

## 6. Durchführung von Evaluationsstudien

Je sorgfältiger die Planungsarbeiten durchgeführt wurden, um so weniger prinzipielle Probleme wird die Durchführung des konkreten Projektes machen. Trotzdem wäre es eine Illusion, außerhalb von sehr kleinen und überschaubaren Projekten von einer reibungslosen Durchführung auszugehen. Zumindest bei längerfristigen Vorhaben sind Schwierigkeiten im Bereich der Organisation zu erwarten, typische Beispiele und Maßnahmen für deren partielle Behebung werden im Abschnitt 6.1 diskutiert. Weniger mit Unsicherheiten behaftet ist die sachgerechte Durchführung der Auswertungsarbeiten, doch muß man auch dort auf eine sorgfältige Fehlerkontrolle und die Einhaltung der Vertraulichkeits- bzw. Datenschutzbestimmungen achten. Bei aller Mühe aber nicht vermeidbar ist das prinzipielle Problem, daß von der unübersehbaren Vielzahl der denkbaren Auswertungen nur ein kleiner Teil erfolgen kann und diese Auswahl stets mit dem Vorwurf von «Manipulation» rechnen muß. Fragen dieser Art werden im zweiten Abschnitt dieses Kapitels besprochen (6.2).

Am Schluß des Evaluationsprojektes steht meist eine schriftliche, oft durch eine mündliche Präsentation ergänzte Berichtlegung an den Auftraggeber, ggf. auch an andere Instanzen (zum Beispiel Öffentlichkeit, Schulleiter). Die damit verbundenen Darstellungs- und Kommunikationsprobleme werden von Akademikern ohne einschlägige Erfahrung oft unterschätzt. Man lebt innerhalb einer universitären Bezugsgruppe in einem für Außenstehende schwer nachvollziehbaren Abstraktionsniveau, einer fachspezifischen Begriffsbildung (die leider nicht nur manchmal unverständlich ist, sondern auch zu Mißverständnissen bei «Laien»

führen kann), und die vor allem im wirtschaftsnahen Bereich sehr elaborierten Darstellungshilfen (Medien bei Vorträgen, optische Auflockerung) werden an der Universität nur selten in vergleichbar intensiver Weise gepflegt. Hinweise zu solchen Fragen gibt Abschnitt 6.3.

### 6.1 Organisatorische Fragen

Die wichtigsten Organisationsprobleme finden sich im Bereich des Personals – sowohl bei den Projektmitarbeitern, als auch bei Auftraggebern oder anderen für die Durchführung erforderlichen Kontaktpersonen (zum Beispiel Lehrer, Verwaltungsbeamte etc.). Zumindest für die eigenen Mitarbeiter kann man einige Schwierigkeiten vermeiden, wenn man rechtzeitig auch deren Einschulungsbedürfnisse und ihre persönlichen Interessen, zum Beispiel im Hinblick auf ihre weitere berufliche Entwicklung, bei der Projektplanung und der Personalführung beachtet (6.1.1). Darüber hinaus muß man sich im klaren sein, daß größere Vorhaben nicht ohne explizite Kontrolltechniken durchgeführt werden können, wenn man vermeiden will, daß sich am Ende einer zum Beispiel zweijährigen Arbeitsphase nicht mehr behebbare Probleme aufürmen und u.U. das ganze Projekt ergebnislos abgebrochen werden muß. Einige Hinweise, die eine entsprechend sorgfältige Einarbeitung natürlich nicht ersetzen können, finden sich im Abschnitt 6.1.2.

Nahezu alle Organisationsprobleme lassen sich lösen, wenn man rechtzeitig an ihr Auftreten denkt, also eine (in den Evaluations-Begriffen gemäß Diagramm II/2) «antizipatorische Input-Evaluation» für das eigene Projekt mit ausreichender Aussagekraft anstellt.

Bei aller Sorgfalt muß trotzdem davon ausgegangen werden, daß sich zum Teil vorhersehbare, zum Teil aber völlig überraschende «Katastrophen» ergeben, die ein bei aller Sorgfalt der Planung doch sehr flexibles Krisenmanagement erfordern. Einige Beispiele finden sich in 6.1.3.

### 6.1.1 Fragen der Personalführung

Die Verhältnisse gestalten sich natürlich unterschiedlich, je nachdem, ob ein Evaluationsprojekt firmen- oder klinikintern, durch Mitarbeiter eines Ministeriums, im Rahmen der Aufgaben eines Universitätsinstituts oder als Auftrag an ein privatwirtschaftlich organisiertes Institut durchgeführt wird (vgl. Diagramm II/5). In großen Firmen und bei öffentlichen Arbeitgebern gibt es meist bewährte, langfristig festgelegte Formen der Personalführung, so daß dort keine besonderen Probleme zu erwarten sind. Anders ist die Situation bei privaten (Klein-)Instituten, die zwar einen erheblichen Arbeitsmarkt bieten (es gibt in der Bundesrepublik Deutschland einige hundert angewandtforschende, sozialwissenschaftliche Institute), aber bezüglich der Personalführung schon wegen des großen bedingten Fehlens entsprechender Spezialisten (kein Institut mit vielleicht 5 bis 10 Mitarbeitern kann sich einen eigenen Personalfachmann leisten) auf diesem Gebiet nicht das professionelle Niveau größerer Institutionen erreichen. Gleichzeitig sind bei kleinen Instituten die Folgen von Fehlern besonders gravierend, da Ausfälle kaum durch andere Mitarbeiter abgedeckt werden können.

Aus der Sicht des Arbeitgebers lassen sich die störendsten Punkte wie folgt beschreiben:

- Es besteht ein außerordentlicher Mangel an fachlich qualifiziertem Personal, das zum Zeitpunkt der Übernahme eines neuen Projektes auch tatsächlich zur Verfügung steht.
- Der Arbeitsmarkt für solche Spezialisten ist relativ klein, ein «Abwerben» von anderen Instituten während dort laufender Projekte problematisch, und arbeitslose berufserfahrene Kräfte sind selten (umgekehrt kann es wegen dem kleinen Stellenmarkt aber durchaus sein, daß auch gute Spezialisten einige Zeit brauchen, um eine neue Stelle zu finden).
- Bei befristeten Projekten wird die Situation noch durch regionale Mobilitätseinschränkungen erschwert, da viele Interessenten zögern, für vielleicht nur ein Jahr über größere Entfernungen hinweg den Wohnort zu wechseln.

Eine Lösung wäre die rechtzeitige Einstellung und vorbereitende Einschulung von Anfängern. Es können sich aber nur wenige private Institute eine solche langfristige Personalplanung finanziell leisten, und überdies wird die praktische Leistungsfähigkeit von Bewerbern unmittelbar nach Abschluß der akademischen Ausbildung skeptisch beurteilt, vielleicht nicht immer zu recht.

Spezialisten unter den Mitarbeitern (insbesondere im EDV-Bereich) können bei längerfristigen Projekten eine erhebliche informelle Machtstellung erlangen, und zwar dann, wenn ihre Detailkenntnisse (zum Beispiel über die Datenorganisation) oder ihre persönlichen Beziehungen zu externen Stellen (zum Beispiel bei der Organisation einer Untersuchung mit zahlreichen Kontaktpartnern aus wichtigen Kliniken) entweder aus Zeit- oder Kostengründen nicht mehr von anderen übernommen werden können (vor allem nicht im Konfliktfall).

Noch schwieriger als fachlich-wissenschaftlich gut ausgebildete Mitarbeiter sind Bewerber zu finden, die darüber hinaus auch Projekt-Managementaufgaben (Abstimmung und Kontrolle der Teilarbeiten, Kontakte mit dem Auftraggeber und den anderen kooperierenden Stellen, flexible Anpassung der Projektarbeit an geänderte Situationen) eigenständig übernehmen können. Dies gilt ganz besonders zu den Konditionen befristeter Projektstellen und der oft an die Vergütung im Öffentlichen Dienst angeglichenen Bezahlung. Eine Einschulung auch begabter Nachwuchskräfte für solche Aufgaben ist kurzfristig schwierig, so rechnet man etwa in der Marktforschung häufig mit etwa zwei Berufsjahren, bis vom Mitarbeiter selbständig auch die Kontakte nach außen übernommen werden können. Aus der Sicht der Bewerber mag diese Zeitvorstellung übertrieben vorsichtig erscheinen; man muß aber auch sehen, daß der Vorgesetzte bzw. Institutsinhaber bei einer «Verärgerung» eines wichtigen Auftraggebers ein erhebliches Risiko eingeht, das bei kleineren Instituten auch die eigene wirtschaftliche

Existenz gefährden kann. Es mag aber auch unsachliche Gründe für diese Skepsis geben; so wertet die Betonung der besonderen Schwierigkeit des Projekt-Managements die Leistung der «erfahrenen» Kräfte gegenüber den Anfängern auf, was in Anbetracht der eher zunehmenden fachlich-wissenschaftlichen Kompetenz der Universitätsabsolventen die erheblichen Einkommensunterschiede zwischen diesen beiden Gruppen subjektiv rechtfertigt. Ob sachlich voll berechtigt oder übertrieben, in jedem Fall muß man mit dem Bestehen dieser Leistungseinschätzung durch potentielle Arbeitgeber rechnen und bei der eigenen Stellensuche und Karriereplanung beachten.

Plötzliche Kündigungen der Projektmitarbeiter können die gesamte Zeitplanung und auch den Kostenrahmen (durch die dann erforderlichen Einschulungsmaßnahmen) sprengen; vertragliche Vereinbarungen sind dagegen nur bedingt wirksam, da ein an seiner Arbeitsstelle nicht mehr tätig sein wollender Mitarbeiter zu einer massiven Störquelle des gesamten Projektes werden kann.

### Organisationsaufgaben

Aus der Sicht der Projektmitarbeiter ist die Arbeitssituation an kleinen Instituten oft problematisch:

- Der Arbeitsplatz wird als unsicher empfunden, vor allem bei befristeten Arbeitsverträgen; Stellenangebote von außen auch während laufender Projekte erscheinen daher besonders attraktiv;
- Die Beschäftigungsdauer sowie die häufige Beschränkung der Arbeit auf typische Spezialistentätigkeit beschränken die Chance zur persönlichen Weiterqualifikation und reduzieren die Möglichkeiten der Befriedigung von Gestaltungsbedürfnissen;
- Es gibt wenig bzw. keine Aufstiegsmöglichkeiten innerhalb des Projektes und später (generell bei kleineren Instituten). Bei Projekten an Universitäten oder vergleichbaren Einrichtungen kann die Projektstelle ein guter Einstieg in ein längerfristig angelegtes Beschäftigungsverhältnis sein, ebenso bei größeren privaten Firmen. Hat aber das beschäftigende private Institut nur zwei oder drei Akademiker als Mitarbeiter, ist eine Beschäf-

tigung über das Projekt hinaus unwahrscheinlich – es sei denn, man ist so gut, daß es aufgrund der eigenen Projekt-(Akquisition-) Fähigkeit zu einer erheblichen Umsatzerweiterung kommt.

- Es gibt wenig Unterstützung durch den Arbeitgeber im Hinblick auf Weiterbildungsmöglichkeiten, die nicht unmittelbar für das konkrete Projekt relevante Kompetenzsteigerungen erbringen.

### Positive Aspekte

Bei diesen Aufstellungen wurden jeweils nur die Negativaspekte betont. Es gibt auch erhebliche positive Anreize durch die Art der Aufgabenstellung, zum Beispiel

- Hohe Identifikationsmöglichkeit mit der eigenen Arbeit als Folge eines überschaubaren Projektes mit erkennbarer eigener Teilleistung;
- Vor allem für Anfänger oft Tätigkeitsanforderungen, die deutlich über den bei Studienabschluß erreichten Fachkompetenzen liegen und entsprechend fordern;
- Die oft hohe Identifikation aller Beteiligten mit den Projektaufgaben und das Wissen, nur bei gemeinsamer Anstrengung wirklichen Erfolg erreichen zu können, führen unter günstigen Voraussetzungen (soziale Kompetenz der Beteiligten, Führungsstil) zu einem sehr angenehmen Betriebsklima innerhalb des Teams.
- Bei guter persönlicher Kompetenz nicht geringe Chancen, allmählich immer selbständigere Bereiche zugewiesen zu bekommen und schließlich selbst Leistungsaufgaben zu übernehmen; diese Chance wird aber meistens nicht vom Vorgesetzten an den Mitarbeiter herangetragen, sondern muß von diesem (manchmal mühsam) erarbeitet werden

Für die Personalführung selbst werden spezielle Instrumente wie sie etwa für Großunternehmen entwickelt wurden (siehe dazu etwa Gabele und Oechsler, 1984; Zander, 1982; Stroebe und Stroebe, 1984) kaum relevant werden, da der unmittelbare persönliche Kontakt formalisierte Instrumente (zum Beispiel schriftliche Personalbeurteilungen) ersetzt. Man sollte sich aber wechselseitig bemühen, die berechtigten Be-

dürfnisse des Partners zu erkennen, auch emotional zu akzeptieren und bei der Durchsetzung der eigenen Ziele so weit wie möglich zu berücksichtigen. Die besonderen Arbeitsbedingungen an kleinen privaten Instituten sollte man bei der persönlichen Stellensuche stark beachten. Viele Absolventen, die in einer stark formell gegliederten großen Organisation sich zu sehr eingeschränkt und zu wenig in ihrer Individualität beachtet fühlen, können bei solchen kleinen Arbeitgebern mehr Entfaltungsmöglichkeiten finden. Wer Sicherheit, starke Arbeitsteilung oder hierarchischen Aufstieg sucht, fühlt sich vermutlich in großen Institutionen wohler.

### 6.1.2 Kontrolltechniken des Projektverlaufes und des Konfliktmanagements

Größere Evaluationsprojekte mit längerer Laufzeit können auch von «genialen» Projektleitern nicht ohne eine regelmäßige Rückmeldung über die Einhaltung von

- Zeitplanung
- Kostenrahmen
- Vereinbarte Qualitätsstandards

durchgeführt werden.

Relativ einfach ist dies bzgl. des Zeitablaufes, wenn bei der Projektplanung ein detaillierter Zeitplan (vgl. Diagramm V/6) erarbeitet wurde. In diesem Fall muß man nur sicherstellen, daß die einzelnen Teilschritte nicht über Gebühr überzogen und dadurch die Pufferzeiten zu früh aufgebraucht werden. Ein wenig schwieriger ist es für den «nur-sozialwissenschaftlich» ausgebildeten Projektleiter, sich auch ein sachgerechtes Bild über die tatsächlich entstandenen Kosten zu erarbeiten. Zwar wird es im Regelfall nicht notwendig sein, umfangreiche formalisierte Kontrolltechniken wie bei Wirtschaftsprojekten üblich (vgl. dazu Ziegenbein, 1984; Bramse-mann, 1978; Wöhe, 1982) einzusetzen, aber eine regelmäßige (evtl. wöchentliche) Gegenüberstellung der bis zum jeweiligen Arbeitsschritt geplanten Ausgaben, der durch Verträge eingegangenen Zahlungsverpflichtungen und den bereits tatsächlich verausgabten Beträgen sollte selbstverständlich sein. Selbst innerhalb

des jeweiligen Projektbudgets, also ohne Berücksichtigung der zugeschlüsselten Gemeinkosten, gar nicht so seltene Fehler sind:

- Keine Berücksichtigung der Lohnnebenkosten (vgl. Kapitel V, Einleitung)
- Kalkulation der Projektarbeiten auf der Basis eines 52-Wochen-Jahres; im Durchschnitt ist wegen Urlaubs- und anderen Ausfallzeiten der deutsche Arbeitnehmer nur 43 Wochen im Jahr produktiv tätig; werden etwa Urlaubszeiten bei der Planung nicht beachtet, kann dies zu einem ganz erheblichen Kostenfaktor (Ablösung des Urlaubsanspruches bzw., sofern überhaupt möglich, die Bezahlung von Ersatzkräften) werden;
- Keine Reserven für die Überbrückung von Ausfällen (vorzeitige Kündigungen, Mutterschaftsurlaub, längere Erkrankungen);
- Keine rechtzeitige Einplanung von (inflationsbedingten) Lohn- und Preissteigerungen.;

Besondere Probleme können auftreten, wenn mit dem Auftraggeber kein Festpreis, sondern zumindest in Teilen der Ersatz der tatsächlichen Aufwendungen verabredet wurde. Dies kann zum Beispiel dann erforderlich sein, wenn sich die Kosten mancher Arbeitsschritte in der Planungsphase nicht realistisch abschätzen lassen oder wesentlich von dem späteren Verhalten des Auftraggebers selbst abhängen. Beispiele sind etwa die teilweise mühevollen und mit Dienstfahrten verbundenen Einholungen von Zustimmungen von Eltern, Lehrern und Schülern bei Schuluntersuchungen, oder die u.U. langwierigen Kommunikationsprozesse zwischen der technischen Entwicklungsabteilung und den Projektmitarbeitern bzgl. der Erarbeitung einer neuen Produktvariante auf der Basis vorläufiger Projektergebnisse im Falle einer formativen Produktevaluation. In solchen Fällen müssen die Bestimmungen des Auftraggebers (zum Beispiel Höchstbeträge für Übernachtungs- und Verpflegungsspesen, Kilometergeld, anzurechnendes Stundenhonorar) zusätzlich zu der eigenen Kalkulation beachtet werden.

Die laufende Kontrolle der dem Auftraggeber zugesicherten Qualitätsstandards kann in formalisierter Form nur für die eher quantitativen Teile der Projektarbeit (Rücklauf von Fragebögen, Verweigerungs-raten von mündlichen In-

terviews, Ausfälle durch fehlende Zustimmung von zum Beispiel Schulen und Kliniken) durchgeführt werden. Hinweise auf entsprechende Verfahren gibt etwa der Bundesminister für Forschung und Technologie, 1977. Die Sicherung der eher qualitativen Anforderung wie Erfolg der Interviewerschulung, Informationsauschöpfung, sicherstellende Kodierung offener Antworten oder umfassend aussagekräftige Datenauswertung bleibt während der Projektarbeit meist dem subjektiven Eindruck des Leiters, der auf der Basis seiner persönlichen Erfahrung das laufende mit erfolgreich abgeschlossenen Projekten vergleichen kann, überlassen.

Eine Übersicht über einige Kontrolltechniken findet sich im Diagramm V/3 (vgl. Abschnitt 5.1.3).

### 6.1.3 Unerwartete Störfälle

Prinzipiell muß man davon ausgehen, daß nahezu jede denkbare Katastrophe im Verlaufe eines längerfristigen Evaluationsprojektes auch tatsächlich eintreten kann. Typische Beispiele:

*Änderung der Zielsetzung des Projektes;* Ursachen dafür kann ein anderer Verwertungszusammenhang sein (vgl. dazu etwa die Szenario-Technik, Diagramm IV/3), ein Personalwechsel beim Auftraggeber mit entsprechend unterschiedlicher Schwerpunktsetzung (als konkretes Beispiel: Plötzliche Aufgabe des vorher als unverzichtbar bezeichneten Grundsatzes, alle Unternehmensteile als Einheit im Hinblick auf die «cooperate identity» aufzufassen und die PR-Maßnahmen auf dieser Grundlage zu optimieren) oder Entfallen der eigentlichen Untersuchungsgrundlage (etwa bei der formativen Evaluation einer Werbekonzeption, wenn von der Konkurrenz ein in jeder Hinsicht überlegenes Produkt überraschend auf den Markt gebracht wird).

*Plötzliches Auftreten nicht bekannter* oder zumindest vom Auftraggeber verschwiegener «Nebenwirkungen»; als ausschließlich der Praxis entnommene Beispiele: Versäumen einer Staatsprüfung einer studentischen Versuchsperson bei der Erprobung der therapeutischen Wirkung eines Psychopharmakons als Folge eines

48-Stunden-Dauerschlafes; eine andere Institution hat die Kooperation mit einem Institut eingestellt, nachdem alle an einem Experiment zur Optimierung von Beipack-Zetteln beteiligten weibliche Versuchspersonen am nächsten Tag wegen Übelkeit arbeitsunfähig waren, obwohl die verschiedenen Beipack-Varianten ausschließlich mit einem Placebo kombiniert gegeben wurden und damit tatsächliche Medikamentenfolgen ausgeschlossen waren.

*Nicht-vorhersehbare Reaktanz;* so mußte etwa die empirische Evaluation eines Studienganges einer Hochschule abgebrochen werden, weil sich die Studenten des Fachbereiches geschlossen weigerten, als Interviewpartner zur Verfügung zu stehen. Die Ursache war, daß der Untersuchungsleiter als Finanzier des Evaluationsprojektes das zuständige Landes-Wissenschaftsministerium (für die Studenten damals ein Feindbild) und nicht, wie es auch faktisch richtig gewesen wäre, die Hochschule selbst als Auftraggeber genannt hatte.

*Überlappung des Untersuchungsdesigns mit anderen Zielsetzungen;* so fanden sich in einer auf zwei Jahre geplanten Evaluation verschiedener Formen von Arbeitsgestaltung am Ende der Projektphase nur noch zwei Personen in der Bedingungskombination, in der sie aufgrund der Untersuchungsplanung hätten sein sollen, die übrigen waren entweder auf eigenen Wunsch (Schichtwechsel, bevorzugte Zusammenarbeit mit bestimmten Kollegen, vielleicht auch Ablehnung der zufällig zuteilten Arbeitsform) oder aufgrund von gemäß Erfordernissen der Arbeitsorganisation erfolgten Änderungen nicht mehr in der entsprechenden Design-Zelle.

*Das Auftreten inakzeptabler Methodenartefakte;* so wurden in einer Studie zu verschiedenen Möglichkeiten der Wohnungsgestaltung auch untersucht, wie hoch die vermutliche Verweildauer in den einzelnen Wohnungsformen für Personen mit verschiedenem Alter, Lebenslaufzyklus und dgl. sein würde; als Folge des verwendeten multipel-linearen Regressionsansatzes ergab sich für alle Personen mit einem Einzugsalter von über 55 eine negative prognostizierte Verweildauer, was infolge des Fehlens rechtzeitiger Qualitätskontrollen erst sehr

knapp vor dem endgültigen Abgabetermin des Schlußberichtes erkannt wurde.

*Auch absolut Unvorhersehbares passiert.* So wurden etwa Teile der Daten einer Schulsystemevaluation durch einen Brand vernichtet (Aurin, 1986). In einem anderen Projekt verweigerte ein befristet Beschäftigter (und aus anderen Gründen in eine Konfliktkonstellation gedrängter) Mitarbeiter die Herausgabe der von ihm erhobenen Patienteninterviews innerhalb einer Therapieevaluation mit der Begründung, er sehe den Vertraulichkeitsschutz der Patientenangaben durch ungenügende Anonymisierung gefährdet, was zu erheblichen Zeitverzögerungen bei der Projektdurchführung führte.

*Unerklärliches;* so mußte etwa bei einer Vergleichsuntersuchung über die kognitive Leistungsfähigkeit der Studenten verschiedener Universitäten ein bestimmter Teilbereich bei der Auswertung unberücksichtigt bleiben, da die Studenten einer großen süddeutschen Universität gerade in diesem Bereich so extrem schlechte Werte zeigten, daß diese Ergebnisse ganz einfach nicht stimmen konnten; die Ursachenforschung (u. a. Fehlverhalten des Versuchsleiters, bewußte Sabotage der Untersuchung durch Studenten etc.) erbrachte keine akzeptable Erklärung;

Keines der hier genannten Beispiele ist erfunden, sofern kein Literaturzitat angegeben ist, kommen sie aus der persönlichen Projektarbeit der Verfasser oder wurden von absolut zuverlässigen Kollegen mündlich berichtet. Man sieht daran vielleicht die Vielfältigkeit der auftretenden Störungen, vielleicht auch einige Strategien zu ihrer Begrenzung. Das Problem ist, daß die Kombination aller denkbaren Verhinderungsmaßnahmen für Störfälle jede Untersuchung wesentlich verteuern und u. U. so aufwendig machen würde, daß sie nicht mehr in Auftrag gegeben werden könnte. Man wird «Katastrophen» nie voll ausschließen können, was es um so wichtiger macht, mit dem Auftraggeber ein möglichst positives Verhältnis herzustellen. Korrektes, entgegenkommendes und auch gegenüber Verhaltensweisen des Auftraggebers «fehlerfreundliches» Verhalten kann viel dazu beitragen, im Notfall dessen Verständnis für unvorhergesehene Störungen zu finden.

## 6.2 Auswertungsarbeiten

Gegenüber den aus dem Studium vertrauten Seminar- bzw. Diplomarbeiten haben größere Evaluationsprojekte den Nachteil, daß sehr viele Detailarbeiten nicht vom Untersuchten selbst durchgeführt oder auch nur annähernd ausreichend supervisiert werden können. Es sind daher für solche Tätigkeiten Kontrollen erforderlich, die gleichzeitig auch eine weitere Motivationshilfe für sorgfältiges Arbeiten sein können (6.2.1). Für viele, gerade vom Inhalt her besonders interessante Evaluationsprojekte stellt sich auch in besonderem Maße die Frage der Vertraulichkeit, und zwar sowohl im Hinblick auf die Interessen der befragten bzw. beobachteten Probanden als auch des jeweiligen Auftraggebers. Hierbei sind neben einer an sich selbstverständlichen professionellen Ethik auch die Rechtsbestimmungen zu beachten (6.2.2). Prinzipiell nicht lösbar ist das Problem, daß durch die spezifizierte Festlegung der Auswertungspläne immer nur ein Teil der denkbaren Vielfalt an Ergebnissen erarbeitet werden kann, was zumindest bei emotional heftig abgelehnten Ergebnissen leicht zum Vorwurf bewußter Manipulation führen kann (6.2.3.).

### 6.2.1 Kontrolle der Datenqualität

Fehler können sich in jeder Phase der Projektarbeit mit Daten einschleichen:

- Bei der Erhebung: zum Beispiel Probleme mit Interviewern bei mündlicher Befragung oder hohe systematische Antwortverweigerungen bei schriftlichen (postalischen) Erhebungen
- Bei der Datenaufbereitung: Kodierfehler, falsche EDV-Eingabe
- Bei der Auswertung: Verwecheln von Variablennummern oder Codes, falsche (Sub-) Dateien

Eine völlige Sicherheit gegen alle solche projektbedingten Fehler kann nicht garantiert werden, vieles läßt sich aber durch organisatorische Maßnahmen erreichen.

### Datenerhebung durch Interview

Interviewer sollten Arbeitsbedingungen erhalten, die eine ehrliche Datenerhebung nahele-

gen und Verfälschungen zumindest nicht provozieren. Neben einer als zu niedrig empfundenen Bezahlung können sich vor allem folgende Bedingungen negativ auswirken:

- Bezahlung der Interviewer nach «erfolgreich» durchgeführtem Interview ohne faires Entgelt für erfolglose Besuche (Proband nicht angetroffen, Antwort verweigert)
- Keine Vorgabe von Namen und Adressen von zu befragenden Personen (da dann der Interviewer weiß, daß kein nachträgliches Kontrollieren des Stattfindens des Interviews möglich ist); sofern sich dies nicht vermeiden läßt, ist eine Lösungsmöglichkeit das nachträgliche Notieren dieser Daten (aus Vertraulichkeitsgründen unabhängig vom eigentlichen Interviewbogen), dies stößt aber häufig auf erhebliche Reaktanz bei den Befragten (man kann zum Beispiel kaum jemand auf der Straße ansprechen und ihn über seine Meinung etwa zu einer politisch relevanten Maßnahme fragen, ihm Anonymität zusichern, und anschließend Name und Adresse aufschreiben).
- Quotenvorgaben; wenn der Interviewer jeweils bestimmte Prozentsätze von zum Beispiel Geschlecht, Alter und Berufsgruppe befragen muß, kann leicht die Situation entstehen, daß er gegen Ende seiner Datensammlung verzweifelt einen 20jährigen männlichen Altersrentner benötigt. Findet er einen solchen nicht, muß er entweder auf einen Teil der bereits durchgeführten Interviews verzichten (die er dann auch nicht abrechnen kann) oder einen Lösungsweg wählen, den man nicht mehr als «empirische Datenerhebung» bezeichnen kann.
- Unbefriedigende Arbeitsbedingungen, fehlerbehaftete Adressenvorgabe (ein hoher Anteil von nicht auffindbaren Personen), schlecht aufgebaute Fragebögen (zum Beispiel mit vielen Verweigerungen aufgrund unverständlich formulierter oder in der Reihenfolge ungünstig positionierter Fragen) oder fehlender Spesensersatz für erfolglose Fahrten.
- Massive Interviewer- bzw. Untersucherleitereffekte sind zu erwarten, wenn sich dieser Personenkreis mit bestimmten Evaluationsergebnissen identifiziert (ein Beispiel dafür sind etwa die Gerüchte, daß bei manchen Schulvergleichsuntersuchungen die für

die Testung bzw. Befragung der Schüler eingesetzten Studenten in den von ihnen subjektiv bevorzugten Schulformen gezielt Hinweise für das optimale Ausfüllen gegeben haben sollen).

Neben der Vermeidung solcher organisatorischer Mängel (wobei man auch die Position «sparsamer» Projektleiter sehen muß: häufig werden Projekte auf der Basis von Ausschreibungen vergeben, was zu knapper Kalkulation zwingt!) sollte man sich um eine gute Interviewerschulung bemühen, die meist von den Projektmitarbeitern selbst zu gestalten ist. Diese sollte die Problematik gefälschter Daten im Hinblick auf die Aussagekraft des gesamten Ergebnisses deutlich machen und versuchen, bei den Interviewern auch intrinsische Motivation und Identifikation mit dem Gesamtprojekt zu erreichen. Auch sollte man die Interviewer auf die Durchführung von Kontrollen hinweisen.

Ein auch bei massiver Schulung nicht völlig ausschließbares Problem sind Interviewereinflüsse (für die dadurch ausgelösten Effekte vgl. etwa König, 1972; Roth, 1984, S. 150 ff.). Wenn es die Größe des Projektes zuläßt, kann man durch ein bewußte Streuung der Interviewermerkmale einen Ausgleich der einzelnen Störeffekte anstreben, jedenfalls ist es ungünstig, einen einzelnen Interviewer zu große Fallzahlen zuzuordnen (in der Praxis relativ bewährt haben sich etwa 10, bei größeren Studien ausnahmsweise auch 15 Interviews als Höchstgrenze). Daraus ergeben sich natürlich erhöhte Aufwendungen für Interviewer-Suche und Schulung, die man aber in Kauf nehmen sollte.

### Interviewerkontrolle

Besonders aussagekräftig sind nachfassende Kontrollen bei den Probanden. Dies sollte nicht unbedingt als «Kontrollanruf» erkennbar sein, günstiger ist ein Anruf bei einer (angeblich) befragten Person mit der Bitte um Aufklärung eines Details oder einer Beurteilung des Interviewerverhaltens. Ein solches Vorgehen ist natürlich nur bei Interviews mit bekannten Probanden, etwa bei vorgegebenen Adressen aus einer Stichprobe, möglich.

Eine zumindest heuristisch wertvolle Hilfe kann auch die Datenanalyse der abgegebenen Interviews sein. Neben Plausibilitätskontrollen (siehe weiter unten) kann auch die Prüfung der Homogenität (erfaßt durch die Varianz) der von einem Interviewer gelieferten Daten im Vergleich zur Gesamtstichprobe (bzw. nach strukturellen Merkmalen ähnlichen Datensätzen) interessante Hinweise geben. Ist die Datenvarianz deutlich geringer, liegt zumindest den Verdacht nahe, daß der Interviewer relativ häufig seine eigene Vormeinung (unmittelbar oder durch Intervieweinflüsse) in die Ergebnisse hat einfließen lassen oder entgegen den üblichen Anweisungen bevorzugt seinen engeren Bekanntenkreis befragt hat, was ebenfalls leicht zu relativ homogenen Antworten führt. Solche Befunde sind aber selbstverständlich nur Hinweise, die nicht zu einem (möglicherweise ungerechtfertigten) Vorwurf führen sollten, sondern vielleicht zu einer gezielten Nachfrage.

### Schriftliche Befragung

Bei der postalischen Befragung sind einige Störeffekte noch massiver als beim persönlichen Interview; so weiß man etwa nie genau, wer den Fragebogen tatsächlich ausfüllt, ob es sich um eine «Gruppenarbeit» der ganzen Familie handelt, und ob nicht einige Fragen mißverständlich formuliert waren. Aus dem letzteren Grund muß man bei der Fragebogenkonstruktion besonders sorgfältig vorgehen.

Das Hauptproblem der postalischen Befragung ist sehr häufig die Rücklaufquote. Es gibt einige Tricks, um schon durch die Gestaltung des Materials diese zu erhöhen (zum Beispiel das Antwortcouvert nicht freistempeln lassen, sondern mit einer Briefmarke versehen), doch wird man trotzdem häufig darauf angewiesen sein, bei den nicht antwortenden Befragten nachzufassen. Dies wird wesentlich dadurch erleichtert, wenn man am Posteingang erkennen kann, welche Personen aus der ursprünglichen Adressenliste geantwortet haben. Diese Identifikationsarbeit kann aber erhebliche Schwierigkeiten machen, wenn die Fragebögen aus Gründen der Anonymität keine Namensangaben tragen.

Einige Lösungsmöglichkeiten:

- Die angeschriebenen Personen werden gebeten, bei absoluter Zusicherung der Anonymität der Auswertung den Absender auf dem Couvert (nicht auf dem Fragebogen) zu vermerken.
- Bei schriftlichen Befragungen in Organisationen wie zum Beispiel Unternehmen werden auf den Fragebögen gerne vor dem Austeilen Code-Nummern vermerkt, die eine Identifikation der jeweiligen Abteilung ermöglichen; selbstverständlich muß dies offiziell geschehen und mit dem Betriebsrat geklärt sein!
- Bei manchen schriftlichen Befragungen findet sich auf dem Rückantwortcouvert eine Adresse, die eine «Projektnummer» enthält; diese Projektnummer ist individuell für jede befragte Person (ein entsprechender Ausdruck, zum Beispiel auf Klebeetiketten oder Endlos-Couverts ist EDV-mäßig unproblematisch) und ermöglicht somit eine Identifikation der antwortenden Personen. Allerdings ist ein solches Vorgehen rechtlich problematisch. Auch fördert es nicht gerade das Vertrauen der Befragten in die Anonymitätsgarantie, wenn mehrfach mit der Begründung «sie haben noch nicht geantwortet» nachgefragt wird, ohne daß den Betroffenen der Kontrollmechanismus für die eingegangene bzw. fehlende Antwort ersichtlich ist. Wertvolle Hinweise für eine rechtlich saubere, allerdings manchmal auch umständliche Regelung finden sich in Lecher, 1988.

### Fehler bei der Datenaufbereitung

Codierarbeiten sind lästig, erfordern aber oft eine erhebliche allgemeine Intelligenz, so daß man neben speziell mit solchen Tätigkeiten beschäftigten Büropersonal bei unregelmäßigem Anfall von Arbeiten auf gutes Aushilfspersonal zurückgreifen muß. Von Mitarbeitern, die nur kurzfristig, ohne Kenntnis der Projektzusammenhänge und oft auch ohne soziale Bindungen an das Projektteam beschäftigt werden, kann man keine ausgeprägte intrinsische Motivation für die Durchführung der (ohnedies meist sehr langweiligen) Codierarbeiten erwarten. Leider ist oft eine erhebliche Fehlerquote zu beobachten, vor allem bei (niedriger) Bezahlung pro durchgearbeitetem Fragebogen. Aus

diesem Grund müssen zur Fehlerkorrektur unbedingt Kontrollen erfolgen. Ideal, aber nur selten finanzierbar sind Doppelarbeiten (das gleiche Ausgangsmaterial wird von zwei Hilfskräften codiert, die Ergebnisse werden verglichen). Eher realisierbar ist eine die Sorgfalt unterstützende Entlohnung, insbesondere die Bezahlung nach Zeit anstatt nach Stück (woraus ein gewisses Eigeninteresse entsteht, die Arbeit nicht allzu schnell zu erledigen) sowie Kontrollen mit einer entsprechenden (auch finanziellen) Verrechnung gefundener Fehler.

Prinzipiell Gleiches gilt für Dateneingabearbeiten an die EDV. Hier ist der Leistungsunterschied zwischen professionellen Mitarbeitern und Aushilfspersonal besonders gravierend (geschulte Kräfte haben etwa die doppelte Geschwindigkeit von ungeschulten), auch die Sorgfalt ist bei Spezialisten höher.

Die früher vor allem im kommerziellen Bereich übliche Gewohnheit, zumindest alle wichtigen Daten über einen «Prüflocher» doppelt einzugeben, ist im Zusammenhang der unmittelbaren Bildschirmeingabe zumindest bei sozialwissenschaftlichen Projekten deutlich zurückgegangen. Ein leider nur partieller Ausweg sind die unten besprochenen Plausibilitätskontrollen an der EDV.

Leider werden die durch Codier- und Dateneingabefehler verursachten Verzerrungen der Ergebnisse von akademisch gebildeten Anfängern meist unterschätzt.

Plausibilitätskontrollen gehen heute nach den professionellen Standards jeder Datenanalyse voraus. Das Grundprinzip ist, daß man sich zunächst überlegt, welche Konfigurationen in den Daten einzelner Personen außerordentlich unwahrscheinlich sind. Meist nimmt man dazu objektive Merkmale, wie Alter, Geschlecht, Art und Dauer der Ausbildung u. ä. Für Evaluationsstudien außerhalb des allgemeinen sozialwissenschaftlichen Bereiches wichtig sind auch die Möglichkeiten, unplausible Testergebnisse zu identifizieren (massive Abweichungen zwischen einzelnen Subtests der gleichen Person, höhere Beantwortungsraten für «schwierige» als für leichte Items) oder auf wahrscheinlich verfälschendes Antwortverhalten zu schließen (Anlegen von «Strickmustern» auf den Antwortblättern, in dem zum Beispiel immer in der Reihenfolge der Antwortalternativen A, B, C, D etc. geantwortet wird; Auswahl bestimm-

ter Antwortalternativen wie etwa immer eine Extremkategorie oder die Unentschlossen-Möglichkeit).

Solche Programme mit Plausibilitätsüberlegungen erkennen Interviewer-, Aufbereitungs- oder Eingabefehler um so besser, je umfangreicher das entsprechende Programm gehalten wird. Allerdings steigt gleichzeitig auch die Wahrscheinlichkeit, ohne Vorliegen eines Fehlers auffällige Datenkombinationen auszuwerfen (schließlich kann es ja tatsächlich den 23jährigen Promovierten geben, der «Rentner» als Beruf angegeben hat, da seine Waisenrente die finanzielle Haupteinnahmequelle darstellt). Solche «auffälligen» Personen dürfen daher nicht sofort ausgeschlossen werden, sondern die Ursache für die auffallenden Kombinationen ist über den Vergleich der Dateneingabe mit den schriftlichen Unterlagen, die richtige Codierung der Unterlagen und schließlich im Extremfall durch Nachfrage zur Kontrolle des Interviewverhaltens nachzuvollziehen. Da dies sehr aufwendig ist, steht man vor dem Dilemma, daß mit steigender Sicherheit der Identifikation von Inplausibilitäten der Arbeitsaufwand für die Klärung der einzelnen Befunde ansteigt. Im allgemeinen wird man sich mit um so größeren Plausibilitätskontrollen zufrieden geben können, je sorgfältiger die Kontrolle der einzelnen Arbeitsschritte vorher erfolgt ist.

## Datenauswertung

Die typischen Fehler hier (insbesondere Verwechslungen von Variablennummern, was besonders leicht geschehen kann, wenn verschiedene Projektmitarbeiter Aufträge an die EDV mit unterschiedlichen Variablen-Ordnungen geben) treten seltener auf als bei der Dateneingabe, können dann aber um so gravierender sein. Neben sorgfältiger Arbeit – wozu auch gleich bezeichnete Unterlagen und das «Wegsperrern» von Vorlagen mit nicht mehr gültiger Numerierung der Variablen gehört – ist die «semantische» Plausibilitätskontrolle zu empfehlen. Jedem (wirklich jedem!) Widerspruch in den Ergebnissen ist nachzugehen, auch wenn dies viel Zeit kostet und rückblickend unnötig erscheint. So kann es ja wirklich sein, daß es zwischen Teilgruppen extreme Varianzunterschiede gibt, die Einteilung nach der Kinder-

zahl die Zusammenhänge zwischen der Beurteilung von Sach- und Personalausstattung von Leitenden Angestellten völlig verändert (konkret in einem Projekt passiert) oder gegenüber der Erwartung völlig entgegengesetzte empirische Befunde auftreten. Trotzdem sollte man in allen solchen und ähnlichen Fällen die Auswertung kontrollieren, ggf. von einem damit noch nicht befaßten Mitarbeiter unabhängig von den bisherigen Eingaben neu durchführen lassen.

Ein besonderes Problem sind «selbstgestrickte» Programme für die Auswertung. Zwar hat die weite Verfügbarkeit publizierter, ausgetesteter Programme die projektspezifische Programmierarbeit eingeschränkt, doch ist auch die Publikation keine Garantie für Fehlerfreiheit. Ein negatives Beispiel ist etwa bei Härtner et al., (1979) zu finden, der Programmfehler wurde erst wesentlich später entdeckt (s. Baumert et al., 1988). Allerdings ist bei «eigenen» Programmen die Fehlerchance im Regelfall besonders hoch, weil weniger Zeit für das Austesten (und manche Fehler zeigen sich nur in seltenen Situationen) verbleibt. In einem konkreten Fall (Überprüfung der Nebenwirkungen eines Pharmakons unter Benutzung eines seltenen statistischen Verfahrens) wurde ein Programmfehler nur dadurch entdeckt, daß die «Signifikanzen» so massiv waren, daß diese Unglaubwürdigkeit zu einem Nachrechnen von Hand führte (es war bei Korrekturarbeiten im Programm ein Statement mit einer Divisionsanweisung irrtümlich gelöscht worden). Kleine Ursachen können erhebliche Auswirkungen haben!

### 6.2.2 Datenschutz

Der gerade für den EDV-Einsatz in Behörden und anderen Verwaltungen potentiell besonders gefährdete Schutz der «Privatsphäre» hat eine massive öffentliche Diskussion verursacht, die auch die Möglichkeiten in Forschungsprojekten wesentlich beeinflusst. Die Wahrung des Persönlichkeitsschutzes ist ein wichtiges Anliegen, gerade auch in sozialwissenschaftlichen Evaluationsprojekten, sollte aber nicht zu einer sachlich nicht nachvollziehbaren Behinderung der Projektarbeit führen, was leider gelegentlich vorkommt. Für einen Juristen zum

Beispiel im Öffentlichen Dienst können aus einer irrtümlichen oder wegen seiner persönlichen Einschätzung «zu großzügigen» Genehmigung von Vorhaben deutliche Nachteile entstehen, die Verhinderung von Projekten bleibt für ihn selbst meist völlig folgenlos. Bei einer solchen Reinforcementstruktur führen schon einfachste subjektive Nutzenüberlegungen (ausführlicher bei Wottawa und Hossiep, 1987, S. 48 ff.) zu einer Bevorzugung restriktiver Auffassungen.

Bei Laien bestehen zum Teil erheblich Mißverständnisse bezüglich des Begriffes «Datenschutz».

Wesentlicher Gesetzesinhalt des Bundesdatenschutzgesetzes (BDSG) ist der Schutz des einzelnen vor den Gefahren der Verbreitung von Daten, die sich auf Personen beziehen. Dabei bezieht es sich, ohne zwischen verschiedenen Arten von Daten zu unterscheiden, auf alle personenbezogenen Daten und regelt die Datenverarbeitung von Behörden und Privatunternehmen.

Im ersten Abschnitt des Gesetzes werden Begriffsbestimmungen, Zuverlässigkeitsvoraussetzungen für die Datenverarbeitung und Rechte der Betroffenen geregelt (§1–§6 BDSG).

Der zweite Abschnitt befaßt sich mit Datenverarbeitung durch öffentliche Verwaltungen (§7–§21 BDSG). Als Überwachungseinrichtung ist der Bundesbeauftragte für den Datenschutz vorgesehen.

Für die Datenverarbeitung im privaten Bereich, insbesondere im Verhältnis zwischen Arbeitgeber und Arbeitnehmer, gilt der dritte Abschnitt (§22–§30 BDSG), während der vierte Abschnitt Sondervorschriften für solche Unternehmen enthält, die Daten für Dritte verarbeiten (§31–§40 BDSG). Der fünfte und sechste Abschnitt des Gesetzes enthalten die Regelungen von Sanktionen sowie Übergangs- und Schlußvorschriften (§41–§47 BDSG). Einige für die Evaluationsarbeit besonders wichtige Paragraphen finden sich im Diagramm VI/1.

Die Problematik «Datenschutz» dürfte an sich für die meisten Evaluationsprojekte keinerlei Schwierigkeiten bereiten, da eine personenbezogene Datenauswertung im Regelfall nicht erforderlich ist.

## Diagramm VI/1

Auszug aus dem Bundesdatenschutzgesetz (BDSG) (vgl. *Kittner*, 1987)

## §2: Begriffsbestimmung

(1) Im Sinne dieses Gesetzes sind personenbezogene Daten Einzelangaben über persönliche oder sachliche Verhältnisse einer bestimmten oder bestimmbarer natürlichen Person (Betroffener).

(2) Im Sinne dieses Gesetzes ist

1. Speichern (Speicherung) das Erfassen, Aufnehmen oder Aufbewahren von Daten auf einem Datenträger zum Zweck ihrer weiteren Verwendung,
2. Übermitteln (Übermittlung) das Bekanntgeben gespeicherter oder durch Datenverarbeitung unmittelbar gewonnener Daten an Dritte in der Weise, daß die Daten durch die speichernde Stelle weitergegeben oder zur Einsichtnahme, namentlich zum Abruf bereitgehalten werden,
3. Verändern (Veränderung) das inhaltliche Umgestalten gespeicherter Daten,
4. Löschen (Löschung) das Unkenntlichmachen gespeicherter Daten, ungeachtet der dabei angewendeten Verfahren.

(3) Im Sinne dieses Gesetzes ist

1. speichernde Stellen jede der in §1 Abs. 2 Satz 1 genannten Personen oder Stellen, die Daten für sich selbst speichert oder durch andere speichern läßt,
2. Dritter jede Person oder Stelle außerhalb der speichernden Stelle, ausgenommen der Betroffene oder diejenige Personen und Stellen, die in den Fällen der Nummer 1 im Geltungsbereich dieses Gesetzes im Auftrag tätig werden,
3. eine Datei eine gleichartig aufgebaute Sammlung von Daten, die nach bestimmten Merkmalen erfaßt und geordnet, nach anderen bestimmten Merkmalen umgeordnet und ausgewertet werden kann, ungeachtet der dabei angewendeten Verfahren. Nicht hierzu gehören Akten und Aktensammlungen, es sei denn, daß Sie durch automatisierte Verfahren umgeordnet und ausgewertet werden können.

## §3: Zulässigkeit der Datenverarbeitung

Die Verarbeitung personenbezogener Daten, die von diesem Gesetz geschützt werden, ist in jeder in §1 Abs. 1 genannten Phasen zulässig, wenn

1. dieses Gesetz oder eine andere Rechtsvorschrift Sie erlaubt oder

2. der Betroffene eingewilligt hat. Die Einwilligung bedarf der Schriftform, soweit nicht wegen besonderer Umstände eine andere Form angemessen ist, wird die Einwilligung zusammen mit anderen Erklärungen schriftlich erteilt, ist der Betroffene hierauf schriftlich besonders hinzuweisen.

## §10: Datenübermittlung innerhalb des öffentlichen Bereichs

(1) Die Übermittlung personenbezogener Daten an Behörden und sonstige öffentliche Stellen ist zulässig, wenn sie zur rechtmäßigen Erfüllung der in der Zuständigkeit der übermittelnden Stelle oder des Empfängers liegenden Aufgaben erforderlich ist. Unterliegen die personenbezogenen Daten einem Berufs- oder besonderem Amtsgeheimnis (§45 Satz 2 Nr.1, Satz 3) und sind Sie der übermittelnden Stelle von der zur Verschwiegenheit verpflichteten Person in Ausübung ihrer Berufs- oder Amtspflicht übermittelt worden, ist für die Zulässigkeit der Übermittlung ferner erforderlich, daß der Empfänger die Daten zur Erfüllung des gleichen Zweckes benötigt, zu dem Sie die übermittelnde Stelle erhalten hat.

## §36: Verarbeitung personenbezogener Daten zum Zweck der Übermittlung in anonymisierter Form

(1) Die in §31 Abs. 1 Satz 1 Nr. 2 genannten Personen, Gesellschaften und andere Personenvereinigungen sind verpflichtet, die gespeicherten personenbezogenen Daten zu anonymisieren. Die Merkmale, mit deren Hilfe bestimmte anonymisierte Daten derart verändern verändert werden können, daß Sie sich auf eine bestimmte Person beziehen oder eine solche erkennen lassen, sind gesondert zu speichern. Diese Merkmale dürfen mit den anonymisierten Daten nicht zusammengeführt werden, es sei denn, daß die dadurch ermöglichte Nutzung der Daten noch für die Erfüllung des Zweckes der Speicherung oder zu wissenschaftlichen Zwecken erforderlich ist.

(2) ...

(3) Bei automatischer Datenverarbeitung ist die Durchführung der in Absatz 1 vorgesehenen Maßnahmen durch entsprechende Vorkehrungen sicherzustellen.

Die European Science Foundation hat bestimmte Grundsätze und Richtlinien für die Verwendung personenbezogener Daten zu Forschungszwecken herausgegeben. Einige Hauptpunkte sollen an dieser Stelle extrakthalt wiedergegeben werden (nähere Erläuterungen bei Lecher, 1988, S. 28 ff.):

- Jede Verarbeitung personenbezogener Daten für Forschungsziele setzt eine ausdrücklich gesetzliche Ermächtigung oder die informierte Einwilligung des Betroffenen voraus.
- Eine informierte Einwilligung liegt vor, wenn die Betroffenen ausdrücklich und eindeutig darüber aufgeklärt worden sind:
  - a) daß die Datenerhebung freiwillig ist;
  - b) daß die Weigerung der Datenerhebung keinerlei Konsequenzen hat
  - c) welchen Zweck das Forschungsprojekt verfolgt;
  - d) wer Auftraggeber der Erhebung ist bzw. wer erhebt;
- Die für Forschungszwecke zur Verfügung gestellten personenbezogenen Daten sollten nicht für andere Zwecke verarbeitet werden, insbesondere nicht für solche, die später für den Betroffenen Konsequenzen haben.
- Die Leiter von Forschungsprojekten, die mit der Verarbeitung personenbezogener Daten verbunden sind, sollten die Verantwortung dafür tragen, daß die gemachten Angaben den aktuellen Datenschutz-Vorschriften auch dem momentan herrschenden Technikstand entsprechen.

Es darf aber nicht übersehen werden, daß es auch datenrechtlich problematische Forschungsstrategien gibt. In der Psychologie gibt es Projekte, bei denen eine volle Aufklärung der Beteiligten erst nach ihrem Abschluß erfolgen kann. In solchen Fällen muß diese Bedingung, in Verbindung mit dem Angebot des jederzeit möglichen Rücktritts, den Betroffenen ausdrücklich erklärt und von ihnen akzeptiert werden.

Probleme können auch entstehen, wenn zum Beispiel aus Gründen einer Längsschnittuntersuchung (mehrfache Befragungen bei der gleichen Person) eine Identifikation der Datensätze gesichert werden muß. Hier liegt die ein-

wandfreieste Methode in der Einholung der Zustimmung der Befragten, doch kann diese in Folge von Befürchtungen der Probanden die Verweigerungsraten deutlich erhöhen.

### Schutz der Privatsphäre

Im Gegensatz zu den Datenschutzbestimmungen im engeren, gesetzlichen Sinn kann der Schutz der Intimsphäre sowie das Problem des «Geheimnisverrates» in Evaluationsstudien sehr relevant werden. Wie man an der Darstellung einiger wichtiger Rahmenbedingungen in Diagramm VI/2 entnehmen kann, ist die Situation für «Berufspsychologen mit abgeschlossener Ausbildung» und andere Sozialwissenschaftler verschieden. Dies kann zu der etwas paradoxen Situation führen, daß ein als Interviewer eingesetzter Sozialwissenschaftler dann unter §203 StGB fällt, wenn der Leiter des Evaluationsprojektes ein Diplom-Psychologe ist, diese zusätzliche Strafbewährung des Schutzes der Intimsphäre aber nicht besteht, wenn der Projektleiter eine andere Vorbildung hat. Dies entspricht in keiner Weise der Intention des Gesetzes, mit diesem Paragraphen sollte an sich die Arbeit der genannten Berufsgruppen erleichtert werden, da es zum Beispiel im Rahmen psychologischer Therapien sicher leichter fällt, eine offene Gesprächsatmosphäre zu erzeugen, wenn der Klient zusätzlich zum persönlichen Vertrauen auch noch auf eine rechtliche Absicherung der Vertraulichkeit der Gesprächsinhalte bauen kann.

### Datenschutz auch bei Änderung der Rahmenbedingungen!

Die Vertraulichkeit von Informationen ist auch ohne böse Absicht besonders dann gefährdet, wenn die befragte bzw. beobachtete Person und der Projektmitarbeiter untereinander verbundenen Sozialnetzen angehören (zum Beispiel Befragung von Studenten durch Studenten des gleichen Faches, Befragungen in räumlicher Nachbarschaft). Der Versuchung zur Verbreitung personenbezogener «Anekdoten» sollte man stets mit dem Hinweis auf die Rechtslage gemäß Diagramm VI/2 schon bei der Schulung der Mitarbeiter entgegenreten.

## Diagramm VI/2

## Wichtige rechtliche Rahmenbedingungen für Psychologen

**Beachtung der Persönlichkeitsrechte der Probanden**

## Grundgesetz (GG)

Artikel 1: Schutz der Menschenwürde

Artikel 2: Freiheitsrechte

Artikel 5: Recht der freien Meinungsäußerung

Artikel 12: Freiheit der Berufswahl

Psychologische Untersuchungen dürfen in dem Umfang geschehen, wie Sie der Würde des Menschen angemessen sind. Das Eindringen in den persönlichkeitsrechtlichen Bereich ist immer dann unproblematisch, wenn es durch die freie Selbstbestimmung des Probanden ermöglicht wird.

## Erhebung im Rahmen eines Dienstvertrages

- Verletzungen der Haupt- und Nebenpflichten des Dienstvertrages führen allgemein zu zivilrechtlichen Sanktionen (z. B. Schadensersatzansprüche)

## Zu beachten sind weiterhin

- Berufsgeheimnis §53 StPO
- Geheimnisverletzung §203 StGB

## Rechtliche Situation des Beurteilens und Bewertens in der Berufsausbildung, hierzu: Betriebsverfassungsgesetz (BetrVG)

- Mitbestimmungsrechte des Betriebsrates beim Beurteilen und Bewerten im Betrieb: §94, §95, §98 BetrVG, §85 BetrV
- Mitwirkungsrecht der Jugendvertretung über Beurteilungsbögen: §66–68, §70, §80 BetrVG

Relativ schwer zu lösen ist das Problem veränderter Zugangsberechtigungen zu Aktenmaterial. Typisch ist hier etwa die Situation in Universitäten, wo in Zusammenhang mit der Besetzung von Professorenstellen externe Gutachten eingeholt werden müssen, die nur dann wirklich aussagekräftig sind, wenn der Gutachtersteller von vertraulicher Behandlung ausgehen kann (bei allen anderen Regelungen würde eine erhebliche Gefahr von Gefälligkeitsgutachten, evtl. auf wechselseitiger Basis, entstehen). Dieses Material wird gemeinsam mit den anderen Personalunterlagen üblicherweise im Dekanat gespeichert, die Bewerber haben dazu keinen Zutritt. Nach Ernennung durch den Minister (es kann ohne weiteres sein, daß nicht der in den Gutachten besonders gut abschneidende Erstplazierte, sondern vielleicht ein gerade noch akzeptabel erscheinender Kandidat schließlich die Position erhält) dauert es oft nicht lange, bis der oder die «Neue» zum Dekan gewählt wird oder eine andere, Zugang zu den Dekanatsunterlagen bedingende Funktion in der akademischen Selbstverwaltung übernimmt. Damit automatisch verbunden ist die Zugangsmöglichkeit auch zu Personalunter-

lagen. Es erfordert ein sehr hohes Maß der Verinnerlichung abstrakter Normen, nicht in den die eigenen Person betreffenden Vorgängen zu schmökern. Strukturell Gleiches gilt für den hierarchischen Aufstieg in Verwaltungen, Schulen, Kliniken und Unternehmen. Für Evaluationsprojekte wichtig kann dieses strukturelle Problem werden, wenn die Bewertung von Maßnahmen eng mit Personen verknüpft ist. Dies trifft zum Beispiel bei Untersuchungen zur Bewertung verschiedener Weiterbildungsseminare zu, wo es bei unzureichender Anonymisierung der Unterlagen durchaus passieren kann, daß nach einem beruflichen Aufstieg einer der bewerteten Seminaranbieter mit Interesse liest, was seine jetzigen Untergebenen damals an Kritikpunkten genannt haben. Überlegungen dieser Art sprechen übrigens dafür, in Zweifelsfällen eher externe Evaluatoren heranzuziehen, als die entsprechenden Informationen firmenintern auszuwerten.

Ein theoretisch bestehendes, nach den verfügbaren allgemeinen Informationen aber nicht allzu gravierendes Problem dürfte der Geheimnisverrat in Kombination mit finanziellen Interessen sein. Man muß an diese Mög-

lichkeit aber bei wirtschaftsorientierten Evaluationsstudien (etwa formative Produkt-evaluation, alle im Zusammenhang mit Marketing-Maßnahmen durchgeführten Studien) denken. Man kann trotz dem Fehlen entsprechender fundierter Gerüchte nicht ausschließen, daß Dinge dieser Art vorkommen (vermutlich hätten im konkreten Fall alle Beteiligten ein erhebliches Interesse, den Vorgang möglichst nicht allzu publik werden zu lassen). Andererseits kann aber angenommen werden, daß das bei sozialwissenschaftlich gestützter Evaluation erhobene Material selten einen so hohen Wert hat, daß sich das Risiko lohnt. Die effektiven informellen Kommunikationskanäle zwischen den Unternehmen machen eine mittelfristige Aufdeckung eines solchen Verhaltens sehr wahrscheinlich, und selbst bei Vermeidung einer Strafanzeige dürfte es sehr schwierig sein, in diesem Berufsfeld später noch Karriere zu machen.

### 6.2.3 Auswertungspläne

Es ist bei einem realistisch großen sozialwissenschaftlichen Evaluationsprojekt so gut wie ausgeschlossen, sämtliche denkbaren Auswertungen durchzuführen. Nimmt man als Beispiel eine Arbeit mit 40 berücksichtigten Variablen (was vor allem bei Verwendung von Codierungsschemata für Interviews oder Verhaltensbeobachtungen, noch stärker als bei Fragebögen, eine sehr kleine Zahl ist), ergeben sich bereits 780 mögliche, an der EDV mit geringem Zeitaufwand berechenbare Korrelationen. Da man jetzt aber jede Variable (als mindestens zweistufigen) Moderator nehmen kann, was im übrigen auch für die wichtige Identifikation von Wechselwirkungen sehr sinnvoll sein kann, bekommt man selbst bei nur einer zweistufigen Einteilung potentieller Moderatoren weitere 29640 Korrelationen. Geht man gar dazu über, mehrere Variablen gleichzeitig als Moderatoren zu verwenden (zum Beispiel die Untersuchung der Fragestellung, ob die Kombination einer bestimmten Alters- und Berufsgruppe einen von der Gesamtstichprobe verschiedenen korrelativen Zusammenhang zwischen einer bestimmten Form der Arbeitsgestaltung und der Arbeitszufriedenheit zeigt), wird mit Leichtigkeit die Millionengrenze für

die Anzahl der statistischen Ergebnisse überschritten. Selbst wenn man den unter methodischen Gesichtspunkten natürlich unsinnigen Trick wählt, an der EDV nur die «sehr signifikanten» Resultate ausdrucken zu lassen, erhält man auch bei rein zufälligem Datenmaterial in diesem Fall ca. 10 000 «statistisch auffällige» Einzelbefunde. Es ist selbstverständlich, daß eine solche Auswertungsstrategie absolut unsinnig wäre.

### Rechtzeitige Auswahl der Auswertungsdetails

Die einzige Alternative zu einer allumfassenden Auswertung ist, von Anfang an gezielt einige der nahezu unendlich vielen Auffälligkeiten als potentiell erwartbar auszuwählen. Dies hat zur Folge, daß man:

- Nur jene Variablen in der Auswertung berücksichtigt, die man subjektiv für «wichtig» hält.
- Vorher mit dem Auftraggeber abklären muß, ob es bzgl. der vermuteten Wichtigkeit von Effekten die gleichen Meinungen hat.
- Sehr sorgfältig prüft, ob man mit dieser Strategie nicht anstelle einer objektiven Befunderhebung nur die eigenen Vorurteile bestärken möchte.
- Stets der Gefahr ausgesetzt ist, daß andersdenkende Leser des Evaluationsberichtes die gezielte Manipulationsabsicht bei der Datenauswertung unterstellen, auch wenn dies im Einzelfall nicht gegeben sein sollte.

Die Gefahr der Bestätigung der Vormeinung entsteht dadurch, daß man eben nur jene Effekte ausweist, die man von Anfang an vermutet hat. So kann man zum Beispiel bei einer Schulsystemuntersuchung die Klassengröße, Details des Lehrerverhalten etc. erheben, ausschließlich mit dem Ziel, diese Variablen als Kovariate zur präziseren Herausarbeitung des «eigentlich wichtigen» Systemunterschiedes heranzuziehen. Man wird in Anbetracht der bei ausreichender Trennschärfe stets widerlegbaren Nullhypothese (vgl. dazu Abschnitt 5.3.1) auf diese Weise auch mit guter Chance einen statistisch bedeutsamen Systemunterschied ermitteln. Ebenso gut könnte man aber auch umge-

kehrt vorgehen und prüfen, ob nicht die aus der subjektiven Sicht nur als Störeffekte aufzunehmenden Variablen um vieles bedeutsamer sind als die eigentlich für wichtig gehaltenen, was man entweder direkt (dann aber in Konfundierung mit dem Systemeffekt) oder unter Ausparialisierung des Beitrages der Systemunterschiede machen kann. Schon aufgrund dieser geänderten Reihenfolge der Aufnahme von Erklärungsvariablen in lineare Modellansätze (s. Abschnitt 5.3.2) sind entsprechende Ergebnisunterschiede zu erwarten, die dann interpretativ entsprechend der eigenen Vormeinung besonders herausgearbeitet werden können.

### Antizipatorische Konsensfindung

Eine theoretisch denkbare Lösung wäre, die Auswahl der ausgewerteten Effekte auf wissenschaftliche Erkenntnisse zu gründen, zum Beispiel dann, wenn bereits aus Vorstudien das Auftreten mancher Zusammenhänge besonders plausibel erscheint. Dieser Weg scheitert bei den meisten Evaluationsstudien aber daran, daß sich aus wissenschaftlich-theoretischen Überlegungen oder auch aus der Fülle der Detailergebnisse von Vorstudien für praktisch jede Zusammenhangsanalyse eine zumindest nachträglich sehr plausibel klingende Begründung geben läßt. Es dürfte daher zweckmäßiger sein, die Auswertungsstrategie vorwiegend nutzenbezogen (natürlich nicht unter völligem Verzicht auf die Kenntnis von wissenschaftlichen Vorergebnissen) aufzubauen und nach Möglichkeit rechtzeitig folgende Schritte einzuleiten:

- Versuch einer Konsensbildung vorwiegend mit Wissenschaftlerkollegen, die eine gegenüber den evaluierten Maßnahmen gegenteilige Voreinstellung haben und/oder anderen wissenschaftlichen «Schulen» angehören.
- Mit den von den Evaluationsergebnissen betroffenen Praktikern denkbare Auswertungsergebnisse durchspielen (im Prinzip analog zur Szenariotechnik bzw. Planspielen, vgl. 4.1.3), und mit diesen diskutieren, welche Einwände sie gegen die aus ihrer Sicht unerwünschten Ergebnisse vorbringen würden. Diese Hinweise auf denkbare alternative Erklärungsansätze können gesammelt und in

das geplante Auswertungsprogramm übernommen werden, so daß man deren relative Bestätigung oder Widerlegung bereits mit in den Abschlußbericht aufnehmen kann.

- Vor allem bei öffentlichkeitswirksamen Evaluationsvorhaben nach Möglichkeit rechtzeitig klären, welche Auswertungslücken vor dem Hintergrund aktueller politischer Auseinandersetzungen in der Öffentlichkeit, also insbesondere bei betroffenen Verbänden und Parteien, auf der Basis des vorläufigen Planes gesehen werden; dies setzt allerdings die Schaffung eines ausreichend dichten Netzes informeller Kontakte voraus und sollte im übrigen niemals ohne Abstimmung mit dem Auftraggeber durchgeführt werden.

Nach Durchführung solcher Vorarbeiten hat man zwar einen vertretbaren Auswertungsrahmen, aber meist auch das Problem, daß die Vielzahl der untersuchten Einzelergebnisse kaum noch rezipierbar ist. Dies macht dann eine nachträgliche Beschränkung bei der Berichterlegung erforderlich, in deren Verlauf vieles an guten Vorarbeiten verlorengehen kann.

## 6.3 Berichterlegung

Das letztlich für die Praxis relevant werdende Ergebnis einer Evaluationsstudie ist nicht das, was im Laufe des Projektes von den dort Beteiligten erfahren wurde, auch nicht das, was in den entsprechenden schriftlichen und mündlichen Berichten enthalten ist; relevant werden höchstens jene Informationen, die bei den Adressaten der Berichte ankommen.

Wie groß die Unterschiede zwischen «gesendeter» und «empfangener» Information sein können, zeigt etwa die Arbeit von Czerwenka et al. (1988) zur Bewertung der Schule durch Schüler. Grundlage waren Aufsätze, in denen Schüler verschiedener Klassenstufen und Schulformen in der Bundesrepublik Deutschland einem «Wesen von einem anderen Stern» schildern sollten, was «Schule» ist. Diese Aufsätze wurden dann unter anderem dahingehend ausgewertet, ob Hinweise für ein eher positives oder negatives Lehrerbild enthalten waren, ob die Schule eher Freude oder keine Freude macht usw. Die Resultate, die auch in den schriftlichen

Pressemitteilungen erläutert wurden, zeigten u.a. ein Überwiegen von «wenig Freude» an der Schule gegenüber «Freude» (besonders deutlich etwa in der 11. Klassenstufe, wo insgesamt 13% positiv und 27% negativ waren). Ähnliche Ergebnisse erbrachten die Aussagen in diesen Aufsätzen über die Lehrer. Die Aufregung über diese Arbeit war vor allem bei den Lehrerverbänden enorm, was bei bloßer Durchsicht des Abschlußberichtes für den Wissenschaftler eigentlich unverständlich bleibt. Erklärbar wird die Aufregung, wenn man die Rezeption dieser Ergebnisse in der Presse liest. Beispiele für (meist groß und fett gedruckte) Überschriften: «Schüler klagen über Leistungsdruck», «Schlechte Noten für das Bildungssystem: Deutsche Schüler haben größten Frust» oder gar «Die Lehrer stellen häufig Feindbild dar» (die letzteren Beispiele übrigens aus dem angesehenen «Münchener Merkur»). Noch krasser formulierte die «Abendzeitung»: «Bayerns Schüler: Unsere Lehrer sind Versager!». Solche Beispiele zeigen vielleicht, welchen (absichtlichen oder tatsächlich auf die Rezeption zurückführbaren) Mißverständnissen die Berichtlegung von Evaluationsstudien ausgesetzt sein kann.

Im Abschnitt 6.3.1 wird zunächst darauf eingegangen, daß jeder Berichtlegung eine sorgfältige Zielgruppenanalyse vorausgehen sollte. Dann anschließend (6.3.2) werden Probleme der Informationsverdichtung behandelt, die einerseits unvermeidlich, andererseits aber gerade wegen der üblichen «projektinternen Sprache» leicht manipulativ wirken können. Im letzten Abschnitt (6.3.3) wird auf einige Fragen bei mündlicher Präsentation verwiesen.

### 6.3.1 Zielgruppenanalyse

Während des Studiums ist es üblich, daß man Berichte ausschließlich für ähnlich vorgebildete Personen verfaßt. Typische Konsequenzen sind:

- Ein sehr komplexer, durchschnittliche kognitive Fähigkeiten der Sprachverarbeitung weit überfordernder Satzbau. Für wissenschaftsinterne Kommunikation ist eine solche Struktur oft sinnvoll, für andere Zielgruppen und insbesondere für Konsumenten von Massenmedien aber dysfunktional.

- Die nahezu selbstverständliche Verwendung von Fachausdrücken, wobei weniger das «Fremdwort» für die Berichtlegung an anders vorgebildete Berufsgruppen stört, sondern das hohe Abstraktionsniveau und die jeweils definitionsabhängige Bedeutung der verwendeten Begriffe und Konzepte.
- Der häufige Gebrauch von einschränkenden Nebenbemerkungen und Konjunktiven; Formulierungen wie «... es könnte sein, wenn das nicht wäre, vielleicht stimmt es doch!» entsprechen meist der wissenschaftlichen Redlichkeit, sind aber für den Laien wegen mangelnder Handlungsrelevanz inakzeptabel.
- Eine unzureichende Beachtung wertender Nebenbedeutungen; so ist eine «negative Schüleräußerung über Lehrer» für den Wissenschaftler eine bestimmte Auswertungskategorie, für den Laien ist «negativ» mit einer Note von Fünf oder Sechs gleichzusetzen.

Lösbar ist die Abstimmung der Berichtlegung auf den Adressaten am einfachsten dann, wenn der Evaluationsbericht nur an eine ganz bestimmte Personengruppe gerichtet ist, zum Beispiel an die Weiterbildungsabteilung eines Unternehmens oder an die verantwortlichen Produktmanager bei einem Packungstest. Wenn man in solchen Situationen

- die auch «unsachlichen», emotionalen Komponenten des Auftrages
- den Verwertungs- bzw. Entscheidungszusammenhang
- den «Sprachstil» der Adressatengruppen, insbesondere deren Fachtermini

beachtet, sollte der Bericht ein Erfolg werden. Selbstverständlich ist, daß man bei der äußeren Form die dortigen Standards einhält; diese sind vor allem in der Wirtschaft bzgl. der optischen und (druck-)technischen Gestaltung um vieles anspruchsvoller als im universitären Bereich.

Bei sehr heterogenen Adressatengruppen, wie sie vor allem bei größeren, öffentlich finanzierten Evaluationen üblich sind (etwa bei einer Untersuchung im Schulbereich: Ministerium, Lehrerschaft, allgemeine Öffentlichkeit, vielleicht auch noch Wissenschaftler) ist die zielgruppenorientierte Berichtlegung besonders schwierig. Man wird natürlich verschiedene

Varianten erarbeiten, muß aber darauf achten, daß sich dabei nicht scheinbare Widersprüche ergeben. Generelle Erfolgsregeln gibt es nicht; in jedem Fall sollte man aber bei den einzelnen Formulierungen oder Bezeichnungen sorgfältig auf mögliche Mißverständnisse achten und jedes Konzept von wenigstens einem Angehörigen der relevanten Teilgruppen (ersatzweise von Kollegen, die in diesen Bereichen besondere Erfahrung haben) vor der Publikation hinsichtlich dieses Aspektes prüfen lassen.

Beispiele für Präsentationsformen gibt Diagramm VI/3.

Diagramm VI/3  
Mögliche Präsentationsformen von  
Evaluationsergebnissen

- Fachbericht
- kommentierte Zusammenfassung
- Veröffentlichung in einer Fachzeitschrift
- populärwissenschaftliche Veröffentlichung
- Pressekonferenz
- Symposium
- Vorstellung in den Medien
- Workshop mit Mitarbeitern in einer Unternehmung
- Poster, Broschüren
- persönliche Diskussion

### 6.3.2 Informationsverdichtung

Eines der ungelösten Probleme sozialwissenschaftlicher Projektforschung ist die notwendige Begrenzung des Umfanges des tatsächlich handlungsrelevant werdenden Berichtes. Bei größeren, vielleicht auch wissenschaftlich interessanten Projekten ist eine umfangreiche Dokumentation in Buchform üblich, die leicht deutlich mehr als 500 Seiten umfassen kann (Beispiele dazu finden sich in Aurin et al. 1986; Seiffge-Krenke, 1981). Der Zwang zur wissenschaftlichen Redlichkeit der Darstellung erfordert dabei umfangreiche Tabellen, die (wegen

der Möglichkeit, daß auch Laien das Material lesen) zusätzlich relativ ausführlich interpretiert werden. Man kann nun von keinem an der Entscheidung interessierten Auftraggeber, schon gar nicht von einem unter Zeitdruck arbeitenden Journalisten oder Laien erwarten, dieses umfangreiche Material durchzuarbeiten. Außer der Zeitproblematik fehlen oft die Detailkenntnisse, und praktisch immer das Interesse (eine Ausnahme findet sich bei advokatorischer Rezeption von Evaluationsergebnissen, siehe unten).

Um überhaupt vom Auftraggeber rezipiert zu werden, muß dieses Material in einen kurz lesbaren Bericht zusammengefaßt werden, der – solange man ihn als Wissenschaftler noch irgendwie vertreten möchte – meist zwischen 80 und 100 Druckseiten aufweist. Da dies noch immer zuviel ist, muß diesem Bericht eine Zusammenfassung beigelegt werden, die aber mit ca. 3 bis 6 Seiten für jemand, der nur wissen möchte, welche von zwei denkbaren Maßnahmenvarianten er realisieren soll, noch immer unnötig lang ist. Möglicherweise entscheidungsrelevant ist meist nur eine ca. einseitige «Kurzdarstellung» oder «Beschlußvorlage».

Dieser Informationsverdichtungsprozeß ist unvermeidlich, er findet auch dann statt, wenn sich wissenschaftlich tätige Evaluatoren den entsprechenden Arbeiten entziehen. In diesem Fall werden die entsprechenden Kurzdarstellungen durch Sachbearbeiter des Auftraggebers erfolgen, was der Übereinstimmung der Kurzdarstellung mit den tatsächlichen Projektaussagen nicht immer dienlich ist.

Mit diesem Informationsverdichtungsprozeß verbunden ist natürlich die kaum vermeidbare Gefahr, die Ergebnisse verzerrt wiederzugeben. Unabhängig von der bestehenden Möglichkeit absichtlicher Manipulation wirken zahlreiche gut untersuchte psychologische Prozesse (etwa selektive Aufmerksamkeit, Halo-Effekte, Dissonanzreduktionsmechanismen) auf den Verfasser ein. Eine ausführlichere Darstellung findet sich bei Wottawa, (1981). Es soll nochmals betont werden, daß diese Mechanismen auch bei größter Sorgfalt nur bedingt einzuschränken sind und man aus entsprechenden Beobachtungen nicht auf absichtliche Verzerrungen der Autoren schließen sollte.

### Nahezu unvermeidbare Angriffspunkte

Die Kombination aus notwendig reduzierter Informationsdarstellung und der hohen Plausibilität von Störfaktoren dabei fordert natürlich, gerade bei emotional kontroversen Ergebnissen, Kritiker mit anderer Meinung heraus. Diese sind im allgemeinen bei ihrer Argumentation bzw. Informationsdarstellung mindestens den gleichen Verzerrungen ausgesetzt wie die eigentlichen Autoren, häufig sogar in stärkerem Maße, da sie subjektiv das Gefühl haben, einer verzerrten Darstellung entgegentreten zu müssen und nur durch Überpointierung ihres Standpunktes bei dem Adressaten eine letztlich ausgewogene Meinung erreichen zu können.

Fast immer zur Abwertung veröffentlichter Evaluationsberichte mögliche Argumentationsfiguren sind:

- Offensichtliche methodische Schwächen (daß es diese bei jedem größeren Vorhaben gibt, wurde zum Beispiel in Abschnitt 5.2.2 begründet).
- Es wurde nicht alles berücksichtigt und ausgewertet, was man unbedingt hätte tun müssen (siehe 6.2.3)
- Die verwendeten Operationalisierungen bzw. Meßmethoden sind fehlerhaft (vgl. dazu Abschnitt 4.2.3).
- Die Autoren widersprechen sich selbst (es ist bei umfangreicheren Berichten sehr unwahrscheinlich, daß sich nicht inhaltlich ähnliche, aber in der Aussagerichtung verschiedene Kleindetailergebnisse und entsprechende Interpretationen finden).
- Die Ergebnisse sind nur singular und nicht verallgemeinerbar (ein oft zutreffendes und im übrigen rhetorisch sehr gutes Argument, da sich große Evaluationsstudien im Regelfall nicht wiederholen lassen).
- Die Ergebnisse sind offensichtlich unsinnig, da sie entweder gesicherten wissenschaftlichen Ergebnissen widersprechen (bei sorgfältiger Suche lassen sich fast immer einige widersprechende Befunde in der Literatur eruieren) oder von angesehenen Experten negativ bewertet wurden (in Anbetracht der Pluralität wissenschaftlicher Meinungen findet sich mit etwas Mühe auch dafür ein Beleg).

Auch hier ist wieder die Schwierigkeit, daß alle diese Kritikpunkte voll zutreffen können, die Möglichkeit von schlechten, elementaren professionellen Standards nicht genügenden Evaluationsprojekten ist natürlich gegeben. Da sich aber die Argumente nahezu bei jeder, auch noch so sorgfältigen Studie (falls diese ausreichend komplex angelegt wurde) verwenden lassen, bleibt es dem Kritiker der Evaluationsprojektkritiker (mit beliebiger Verallgemeinerung dieses Meta-Kritiker-Begriffes) überlassen, die Angemessenheit der Gegendarstellung zu bewerten. Im praxisbezogenen Evaluationsbereich greifen die dafür in der Wissenschaft etablierten Mechanismen (eine breite, vielfältige und relativ wenig interessenbezogene Diskussion, ein allmähliches «Vergessenwerden» unzutreffender Interpretationen) aufgrund des meist bestehenden Zeitdruckes nicht. Lösungen für dieses Problem können nicht angeboten werden – leider.

### 6.3.3 Mündliche Präsentation

Für diese, besonders wichtige Form der Ergebnisdarstellung gilt zunächst ebenso wie für den schriftlichen Bericht eine besondere Betonung der «Zuhörerfreundlichkeit». Die technischen Kompetenzdefizite in der mündlichen Repräsentation sind bei vielen Studienabgängern noch gravierender als bei schriftlichen Darstellungen, selbst für die Benutzung üblicher Medien fehlt häufig sowohl theoretisches Wissen als auch praktische Übung. Selbstverständlich sollte die freie, gegebenenfalls mediengestützte (Dias, Overhead-Folien) Rede sein, das Vorlesen vorbereiteter Manuskripte ist für die Zuhörer meist trostlos. Das Schreiben eines Textes, der bei mündlichem Vortrag die Verständlichkeit der freien Rede erreicht, ist eine ausgesprochene Kunst, die nur wenige beherrschen. Zumindest elementare Grundsätze der Rhetorik sollte man beachten, sie finden sich zum Beispiel in Hirsch, 1985, Schuh und Watzke, 1983.

Im übrigen dürfte (abgesehen von langfristigen, geistesgeschichtlich bedingten Meinungenwelen) dem Ansehen der Wissenschaft kaum etwa so geschadet haben, wie die kontroverse öffentliche Diskussion zu Evaluationsthemen, zum Beispiel bzgl. gesetzlicher Regelungen, der Energie-

versorgung oder der Schulorganisation. Die Unterstellung simplen Käuflichkeit wie etwa durch den damaligen Vorsitzenden der Gewerkschaft Erziehung und Wissenschaft (Frister, 1972) ist zwar für alle an Evaluationsvorhaben beteiligten Wissenschaftler eine Zumutung, aber in Anbetracht der dargelegten unvermeidbaren Probleme bei Evaluationsprojekten und dem fehlenden Kenntnisstand über diese Schwierigkeit auch bei sich selbst kompetent fühlenden Laien psychologisch verständlich. Im übrigen hat sicher auch der manchmal bedauerlich stil von «wissenschaftlichen» Diskussionen in der Öffentlichkeit ganz wesentlich dazu beigetragen, das Ansehen (nicht nur, aber besonders) der Sozialwissenschaften zu schädigen. Tatsächlich führen verschiedene Formen der mündlichen Darstellung oder Diskussion von Evaluationsergebnissen zu Rollenzwängen, die eine im Extrem selbst nicht mehr so ganz gerechtfertigt erscheinende Überpointierung von Darstellungen nahelegen. Wenn man zu einer Veranstaltung schon als «Anwalt» für eine bestimmte Position gemeinsam mit einem Kollegen, der die gleiche Rolle für eine andere Auffassung

übernimmt, eingeladen wird und zusätzlich vielleicht noch in dem jeweiligen sozialen Netzwerk der zu einer bestimmten Meinung neigende Gruppe eingebunden ist, ist eine «gefärbte» Betrachtungsweise sehr naheliegend. Da der Kollege den analogen Zwängen unterliegt, kann man sich wechselseitig so weit steigern, daß zum Schluß die Glaubwürdigkeit tatsächlich fraglich wird und der Stil der Äußerungen vielleicht manchmal an der Grenze dessen liegt, was man vor sich selbst gerne vertritt. Interessant ist das Phänomen, das nach dem offiziellen Teil solcher Veranstaltungen durchaus ein emotional herzliches Verhältnis zu dem Kontrahenten bestehen kann und man sich eigentlich darüber einig ist, daß man jetzt ein bißchen übertrieben hat.

Es ist schade, daß sich weder bei öffentlich-politischen noch bei kleineren, etwa für einen bestimmten Konzern oder eine bestimmte Klinik relevanten Evaluationsvorhaben ein berufsethisch begründeter Konsens über akzeptable Verhaltensweisen gerade bei der mündlichen Berichterlegung ergeben hat. Hier bleibt noch einiges zu tun.

Übersicht Kapitel 6:  
Durchführung von Evaluationsstudien

6.1

Organisatorische Fragen

Fragen der Personalführung

**Großfirma als Arbeitgeber**

- meist professionelles Personalmanagement & strukturierte Hierarchie

**Kleinere Unternehmen & private Kleininstitute**

*aus der Sicht des Arbeitgebers*

- Mangel an fachlich qualifiziertem Personal
- Informelle Machtstellung von Spezialisten
- Unerwartete Kündigungen oder Ausfälle
- Mangel an Personal zur Übernahme von Projekt-Management-Aufgaben

*aus der Sicht der Mitarbeiter*

- Unsicherheit des Arbeitsplatzes
- Beschränkung auf Spezialistentätigkeit
- geringe oder keine Aufstiegschancen
- geringe Unterstützung bei der individuellen Weiterbildung

Kontrolltechniken des Projektverlaufes

**Zeitplanung**

- detaillierte Zeitplanung bei der Projektplanung

**Kostenplanung**

- Gegenüberstellung der geplanten Kosten mit verausgabten Beträgen

**Qualitätsstandard**

- laufende Kontrolle der dem Auftraggeber zugesicherten Qualitätsstandards

- Änderung der Zielsetzung des Projektes
- Überlappungsprozesse

**Krisenmanagement**

*Typische Beispiele*

- Unbekannte oder vom Auftraggeber verschwiegene Nebenwirkungen
- Methodenartefakte

- Nicht vorhersehbare Reaktion
- Personalausfälle
- absolut Unvorhersehbares

6.2

Auswertungsarbeiten

**Datenerhebung**

*Probleme bei der Interviewerhebung*

- Schlechte Arbeitsbedingungen, fehlerhafte Adressenvorgaben, schlecht aufgebaute Fragebögen, Interviewer- bzw. Untersuchungsleitereffekte, Quotenvorgaben, Bezahlung nach erfolgreich durchgeführten Interviews

*Hauptproblem bei der postalischen Befragung*

- Rücklaufquote

*Vorbeugende Maßnahmen*

- Absender der Probanden auf dem Couvert
- Ausstellen von Code-Nummern auf den Fragebögen
- Rückcouvert mit Projektnummern

**Datenaufbereitung**

*Probleme*

- Kodierfehler
- fehlerhafte Dateneingabe

*Vorbeugende Maßnahmen*

- Doppelarbeiten
- Bezahlung nach Zeit und nicht nach «Stückzahl»
- Verrechnung gefundener Fehler
- Plausibilitätskontrollen

**Datenauswertung**

*Probleme*

- Verwechslung von Variablennummern oder Codes
- Selbstentwickelte Programme

*Vorbeugende Maßnahmen*

- sorgfältige Arbeit
- semantische Plausibilitätskontrolle

**Datenschutz**

- Beachtung des Bundesdatenschutzgesetzes und der grundsätzlichen gesetzlichen Richtlinien zum Schutz der Privatsphäre

6.3

Berichtlegung

*Zielgruppenanalyse*

Abstimmung der Berichtlegung mit den Adressaten unter Beachtung von:

- Sprachstil der Adressatengruppen
- Verwertungs- bzw. Entscheidungszusammenhang
- unsachliche – emotionale Komponenten des Auftraggebers

**Informationsverdichtung**

*Probleme*

- Absichtliche Manipulation
- Psychologische Prozesse, z.B. Selektive Aufmerksamkeit, Halo-Effekt, Dissonanzeffekte

**Mündliche Präsentation**

- Beachtung unterschiedlicher Präsentationsmöglichkeiten von Evaluationsergebnissen
- Unterstützung der Präsentation durch Medien bzw. Moderationstechniken
- Vorteil advokatorischer Darstellung