesen Qualitätsprinzipien orientiert, umfaßt die drei folgenden Komponenten:

1. Präzisierung des Handlungskonzepts in Form einer Programmtheorie,
2. Monitoring des Implementationsprozesses und
3. Wirkungsanalyse.

Jede dieser Komponenten hat ihre eigene Aufgabe: Die Programmtheorie beschreibt den vorgesehenen Wirkungsmechanismus, das Implementations-Monitoring überprüft, ob der Wirkungsmechanismus ausgelöst ist und programmgemäß abläuft, und die Wirkungsanalyse schließlich untersucht, ob und wie gut er funktioniert hat.

Voraussetzung für die praktische Anwendbarkeit dieses Evaluationskonzepts ist eine mit der Programmplanung beginnende, kontinuierliche Kooperation zwischen Programmentwicklern/Projektmitarbeitern und Evaluationsforschern.

**PROGRAMMTHEORIEN**

3

# 3.

## PROGRAMMTHEORIEN

„*Es gibt nichts, was so praktisch wäre, wie eine gute Theorie.*“<sup>1</sup> Dieser Merksatz stand bereits am Anfang der sozialwissenschaftlichen Evaluationsforschung. Er wird Kurt Lewin zugeschrieben, dessen sozialpsychologische Forschungen während der 30er und 40er Jahre in den USA wichtige Grundlagen für soziale Interventionsprogramme und deren Erforschung geschaffen haben.<sup>2</sup> Dieser Anregung folgten später jedoch nicht allzu viele Evaluationsstudien, so daß gelegentlich der atheoretische Charakter der Evaluationsforschung beklagt wurde (Lipsey et al., 1985; Shadish/Epstein, 1987; Wortman, 1983). Hierfür wurde vor allem die methodische Doktrin verantwortlich gemacht, derzufolge Programmeffekte durch (randomisierte) Kontrollgruppen-Experimente nachzuweisen seien. Der Vergleich von Ergebnissen der Interventionsgruppe mit denen einer Kontrollgruppe führte in der Regel dazu, daß man sich darauf beschränkte, nur die Wirkungen der Intervention zu betrachten, während die internen Vorgänge des eigentlichen Interventionsprozesses als „black box“ behandelt wurden und somit weitgehend außer Betracht blieben (Lipsey, 1993; Pawson/Tilley, 1997). Vielfach wurden auch Sozialprogramme in Gang gesetzt, ohne viel darüber nachzudenken, auf welche Weise und unter welchen Umständen die erwarteten Wirkungen zustande kommen würden (Sechrest/Figueredo, 1993).

Seit Mitte der 80er Jahre tritt wieder stärker die Idee in den Vordergrund, daß man sich bei der Programmentwicklung und Forschungsplanung mehr von theoretischen Überlegungen leiten lassen solle. Ziel ist die theoriegeleitete Evaluation (vgl. *Abschnitt 2.1.*). Die Evaluationsziele und der Evaluationsplan leiten sich aus einer Programmtheorie ab, die in der Konzeptionsphase des Programms erarbeitet wird und gewissermaßen als Leitlinie dient für die Festlegung der Interventionsmaßnahmen, für deren Implementation in einem spezifischen sozialen Kontext, für die Auswahl der in der Evaluationsstudie zu untersuchenden Forschungsfragen und nicht zuletzt für die Qualitätsverbesserung des Programms. Eine solche theoriegeleitete Orientierung wurde in den letzten zehn Jahren mehr und mehr auch zu einer Maxime der Gesundheitsförderungspraxis.<sup>3</sup>

---

1 Zitiert nach H.E. Lück (1996). Siehe auch Marrow (1969).

2 Dies sind vor allem die Forschungsarbeiten, die Kurt Lewin nach seiner Emigration aus Deutschland in die USA dort in den 30er und 40er Jahren geleitet hat. Am bekanntesten sind die Feldexperimente über die Wirkungen von Erziehungsstilen auf das Sozialverhalten von Kindern in den Jahren 1937 und 1938 (Lewin/Lippitt/White, 1939) und die 1942 durchgeführten Experimente über die Wirkungen unterschiedlicher Kommunikationsstrategien in Kleingruppen auf die Änderung von Ernährungsgewohnheiten (Lewin, 1951 [dt. 1963]).

3 Siehe hierzu z.B. Green/Kreuter (1991) oder Bartholomew/Parcel/Kok (1998) sowie den Kommentar zu letzterem von C. A. Heaney (1998).

# 3.1. ZUM THEORIEVERSTÄNDNIS



Unter einer Theorie verstehen die Vertreter der theoriegeleiteten Programmevaluation – nicht anders als empirisch orientierte Sozialwissenschaftler generell – eine Menge von Annahmen, Prinzipien und Aussagen, mit denen sich spezifische soziale Handlungsweisen (Aktivitäten, Einstellungen) und deren Veränderung erklären lassen. Zum Teil enthalten solche Theorien relativ allgemein und abstrakt formulierte Annahmen, wie z.B. die bei der Planung von Aufklärungskampagnen häufig zugrundeliegende Annahme, daß Informationsvermittlung zu Einstellungswandel und Einstellungswandel wiederum zu Verhaltensänderung führt. Vor allem aber gehören zu einer Theorie spezifische, realitätsbezogene Aussagen oder Hypothesen über die Wirkungen von ganz bestimmten *unabhängigen (Kausal-)Variablen* auf bestimmte *abhängige Variable*. Diese Forschungshypothesen können darüber hinaus *intermediäre Variable* enthalten, die erklären, wie eine unabhängige Variable eine abhängige Variable beeinflusst.<sup>4</sup>

Diese formale Beschreibung dessen, was unter einer Theorie verstanden werden soll, läßt sich durch einige inhaltsbezogene Kriterien ergänzen, die vor allem für Programmtheorien von Bedeutung sind (Chen, 1990:40ff.):

- Programmtheorien sind *ergebnisorientiert*: Die Theorie grenzt ein, welche Programmergebnisse grundsätzlich erreicht werden können und ermöglicht damit die Entscheidung, welche Interventionsergebnisse unter den jeweiligen Kontextbedingungen und Ressourcen im einzelnen erreicht werden sollen.
- Programmtheorien sind *handlungsorientiert*: Die unabhängigen (Kausal-)Variablen beschreiben in erster Linie ganz bestimmte Aktivitäten (Einsatz von Ressourcen, Interventions- und Implementationstätigkeiten).
- Programmtheorien sind *interventionsorientiert*: Sie sind nicht so sehr in abstrakten Konzepten formuliert, sondern beschreiben in konkreten Begriffen, wie der Wirkungsmechanismus – unter ganz bestimmten Kontextbedingungen – abläuft, also wie die Interventionsaktivitäten Schritt für Schritt zu den erwarteten Ergebnissen führen.

---

<sup>4</sup> In der Sprache der Evaluationsforschung wird der ursächliche Faktor als unabhängige Variable, das verursachte Problem als abhängige Variable bezeichnet. Die zwischen unabhängige und abhängige Variable tretenden Variablen heißen intervenierende Variable. In diesem Arbeitspapier wird jedoch statt von intervenierenden Variablen von intermediären Variablen gesprochen. Damit soll deutlich gemacht werden, daß intervenierende Variable nicht unbedingt mit (gesundheitsfördernden) Interventionen gleichzusetzen sind.

## 3.2. THEORIEBILDUNG: QUELLEN

Als Programmtheorien können vorhandene *sozialwissenschaftliche* Theorien verwendet werden; allerdings ist dies nur in seltenen Fällen möglich, da für die Gesundheitsförderung – wie für die meisten anderen Anwendungsbereiche von Sozialprogrammen auch – nur wenige realitätsbezogene und empirisch abgesicherte Theorien verfügbar sind, die sich unmittelbar auf den Gegenstandsbereich von spezifischen Gesundheitsförderungsprogrammen anwenden lassen.

Einzelne Theorien aus den Grundlagendisziplinen der Gesundheitsförderung können jedoch als Ausgangspunkt oder Bezugsrahmen für die Erarbeitung von Interventionsstrategien und deren Wirkungsmechanismen dienen. So wurde in dem „Michigan Jobs Program“, das im folgenden häufiger als Beispiel verwendet wird, ein Kursprogramm zur Verbesserung der praktischen und psychischen Kompetenz für die Stellensuche entwickelt, das sich auf moderne lerntheoretische Konzepte stützt (Caplan/Vinokur/Price, 1997). Die Planer von Aufklärungskampagnen wählen als theoretischen Ausgangspunkt gelegentlich das Modell des Einstellungswandels von McGuire (1989).<sup>5</sup>

Eine weitere Quelle für die Theoriebildung sind Forschungssynthesen oder Meta-Analysen vorliegender Evaluationsstudien oder anderer inhaltlich relevanter Forschungen. Die Ergebnisse solcher Übersichten sind besonders dann nützlich, wenn sie erkennen lassen, ob und wie die dort behandelten Interventionsstrategien wirken. So ergab z.B. eine im Auftrag der Bundeszentrale für gesundheitliche Aufklärung durchgeführte Forschungssynthese von Studien zur Suchtprävention eine ganze Reihe von Hinweisen für die Brauchbarkeit des Lebenskompetenzansatzes (Künzel-Böhmer/Bühringer/Janik-Konecny, 1993).

Ein weiteres Beispiel für Forschungssynthesen zur Wirksamkeit von Gesundheitsförderungsinterventionen ist die von der Europäischen Kommission geförderte und von der International Union for Health Promotion and Education durchgeführte Studie zur Effektivität von Gesundheitsförderung und Gesundheitserziehung.<sup>6</sup>

Bei besonders umfangreichen und kostenaufwendigen Programmen empfiehlt es sich, einzelne Theorieteile, wie z.B. Hypothesen über die Wirksamkeit bestimmter Interventionskomponenten, vor Beginn der Implementation in *vorbereitenden Untersuchungen* an der Realität zu überprüfen.

---

<sup>5</sup> Siehe auch J. Töppich, 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: Tagungsbericht, 1996).

<sup>6</sup> Ein Liste von 16 Einzelberichten zu einzelnen Feldern der Gesundheitsförderung findet sich in dem zusammenfassenden Bericht zu diesem Projekt von C. A. Veen (1995).

Die wichtigste Quelle der Theoriebildung sind jedoch die Konzepte und das *professionelle Wissen* der Programmentwickler, Projektmitarbeiter und anderer wichtiger Stakeholder. In deren generellen Zielvorstellungen und konzeptionellen Grundideen zum Programm sowie in deren Vorstellungen über konkrete Verfahrensweisen und Routinen sind implizit theoretische Annahmen oder Hypothesen über die Wirkungen bestimmter Maßnahmen oder über die Ursachen spezifischer Outcomes enthalten. Eine Programmtheorie zu formulieren bedeutet, diese impliziten Handlungstheorien explizit zu machen.

# 3

## 3.3. THEORIEBILDUNG: ARBEITSSCHRITTE

Die Ausarbeitung einer Programmtheorie kann in mehreren Arbeitsschritten vor sich gehen:

- Problemanalyse,
- Festlegung der effektiven Interventionskomponenten,
- Definition der Outcomes,
- Beschreibung der Wirkungsmechanismen.

Bei der praktischen Erarbeitung einer Theorie lassen sich die Arbeitsschritte nicht immer in eine feste Reihenfolge bringen, und man muß bei jedem Arbeitsschritt die durch die übrigen Arbeitsschritte gesetzten Rahmenbedingungen jeweils mitbedenken. Am Anfang steht jedoch in der Regel die Problemanalyse.

### PROBLEMANALYSE

Problemanalysen für Programme der Gesundheitsförderung beschreiben die nichtmedizinischen Ursachen oder Entstehungsprozesse gesundheitlicher Probleme. Dies sind beispielsweise Probleme, die sich auf die Lebenssituation bestimmter Bevölkerungsgruppen zurückführen lassen, wie z.B. eine besonders hohe Prävalenz bestimmter Krankheiten oder besondere soziale Voraussetzungen für die Entwicklung und Erhaltung der Gesundheit. Oder es handelt sich um Probleme, die sich als Konsequenz allgemeiner sozialer Probleme ergeben, wie beispielsweise die Beeinträchtigung der psychischen Gesundheit als Folge von Arbeitslosigkeit.

Diese unerwünschten Zustände wird die Problemanalyse möglichst genau darstellen, und sie wird – möglichst auf der Basis epidemiologischer Daten – aufzeigen, welchen Umfang das Problem hat, wo und bei welchen Problemgruppen es auftritt und welche Konsequenzen es sonst noch hat. Weiter kann untersucht werden, warum sich das Problem nicht von selbst löst oder begrenzt, und wie es sich – ohne Intervention – in Zukunft weiterentwickeln wird.

Besonders wichtig für die Auswahl einer Interventionsstrategie und die Ausarbeitung einer Programmtheorie ist es, die Ursachen des Problems herauszufinden und möglichst genau zu beschreiben, wie der problematische und unerwünschte (Gesundheits-)Zustand normalerweise – ohne Einwirkung einer Intervention – zustande kommt. Dabei kommt es nicht allein darauf an, die grundlegenden sozialen Ursachen zu bestimmen. Die Programmanalyse sollte klären, ob zwischen der Ursache und ihren unerwünschten Folgen intermediäre Variable identifiziert werden können, mit deren Hilfe sich verstehen läßt, auf welche Weise die soziale Ursache zu dem Problem führt. Dies ist vor allem deshalb wichtig, weil Erkenntnisse über die Art und Weise, wie das Problem entsteht, die Entscheidung erleichtert, auf welche Weise interveniert werden muß, um eine realisierbare Veränderung des unerwünschten Zustands zu erreichen. Wünschenswert ist, wenn sich die Problemanalyse auf empirische Forschungsergebnisse stützen kann.

Ein Beispiel für eine solche Problemanalyse sind Untersuchungen zur Arbeitslosigkeit in der Region von Detroit, USA, während der Rezession Anfang der 80er Jahre. Deren Ergebnisse wurden später als Grundlage für die theoriegeleitete Planung und Evaluation des „Michigan Jobs Program“ verwendet, das im weiteren Verlauf dieses Kapitels eben wegen dieses theoriegeleiteten Ansatzes (und seiner nachgewiesenen Wirksamkeit) mehrfach als Beispiel herangezogen wird.

Ausgangspunkt für die Problemanalyse dieses Programms war die Beobachtung, daß der Verlust des Arbeitsplatzes bei vielen Arbeitslosen zu Beeinträchtigungen der psychischen Gesundheit führte, z.B. in Form von Angst oder Depressivität. Umfragen bei Beschäftigten, Arbeitslosen und wiederbeschäftigten ehemaligen Arbeitslosen bestätigten diese Beobachtungen. Darüber hinaus zeigten diese Umfragen, daß die Symptome bei einer Wiederbeschäftigung wieder verschwanden. Es wurde eine ganze Reihe möglicher intermediärer Variablen untersucht, um erklären zu können, auf welche Weise Arbeitslosigkeit zu Depressionen führt. Nachweisen ließ sich die folgende Kausalkette: Verlust des Arbeitsplatzes führt bei andauernder Arbeitslosigkeit zu einer ungünstigen finanziellen Situation. Die dadurch erzeugte generelle Streßbelastung wiederum verursacht dann die Beeinträchtigungen der psychischen Gesundheit. Damit war mit dem finanziell bedingten Streß die wichtigste intermediäre Variable identifiziert. Eine weitere, etwas schwächer wirkende intermediäre Variable war eine herabgesetzte Fähigkeit von Arbeitslosen, zusätzliche Lebenskrisen zu bewältigen, was dann wieder die psychische Gesundheit verschlechterte. Andere potentiell intermediäre Faktoren, die häufig als Folge von Arbeitslosigkeit auftreten, wie z.B. Ehekonflikte oder reduzierte Sozialkontakte, hatten keinen Einfluß. Es ließen sich allerdings einige modifizierende Faktoren, wie soziale Unterstützung, Selbstwertgefühl oder die Fähigkeit zur Bewältigung persönlicher Krisen, feststellen, deren Vorhandensein die psychischen Auswirkungen des Arbeitsplatzverlustes abschwächten (Kessler/Turner/House, 1987; 1988. Siehe Abbildung 1).

Das bei der Problemanalyse gefundene einfache, aber empirisch begründete Modell des Zusammenhangs zwischen Arbeitsplatzverlust und psychischer Gesundheit läßt erkennen,

worin die wirkungsvollste Interventionsmöglichkeit besteht: in einer Veränderung der wirksamsten intermediären Variablen, d.h. die psychische Situation von Arbeitslosen wird am ehesten durch eine Veränderung ihrer finanziellen Situation verbessert.

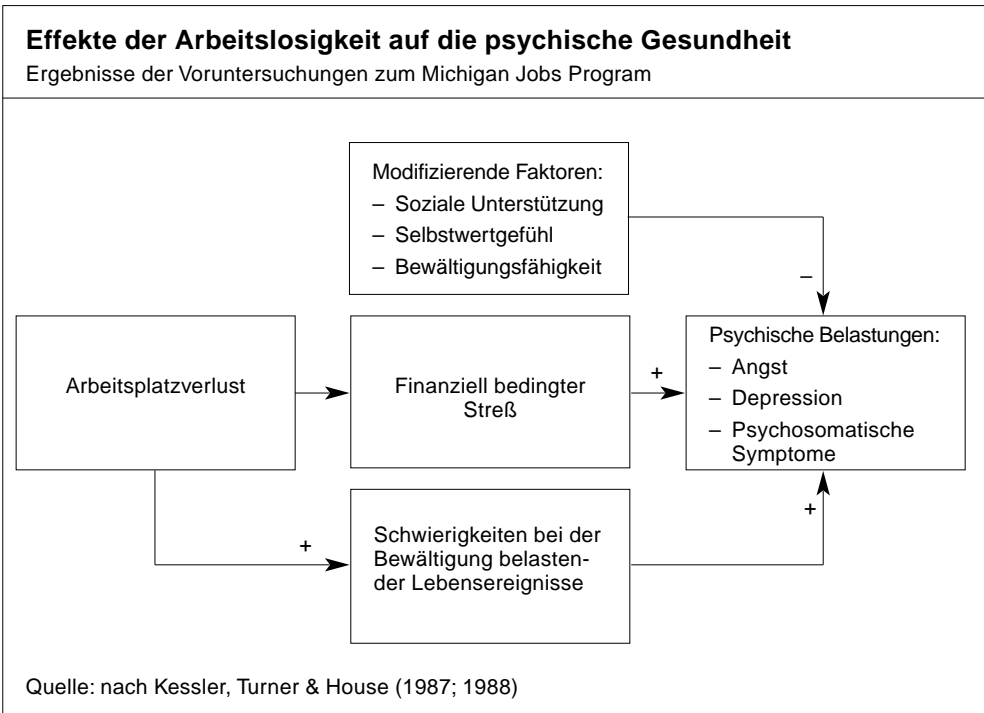


Abb. 1

Diese Interventionsmöglichkeit war dem Projektträger, dem Michigan Prevention Research Center am Institut für Sozialforschung der Universität Michigan, nicht gegeben. Deshalb wurde ein Programm gewählt, das die Wahrscheinlichkeit einer möglichst frühzeitigen Wiederbeschäftigung erhöhte, damit die Streßsituation gar nicht erst wirksam werden und die psychische Gesundheit verschlechtern konnte. Das Programm bestand aus einem Kursangebot für Arbeitslose, durch das deren praktische und psychische Kompetenzen für die Arbeitsplatzsuche möglichst unmittelbar nach dem Verlust des Arbeitsplatzes verbessert werden sollte.

Dieses Beispiel zeigt die wichtigste Aufgabe der Programmanalyse: Die Entwicklung einer Theorie über die Ursachen des Problems, die die Wirkungskette von ursächlichen und intermediären Variablen aufzeigt, so daß sich erkennen läßt, an welcher Stelle dieses Ursachensystems man - unter Berücksichtigung der gegebenen Ressourcen – eingreifen muß, um die Problemsituation effektiv zu verbessern.

## FESTLEGUNG DER EFFEKTIVEN INTERVENTIONSKOMPONENTEN

Die Festlegung der effektiven Interventionskomponenten bedeutet die Spezifizierung der gewählten Interventionsstrategie: Die einzelnen Interventionsaktivitäten und Dienstleistungen, die das Programm anbieten soll, werden unter Berücksichtigung der finanziellen und personellen Ressourcen festgelegt und beschrieben. Besondere Bedeutung hat dabei die Festlegung derjenigen Komponenten und Einzelmaßnahmen, die im Hinblick auf die angestrebten Outcomes am effektivsten erscheinen.

Als Beispiel für die Festlegung von Interventionskomponenten kann wiederum das „Michigan Jobs Program“ herangezogen werden (Price/Vinokur, 1995). Die Interventionsstrategie sah ein Kursprogramm vor für Arbeitsuchende, die nicht länger als dreizehn Wochen arbeitslos waren; darin konnten sie lernen, wie man sich bewirbt, d.h. also die praktischen Fertigkeiten für Stellensuche, Bewerbungsschreiben und Vorstellungsgespräch. Entscheidend war jedoch der Aufbau und die Stärkung eines Verhaltensrepertoires zur psychischen Bewältigung der Arbeitsuche.

Das generelle Interventionskonzept wurde für den Kurs in fünf einzelne Lernziele ausdifferenziert:

1. erfolgreiche Rekrutierung der Teilnehmer,
2. Schaffung von Vertrauen in die Trainer und die übrigen Teilnehmer des Kurses,
3. Steigerung der Motivation zur Anwendung von wirksamen Handlungsstrategien bei der Arbeitsuche,
4. Verbesserung von prozeduralem Wissen und von Bewältigungskompetenz,
5. Transfer des Gelernten.

Das didaktische Konzept zur Erreichung dieser Lernziele beruhte auf einer zielgerichteten Anwendung moderner psychologischer Lerntheorien, aus denen vier generelle Lernstrategien abgeleitet wurden:

1. aktives Lernen,
2. Modell-Lernen,
3. abgestufte Exposition der Lerninhalte,
4. Verstärkung von Selbstwirksamkeit und Rollenübernahme.

Lernziele einerseits und Lernstrategien andererseits bildeten dann wiederum die Basis für die Festlegung der eigentlichen Interventionsaktivitäten, d.h. von konkreten Verhaltensweisen der Trainer gegenüber den Kursteilnehmern. Mit Hilfe einer Matrix aus den fünf Lernzielen und den vier Lernstrategien wurden 20 solcher effektiven Interventionsaktivitäten ausgearbeitet (Caplan/Vinokur/Price, 1997). Diese wiederum waren die Grundlage für genaue Handlungsanweisungen, die den Trainern für ihre Arbeit an die Hand gegeben wurden.

Eine solche detaillierte Festlegung konkreter Interventionskomponenten sollte möglichst durch eine Einschätzung der Interventionsstärke, also der Intensität oder der „Dosierung“

der Intervention, ergänzt werden, denn die meisten Interventionsinstrumente und -verfahren erzielen nur dann die intendierten Wirkungen, wenn sie häufig genug und/oder über einen ausreichenden Zeitraum hinweg eingesetzt werden (Sechrest et al., 1979).

Mögliche Dimensionen und Indikatoren der Interventionsstärke sind:

- Kontaktintensität (Dauer einer Beratungssitzung, Länge eines Informationstextes, Dauer einer Gruppendiskussion),
- Häufigkeit des Einsatzes einer Komponente pro Zeiteinheit (z.B. Zahl der Kursveranstaltungen während des Interventionszeitraums, Zahl von TV-Spot-Ausstrahlungen pro Monat),
- Zeitdauer (z.B. Laufzeit einer Kampagne oder eines Kursangebots).

Die Festlegung der Interventionsstärke in der Programmtheorie ist vor allem auch deshalb wichtig, weil sich daraus die besonders kostenrelevanten Entscheidungen über den Einsatz von Personal und anderen Ressourcen ableiten lassen.

## DEFINITION DER OUTCOMES

Die Definition der Outcomes ist häufig der schwierigste Schritt der Theoriebildung: Er kommt nicht ohne einen oft langwierigen Diskussionsprozeß aus, an dessen Ende Politiker, Geldgeber, Programmentwickler, Evaluationsforscher und manchmal auch noch weitere Stakeholder, wie Klienten oder Mitarbeiter, eine gemeinsam getragene Entscheidung über ein Programmresultat treffen müssen, das zum einen die Interessen aller Beteiligten berücksichtigt und zum anderen eindeutig, konkret und realitätsbezogen formuliert ist.

Diese Diskussion ist oft deshalb problematisch, weil sie häufig auf der Grundlage von relativ abstrakten Zielvorstellungen oder bloßen Absichtserklärungen über den Einsatz bestimmter Maßnahmen geführt wird. Diese Ziele oder Maßnahmen werden häufig in sehr allgemeinen Begriffen beschrieben, wie z.B.: „Die Entwicklung eines unterstützenden, familienzentrierten und kompetenzschaffenden Interventionssystems für Familien und Kinder.“ In dieser Situation ist es notwendig, die allgemeine „maßnahmenbezogene“ Begrifflichkeit in eine spezifische „klientenbezogene“ Aussage umzuformulieren, aus der erkennbar wird, bei wem was geschehen soll. In dem genannten Beispiel könnte eine solche klientenorientierte Formulierung etwa folgendermaßen lauten: „Eltern werden (als Ergebnis des Programms) auf adäquate Weise für ihre Kinder sorgen“ (Patton, 1997:157). Eine solche Aussage wäre Ausgangspunkt für den weiteren Abstimmungsprozeß zur Präzisierung des Outcomes. So muß weiter entschieden werden, um welche Art von Eltern es sich handelt, welche Tätigkeiten mit „um Kinder sorgen“ gemeint sind und was in diesem Zusammenhang „adäquat“ bedeutet. Die Abstimmung geht hin bis zur Operationalisierung der Outcomes, indem festgelegt wird, mit welchem Verfahren Daten über das jeweilige Outcome erhoben werden, nach welchem Kriterium entschieden wird, ob das

Ergebnis eingetreten ist, und auf welche Weise diese Entscheidung getroffen werden soll. Unter Umständen lassen sich auf jeden Fall zu erreichende Minimum-Werte oder das Maximum spezifizieren, das unter günstigen Umständen erreicht werden kann. Wichtig ist auch, den Zeitraum festzulegen, innerhalb dessen das Outcome eintreten soll. Generell kommt es also darauf an, die Outcomes so präzise wie möglich zu definieren.

Schwierigkeiten bereitet manchmal auch die adäquate inhaltliche Definition der Outcomes. Bei Gesundheitsförderungsprojekten liegt es nahe, vor allem *Health Outcomes* als Zielvariable des Programms auszuwählen, also Messungen von gesundheitlichen Merkmalen und Zuständen, wie sie in epidemiologischen Studien untersucht werden. Es gibt Projekte, bei denen Health Outcomes von der Intervention direkt beeinflusst werden, so daß sie dann auch als Indikator zur Messung der Wirksamkeit herangezogen werden können. Meistens entstehen jedoch beim Wirkungsnachweis mit Hilfe von Health Outcomes erhebliche praktische Probleme. Selbst beim „Michigan Jobs Program“, bei dem das Health Outcome „psychische Gesundheit“ in einer relativ engen Beziehung zur Intervention stand, mußte sich die Evaluationsstudie mit mehreren Meßzeitpunkten über zweieinhalb Jahre erstrecken, da die Wiederherstellung der psychischen Gesundheit das Ergebnis einer Arbeitsuche war, die in vielen Fällen erst nach einem längeren Zeitraum erfolgreich war. Oft ist die Wirkungskette zwischen Intervention und Health Outcomes jedoch noch erheblich länger, und letztere sind in der Regel durch eine ganze Reihe weiterer Variablen mitbeeinflusst, so daß sich die direkten Auswirkungen der Intervention kaum noch messen lassen. Es wird deshalb empfohlen, als Outcome eine Variable zu wählen, die zum einen unmittelbar durch das Gesundheitsförderungsprogramm bewirkt wird und zum anderen ihrerseits einen Einfluß auf das Health Outcome hat (McKinnon/Dwyer, 1993; Tones, 1992). Wie Intervention, Interventionsergebnis und Health Outcome kausal miteinander zusammenhängen, sollte in der Programmtheorie deutlich gemacht werden und möglichst durch vorliegende Forschungsergebnisse begründbar sein.

Als ein weiteres Kriterium für die inhaltliche Definition von Outcomes wird vorgeschlagen, daß sie einen Nutzen für die Klienten beschreibt, daß also durch die Intervention Einstellungen, Verhaltensweisen oder Lebensverhältnisse von Klienten verbessert oder verändert werden. Ein Outcome ist etwas, was Programmteilnehmer als Reaktion auf die Programmdienstleistungen tun, haben oder sind. Die bloße Teilnahme an einem Programm ist noch kein Outcome, und die Teilnehmerzahl ist keine Outcome-Variable (Plantz/Greenway/Hendricks, 1997).

Erleichtert wird die Definition der Outcomes durch die Problemanalyse: Outcome-Variable entsprechen den unerwünschten Zuständen, die durch die Intervention geändert werden sollen. Diese zu verändernden Variablen sind sehr viel einfacher und eindeutiger definierbar, wenn die Problemanalyse erkennen läßt, durch welche sozialen Einflußgrößen und Regelmäßigkeiten sie zustande kommen. Im Falle des „Michigan Jobs Program“ wurde durch die Problemanalyse empirisch bestätigt, daß die Verbesserung der durch den Arbeitsplatzverlust beeinträchtigten psychischen Gesundheit das zentrale Resultat des Pro-

gramms sein müßte. Zusätzlich war diese Variable durch die Voruntersuchungen bereits präzise definiert: Zum einen waren die Einzeldimensionen (Angst, Depression, psychosomatische Symptome) der Beeinträchtigungen der psychischen Gesundheit ausdifferenziert, und zum anderen waren auch schon die Indikatoren und Forschungstechniken ausgewählt, mit denen diese begrifflichen Dimensionen operationalisiert werden konnten.

Die Definition der Outcomes kann ihren Ausgangspunkt natürlich auch in der Programmtheorie haben. Sind die Interventionskomponenten festgelegt und die Wirkungsmechanismen beschrieben, lassen sich auch die möglichen Outcomes erkennen, die unter diesen Voraussetzungen prinzipiell erreicht werden können; dies ist um so eher möglich, je genauer die Programmtheorie ausgearbeitet ist. Aus der Programmtheorie des „Michigan Jobs Program“ läßt sich neben der psychischen Gesundheit ein weiteres Outcome des Programms ableiten: die Wiederbeschäftigung (innerhalb eines definierten Zeitraums), da dies der entscheidende Mechanismus der Interventionsstrategie ist, der die psychische Gesundheit erhält oder wiederherstellt.

Mit Hilfe der Programmtheorie läßt sich auch erkennen, inwieweit die Intervention nichtintendierte Wirkungen hervorrufen würde. Solche nicht erwünschten Outcomes und deren mögliche Ursachen zu spezifizieren ist deshalb wichtig, weil sich so eventuell Möglichkeiten erkennen lassen, wie unerwünschte Folgen der Intervention vermieden werden können. Wenn diese nicht zu vermeiden sind, ist es oft wichtig, daß sie in die Bewertung des Programms einbezogen werden können. Bei politisch kontroversen Programmen stehen nichtintendierte (unerwünschte) Wirkungen häufig sogar im Mittelpunkt der Diskussion. Die Evaluationsstudie hat dann oft die Aufgabe, zu klären, ob das Programm sich tatsächlich so negativ auswirkt, wie von manchen erwartet, oder nicht.

## BESCHREIBUNG DER WIRKUNGSMECHANISMEN

Bei der Beschreibung der Wirkungsmechanismen kommt es darauf an, den Prozeß aufzuzeigen, der die Beziehung zwischen den effektiven Interventionskomponenten und den Outcomes herstellt. Dies geschieht, indem zwischen den unabhängigen Variablen der Intervention und den abhängigen Outcome-Variablen eine Kausalkette ausgearbeitet und aufgezeichnet wird. Diese Kausalkette definiert möglichst alle Variablen, die als Verbindungsglieder zwischen den Komponenten und Dienstleistungen der Intervention und den Outcomes wirksam sind.

In Abbildung 2 ist eine solche Wirkungskette für das „Michigan Jobs Program“ skizziert: Die Abbildung zeigt zunächst eine Übersicht über die inhaltlichen Dimensionen der Intervention, d.h. über die verschiedenen Lernstrategien einerseits und die angestrebten Lernziele andererseits. (Die einzelnen Interventionskomponenten, die sich, wie oben beschrieben, aus den jeweiligen Kombinationen von Lernzielen und Lernstrategien ergeben, sind nicht im einzelnen dargestellt.) Die Intervention bewirkt insgesamt die Veränderung

einer ganzen Reihe von Einstellungs- und Selbstwahrnehmungsvariablen. Gelingt es, diese Einstellungsveränderungen zu erzeugen, hat dies Auswirkungen auf weitere intermediäre Variable, die das Verhalten bei der Stellensuche beschreiben. Die Stellensuche wird, so besagt die Programmtheorie, bei Vorhandensein der entsprechenden Einstellungen effektiver und mit mehr Beständigkeit verfolgt. Das Verhalten der Kursteilnehmer bei der Arbeitsuche wiederum führt mit höherer Wahrscheinlichkeit zum angestrebten Outcome einer Wiederbeschäftigung, und zwar möglichst an einem Arbeitsplatz, der der verlorenen Stelle mindestens gleichwertig ist. Als Folge der Wiederbeschäftigung war – gemäß den Voruntersuchungen zur Problemanalyse – als weiteres Outcome die Verbesserung der psychischen Gesundheit zu erwarten.

Daneben ist ein weiterer Wirkungsmechanismus denkbar: Die mit dem Kursprogramm erzeugten Einstellungsänderungen, wie z.B. die Verbesserung des Selbstwertgefühls, erleichtern die Bewältigung von Streßbelastungen, wenn infolge länger dauernder Arbeitssuche finanzielle Einbußen entstehen. Diese größere Bewältigungsfähigkeit bedeutet, daß die psychische Gesundheit weniger stark beeinträchtigt wird.

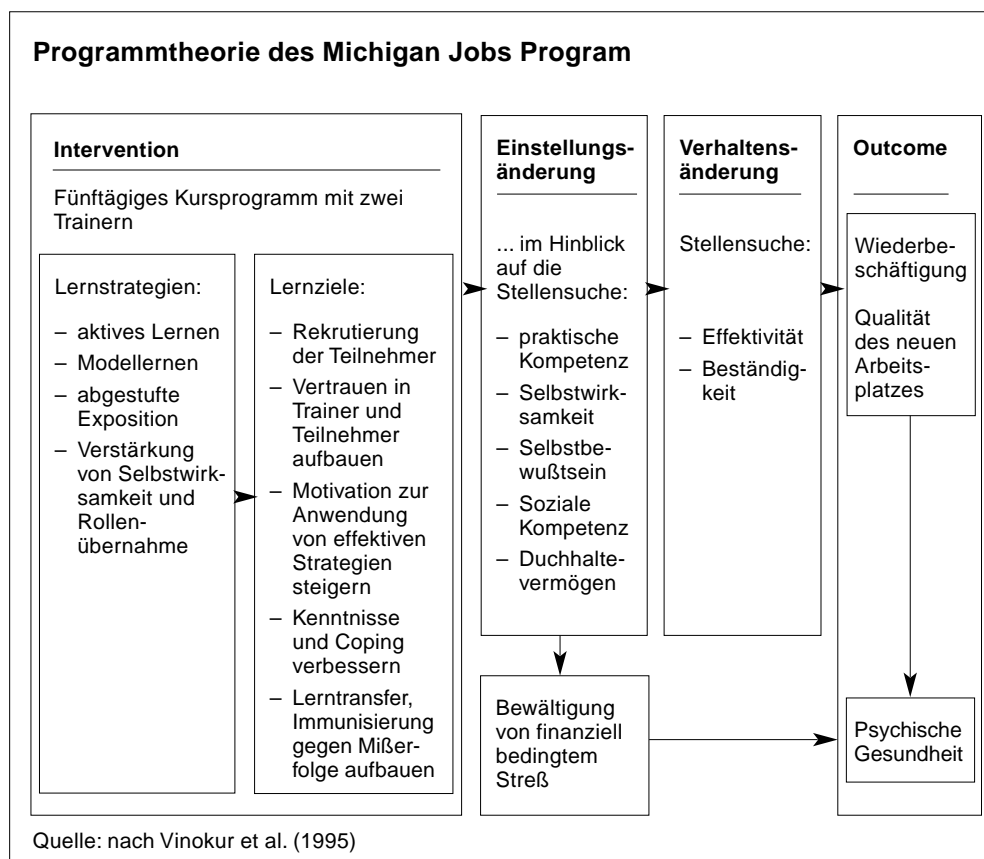


Abb. 2

Bei manchen Gesundheitsförderungsprogrammen gehören einige der zuvor beschriebenen Arbeitsschritte der Theorieentwicklung ohnehin zur Konzeptbildung und Programmplanung. Die Ausarbeitung einer Programmtheorie geht jedoch darüber hinaus: Es sollten alle Arbeitsschritte umfassend berücksichtigt werden, und die dabei erarbeiteten Begriffsklärungen, spezifizierten Aussagen und Zusammenhänge sollten so systematisiert und organisiert werden, daß als Ergebnis am Ende ein (schriftlich) ausformuliertes Modell des Programms vorliegt, nach dem sich Programmimplementation, Evaluation und nicht zuletzt auch die Qualitätssicherung des jeweiligen Projekts ausrichten können.

# 3

## 3.4. PROGRAMMTHEORIE UND QUALITÄTSSICHERUNG

Der praktische Nutzen einer guten Theorie, wie ihn der bereits zitierte Merksatz von Lewin verspricht, besteht vor allem darin, daß eine realitätsbezogene, empirisch orientierte Theorie ein Modell der praktischen Vorgänge des Programms ergibt. Mit diesem Modell läßt sich der Ablauf des Programms vor seinem eigentlichen Beginn gewissermaßen simulieren, d.h. es lassen sich realistische und spezifische Voraussagen über die genauen Umstände machen, unter denen ganz bestimmte Wirkungen erwartet werden können. Je früher ein solches Modell erarbeitet wird, um so besser kann es von praktischem Nutzen sein – vor allem für die Qualitätssicherung im Sinne der im vorangegangenen Kapitel (siehe *Abschnitt 2.4.*) aufgeführten fünf Qualitätsorientierungen:

- Die frühzeitige theoretische „Simulation“ des Programms hat *präventiven* Charakter im Sinne der Fehlervermeidung: Bevor viel investiert ist, läßt sich erkennen, ob das Programm überhaupt theoretisch funktionieren kann, ob es in dem vorgesehenen Kontext realisierbar ist oder ob möglicherweise Fehler entstehen können, die sich durch eine veränderte Planung bereits vor Beginn des Programms vermeiden lassen. Die explizite Ausformulierung der dem Interventionsansatz zugrundeliegenden Annahmen schafft die Voraussetzung dafür, logische Inkonsistenzen aufzudecken und darüber nachzudenken, ob überhaupt eine kausale Beziehung zwischen der beabsichtigten Intervention und deren Wirkungen besteht.
- Die Erarbeitung, Diskussion und Veränderung der Programmtheorie bedeutet mehr „Planungsqualität“. Im Prozeß von der Planung bis zur Überprüfung der Programm-ergebnisse findet Qualitätssicherung soweit wie möglich am Anfang dieses Prozesses statt. Damit ist die Forderung der Qualitätsmanagement-Theorie nach *Prozeborientierung* realisiert, nämlich daß die Qualität von Produkten oder Dienstleistungen vor allem durch eine hohe Qualität der Entstehungs- und Erarbeitungsprozesse erreicht werden sollte.

- Auch der Zyklus der *kontinuierlichen Verbesserung* beginnt bereits bei der Erarbeitung der Programmtheorie. Vor allem schafft eine präzise ausformulierte Programmtheorie die Möglichkeit, im voraus die Effektivität des Programms einzuschätzen und gegebenenfalls Verbesserungen zu planen, die zu mehr Wirksamkeit führen. Eine Möglichkeit zur Verbesserung der Wirksamkeit besteht darin, die Implementation so vorzubereiten, daß die Vorgaben der Programmtheorie mit möglichst großer „Wiedergabetreue“ realisiert werden. Vor allem aber kann die Stärke oder „Dosierung“ der Interventionsaktivitäten gesteigert werden, wenn zu vermuten ist, daß deren Wirksamkeit nicht ausreicht (Sechrest et al., 1979).
- Im Hinblick auf die *Stakeholder-Orientierung* bedeutet allein die größere Transparenz bereits einen wichtigen Schritt in Richtung Qualitätssicherung. Eine ausformulierte Programmtheorie ist eine umfassende und eindeutige Vorlage für die Verhandlungen über Outcomes und Ressourcen des Programms oder für die Information von Mitarbeitern oder Klienten.
- Vor allem aber ermöglicht eine Programmtheorie, frühzeitig im Sinne der *Outcome-Orientierung* und der finanziellen Verantwortlichkeit zu klären, ob das Programm überhaupt realisierbar ist: Sie ist nicht nur ein Instrument, mit dem sich voraussagen läßt, ob die erwarteten Outcomes theoretisch eintreten können, sondern sie kann auch als Grundlage für eine Kostenschätzung dienen. Beides sind Voraussetzungen für die Aufstellung einer Kosten-Nutzen-Bilanz, mit deren Hilfe sich einschätzen läßt, ob das Projekt mit den gegebenen Mitteln unter den gegebenen Umständen machbar ist. Und es läßt sich entscheiden, ob die zu erwartenden Outcomes und die verfügbaren Mittel in einem Verhältnis zueinander stehen, das eine Realisierung des Projekts rechtfertigt.

Für die *Evaluationsforschung* besteht die praktische Bedeutung von Programmtheorien vor allem darin, daß sie die Evaluationsforscher dazu anleiten, die richtigen Fragen zu stellen, deren Antworten eine dem Programm angemessene Bewertung und Qualitätsverbesserung erlauben. Eine konkrete Beschreibung der Zusammenhänge zwischen effektiven Interventionskomponenten, Wirkungsmechanismen und Outcomes ist gleichzeitig eine Zusammenstellung der möglichen Forschungshypothesen für das Evaluationsprojekt. Sie bestimmt die Auswahl der deskriptiven Variablen und Indikatoren und ermöglicht die treffsichere Auswahl derjenigen Personen, die über Programmablauf und Programmeffekte am zuverlässigsten Auskunft geben können; sie ist eine wichtige Leitlinie bei der Konstruktion der Erhebungsinstrumente und bei der Entscheidung über die Auswertungsverfahren.

## 3.5. ZUSAMMENFASSUNG

Die Evaluationsforschung orientiert sich mehr und mehr am Konzept der theoriegeleiteten Evaluation: Ziele, Plan und Methodik der Evaluation leiten sich aus einer *Programmtheorie* ab. Diese beschreibt die Zusammenhänge zwischen den effektiven Interventionskomponenten, Wirkungsmechanismen und den Outcomes.

Eine Programmtheorie ist somit ein systematisch ausgearbeitetes, realitätsbezogenes formuliertes Konzept des Programms. Dabei kommt es vor allem darauf an, die oft nur impliziten Handlungstheorien der Programmplaner explizit zu machen. Sie läßt sich damit als Handlungsanleitung verwenden, die die einzelnen Interventionsmaßnahmen festlegt und vorgibt, wie diese in dem jeweiligen sozialen Kontext implementiert werden sollen.

Wenn die Programmtheorie in der ersten Konzeptionsphase und unter Beteiligung der wichtigen Stakeholder erarbeitet wird, läßt sie sich für eine möglichst frühzeitige, „präventive“ Qualitätssicherung nutzen, vor allem, indem bei der Erarbeitung Fragen über die Realisierbarkeit des Programms gestellt werden:

- Ist das Programm logisch möglich, d.h. existiert eine Kausalbeziehung oder ein Wirkungsmechanismus zwischen Intervention und Outcome?
- Sind die Interventionsmaßnahmen ausreichend effektiv, wirken die Interventionen stark genug, um die angestrebten Outcomes hervorzubringen?
- Ist das Programm bei den gegebenen Ressourcen durchführbar?

Aus den Antworten lassen sich erste Qualitätsverbesserungen ableiten:

- Planung von „realistischen“ Interventionen, Wirkungsmechanismen und Outcomes,
- Verstärkung der Interventionsmaßnahmen,
- Abstimmung von Interventionsstrategie und Ressourcen, gegebenenfalls Einstellung des Vorhabens.

Eine frühzeitige Erarbeitung der Programmtheorie ermöglicht also, daß der Zyklus der kontinuierlichen Verbesserung bereits in der Planungsphase beginnt.

Der erste Schritt der Theoriebildung ist in der Regel eine Problemanalyse. Eine Problemanalyse für Programme der Gesundheitsförderung beschreibt die nichtmedizinischen Ursachen oder Entstehungsprozesse gesundheitlicher Probleme. Sie präzisiert die unerwünschten Outcomes, definiert die (Ziel-)Gruppen, in denen das Problem auftritt, und spezifiziert die Mechanismen und intermediären Variablen zwischen den grundlegenden sozialen Ursachen und dem Problem. Dies erfolgt möglichst so detailliert, daß sich effektive und realisierbare Interventionsmöglichkeiten erkennen lassen.

Die Programmtheorie sollte die Interventionen des Programms möglichst detailliert beschreiben und vor allem die eigentlich effektiven Interventionskomponenten herausarbeiten. Dazu gehört auch eine Einschätzung der Effektstärke, d.h. wie stark die Effekte der Intervention sein müssen, um die beabsichtigten Outcomes hervorzubringen.

Die Programmtheorie gibt darüber hinaus an, welche Outcomes erreicht werden sollen, aber auch, welche als nichtintendierte Nebenwirkungen der Intervention anzusehen sind. Die Outcomes sollten so präzise wie möglich definiert werden, indem angegeben wird, welche Einstellungen, Verhaltensweisen oder Lebensverhältnisse von Klienten verbessert oder verändert werden sollen, in welchem Zeitraum dies geschehen kann und wie die Outcomes operationalisiert und gemessen werden können.

Vor allem beschreibt die Programmtheorie die Wirkungsmechanismen – den Prozeß, der Schritt für Schritt die Wirkungen der effektiven Interventionskomponenten auf die Outcomes überträgt. Dies geschieht, indem zwischen den unabhängigen Variablen der Intervention und den abhängigen Outcome-Variablen die intermediären Variablen definiert werden, die als Verbindungsglieder zwischen beiden wirksam sind.

**MONITORING VON  
IMPLEMENTATIONSPROZESSEN**



# 4.

## MONITORING VON IMPLEMENTATIONSPROZESSEN

Es gibt zwei Gründe, weshalb Programme scheitern können<sup>1</sup>:

1. Der Wirkungsmechanismus funktioniert nicht so, wie es die Programmtheorie voraussagt (Konzeptionsfehler).
2. Die Art und Weise, wie die Konzeption in einem bestimmten sozialen Kontext realisiert wurde, hat nicht dazu geführt, daß ein (möglicherweise funktionierender) Wirkungsmechanismus überhaupt in Gang gesetzt wurde (Implementationsfehler).

Daraus ergeben sich für die Programmevaluation zwei Aufgaben. Sie muß zum einen den Implementationsprozeß untersuchen und Informationen sammeln, mit denen sich dokumentieren läßt, ob und wie das Programmkonzept in einem konkreten sozialen Kontext realisiert wurde. Ist nachgewiesen, daß der Implementationsprozeß funktioniert, ist damit die notwendige Voraussetzung für den zweiten Teil der Evaluationsstudie erfüllt, also für die Wirkungsanalyse. Diese muß untersuchen, ob der Wirkungsmechanismus zu den vorausgesagten Outcomes geführt hat. In diesem Kapitel wird die Evaluation des Implementationsprozesses behandelt.

Zunächst soll noch einmal darauf hingewiesen werden, daß hier nicht dafür plädiert wird, sich ausschließlich mit dem Implementationsprozeß zu beschäftigen und auf eine Evaluation der Ergebnisse zu verzichten. Beide Aufgaben der Programmevaluation werden als gleichermaßen wichtig angesehen.

### 4.1. MONITORING

Ein für die Evaluation des Implementationsprozesses besonders gut geeignetes Instrument ist ein Monitoring. Dies bedeutet die laufende Beobachtung des Programmverlaufs mit Hilfe eines systematisch zusammengestellten Satzes von Daten, die regelmäßig und routinemäßig erhoben werden. Dazu gehören beispielsweise Teilnehmerzahlen, Daten über die Teilnehmer (soziodemographische Merkmale, Wahrnehmung und Akzeptanz von Interventionsleistungen), Daten über Mitarbeiteraktivitäten (Teilnehmerkontakte, Akzeptanz und Befolgung von Implementationsvorgaben) oder Programmkosten.

---

<sup>1</sup> Auf diese Unterscheidung hat zuerst E. A. Suchman (1967) hingewiesen.

Wenn die erhobenen Daten während des Implementationsprozesses laufend aufbereitet, ausgewertet und rückgekoppelt werden, entsteht auf diese Weise ein programmbezogenes Management-Informationssystem.

Die regelmäßig abrufbaren Daten dieses Informationssystems lassen sich dazu verwenden, frühzeitig Abweichungen und Schwierigkeiten im Implementationsprozeß zu entdecken, um so möglichst noch während des laufenden Programms für Lösungen zur Qualitätsverbesserung sorgen zu können.

Durch ein Monitoring des Implementationsprozesses läßt sich vor allem folgendes erreichen:

- Es stellt Verlaufsdaten über den Ablauf der Intervention zur Verfügung: zu Inputs, Dienstleistungen, Teilnehmerreaktionen.
- Es kontrolliert die Übertragungsqualität, also die Genauigkeit, mit der die Programmtheorie in die Realität übertragen wird.
- Es dokumentiert Veränderungen des ursprünglichen Konzepts, die im Verlauf des Implementationsprozesses notwendig wurden.
- Seine Ergebnisse ermöglichen eine bessere Interpretation und Einschätzung der gemessenen Programmwirkungen.
- Es erhöht die Qualität der Daten über den Implementationsprozeß: Zum einen erzielt die Datenerhebung in größtmöglicher zeitlicher Nähe zum Interventionsgeschehen eine hohe Meßgenauigkeit, zum anderen erlauben Wiederholungsmessungen eine bessere Einschätzung von Zufallsvariationen im Zeitverlauf und damit eine Verbesserung der Zuverlässigkeit von Daten und Ergebnissen.

Die Planung eines Monitoringsystems muß eine ganze Reihe von inhaltlichen wie auch erhebungs- und aufbereitungstechnischen Anforderungen berücksichtigen, deren Zielsetzungen sich teilweise widersprechen: So muß einerseits eine möglichst umfassende Auswahl inhaltlich relevanter Indikatoren getroffen werden, andererseits darf der Erhebungsaufwand weder die Mitarbeiter und Teilnehmer noch das Budget übermäßig beanspruchen. Es muß zum einen dafür gesorgt werden, daß durch wiederholte Datenerhebungen Veränderungen im Interventionsgeschehen möglichst zuverlässig wiedergegeben werden, zum anderen muß vermieden werden, daß sich durch die häufige Datenerhebung während des Implementationsprozesses auf seiten der befragten Teilnehmer oder Mitarbeiter die Wahrnehmung des Interventionsgeschehens so verändert, daß sich dies verzerrend auf nachfolgende Outcome-Messungen auswirkt. Erleichtert wird die Planung eines Monitoringsystems durch eine inhaltsreiche und detailliert ausgearbeitete Programmtheorie (siehe *Kapitel 3*).

In inhaltlicher Hinsicht bedeutet die Evaluation des Implementationsprozesses die Beantwortung einiger spezifischer Forschungsfragen, die – je nach Programm – einmal stärker und einmal weniger stark berücksichtigt werden müssen. In den folgenden Abschnitten werden diese zentralen Forschungsfragen der Implementationsevaluation jeweils kurz beschrieben.

## 4.2. ANALYSE DES PROGRAMMKONTEXTS

Noch vor Beginn der eigentlichen Implementation sollte eine sorgfältige Überprüfung des sozialen Kontexts erfolgen, in dem das Programm durchgeführt wird. Hierbei wird untersucht, inwieweit die Programmkonzeption mit der sozialen Realität des Kontexts vereinbar ist, in dem die Implementation stattfinden soll. Es ist eine ganze Reihe von Umständen denkbar, die eine Implementation erschweren – sei es, daß die Qualifikation der Projektmitarbeiter für bestimmte Interventionsaktivitäten nicht ausreicht, sei es, daß sich die Interventionen nicht in die Lebensweise der Klienten einfügen lassen, oder sei es, daß sich die öffentliche Meinung gegen das Projekt gerichtet hat, und vieles andere mehr.

Zur Systematisierung möglicher Kontexteinflüsse mag eine Checkliste beitragen, die Evaluationsforscher als Grundlage für die Evaluation von Implementationsprozessen vorgeschlagen haben (Chen/Rossi, 1983; Chen, 1990):

### **Zielgruppen (Klienten, Teilnehmer, Nutzer):**

- soziale Merkmale (soziale Situation, Lebensweise),
- Wertvorstellungen, Einstellungen,
- Compliance (Kooperationsbereitschaft),
- Aufnahmebereitschaft und -fähigkeit,
- Definition und Auswahlregeln für die Abgrenzung von Zielgruppen.

### **Mitarbeiter des Projektteams:**

- Qualifikation,
- spezifische Qualifizierung für das Projekt,
- Motivation,
- Fähigkeit zum Kontakt (Rapport) mit den Teilnehmern.

### **Umstände der Leistungserbringung:**

- technische Rahmenbedingungen,
- Organisation des Settings,
- Akzeptanz der Intervention im Setting.

### **Organisation der Implementation:**

- Flexibilität/Rigidität der Organisationsstruktur,
- Routinisierung von Arbeitsvorgängen,
- Qualitätskontrolle von Arbeitsvorgängen,
- Kapazität.

### **Beziehung zu anderen Organisationen:**

- kooperierende Organisationen,

- konkurrierende Organisationen,
- Abstimmungsprozesse und -verfahren.

**Soziale Umwelt der Teilnehmer:**

- soziale Beziehungen der Zielgruppen oder Klienten,
- negative Einflüsse auf die Teilnehmer,
- soziale Unterstützung der Teilnehmer.

**Makro-Kontext (in der Region oder im Mitgliedsland):**

- nationale oder regionale Rahmenbedingungen,
- politische Einflüsse,
- soziales Klima.

Es ist denkbar, eine vergleichbare Checkliste zu entwickeln, die genauer auf die Probleme und Interventionsstrategien von Gesundheitsförderungsprogrammen zugeschnitten ist.

Auf der Grundlage einer solchen Checkliste kann erarbeitet werden, welche konkreten Ausprägungen die dort aufgeführten Dimensionen in dem jeweiligen Kontext haben, und es lassen sich Hypothesen formulieren, welche Kontextbedingungen in welcher Weise die Intervention erschweren oder erleichtern. Diese Hypothesen können in einem eigenständigen Forschungsprojekt überprüft werden. Bei den meisten Projekten reichen jedoch vorhandene Informationen oder leicht zugängliche Informationsquellen aus. Solche Informationsquellen wären:

- Organisationswissen und Arbeitserfahrungen des Projektteams über Klienten und Settings,
- sozialwissenschaftliches Expertenwissen über kontexttypische soziale Strukturen, Prozesse und Regelmäßigkeiten,
- Analysen von amtlichen Daten und Verwaltungsdokumenten,
- Interviews mit typischen Klienten, Multiplikatoren und anderen Stakeholdern, die über bestimmte Ausschnitte des sozialen Kontexts besonders gut informiert sind,
- Ergebnisberichte von Evaluationsstudien mit ähnlichen Programmen in vergleichbaren Kontexten.

Manchmal läßt sich die zukünftige Entwicklung bestimmter Kontextbedingungen nur schwer voraussagen. Oft ist es notwendig, während des Programmverlaufs den Einfluß einer bestimmten Kontext-Variablen zu kontrollieren: In diesen Fällen wird man geeignete Indikatoren entwickeln und deren Veränderungen im Monitoring beobachten.

## 4.3. EVALUATION VON PROGRAMMKOMPONENTEN

Eine Evaluation einzelner Programmkomponenten sollte ebenfalls möglichst vor Beginn der eigentlichen Implementationsphase vorgenommen werden. Dies kann vor allem dann notwendig sein, wenn das Programm deutlich abgrenzbare Programmteile umfaßt, wie z.B. TV-Spots, Broschüren oder andere Medien, aber auch spezielle Beratungsformen oder andere Dienstleistungsangebote, die relativ unabhängig vom übrigen Programm geplant und entwickelt wurden.

Komponenten-Evaluationen können im vorgesehenen Kontext durchgeführt werden, um auch hier wieder die Vereinbarkeit mit den spezifischen Kontextbedingungen zu überprüfen (z.B. die Verständlichkeit einzelner Medien oder die Akzeptanz bei bestimmten Klientengruppen). Oder es läßt sich untersuchen – gegebenenfalls in Form von Laborexperimenten –, ob mit anderen Interventionsaktivitäten besondere Wechselwirkungen bestehen, z.B. ob sich die Informationswirkungen mehrerer unterschiedlicher Medien gegenseitig verstärken oder aufheben. Auch im Rahmen eines Monitorings lassen sich Komponenten-Evaluationen durchführen, z.B. wenn im Rahmen einer Lernstrategie eine bestimmte Argumentationsweise der Trainer als Einzelelement betrachtet und beobachtet wird, ob sie sich im Zeitverlauf verändert.

Die eigenständige Evaluation von einzelnen Programmkomponenten wird vor allem deshalb vorgeschlagen, weil ihre Ergebnisse vergleichsweise leicht verallgemeinert werden können, so daß sich aus der Synthese dieser Ergebnisse Standards für die Konstruktion und Anwendung der jeweiligen Komponenten erarbeiten lassen (Bickmann, 1985). Ansätze hierfür gibt es bei der Gestaltung von TV-Spots zur Gesundheitsaufklärung.

## 4.4. KONTROLLE DER ÜBERTRAGUNGSQUALITÄT

Bei der Kontrolle der Übertragungsqualität geht es um die Frage, inwieweit eine konzeptionsgetreue Implementation stattfindet, d.h. ob die Konzeption in diejenigen Interventionsaktivitäten übersetzt wird, mit denen sich die vorgesehenen Wirkungsmechanismen auslösen lassen (Sechrest et al., 1979).

Dieser Übersetzungsvorgang wird erheblich erleichtert, wenn ein Implementationshandbuch vorliegt, in dem die einzelnen Interventionsaktivitäten beschrieben sind, und das die Mitarbeiter anleitet, wie sie sich in welchen Situationen und bei welchen Klienten verhalten sollen. Ein solches Handbuch bietet zudem die Möglichkeit, die Übertragungs-

qualität bereits durch eine programm- und kontextbezogene Qualifizierung der Projektmitarbeiter zu verbessern. Es erleichtert jedoch auch, Abweichungen und Fehler festzustellen, die bei der Realisierung der Konzeption in einer bestimmten sozialen Umwelt entstehen, und es ermöglicht die Erarbeitung projektbezogener Kriterien für Konzeptionstreue, die im Monitoring herangezogen werden können.

## 4.5. MONITORING UND PROGRAMMEVALUATION

Das Monitoring läßt sich vor allem auch zur systematischen Informationsbeschaffung für die Analyse der Programmwirkungen einsetzen, also zur Überprüfung von Hypothesen über die Zusammenhänge zwischen effektiven Interventionskomponenten, Wirkungsmechanismen und Outcome-Variablen. Die unabhängigen Variablen zur Messung der Interventionsinputs lassen sich z.B. aus Angaben zur Häufigkeit, Reichweite oder Dauer der eingesetzten Interventionskomponenten konstruieren. Für Variablen, die die Wirkungsmechanismen abbilden, eignen sich insbesondere Befragungs- oder Beobachtungsdaten, die an den Schnittstellen der Kommunikationsabläufe des Programms erhoben werden. Solche Schnittstellen entstehen z.B. bei der Zusammenarbeit einzelner Organisationseinheiten des Programms, in Situationen, in denen Mitarbeiter und Klienten etwas miteinander zu tun haben, bei der Nutzung von Medien des Programms oder durch Kontakte von Klienten mit Menschen außerhalb des Programms. Dort kann jeweils eine ganze Reihe von Voraussetzungen für die Wirksamkeit überprüft werden, z.B. Wahrnehmung, Verständnis und Akzeptanz von Kommunikationsinhalten.

Auch die Outcome-Variablen lassen sich in das Monitoringsystem einbauen. Dies setzt voraus, daß die Erhebung der entsprechenden Daten möglichst frühzeitig geplant, organisiert und im Verlauf des gesamten Programms stattfinden kann. Auf diese Weise lassen sich die Veränderungen der Outcome-Variablen zeitlich parallel zu den Interventionsinputs beobachten.

Die Messung dieser analytischen Variablen bedeutet nicht unbedingt, daß ein zusätzlicher Aufwand für die Datenerhebung notwendig ist. Vielfach können hierfür Daten genutzt werden, die für die Beobachtung des Implementationsprozesses oder für das Projektmanagement ohnehin erhoben werden.

## 4.6. MONITORING UND QUALITÄTSSICHERUNG

Ein Monitoringsystem ist ein wichtiges Instrument für die Qualitätssicherung. So bedeuten frühzeitig beginnende Datensammlungen und Untersuchungen zum sozialen Kontext, zu einzelnen Programmkomponenten und zur Übertragungsqualität eine weitere Chance für die Verbesserung eines Programms, bevor es vollständig implementiert ist. Durch diese Verbesserungen läßt sich frühzeitig die Effektivität des Programms steigern: Interventionsinputs, die sich anfangs als zu schwach erwiesen haben, lassen sich verstärken, die Übertragungsqualität läßt sich präzisieren, und man kann versuchen, verzerrende und konfundierende Einflüsse des Kontexts zu kontrollieren.

Die kontinuierliche Datenerhebung und Beobachtung des Programmverlaufs machen das Programm-Monitoring mit bestimmten zentralen Vorstellungen des Total Quality Managements vergleichbar. Dieses will Qualität vor allem dadurch erreichen, daß es die Arbeitsprozesse ständig beobachten, gegebenenfalls genauer einstellen oder umstellen läßt, um auf diese Weise die Qualität der Endprodukte zu verbessern oder zu erhalten. Ähnlich ließe sich auch das Monitoring von Gesundheitsförderungsprogrammen zur Qualitätskontrolle einsetzen, indem anhand der Beobachtungsergebnisse wichtiger Outcome-Variablen oder intermediärer Variablen regelmäßig geprüft wird, ob deren Verlauf bestimmten vorher festgelegten Erwartungen entspricht.

Im Total Quality Management wird diese Art der Qualitätskontrolle durch ein Qualitätsbewußtsein der Mitarbeiter unterstützt, das darauf ausgerichtet ist, nach qualitätsmindernden Abweichungen und Fehlern in den Prozeßabläufen zu suchen. Um diese festzuhalten und zu analysieren, wird eine Reihe von Informationsinstrumenten verwendet, wie z.B. Fehlersammlisten, Säulendiagramme zur Darstellung von Fehlerhäufigkeiten, einfache statistische Auswertungstechniken oder Ursache-Wirkungs-Diagramme, mit deren Hilfe die Erkenntnisse aus der Fehleranalyse in Programmlösungen und Verbesserungen umgesetzt werden können. Diese qualitätsorientierte Prozeßbeobachtung durch die Mitarbeiter hat wiederum Ähnlichkeiten mit dem Evaluationsinstrument der Selbstevaluation. Dabei führen die Mitarbeiter des Programms selber eine Erhebung, Auswertung und Aufbereitung von Daten durch, indem sie z.B. ihre eigenen Arbeitsprozesse oder Interaktionen mit den Klienten beobachten und dokumentieren und diese Ergebnisse an Programmzielen messen und bewerten. Der besondere Vorteil der Selbstevaluation besteht darin, daß sie bei den Mitarbeiterinnen und Mitarbeitern eines Projekts die Aufmerksamkeit für Schwierigkeiten und Probleme der Implementation weckt und die Motivation erhöht, über Problemlösungen nachzudenken. Voraussetzung ist eine intensive Schulung und Supervision. Vor allem muß sichergestellt sein, daß die Selbstevaluation mit den zeitlichen, kognitiven und emotionalen Anforderungen der Arbeit vereinbar ist.<sup>2</sup>

---

2 Zu einem Projekt zur Nutzung der Selbstevaluation in der Kinder- und Jugendhilfe siehe M. Heiner (1996).

Des Weiteren kann das Monitoring als Informationsgrundlage für andere Instrumente des Qualitätsmanagements genutzt werden, z.B. für die Erstellung von Audits, also für Tätigkeitsberichte, die Tätigkeiten und Leistungen auflisten und eine Einschätzung ermöglichen sollen, inwieweit die Qualitätsanforderungen erfüllt wurden. Audits können wiederum als Grundlage für eine Qualitätszertifizierung dienen.<sup>3</sup>

## 4.7. ZUSAMMENFASSUNG

Die Evaluation des Implementationsprozesses dokumentiert den Programmverlauf und weist nach, daß das Programm unter den Bedingungen, unter denen es eingesetzt wurde, realisiert werden konnte.

Ein besonders geeignetes Instrument für diese Aufgabe ist ein Monitoring, d.h. die kontinuierliche Beobachtung des Programmverlaufs mit Hilfe eines systematisch zusammengestellten Satzes von Daten, die regelmäßig und routinemäßig erhoben werden. Neben den Daten zur administrativen Abwicklung sind dies Informationen zum Einfluß von Kontextfaktoren, zur Wirksamkeit und Leistung einzelner Programmkomponenten und zur konzeptgetreuen Übertragungsqualität.

Das Monitoring eignet sich des Weiteren als Erhebungsinstrument für die Variablen, die für die Wirkungsanalyse benötigt werden: also für die Stärke der effektiven Interventionskomponenten, für die Variablen zur Wirkungsweise des Programms sowie für Outcome-Variable, die sich bereits während des Implementationsprozesses messen lassen.

Entscheidend ist, daß die erhobenen Daten während des Implementationsprozesses laufend aufbereitet und ausgewertet werden. Diese regelmäßig abrufbaren Daten lassen sich dazu verwenden, frühzeitig Abweichungen und Schwierigkeiten im Implementationsprozeß zu entdecken, damit möglichst noch während des laufenden Programms für Lösungen zur Qualitätsverbesserung gesorgt werden kann.

Durch diese Verbesserungen läßt sich frühzeitig die Effektivität des Programms steigern: Interventionsinputs, die sich anfangs als zu schwach erwiesen haben, lassen sich verstärken, die Übertragungsqualität läßt sich präzisieren, und man kann versuchen, verzerrende und konfundierende Einflüsse des Kontexts zu kontrollieren.

---

<sup>3</sup> Siehe hierzu das internationale Normenwerk zu Qualitätssicherungssystemen DIN ISO 9000 bis 9004.



**WIRKUNGSANALYSE**

5

# 5.

## WIRKUNGSANALYSE

Die Wirkungsanalyse überprüft die Programmeffekte. Sie stellt fest, ob die erwarteten Outcomes tatsächlich eingetroffen sind, und untersucht, inwieweit sie als Ergebnis der Intervention und ihrer Wirkungsmechanismen angesehen werden können.

Die Messung und Analyse der Programmwirkungen („summative“ Evaluation) sind ebenso wichtig wie die Untersuchung von Implementationsprozessen („formative“ Evaluation). Darauf soll auch in diesem Zusammenhang noch einmal hingewiesen werden, weil in den letzten Jahren immer wieder die Ansicht vertreten wurde, eine summative Evaluation sei nicht unbedingt notwendig, solange eine formative Evaluationsstudie stattfinde (Sechrest/Figueroa, 1993). Tatsächlich wird die Wichtigkeit der formativen Evaluation in der Evaluationsliteratur immer wieder besonders hervorgehoben, und auch in der Qualitätsdiskussion wird auf die besonders große Bedeutung der Qualitätsmessung von organisatorischen Abläufen und Prozessen hingewiesen (Brook/McGlynn/Cleary, 1996). In erster Linie sollen diese Hinweise jedoch verdeutlichen, daß die Prozeßqualität eine besondere Rolle für die Qualität des Produkts spielt und deshalb diese Herstellungsprozesse bei Qualitätsmessungen oder Evaluationen nicht außer acht gelassen werden sollten. Es muß jedoch trotzdem belegt werden, ob diese Prozesse die erwarteten Resultate erbracht haben oder nicht.

Für die Qualitätssicherung ist die Wirkungsanalyse von besonderer Bedeutung:

- Sie untersucht, ob das Programm effektiv war (Outcome-Orientierung).
- Sie untersucht weiter, ob der Wirkungsmechanismus so funktioniert hat, wie aufgrund des Programmkonzepts zu erwarten war, oder ob das Programm in seiner Wirkungsweise verbessert werden muß (ständige Verbesserung).
- Sie versucht die Frage zu beantworten, ob Qualität im Sinne der Klientenorientierung erreicht wurde.
- Mit der Beantwortung dieser Fragen schafft sie die Grundlage für die Entscheidung, ob sich die eingesetzten Mittel rechtfertigen lassen.

Dieses Arbeitspapier muß sich auf einen kurzen Überblick über die Planung von Wirkungsanalysen beschränken. Für die Detailplanung steht ein breiter Fundus von Forschungstechniken, Kontrollstrategien sowie von analytischen und statistischen Verfahren zur Verfügung, die hier nicht im einzelnen beschrieben werden können. Zur praktischen Evaluationsforschung existiert eine umfangreiche Literatur, und es dürfte in den Mitgliedstaaten der EU überall ausreichende Beratungsmöglichkeiten zu den Methoden und Forschungstechniken der empirischen Sozial- und Evaluationsforschung geben.

Vor allem soll hier noch einmal verdeutlicht werden, daß Evaluationsplanung, wie die Planung anderer Projekte auch, ein komplizierter Entscheidungsprozeß ist, der eine Viel-

zahl von Bedingungen berücksichtigen muß. Diese Komplexität hat zum einen den Vorteil, daß im Prinzip „maßgeschneiderte“ Wirkungsanalysen möglich sind, die auf die besonderen Aufgaben und Bedingungen sowie auf die jeweiligen Stakeholder-Interessen abgestellt sind. Zum anderen gibt es aber auch enorme Fehlermöglichkeiten, mit dem Risiko, falsche Schlußfolgerungen über die Qualität der zu untersuchenden Programme zu ziehen.

## 5.1. PLANUNG VON WIRKUNGSANALYSEN

Für die praktische Realisierung einer Wirkungsanalyse muß ein dem jeweiligen Programm und seinen Kontextbedingungen entsprechendes spezifisches Untersuchungskonzept entwickelt werden. Darin ist festgelegt, was wie untersucht werden soll.

In einem Abstimmungsprozeß mit den Stakeholdern werden die generellen Evaluationsziele bestimmt, es werden die zu messenden Outcome-Variablen ausgewählt und Forschungshypothesen über die Zusammenhänge zwischen effektiven Interventionskomponenten, Wirkungsmechanismen und Outcomes formuliert.

Diese Untersuchungsziele und Forschungsfragen bilden wiederum die Vorgaben für die Entscheidung, mit welchen Methoden und Forschungstechniken die einzelnen Fragestellungen untersucht werden sollen. Grundsätzlich müssen dabei zwei methodische Probleme gelöst werden: Es müssen Verfahren zur Messung der Outcomes sowie der übrigen Variablen gefunden oder konstruiert werden, und es muß ein Forschungsdesign festgelegt werden, das die Datenerhebung und Datenanalyse steuert.

### MESSVERFAHREN

Bei der Konstruktion von Meßverfahren müssen wiederum die folgenden vier methodischen Probleme gelöst werden:

1. die Auswahl der deskriptiven Variablen, d.h. der Indikatoren, welche die in den Forschungshypothesen verwendeten Begriffe mit möglichst großer inhaltlicher Gültigkeit wiedergeben,
2. die Auswahl oder Konstruktion von Erhebungsinstrumenten,
3. die Auswahl von Untersuchungseinheiten, bei denen die Informationen zur Messung der Variablen gewonnen werden können,
4. die Festlegung von Verfahren der Datenanalyse.

## FORSCHUNGSDESIGN

Die Aufgabe des Forschungsdesigns besteht in der Organisation der Daten: Welche Daten werden wann und bei wem erhoben, welche Daten sollen mit welchen statistischen Verfahren mit welchen anderen Daten verglichen werden? Ziel des Forschungsdesigns ist es, die Gültigkeit von Schlußfolgerungen aus den Forschungsergebnissen zu gewährleisten.

Bei der Evaluation von Sozial- und Gesundheitsprogrammen muß die Analyse der Daten in der Regel die Schlußfolgerung überprüfen, daß sich die Outcome-Variablen nach Einwirkung der Intervention verändert haben. Solche Veränderungen werden am besten mit Hilfe eines Längsschnittdesigns untersucht, bei dem die Outcome-Variablen mindestens einmal vor der Intervention gemessen werden und mindestens einmal zu einem späteren Zeitpunkt, für den die Wirkung der Intervention erwartet wird (Vorher-Nachher-Design).

Zum Nachweis, daß die Veränderung der Outcome-Variablen kausal durch die Intervention hervorgerufen wurde, dienen Forschungsdesigns, die auf dem methodischen Prinzip der *Kontrolle* beruhen, also darauf, daß die Variation bestimmter Variablen bewußt herbeigeführt, eingeschränkt oder konstant gehalten wird.

Um die bei Kausalanalysen notwendige Klarheit darüber zu schaffen, daß die unabhängige Kausalvariable der abhängigen Outcome-Variablen zeitlich vorgeht, wird erstere kontrolliert, indem sie „experimentell“ hergestellt wird. In diesem (methodischen) Sinne werden Programme als „Experimente“ betrachtet: Die Intervention ist die zielgerichtet und systematisch herbeigeführte unabhängige Variable.

## KONTROLLE EXTERNER STÖRFAKTOREN

Zur Vermeidung von Fehlschlüssen muß darüber hinaus der Einfluß externer Störfaktoren kontrolliert werden: Veränderungen der Outcome-Variablen in Vorher-Nachher-Anordnungen lassen sich nicht ohne weiteres als Nachweis von Interventionseffekten interpretieren, wenn nicht ausgeschlossen werden kann, daß andere, externe Ereignisse zwischen Vorher- und Nachher-Messung einen Einfluß ausgeübt haben. So könnte beispielsweise eine öffentliche Diskussion über das Rauchen möglicherweise die Nachher-Messungen von Ergebnissen eines Nichtraucherurses beeinflussen, so daß nicht eindeutig ist, ob die gemessenen Effekte auf das Nichtrauchertraining oder auf die öffentliche Diskussion zurückzuführen sind. Solche externen Ereignisse oder Entwicklungen, die zwischen Vorher- und Nachher-Messungen auftreten, werden als Zeiteinflüsse (*history*) bezeichnet.

Weitere Einflußfaktoren auf die Gültigkeit von Längsschnittuntersuchungen sind – von der Intervention unabhängige – Veränderungen in der Persönlichkeit der Klienten (*maturation*) oder statistische Regressionseffekte.

Darüber hinaus können externe Störfaktoren durch die verwendeten Forschungsverfahren entstehen: Zum einen sind dies bei den Befragten wirkende Meßeffekte, wenn die Vorher-Messung die nachfolgenden Messungen beeinflusst (*testing*). Zum anderen können sich zwischen Vorher- und Nachher-Messung die Meßinstrumente verändern; dies wäre z.B. der Fall, wenn die Mitarbeiter bei einer Selbstevaluation durch ihre Beobachtungen des Klientenverhaltens dieses im Zeitverlauf zunehmend in eine für die Bewertung des Programms vorteilhafte Richtung deuten würden (*instrumentation*).<sup>1</sup>

Externe Störfaktoren können mit Hilfe von Kontrollgruppen kontrolliert werden, bei denen keine Intervention stattfindet, aber ebenfalls Vorher- und Nachher-Messungen durchgeführt werden. Falls externe Störfaktoren existieren, tritt ihre Wirkung in beiden Gruppen auf, und der Einfluß der Intervention kann durch den Vergleich von Interventionsgruppen- und Kontrollgruppenergebnissen nachgewiesen werden.

Die Einrichtung von Kontrollgruppen bringt jedoch weitere Interpretationsprobleme mit sich: So erschweren Unterschiede in der Zusammensetzung von Interventions- und Kontrollgruppe eine eindeutige Interpretation der Ergebnisse. Dies wiederum kann jedoch durch eine zufallsgesteuerte Verteilung von Versuchspersonen auf die beiden Gruppen kontrolliert werden.

In den Sozialwissenschaften wird das randomisierte Kontrollgruppen-Experiment mit Vorher-Nachher-Messung häufig als das Instrument angesehen, mit dem sich Kausalwirkungen am ehesten nachweisen lassen. Es vereint mehrere Kontrollmöglichkeiten in einem Forschungsdesign:

- die Kontrolle der unabhängigen Variablen durch die Intervention,
- den Vergleich der Experimentalgruppe mit einer Kontrollgruppe mit gleichen Ausgangsbedingungen bei der Vorher-Messung,
- die Kontrolle externer Faktoren durch den Vergleich von Vorher- und Nachher-Messung bei Interventions- und Kontrollgruppe,
- die statistische Kontrolle weiterer externer Einflüsse durch die randomisierte Verteilung der Versuchspersonen auf Interventions- und Kontrollgruppe.

## KONTAMINIERENDE EFFEKTE

Dieses Forschungsdesign kontrolliert jedoch nicht alle insgesamt möglichen Störfaktoren. Dies gilt z.B. für eine Gruppe störender Einflüsse, die sich unter der Bezeichnung „Kontamination“ zusammenfassen lassen (Mohr, 1995). Hierbei handelt es sich um Verzer-

<sup>1</sup> Die Terminologie stammt aus der klassischen Systematisierung von Einflußfaktoren auf die Gültigkeit von Ergebnissen aus experimentellen und quasi-experimentellen Versuchsanordnungen von D. T. Campbell und J. C. Stanley (1966).

rungen in den Ergebnissen der Nachher-Messungen, die durch die Intervention und/oder durch die Versuchsanordnung des Evaluationsprojekts entstehen. Eine unzureichende Implementation kann beispielsweise zu solchen Verzerrungen führen; oder psychische Identifikationsprozesse von Teilnehmerinnen oder Teilnehmern mit dem Interventions-team führen zur erwartungskonformen Beantwortung bei der Nachher-Messung; oder in der Kontrollgruppe scheiden ganz bestimmte Teilnehmer im Verlauf der Intervention aus, weil sie – im Unterschied zu den Teilnehmern der Interventionsgruppe – nicht besonders zur weiteren Teilnahme motiviert werden, so daß sich die beiden Vergleichsgruppen nachträglich unterschiedlich zusammensetzen. Darüber hinaus kann der gemessene Unterschied zwischen Interventions- und Kontrollgruppe kontaminiert sein, wenn es nicht gelingt, die Mitglieder beider Gruppen streng voneinander zu trennen, und dadurch zu verhindern, daß sie sich über ihre Erfahrungen mit der Intervention austauschen.

Diese kontaminierenden Effekte werden nicht durch Kontrollgruppen-Designs kontrolliert. Im Gegenteil, teilweise entstehen sie erst durch die Einrichtung von Kontrollgruppen. Vermeiden lassen sie sich durch Qualitätsverbesserung bei der Intervention und durch Abschirmung von Kontexteinflüssen.

## **ERWARTUNGEN DER STAKEHOLDER**

Neben diesen methodischen Überlegungen zur Messung und Kontrolle der Variablen muß die Planung von Wirkungsanalysen Erwartungen und Kriterien berücksichtigen, die sich aus den Interessen der Stakeholder ergeben. So muß der Untersuchungsplan oft berücksichtigen, daß Stakeholder nicht nur den Nachweis der Wirksamkeit erwarten, sondern auch noch anderen Nutzen aus Evaluationsstudien ziehen möchten, wie z.B. die allgemeine Erweiterung ihres Wissens über den Untersuchungsgegenstand.

## **FORSCHUNGSETHISCHE PROBLEME**

Von besonderer Bedeutung sind forschungsethische Probleme, wie sie immer auftreten, wenn Menschen ausgewählt, befragt oder beobachtet werden. In der Evaluationsforschung werden forschungsethische Fragen vor allem im Zusammenhang mit der Verwendung von Kontrollgruppen diskutiert: Darf den Mitgliedern der Kontrollgruppe die Verbesserung von Lebenschancen verweigert werden, die durch die Intervention erzielt werden soll? Dürfen Placebo-Effekte kontrolliert werden, indem den Mitgliedern der Kontrollgruppe eine scheinbare soziale Verbesserung vorgespielt wird, die jedoch aus methodischen Gründen praktisch wirkungslos sein muß? Kontrollgruppen, die einzig und allein den Zweck haben, das Placebo-Problem zu kontrollieren oder zu messen, werden für die Evaluation von Sozialprogrammen oft abgelehnt, weil die Vorspiegelung einer sozialen Verbesserung, die in Wahrheit nicht eintritt, als unethisch gilt (Young, 1966).

Darüber hinaus muß berücksichtigt werden, daß Mitarbeiter oder Klienten durch die Datenerhebung übermäßig belastet oder die organisatorischen Abläufe bei der Durchführung des Programms beeinträchtigt werden könnten.

## **FINANZIELLE ASPEKTE**

Nicht zuletzt spielt auch die finanzielle Verantwortlichkeit eine Rolle: Generell gilt die Regel, daß Evaluationskosten immer nur einen begrenzten Anteil an den Gesamtprogrammkosten umfassen dürfen. Für Gesundheitsförderungsprojekte wird hin und wieder ein Richtwert von zehn Prozent genannt. Abweichungen nach unten ergeben sich bei sehr hohen Programmkosten, die mit relativ kostengünstigen Evaluationsprojekten sehr gut erforschbar sind. Abweichungen nach oben sind notwendig, wenn ein Zehnprozentanteil eine akzeptable Überprüfung der Wirksamkeit nicht mehr gewährleistet, und/oder wenn die Entscheidung über Weiterführung oder Einstellung des Programms erhebliche soziale Konsequenzen hätte.

## **5.2. QUALITÄTSSICHERUNG DER EVALUATIONSFORSCHUNG**

Ob sich die Qualität eines Programms nachweisen läßt, hängt von der Qualität seiner Evaluationsstudie ab. Schlechte Forschung erhöht die Wahrscheinlichkeit, ein in Wahrheit völlig wirkungsloses Programm effektiv erscheinen zu lassen, und es besteht die Gefahr, daß die Klienten späterer Anwendungen des Programms über dessen Wirksamkeit getäuscht werden. Noch folgenschwerer ist es vielleicht – vor allem im Hinblick auf die ökonomische Effizienz –, wenn schlechte Forschung dazu führt, daß eine tatsächlich vorhandene Wirkung nicht entdeckt und nachgewiesen wird und infolgedessen ein an sich wirkungsvolles Gesundheitsförderungsprogramm nicht weiter eingesetzt wird.

Die Wirksamkeitsmessung setzt also voraus, daß Untersuchungsplan und Erhebungsmethoden der Evaluationsstudie so angelegt werden, daß sie zutreffende, gültige und zuverlässige Ergebnisse ergibt. Dies erfordert auch für die Evaluationsforschung selbst eine Qualitätssicherung. Einige Möglichkeiten einer solchen Qualitätssicherung werden im folgenden aufgezeigt.

## **PARADIGMENDISKUSSION**

In der Vergangenheit wurden verschiedene „Paradigmen“ mit dem Anspruch vorgeschlagen, es handele sich dabei um die jeweils „richtige“ Vorgehensweise für die Evaluation von Sozialprogrammen. Im wesentlichen waren es zwei Paradigmen (nach Van de Vall, 1993):

- Das erste Paradigma forderte als Untersuchungsdesign randomisierte Kontrollgruppen-Experimente und als Erhebungsverfahren quantitative Auswahl- und Meßtechniken. Es hatte vorrangig das Ziel, Programmwirkungen zu messen und konzentrierte sich somit auf die summative Evaluation.
- Das zweite Paradigma setzte dem ein Untersuchungsdesign entgegen, das komparative Methoden durch den Vergleich relevanter Teilgruppen vorsah, und qualitative Erhebungsmethoden forderte, die vor allem den Implementationsprozeß nachvollziehen sollten. Es konzentrierte sich somit auf die formative Evaluation.

Die beiden Untersuchungsstrategien wurden als Gegensatz empfunden und lange Zeit kontrovers diskutiert. Die Diskussion verlief jedoch weitgehend ergebnislos und ist in der letzten Zeit deutlich abgeklungen (Patton, 1997:290ff.).

Die Vorgabe solcher umfassender Forschungsparadigmen bringt erhebliche praktische Nachteile mit sich, denn normalerweise ist es kaum möglich, ein Programm ohne Abstriche bei der Interventionsstrategie an die methodischen Anforderungen umfassender Untersuchungspläne – z.B. an die einer experimentellen Versuchsanordnung – anzupassen. Denn dies würde bedeuten, daß ein Programm eingesetzt würde, das mit Rücksicht auf einen angeblich „richtigen“ Forschungsansatz seine Interventionsmöglichkeiten nicht vollständig ausschöpft. Im umgekehrten Fall besteht die Gefahr, daß Evaluationsstudien durchgeführt werden, die nicht programmadäquat konzipiert sind und somit zu falschen Bewertungen führen. Die Starrheit vorgegebener Forschungsansätze steht im Widerspruch zur Vielfalt praktischer Programme.

## **STANDARDISIERUNG**

Von einer Standardisierung wäre vor allem die Einhaltung von Mindeststandards bei der Forschung und die Vergleichbarkeit unterschiedlicher Programmevaluationen zu erwarten. Aber auch hier stellt sich die Frage, ob die Vorgabe bestimmter Forschungsverfahren für die Evaluationsforschung nicht die Wahrscheinlichkeit von inadäquaten Evaluationsansätzen und damit von Fehlbewertungen erhöht. Beim gegenwärtigen Stand der Gesundheitsförderung und der Evaluationsforschung Standards für Forschungsverfahren zu setzen, führt unter Umständen zu weniger statt zu mehr Qualitätssicherung bei der Evaluationsforschung.

Eine Standardisierung ist jedoch nicht prinzipiell ausgeschlossen, und es lassen sich auch erste Ansätze dafür erkennen: So werden Erhebungsinstrumente standardisiert, indem bewährte Meßinstrumente, Skalen oder Fragebogenfragen – z.B. für Indikatoren des Gesundheitsverhaltens oder zu Krankheitsrisiken – unter Angabe von Gütekriterien dokumentiert und u.a. für die Verwendung in Evaluationsstudien empfohlen werden.<sup>2</sup> Dies

---

<sup>2</sup> Einige für die Gesundheitsförderung relevante Beispiele sind: A. Bowling (1995; 1996) und Riemann /Gerber, (1997)

bedeutet, daß man sich auf die Qualität dieser Meßinstrumente relativ gut verlassen kann, aber in der Regel überprüft werden muß, ob sie auch in dem jeweiligen Kontext (Setting, Zielgruppe) funktionieren.

Die Erforschung der Anwendbarkeit von Meßinstrumenten unter unterschiedlichen Bedingungen ist ein zeitaufwendiger und kostspieliger Prozeß. Beispiele hierfür sind die umfangreichen Forschungen zum Kohärenzgefühl des Salutogenese-Modells von Aaron Antonovsky<sup>3</sup> oder die Dokumentations- und Forschungsarbeit von Ruut Veenhoven (1992) zur Lebenszufriedenheit (Happiness).

Aber nur wenn die Standardisierung auf breit angelegten Forschungssynthesen beruht, kann sie zu einer höheren Qualität der Evaluationsforschung beitragen. Voraussetzungen hierfür sind die Dokumentation abgeschlossener Evaluationsstudien nach möglichst einheitlichen Deskriptoren, die Sammlung und Archivierung der Studien und vor allem vergleichende Analysen, in denen untersucht wird, welche Meßinstrumente und Forschungsdesigns sich am besten für welche Gesundheitsförderungsansätze in welchen sozialen Kontexten oder Settings eignen würden. Diese Ergebnisse bildeten dann die Grundlage für einen Abstimmungsprozeß zwischen Programmentwicklern und Evaluationsforschern, deren Aufgabe es wäre, konsensfähige Forschungsstandards vorzuschlagen.

## **BERUFLICHE QUALIFIZIERUNG**

Als weitere Qualitätssicherungsstrategie bietet sich an, die Qualifikation der Forscher auf einem möglichst hohen und aktuellen Niveau zu halten. Dies kann zum einen durch Fortbildung der Evaluationsforscher geschehen, etwa nach dem Vorbild der Umfrageforschung, für die in den letzten Jahrzehnten eine Reihe von speziellen Forschungs- und Ausbildungsinstitutionen eingerichtet wurden. Eine weitere Möglichkeit bestünde im Ausbau einer Evaluationsforschungs-Infrastruktur, die fachgerechte Evaluationsforschung anbieten könnte. Voraussetzung wäre eine ausreichende Nachfrage nach Evaluationsforschung generell – nicht nur für die Gesundheitsförderung –, so daß ein leistungsfähiger und effizienter Markt entstehen könnte.

## **FEHLERKONTROLLE**

Das in der sozialwissenschaftlichen Forschung derzeit bevorzugte Qualitätssicherungsverfahren ist das der Fehlerkontrolle. Es besteht darin, mögliche Fehlerquellen zu antizipieren und im Forschungsdesign wirksame Kontrollmöglichkeiten dafür zu schaffen.

<sup>3</sup> Ein Überblick hierzu findet sich in J. Bengel, R. Strittmatter und H. Willmann, Was erhält Menschen gesund? Antonovskys Modell der Salutogenese. (Forschung und Praxis der Gesundheitsförderung Bd. 6; Köln: Bundeszentrale für gesundheitliche Aufklärung, 1998).

Zur Entscheidung, welche Fehler kontrolliert werden sollen, werden Fehlertheorien und Fehlersystematiken herangezogen, die in den verschiedenen Grundlagenwissenschaften der Evaluationsforschung entwickelt wurden. Eine fachübergreifende, vollständige und übersichtliche Standardsystematik existiert jedoch nicht. Einige der vorliegenden Fehlersystematiken beziehen sich eher auf Meßfehler, andere eher auf Analyse- oder Designfehler. Die Terminologien unterscheiden sich erheblich, und die Vielzahl der aufgezählten Fehlermöglichkeiten erschwert die Entscheidung, welche Fehler überprüft werden sollen.<sup>4</sup>

Die Maßnahmen zur Fehlerkontrolle während der Planung des Forschungsdesigns erstrecken sich in der Regel darauf, zusätzliche Variable zu messen, zusätzliche Vergleichsgruppen einzuführen und zusätzliche Fälle zu erheben.

So kann die Stichprobenvarianz reduziert und damit die Präzision der Datenerhebung verbessert werden, indem die Fallzahl des Samples erhöht wird. Die Varianz von Befragungsinstrumenten kann verringert und damit ihre Zuverlässigkeit verbessert werden, indem man jeweils mehrere Einzelfragen oder Items für die zu messenden Variablen stellt und diese gegebenenfalls zu Skalen oder Tests zusammenstellt. Ebenso läßt sich die „inhaltliche Treffsicherheit“ oder Gültigkeit von Meßinstrumenten durch Vergleiche mit anderen Meßinstrumenten kontrollieren, die inhaltlich und methodisch variieren. Ob Vergleichsgruppen im Hinblick auf ihre Zusammensetzung vergleichbar sind, läßt sich durch die Erhebung zusätzlicher Merkmale der Mitglieder kontrollieren. Bei experimentellen Versuchsanordnungen werden mögliche systematische Fehler oft durch zusätzliche Vergleichsgruppen kontrolliert.

Generell besteht die Qualitätssicherungsstrategie der Fehlerkontrolle darin, Vorkehrungen zu treffen, damit die Fehler meßbar werden. Auf diese Weise kann im nachhinein nachgewiesen werden, daß die Ergebnisse nicht (oder nur geringfügig) durch diese Fehler verzerrt sind. Wenn Fehler auftreten, werden sie bei der Datenanalyse als eigenständige Größen isoliert und bei den substantiellen Ergebnissen entsprechend berücksichtigt. Dies geschieht mit Hilfe komplexer multivariater Analyseverfahren.

Für die sozialwissenschaftliche Forschung generell ist die Fehlerkontrolle sicherlich ein wichtiges Verfahren zur Absicherung ihrer Ergebnisse. Für die Evaluation von Gesundheitsförderungsprojekten ist dieses Verfahren meistens nicht geeignet: Häufig verbietet sich die Anwendung bestimmter Kontrolltechniken – z.B. von großen Stichproben oder experimentellen Versuchsanordnungen –, weil sie zuviel kosten, oder weil sie praktische Schwierigkeiten bei der Implementation des Programms schaffen. Zudem ist keineswegs garantiert, daß alle wirklich relevanten Fehler berücksichtigt werden. Hinzu kommt, daß die Ergebnisse multivariater Analysen wenig anschaulich und schwer darstellbar sind.

---

4 Eine in der Evaluationsforschung häufig verwendete Fehlersystematik ist die von Campbell und Stanley (1966) sowie Cook und Campbell (1979) entwickelte Zusammenstellung von Einflußmöglichkeiten auf die Validität (threats to validity) von Ergebnissen experimenteller und quasi-experimenteller Versuchsanordnungen. Diese Fehlersystematik zielt größtenteils auf die Kontrolle systematischer Verzerrungen und umfaßt mittlerweile 33 Fehlermöglichkeiten.

Diese Nachteile ergeben sich nicht, wenn man Fehler von vornherein zu vermeiden versucht statt sie im nachhinein zu messen. So geht die Qualitätssicherungsstrategie der theoriegeleiteten Wirkungsanalyse<sup>5</sup> vor, die im nächsten Abschnitt in ihren Grundzügen erläutert wird.

## 5.3. THEORIEGELEITETE WIRKUNGSANALYSE

Die theoriegeleitete Wirkungsanalyse setzt zur Fehlervermeidung zwei grundlegende Instrumente ein: zum einen die Programmtheorie, die gewissermaßen als methodisches Verfahren verwendet wird (Lipsey, 1993), zum anderen das Monitoring des Implementationsprozesses. Die Qualitätssicherung der Wirkungsanalyse nutzt also dieselben beiden Elemente, die auch für die Qualitätssicherung des Programms eingesetzt werden.

### PROGRAMMTHEORIE

Die Programmtheorie dient zunächst einmal als Leitlinie für die zuverlässige und inhaltlich gültige *Messung der Variablen*, die in der Wirkungsanalyse verwendet werden sollen:

- Die Theorie bestimmt die Auswahl der deskriptiven Variablen. Sie beschreibt, wie die Variablen miteinander im Zusammenhang stehen, und legt damit fest, welches die Outcome-Variablen sind, welche unabhängigen Variablen die Intervention beschreiben und welche als intermediäre Variablen zu verstehen sind.
- Die Theorie präzisiert die inhaltliche Bedeutung. Sie definiert die Variablen und legt damit fest, was sie bedeuten; darüber hinaus spezifiziert der theoretische Kontext der Variablen, was sie nicht bedeuten.
- Die Präzisierung der Bedeutung erlaubt eine präzisere Operationalisierung. Die Festlegung inhaltlicher Dimensionen und die Auswahl treffsicherer Indikatoren wird erleichtert. Wenn die Theorie darüber hinaus etwas über die Zeit aussagt, nach der die Wirkungen des Programms erwartet werden können, lassen sich die Zeitpunkte bestimmen, zu denen zuverlässige Messungen möglich sind. Aufgrund von Aussagen über die vermutete Effektstärke der Intervention läßt sich des weiteren abschätzen, wie groß die Sensitivität der Meßinstrumente sein muß, um die Programmeffekte möglichst genau messen zu können.
- Dies alles verbessert wiederum die Detailplanung für die Erhebungsverfahren. Man kann genauer bestimmen, welche Informationen gesammelt werden müssen, wer diese

<sup>5</sup> Dieser Begriff stammt von H. T. Chen (1990).

Informationen mit den geringsten Schwierigkeiten und mit der größten Wahrheits-treue geben kann, und mit welchen Instrumenten die Informationen am sichersten gewonnen, festgehalten und aufbereitet werden können.

Insgesamt zielt die theoriegeleitete Messung auf eine Verbesserung der *Konstruktvalidität*: Es wird versucht, die zu untersuchenden Variablen möglichst genau begrifflich zu bestimmen und voneinander abzugrenzen und sie so zu operationalisieren, daß wirklich das gemessen wird, was gemessen werden soll, und zwar möglichst eindeutig und ohne inhaltliche Überschneidung der einzelnen Variablen.

Die Konstruktvalidität der unabhängigen (Interventions-)Variablen hängt zusätzlich von der Qualität der Intervention ab. Die Vorkehrungen zur Verbesserung der Effektivität der Interventionen bedeuten gleichzeitig, daß die (unabhängige) Kausalvariable eindeutiger wirksam wird, d.h. möglichst wenig verzerrt durch unklare, widersprüchliche oder sonstige unwirksame Interventionsaktivitäten und weitgehend unbeeinflusst von gegenläufig wirkenden Kontextbedingungen. Eine effektive Intervention führt dazu, daß kontaminierende Faktoren möglichst von vornherein ausgeschlossen werden.

Alle diese methodischen Verbesserungen stärken nicht nur die Meßeigenschaften der Instrumente, sondern sie verbessern auch die *Nachweisbarkeit der Programmwirkungen*:

- Die verbesserten Meßeigenschaften und die Minimierung kontaminierender Einflüsse führen zu mehr inhaltlicher Eindeutigkeit und größerer Präzision der einzelnen Outcome-Messungen. Dies bedeutet, daß die festgestellten Unterschiede zwischen Vorher- und Nachher-Messung mit größerer Sicherheit als Ergebnis der Programmintervention interpretiert werden können.
- Gleichzeitig reduziert sich der Einfluß externer Störfaktoren: Durch eine effektive und zielgenaue Implementation einerseits und eine hohe Sensitivität bei der Messung der Outcome-Variablen andererseits fällt der Unterschied zwischen Vorher- und Nachher-Messungen so deutlich aus, daß externe Störfaktoren sich nur vergleichsweise schwach auswirken können.

Welche externen Störfaktoren überhaupt eine Rolle spielen könnten, kann ebenfalls durch eine gut ausgearbeitete Programmtheorie beantwortet werden. Vor allem ist es möglich, nach spezifischen, inhaltlich beschreibbaren Einflußfaktoren zu suchen. Dies geschieht, indem aus der Programmtheorie und aus zusätzlichen Kenntnissen über den Kontext „rivalisierende Hypothesen“ abgeleitet werden, mit denen Unterschiede zwischen Vorher- und Nachher-Messungen (abgesehen von den Interventionseffekten) ebenfalls erklärt werden könnten. So kann bereits bei der Planung des Programms nach Möglichkeiten gesucht werden, den Einfluß dieser spezifischen Störfaktoren bei der Intervention von vornherein abzuschirmen oder sonstwie zu kontrollieren. Oder man kann nach (kostengünstigen) Datenquellen suchen, anhand derer die rivalisierenden Hypothesen über-

prüft werden können. Noch wichtiger ist die (durch Programmtheorie und Kontextanalyse begründete) Erkenntnis, daß wahrscheinlich keine nennenswerten externen Störfaktoren auftreten werden. In diesem Fall spricht vieles dafür, auf ein Kontrollgruppen-Design zu verzichten, vor allem dann, wenn dieses mit höheren Kosten oder praktischen Schwierigkeiten für die Programmimplementation verbunden wäre.

## **MONITORING DES IMPLEMENTATIONSPROZESSES**

Das Monitoring des Implementationsprozesses bietet weitere Möglichkeiten, die Nachweisbarkeit der Programmwirkungen zu verbessern. Voraussetzung hierfür ist, daß die Outcome-Variablen möglichst frühzeitig in das Monitoringsystem aufgenommen und kontinuierlich gemessen werden.

Zunächst jedoch verbessert die kontinuierliche Messung des Monitorings die Qualität der Daten: Die Datenerhebung in größtmöglicher zeitlicher Nähe vor, während und nach der Intervention ermöglicht eine hohe Meßgenauigkeit; gleichzeitig erlauben die Wiederholungsmessungen eine bessere Einschätzung von Zufallsvariationen im Zeitverlauf und damit eine Verbesserung der Zuverlässigkeit von Daten und Ergebnissen.

Weiter kann man im frühen Verlauf der Kurve für die Outcome-Variablen feststellen, ob die vorangegangenen Messungen die nachfolgenden verzerren (testing, instrumentation). Bei Zeitreihen-Messungen werden diese Störfaktoren sichtbar, bevor sie die Interventionseffekte überlagern. Es kann überprüft werden, ob diese Einflüsse überhaupt auftreten, und wenn sie erkennbar sind, lassen sie sich von den Effekten der Intervention trennen.

Vor allem aber bietet das Zeitreihen-Design die Möglichkeit, bestimmte inhaltliche Hypothesen hinsichtlich der Wirksamkeit des Programms zu überprüfen und zu belegen, daß es wirksam ist. Ändert sich der Verlauf der Kurve für die Outcome-Variablen während der Intervention deutlich in die erwartete Richtung, so spricht das für die Wirksamkeit.

Weiter lassen sich Zusammenhangs-Hypothesen über den Wirkungsmechanismus testen. Bestätigt sich, daß die Interventionsaktivitäten bestimmte Variablen des Wirkungsmechanismus verändern und diese wiederum mit den Outcome-Variablen in Zusammenhang stehen, so kann das ebenfalls als Beleg für die Wirksamkeit der Intervention gewertet werden.

## 5.4. ZUSAMMENFASSUNG

Die Wirkungsanalyse überprüft die Programmeffekte und untersucht, inwieweit diese als Ergebnis der Intervention und ihrer Wirkungsmechanismen angesehen werden können. Die Ergebnisse sollten so aufbereitet werden, daß sich daraus Verbesserungen für die Interventionsstrategie ableiten lassen.

Die Planung von Wirkungsanalysen kann sich am Modell der theoriegeleiteten Wirkungsanalyse orientieren, die von folgenden forschungsstrategischen Vorgaben angeleitet ist:

- Der Untersuchungsplan sollte „maßgeschneidert“ auf das Programm zugeschnitten sein und deshalb möglichst auf der Grundlage einer Programmtheorie erarbeitet werden. Vorgefertigte Forschungsstandards („Forschungsparadigmen“) erhöhen das Risiko, aufgrund einer inadäquaten Wirkungsanalyse zu Fehlschlüssen über die Wirksamkeit des Programms zu kommen.
- Das Programm sollte durch die Evaluationsstudie möglichst wenig beeinträchtigt werden. Wenn bei der Programmplanung aufwendige Verfahren zur Fehlerkontrolle – z.B. randomisierte experimentelle Versuchsanordnungen – zu sehr berücksichtigt werden, kann dies die Wirksamkeit und Effizienz des Programms gefährden.
- Fehler von vornherein zu vermeiden, hat Priorität vor aufwendigen Verfahren der nachträglichen Fehlerkontrolle und Fehlermessung.
- Vor allem sollten bei der Auswahl oder Konstruktion der Meßinstrumente Fehler vermieden werden. Eine begrifflich klar ausformulierte Programmtheorie, eine qualitätsgesicherte Intervention und kontinuierliche Messungen mit Hilfe eines Monitorings sind hierbei hilfreich.
- Das grundlegende Forschungsdesign für den Nachweis der Wirksamkeit ist eine Längsschnittstudie mit Vorher-Nachher-Messungen. Dieses Design wird verstärkt durch eine hohe Qualität der Meßinstrumente, durch kontinuierliche Zeitreihen-Messungen im Rahmen des Monitorings und durch die empirische Überprüfung der aus der Programmtheorie abgeleiteten Hypothesen bezüglich der Wirksamkeit der Intervention.

Generell beruht diese Qualitätssicherungsstrategie der theoriegeleiteten Wirkungsanalyse darauf, daß zwischen der Qualitätssicherung des Programms und der Qualität der Evaluationsforschung ein direkter Zusammenhang besteht: Je besser ein Sozial- oder ein Gesundheitsprogramm (theoriegeleitet) geplant und (qualitätsgesichert) implementiert ist, desto eher ist es möglich, treffsichere und präzise Meßinstrumente und Untersuchungspläne für die Analyse der Wirksamkeit zu entwickeln.

**RESÜMEE UND  
HANDLUNGSEMPFEHLUNGEN**



# 6.

## RESÜMEE UND HANDLUNGSEMPFEHLUNGEN

Das vorliegende Arbeitspapier untersucht, wie die Programmevaluation für die Qualitätssicherung genutzt werden kann. Es beschreibt ein qualitätsorientiertes Evaluationskonzept, das vor allem darauf ausgerichtet ist, in allen Phasen eines Projekts Informationen zu sammeln und rückzukoppeln, um auf diese Weise die Qualität der jeweils nachfolgenden Projektphasen präventiv sicherzustellen oder zu verbessern.

Aus diesem Evaluationskonzept sind eine Reihe von Handlungsempfehlungen und Planungsprinzipien ableitbar, die sich als Leitlinien für die Planung von Gesundheitsförderungsprojekten und deren Evaluationsstudien verstehen lassen. Diese Handlungsempfehlungen werden in diesem abschließenden Kapitel noch einmal zusammengestellt, wobei sich diese Zusammenstellung an den drei überwiegend verwendeten (und in *Abschnitt 2.3.* näher beschriebenen) Qualitätsperspektiven orientiert:

- Qualität aus der Experten-Perspektive (Qualitätssicherung der Interventionsentwicklung),
- Qualität aus der Stakeholder-Perspektive,
- Qualität aus der Sicht der wirtschaftlichen Verantwortlichkeit.

Diese Darstellung soll noch einmal verdeutlichen, daß Programmevaluation hier nicht einfach als eine nach Beendigung der Intervention stattfindende Qualitäts- oder Erfolgskontrolle verstanden wird; vielmehr kann sie als integraler Teil des Projektmanagements von Anfang an zur Qualitätssicherung des gesamten Programmablaufs beitragen. Qualitätssicherung durch Programmevaluation ist damit nicht die alleinige Aufgabe von Evaluationsexperten, sondern fällt in die Verantwortlichkeit aller für das Projekt oder Programm zuständigen Entscheidungsträger: Das sind die Fachleute für Prävention und Gesundheitsförderung, die politisch oder finanziell Verantwortlichen sowie andere Stakeholder und vor allem das für die Koordination der verschiedenen (Qualitäts-)Perspektiven zuständige Projektmanagement.

### QUALITÄTSSICHERUNG DER INTERVENTIONSENTWICKLUNG

Die Qualitätssicherung der Interventionsentwicklung orientiert sich an der Qualitätsstrategie des kontinuierlichen Verbesserungsprozesses (KVP), in die sich die Vorgehensweise der Programmevaluation ohne weiteres einpassen läßt. Beide beruhen auf demselben Prinzip, nämlich der Rückkopplung von Informationen über Verfahren, Prozesse und Ergebnisse zu den Entscheidungsprozessen des Programms.

Im folgenden werden die wichtigsten Rückkopplungsmöglichkeiten beschrieben, die bei den drei grundlegenden Phasen eines Programms möglich sind:

- bei der Konzeptentwicklung,
- bei der Implementation und
- bei der Wirkungsüberprüfung.

Alle drei Rückkopplungen zusammen können zu einem geschlossenen System der Qualitätssicherung von (Gesundheitsförderungs-)Programmen integriert werden.

### **Konzeptionsphase**

Die erste Rückkopplungsmöglichkeit bietet die Konzeptionsphase, wobei die Qualitätssicherung der Konzeptionsentwicklung durch die Erarbeitung und Ausformulierung einer interventionsspezifischen Programmtheorie – im Sinne der sich in den letzten Jahrzehnten zunehmend durchsetzenden Strategie der „theoriegeleiteten Evaluation“ – gestützt werden kann.

Als systematisch ausgearbeitetes, realitätsbezogen formuliertes Konzept des Programms beschreibt die Programmtheorie die genauen Zusammenhänge zwischen den Aktivitäten, aus denen die Intervention besteht – den effektiven Programmkomponenten –, den dadurch in Gang gesetzten Wirkungsmechanismen und den hierdurch wiederum bewirkten konkreten und prinzipiell meßbaren Outcomes, und dies unter Berücksichtigung des Kontexts, in dem das Programm implementiert werden soll. Eine Programmtheorie beschränkt sich demnach nicht darauf, schematisch Ziele, Zielgruppen und generelle Interventionsmaßnahmen oder -methoden aufzulisten. Es kommt vielmehr vor allem darauf an, die zunächst oft nur sehr allgemeinen Konzeptideen zu spezifischen Aussagen über (Kausal-)Zusammenhänge zu präzisieren, und die häufig nur impliziten Vorstellungen der Programmentwickler zu expliziten Handlungsvorgaben für konkrete Interventionsvorgänge auszuformulieren.

In der Phase der Konzeptentwicklung dient die Programmtheorie der frühzeitigen „präventiven Evaluation“ des Programmkonzepts, indem die Aussagen der Theorie mit vorhandenem Wissen und relevanten Forschungsergebnissen konfrontiert werden.

Auf diese Weise lassen sich Fragen nach der Realisierbarkeit und voraussichtlichen Wirksamkeit des Konzepts stellen und beantworten, wie z.B.:

- Kann das Programm überhaupt funktionieren, oder müssen andere oder zusätzliche effektive Programmkomponenten eingesetzt werden?
- Ist das Programm in dem vorgesehenen Kontext praktikabel oder muß es besser auf Mitarbeiter, Klienten oder spezifische Kontexteinflüsse eingestellt werden?
- Ist die Interventionsstärke groß genug oder muß sie verstärkt werden, damit die beabsichtigten Outcomes auch tatsächlich erreicht werden?

Die Beantwortung dieser Fragen ermöglicht eine erste grundlegende Qualitätsverbesserung:

- die Vermeidung grundsätzlicher Konzeptionsfehler,
- eine realistische, auf die spezifischen Kontextbedingungen ausgerichtete Umsetzung des Konzepts und
- ausreichend stark angelegte Interventionsmaßnahmen.

Eine qualitätsgesicherte Programmtheorie ist ihrerseits eine wichtige Grundlage für die Qualitätssicherung des Implementationsprozesses, da sich durch eine Präzisierung der Handlungsvorgaben eine möglichst konzeptgetreue Übertragungsqualität erreichen läßt.

### **Implementationsphase**

Die zweite Rückkopplungsmöglichkeit ergibt sich in der Implementationsphase. Eine durch ein Monitoring erstellte systematische Datensammlung während des Implementationsprozesses stellt Informationen bereit, mit deren Hilfe Schwierigkeiten im Programmablauf entdeckt und Fehler somit frühzeitig vermieden werden können.

Ein Monitoring des Implementationsprozesses bedeutet die kontinuierliche Beobachtung des Programmverlaufs mit Hilfe eines systematisch zusammengestellten Satzes von Daten, die regelmäßig und routinemäßig gemessen werden. Eine wichtige Voraussetzung für die Auswahl der erforderlichen Indikatoren ist wiederum eine gute Programmtheorie. Erhoben werden – neben Daten zur administrativen Abwicklung – Informationen zum Einfluß von Kontextfaktoren, zur Wirksamkeit und Leistung einzelner Programmkomponenten und zur konzeptgetreuen Übertragung der Implementationsvorgaben.

Entscheidend ist dabei, daß die Daten während des Implementationsprozesses laufend aufbereitet und ausgewertet werden. Anhand dieser regelmäßig abrufbaren Informationen lassen sich Abweichungen und Schwierigkeiten während des laufenden Programms entdecken, so daß noch in der Implementationsphase Lösungen gesucht und gefunden werden können. Vor allem läßt sich frühzeitig die Effektivität des Programms steigern: Interventionsinputs, die sich anfangs als zu schwach erwiesen haben, lassen sich verstärken, die Übertragungsqualität läßt sich präzisieren, und man kann versuchen, störende Einflüsse des Kontexts abzuschirmen und zu kontrollieren.

### **Wirkungsüberprüfung**

Die dritte Rückkopplungsmöglichkeit nutzt die Ergebnisse der Wirkungsanalyse. Diese kann vor allem dann zur Verbesserung der Qualität von (künftigen) Interventionen beitragen, wenn untersucht und nachgewiesen werden kann, wie die festgestellten Wirkungen zustande gekommen sind oder – wenn keine oder nur geringe Wirkungen entstanden sind – wodurch die Effekte verhindert wurden bzw. weshalb sie so schwach ausgefallen sind. Die Wirkungsanalyse beruht auf einem Forschungsplan, der eigens für das jeweilige Programm aufgestellt werden muß. Der Forschungsplan legt zum einen fest, wie die Programm-Outcomes gemessen werden sollen; zum anderen gibt er an, wie untersucht werden soll, ob die Outcomes als Effekte der Intervention und ihrer Wirkungsmechanismen angesehen werden können.

Bei der Aufstellung des Forschungsplans empfiehlt es sich, auf die Messung der Variablen besonders große Sorgfalt zu verwenden – sowohl auf die Messung der Outcome-Variablen als auch der Variablen, welche die Interventionen und die Wirkungsmechanismen repräsentieren.

Eine hohe Meßqualität ist deshalb so wichtig, weil die Untersuchungen zum Nachweis der Wirksamkeit um so genauer sind, je treffsicherer und zuverlässiger die einzelnen Variablen gemessen werden. Die Meßqualität läßt sich folgendermaßen verbessern:

- Eine konsequente Qualitätssicherung des Implementationsprozesses präzisiert die Interventionen und die Wirkungsmechanismen und damit die entsprechenden Variablen.
- Eine möglichst weitgehend ausgearbeitete Programmtheorie ermöglicht genaue Definitionen dessen, was gemessen werden soll, und erleichtert die Konstruktion von Meßinstrumenten.
- Die Messung der Variablen zeitgleich mit den Programmabläufen – z.B. im Rahmen des Monitorings – ermöglicht genauere Daten.
- Die Konstruktion von Meßinstrumenten, die den Informationsstand und die Auskunftsbereitschaft der jeweiligen Befragten berücksichtigen, erhöht die Validität und Zuverlässigkeit von Messungen.

Das Untersuchungsdesign zum Nachweis der Wirksamkeit sollte an den Realitäten des Programms ausgerichtet sein und nicht an bestimmten methodischen Vorgaben (wie z.B. an der Forderung nach der Anwendung randomisierter Kontrollgruppenversuche). Es sollte vor allem die durch die Intervention hervorgerufenen Veränderungen sichtbar machen. Ein solches programmadaquates Untersuchungsdesign läßt sich durch folgende Maßnahmen herstellen:

- Messung von Veränderungen der Outcomes und anderer wichtiger Variablen im Zeitverlauf, am besten im Rahmen des Monitorings, mit mehreren Meßzeitpunkten vor und nach der Intervention,
- empirische Überprüfung von spezifischen, aus der Programmtheorie abgeleiteten Hypothesen bezüglich der Wirksamkeit der Interventionen,
- Untersuchung nur derjenigen „rivalisierenden“ Hypothesen zu möglichen alternativen Erklärungen für das Auftreten der Outcomes, die nach der zugrundeliegenden Programmtheorie, unter den gegebenen Kontextbedingungen, bei der gegebenen Qualität der Implementation und der Meßqualität als wirklich wichtig angesehen werden müssen.

Generell beruht diese Qualitätssicherungsstrategie darauf, daß zwischen der Qualitätssicherung des Programms und der Qualität der Evaluationsforschung ein direkter Zusammenhang besteht: Je besser ein Sozial- oder ein Gesundheitsprogramm (theoriegeleitet) geplant und (qualitätsgesichert) implementiert ist, desto eher ist es möglich, treffsichere und präzise Meßinstrumente und Untersuchungspläne für die Analyse der Wirksamkeit zu entwickeln und anzuwenden.



## QUALITÄTSSICHERUNG AUS DER STAKEHOLDER-PERSPEKTIVE

Von grundlegender Bedeutung für die Qualitätssicherung von Gesundheitsförderungsprojekten ist eine enge Kooperation von Projektentwicklern und Evaluationsfachleuten, die mit der Vorbereitung des Projekts beginnt und mindestens bis zur abschließenden Bewertung der Ergebnisse des Projekts anhält.

Eine solche möglichst frühzeitig beginnende projektbegleitende Kooperation ist allein schon wegen des präventiven Qualitätssicherungsansatzes notwendig. Darüber hinaus lassen sich bereits bei der Konzeptentwicklung Synergieeffekte realisieren, wenn Evaluationsforscher in dieser Projektphase mitarbeiten: Sie können bei der Erarbeitung der Programmtheorie theoretisches und empirisches Wissen im Hinblick auf die Bildung konkreter, realitätsbezogener Hypothesen einbringen, während sie ihrerseits besser zu verstehen lernen, wie das Programm funktioniert und was es erreichen soll.

Für eine konsensgetragene Festlegung und eindeutige Definition der Interventions-Outcomes sollte ebenfalls eine möglichst frühzeitige Kooperation zwischen allen wichtigen Stakeholdern hergestellt werden. Dies ist vor allem notwendig, um möglichst frühzeitig einen Konsens über die Kriterien zu erzielen, nach denen entschieden werden soll, ob die beabsichtigten Wirkungen eingetreten sind oder nicht.

Auch bei der Aufstellung eines Monitoringsystems zur kontinuierlichen Beobachtung des Implementationsprozesses sollten Evaluationsfachleute, Projektmanagement und Projektmitarbeiter zusammenarbeiten.

Die Aufgabe dieser Zusammenarbeit besteht darin, auf der Grundlage einer gemeinsamen Analyse der Arbeits- und Interaktionsprozesse ein System von Indikatoren festzulegen, die möglichst zuverlässig und inhaltsreich die Vorgänge und Ergebnisse der Interventionsprozesse wiedergeben und dies zu möglichst geringen Kosten für Erhebung und Aufbereitung der Daten. Die Kooperation mit den Mitarbeitern kann vermeiden, daß deren Arbeit durch die Datenbeschaffung für das Monitoring so sehr belastet ist, daß dadurch unter Umständen die Qualität der Implementation leiden könnte.

Vergrößert wird die Qualität, wenn die Monitoring-Daten den Mitarbeitern zugänglich sind: Dies schafft Transparenz und verstärkt die Bereitschaft zur Qualitätsverbesserung. Ebenso müssen bei einer Selbstevaluation der Implementation durch die Projektmitarbeiter die Vorteile einer größeren Transparenz und eines stärkeren Qualitätsbewußtseins abgewogen werden gegen die Gefahr von Arbeitsbelastung und den sich daraus ergebenden Qualitätseinbußen.

Über den Forschungsplan für die Wirkungsanalyse muß gemeinsam von Projektentwicklern und Evaluationsforschern entschieden werden. Priorität hat dabei das Programm. So sollte das Programmkonzept durch die Evaluationsstudie möglichst wenig beeinträchtigt werden. Aufwendige Verfahren zur Fehlerkontrolle, wie z.B. randomisierte experimentelle Versuchsanordnungen, können die Wirksamkeit und Effizienz des Programms gefährden.

## **QUALITÄTSSICHERUNG AUS DER PERSPEKTIVE DER WIRTSCHAFTLICHEN VERANTWORTUNG**

Das hier vorgeschlagene qualitätsorientierte Evaluationskonzept erleichtert Planung und Einsatz der Finanzmittel für Gesundheitsförderungsprojekte.

Die Programmtheorie ermöglicht das Maß an Transparenz, das für eine präzise Planung eines effizienten Einsatzes der Finanzmittel nötig ist. Mit einer genauen Beschreibung der einzelnen effektiven Interventionselemente und des dafür erforderlichen Personaleinsatzes bietet sie eine besonders gute Basis für die Entscheidung darüber, wie die knappen Mittel verwendet werden müssen oder wo das Konzept aus Einsparungsgründen geändert werden kann. Eine bereits während der Konzeptentwicklung auf der Grundlage der Programmtheorie vorgenommene Machbarkeitsprüfung läßt frühzeitig und nicht erst nach Abschluß des Projekts erkennen, ob es sich bei den vorhandenen finanziellen Mitteln lohnt, das Projekt überhaupt fortzusetzen.

Die in dem Evaluationskonzept vorgeschlagenen Maßnahmen der Qualitätssicherung ermöglichen eine Steigerung der Produktivität und Effizienz sowohl der Intervention als auch der Evaluation.

So können die zukünftigen Kosten einer Intervention reduziert werden, wenn sich mit Hilfe der Evaluationsergebnisse nachweisen läßt, daß eine bestimmte Wirkung auch mit einer geringeren Effektstärke erreicht werden kann. Eine gut geplante und konzeptgetreu ablaufende Implementation wiederum ermöglicht ein sparsames Evaluationsdesign, das auf umfangreiche und teure Kontrollverfahren (z.B. auf randomisierte Kontrollgruppen-Designs) verzichten kann, weil sich bei implementierten Interventionen die Wirkungen eindeutiger messen und die Wirkungsmechanismen verlässlicher nachweisen lassen.

Insgesamt führt Qualitätssicherung durch Evaluationsforschung zu mehr Effizienz in der Gesundheitsförderung, vor allem weil sie falsche Schlußfolgerungen über die Wirksamkeit von Programmen verhindert: Es reduziert sich einmal die Wahrscheinlichkeit, daß in Wahrheit völlig wirkungslose Gesundheitsförderungsprojekte als effektiv angesehen und weiterhin finanziert werden, und auf der anderen Seite steigen die Chancen, daß wirkungsstarke Interventionen auch tatsächlich als solche erkannt und für die Verbesserung der gesundheitlichen Lage der Menschen eingesetzt werden.

6





**BIBLIOGRAPHIE DER BENUTZTEN  
UND ZITIERTEN LITERATUR**

# 7.

## BIBLIOGRAPHIE DER BENUTZTEN UND ZITIERTEN LITERATUR

- 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. (1998): Was erhält Menschen gesund? Antonovskys Modell der Salutogenese. (Forschung und Praxis der Gesundheitsförderung, Bd. 6). Köln: Bundeszentrale für gesundheitliche Aufklärung (Hrsg.).
- 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. (2. Auflage) 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 (Hrsg.) (1997): Quality management in workplace health promotion. Tagungsbericht (Tb 81) zum Workshop November 1996 in Erkner bei Berlin. Dortmund/Berlin.
- Bundeszentrale für gesundheitliche Aufklärung (Hrsg.) (1994): Qualitätsraster Adipositas zur Beurteilung von Maßnahmen gegen Übergewicht. Köln.
- Campbell, D. T. / Stanley, J. C. (1966): Experimental and quasi-experimental designs for research. Chicago: Rand McNally. (Reprint aus Gage, N. L. [Hrsg.] [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. (Hrsg.): Primary prevention works. Newbury Park, CA: Sage.
- Chelimsky, E. (1997): The coming transformations in evaluation. In: Chelimsky, E. / Shadish, W. R. (Hrsg.): 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). Genf: WHO.
- Heaney, C. A. (1998): Intervention Mapping and the new health promotion. *Health Education and Behavior*, 25 (5), 564–568.
- Heiner, M. (Hrsg.) (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, Bd. 20). Baden-Baden: Nomos.
- Lethbridge J. / Parish, R. / Hagar, S. (1996): Quality and health promotion – developing concepts and principles. Discussion document for the meeting in Ormoz, Slowenien, 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. (Hrsg.): *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. (Hrsg.) (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. (Hrsg.): *Public communication campaigns*. (2. Auflage). 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*. (2. Auflage). 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*. (3. Auflage). 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. (Hrsg.): *Qualität in der Gesundheitsförderung*. Wien: Facultas-Universitätsverlag.
- Plantz, M. C. / Greenway, M. T. / Hendricks, M. (1997): Outcome measurement: showing results in the non-profit sector. In: Newcomer, K. E. (Hrsg.): *Using performance measurement to improve public and non-profit 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. (Hrsg.): *Employees, Careers and Job Creation*. San Francisco: Jossey-Bass.

- Riemann, K. / Gerber U. (1997): Standardisierung von Fragestellungen zum Rauchen. (Forschung und Praxis der Gesundheitsförderung, Bd. 1). Köln: Bundeszentrale für gesundheitliche Aufklärung (Hrsg.).
- Rossi, P. H. / Freeman, H. E. (1993): Evaluation: a systematic approach. (3. Auflage). Beverly Hills, CA: Sage.
- Saan, H. (1997): Quality revisited. In: Macdonald, G. (Hrsg.): Development and Measurement of Quality in Health Promotion. International Journal of Health Promotion and Education, 4 (2).
- Sechrest, L. (Hrsg.) (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: Tagungsbericht.
- 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/München: Juventas 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 Rotterdamexperimentation: 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. (Hrsg.): Job stress interventions (S. 125–138). Washington, D.C.: American Psychological Association.
- Weiss, C. H. (1972): Evaluating educational and social action programs. In: Weiss, C. H. (Hrsg.): 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, Kopenhagen: WHO/EURO.

**In der Fachheftreihe „Forschung und Praxis der Gesundheitsförderung“  
sind bereits erschienen:**

**Band 1 – *Standardisierung von Fragestellungen zum Rauchen***

Ein Beitrag zur Qualitätssicherung in der Präventionsforschung von Klaus Riemann und Uwe Gerber im Auftrag der BZgA.

Bestellnr.: 60 600 000

**Band 2 – *Geschlechtsbezogene Suchtprävention***

Praxisansätze, Theorieentwicklung, Definitionen.

Abschlußbericht eines Forschungsprojekts von Peter Franzkowiak, Cornelia Helfferich und Eva Weise im Auftrag der BZgA.

Bestellnr.: 60 602 000

**Band 3 – *Gesundheit von Kindern***

Epidemiologische Grundlagen. Eine Expertentagung der BZgA.

Bestellnr.: 60 603 000

**Band 4 – *Prävention durch Angst?***

Stand der Furchtappellforschung. Eine Expertise von Jürgen Barth und Jürgen Bengel im Auftrag der BZgA.

Bestellnr.: 60 604 000

**Band 5 – *Prävention des Ecstasykonsums***

Empirische Forschungsergebnisse und Leitlinien. Dokumentation eines Statusseminars der BZgA vom 15. bis 17. September 1997 in Bad Honnef.

Bestellnr.: 60 605 000

**Band 6 – *Was erhält Menschen gesund?***

Antonovskys Modell der Salutogenese – Diskussionsstand und Stellenwert. Eine Expertise von Jürgen Bengel, Regine Strittmatter und Hildegard Willmann im Auftrag der BZgA.

Bestellnr.: 60 606 000

**Band 7 – *Starke Kinder brauchen starke Eltern***

Familienbezogene Suchtprävention – Konzepte und Praxisbeispiele

Bestellnr.: 60 607 000

**In Kürze erscheint:**

**Band 9 – *Die Herausforderung annehmen***

Aufklärungsarbeit zur Organspende im europäischen Vergleich.

Bestellnr.: 60 609 000



**Bundeszentrale  
für  
gesundheitliche  
Aufklärung**

**ISBN 3-933191-26-2**