

**Hans Günther Roßbach**

**Dr. Bettina Paetzold**

DOI: 10.20378/irb-57695

# **Evaluation von Fortbildungsver- anstaltungen für Erzieher: Zur Ent- wicklung und Er- probung eines Fragebogens**

**Hochschulschriften**

**Lit**

**CIP-Kurztitelaufnahme der Deutschen Bibliothek**

**ROSSBACH, HANS-GÜNTHER:**

**Evaluation von Fortbildungsveranstaltungen  
für Erzieher: zur Entwicklung und Erprobung  
eines Fragebogens / von Hans-Günther Ross-  
bach. - Münster : Lit-Verlag, 1981.**

**(Hochschulschriften : Erziehungswiss.)**

**ISBN 3-8866-090-4**

**NE: Hochschulschriften / Erziehungswissenschaft**

**Lit Verlag Vorländerweg 48 4400 Münster**

**Alle Rechte vorbehalten**

**Printed in Germany**

# INHALTSVERZEICHNIS

	<u>Seite</u>
Einleitung	1
1. Notwendigkeit von Fortbildung im Vorschulbereich	5
2. Grundzüge der Evaluation von Fortbildungsveranstaltungen	15
2.1 Evaluation als Optimierung von Fortbildung	15
2.2 Der Ansatz der "zielfreien" Evaluation	21
2.3 Kriterien des Erfolgs	24
2.4 Die Funktion einer Literaturanalyse	33
2.5 Das Verhältnis der Evaluation zu ihren Adressaten	35
2.6 Zur Auswahl eines Evaluationsinstruments	39
2.7 Zusammenfassung: Leitlinien der Evaluation von Erzieherfortbildung	43
3. Erfolgskriterien von Fortbildungsveranstaltungen	48
3.1 Reziproke Kommunikation als regulatives Prinzip für Fortbildung	48
3.2 Literaturanalyse von Kriterien für erfolgreiche Fortbildungsveranstaltungen	56
3.2.1 Erfolgskriterien im Bereich Thematik	56
3.2.2 Erfolgskriterien im Bereich Methodik	68
3.2.3 Erfolgskriterien im Bereich Interaktionsprozesse	89
3.2.4 Erfolgskriterien im Bereich situativer/organisatorischer Bedingungen	97
3.2.5 Erfolgskriterien im Bereich der Produkte von Fortbildung	99
4. Evaluation durch einen Fragebogen bestehend aus Ratingskalen	102
4.1 Vorteile von Ratingverfahren	102
4.2 Zur Gestaltung des Fragebogens	109
4.3 Mögliche Verzerrungen bei der Verwendung von Ratingskalen	115
5. Ergebnisse der ersten Erprobungsphase	127
6. Ergebnisse der zweiten Erprobungsphase	146
6.1 Analyse des Fragebogens und interne Beziehungen der Items	147
6.1.1 Itemstatistiken und Reliabilität	148
6.1.2 Polungseffekte	152

	<u>Seite</u>
6.1.3 Faktorenanalysen	154
6.1.4 Regressionsanalysen	163
6.1.5 Diskriminanzanalysen	170
6.1.6 Skalenanalysen	173
6.2 Vergleich zwischen Teilnehmer- und Referenten-/Leiterbeurteilung	184
6.3 Experten- und Fortbildungsleiterbefragung	188
6.4 Abhängigkeit der Teilnehmerbeurteilung von externen Einflußgrößen	196
6.4.1 Individualebene	197
6.4.2 Veranstaltungsebene	201
6.5 Zusammenfassung der Ergebnisse	212
7. Evaluation von Erzieherfortbildungsveranstal- tungen: Zusammenfassung und Ausblick	220
LITERATURVERZEICHNIS	227
ANHANG	246



## Einleitung

Zentraler Gegenstand der Praxis und Theorie der vorschulischen Erziehung ist die pädagogisch-sozialisatorische Interaktion zwischen Erzieher und Kind. Eine solche Sichtweise rückt die Person und Identität des Erziehers, seine Aus- und Fortbildung in den Mittelpunkt des praktischen und wissenschaftlichen Interesses (vgl. KOSSOLAPOW 1978; PEUKERT 1979). Quer durch alle theoretischen Positionen ist dabei die Bedeutung der Fortbildung der schon in der Praxis tätigen Erzieher unumstritten, und vielfältiges, oft konzeptionell unterschiedliches Bemühen gilt dieser Aufgabe. Einen besonderen Stellenwert bei der Analyse der Fortbildungsarbeit nimmt die Evaluation ein, die die praktische Arbeit gerade dadurch fördern kann, daß sie Informationen zur Verbesserung und Optimierung des pädagogischen Geschehens bereitstellt. Obwohl eine solche Evaluation oft postuliert wird, überrascht, daß konkrete und praktikable Konzepte zur Überprüfung der Auswirkungen von Erzieherfortbildungsveranstaltungen kaum entwickelt und erprobt vorhanden sind. Es ist deshalb das Ziel der vorliegenden Arbeit - die die Evaluation nicht als etikettierenden Selbstzweck, als (Deck-) Mantel "wissenschaftlicher Dignität" oder als die Fortbildungsarbeit bedrohende Instanz betrachtet -, ein für die Funktion der Verbesserung von Fortbildung brauchbares und nützlich Evaluationsdesign und -instrument zu entwickeln.

Primär verfolgt die Arbeit dabei die Frage, ob in der Praxis an einigen Stellen verwendete, aber nicht in ihrer Qualität überprüfte Fragebögen Informationen für eine potentielle Verbesserung der Fortbildungsarbeit bereitstellen können. Bei diesen Fragebögen werden den teilnehmenden Erziehern Aussagen zu verschiedenen Aspekten der besuchten Veranstaltung vorgelegt; auf einer beigefügten Skala sollen sie dann einschätzen, inwieweit die

Aussagen zutreffen oder nicht. Ein solches Evaluationsinstrument bedarf einer ausführlichen Begründung und Einordnung in einen Auswertungskontext sowie einer empirischen Überprüfung seiner Qualität. Die im Verlauf der Arbeit zu sammelnden Informationen sollen Hinweise für eine weitere mögliche Verwendung eines entsprechenden Fragebogens erbringen.

Der Aufbau der Arbeit folgt dem hier vorgestellten Problemaufriß. Kapitel 1 systematisiert verschiedene Legitimationen von Erzieherfortbildung und vergleicht einige Formen ihrer Gestaltung. Als Gegenstand der Evaluation werden traditionelle, fest eingespielte, mehrtätige Fortbildungsveranstaltungen ausgewählt, die - trotz teilweise heftiger Kritik - auch in neueren, sich alternativ verstehenden Fortbildungsmodellen einen ausgewiesenen Stellenwert haben.

Kapitel 2 versucht ein Modell zur Evaluation von Fortbildungsveranstaltungen für Erzieher zu entwickeln und zu begründen. Dabei wird sich an das Konzept der "zielfreien" Evaluation von SCRIVEN angelehnt. In 12 Leitlinien werden die Grundgedanken für das in den weiteren Kapiteln beschriebene Vorgehen festgelegt. Zentrale Vorstellung des Evaluationsmodells ist die gemeinsame Auswertung der Lernerfahrungen durch Teilnehmer, Referenten und Fortbildungsleiter, eine "Selbstevaluation", in der die Teilnehmer gleichberechtigte Partner der Veranstalter sind und der zu entwickelnde Fragebogen ein Distanz schaffendes, objektivierendes Hilfsmittel darstellt. Evaluation und das Ziel der Optimierung von Fortbildung verlangen zwingend, daß im Verlauf der Arbeit der konkrete Nutzen der evaluativen Informationen erhoben wird.

Kapitel 3 untersucht ausführlich in einer Literaturanalyse, an welchen Meßlatten des Erfolgs die Güte einer Fortbildungsveranstaltung abgelesen werden kann. Die speziellen Arbeitsbedingungen für den Erzieher im Vorschul-

bereich - pädagogische, asymmetrische Interaktionen - erfordern vom Erzieher in seinem Alltagshandeln die Herstellung einer intergenerationellen Reziprozität in seinem Verhältnis zum Vorschulkind. Dieses normative Prinzip von Erzieher-Kind-Interaktion läßt sich zwar nicht direkt auf die Fortbildung als einem Handeln zwischen voll interaktionsfähigen Partnern verallgemeinern, stellt aber ein oberstes, regulatives, alle anderen Erfolgskriterien kontrollierendes Postulat an die Fortbildungsprozesse: Erzieher müssen in ihnen gleichberechtigte Partner der Veranstalter und Subjekte selbstbestimmter Lernprozesse sein.

Kapitel 4 diskutiert die Auswahl des Evaluationsinstruments. Der angestrebte Evaluationskontext - gleichberechtigte Partnerschaft, Selbstevaluation unterstützt durch ein Distanz schaffendes Instrument - spricht für die Verwendung eines Systems von Ratingskalen, welches die Informationen bezüglich der Erfolgskriterien erhebt und dessen Einsatz und Gestaltung den ausfüllenden Teilnehmer als Partner ernst nimmt. Ein solches Instrument wird hier entwickelt; mögliche Schwächen und Verzerrungen einer Informationserhebung mittels Ratingskalen bedingen eine ausführliche Erprobung, deren programmatische Grundlinien ebenfalls in diesem Kapitel entwickelt werden.

Die Ergebnisse einer ersten Erprobungsphase (Kapitel 5) führen zu einer revidierten Fassung des Fragebogens, dessen technische und evaluative Qualität in Kapitel 6 ausführlich untersucht werden. Ein besonderes Augenmerk wird auf die Validität gelegt, da Fragebogeninformationen ein Auswertungsgespräch nur dann sinnvoll vorbereiten helfen, wenn sie Veranstaltungsrealität reflektieren.

Kapitel 7 faßt den Gang der Arbeit kurz zusammen und versucht, Perspektiven für weitere Anwendungsmöglichkeiten des entwickelten Fragebogens sowohl in der praktischen Auswertungsarbeit als auch für die Erforschung von Fortbildungsprozessen aufzuzeigen.

Die vorliegende Arbeit konnte nur durch vielfältige Anregungen und Unterstützungen entstehen. Besonders herzlich möchte ich den Erziehern, Referenten sowie Leitern der untersuchten Fortbildungsveranstaltungen für die bereitwillige Teilnahme und auch für die kritischen Stellungnahmen danken. Viele Anregungen empfing die Arbeit von den Mitarbeitern des Arbeitsbereichs Vorschulerziehung des Instituts für Erziehungswissenschaft der Universität Münster. Frau Prof. Dr. L. Kossolapow hat sie mit großer Sorgfalt betreut und durch ihre Hinweise und Ratschläge erst ermöglicht. Herr Prof. Dr. W. Tietze war ein ständiger Ansprechpartner, dessen kritische Gedanken mich ermutigten und die Arbeit vorantrieben.

## 1. Notwendigkeit von Fortbildung im Vorschulbereich

Die Begriffe "Erwachsenenbildung", "Fortbildung" und "Weiterbildung" werden sehr unterschiedlich gebraucht. Im Strukturplan für das Bildungswesen spricht der DEUTSCHE BILDUNGSRAT von Weiterbildung als dem übergeordneten Begriff für den quartären (nachschulischen) Bereich. Gemeint sind damit alle Formen organisierten Lernens der Erwachsenen nach Eintritt in die volle Erwerbstätigkeit, d.h. Aktivitäten der beruflichen Fortbildung, der Berufsumschulung sowie der nicht primär berufsbezogenen, sondern eher allgemein orientierten Erwachsenenbildung (vgl. RAAPKE 1974, S. 625). Andere Autoren setzen Erwachsenenbildung als Oberbegriff für die verschiedenen Bildungsbemühungen, dem dann Fort- und Weiterbildung - oft mit verschiedenen Funktionen - untergeordnet werden (vgl. HEYER 1977, S. 40-44)<sup>1</sup>. Beide Begriffe beziehen sich dabei auf berufliche Lernprozesse. Fortbildung meint dann oft jene berufsbegleitenden Maßnahmen, die dazu dienen, die berufsspezifische Handlungsfähigkeit der Fachkraft zu erhalten oder zu erweitern, wobei damit der berufliche Status oder die Bezahlung nicht verändert wird. Weiterbildung dagegen zielt auf eine berufliche Höherqualifizierung oder auf die Qualifizierung für Berufsfelder, die von jenen verschieden sind, für die in der Ausbildung die Befähigung erworben wurde (vgl. z.B.

---

<sup>1</sup> HEYER diskutiert insgesamt 10 verschiedene Begriffsschattierungen. Eine ausführliche Diskussion der Begriffe Erwachsenenbildung und Weiterbildung findet sich bei FISCHER. In den letzten 20 Jahren läßt sich eine Pendelbewegung der Begrifflichkeit erkennen. Das Gutachten des Deutschen Ausschusses von 1960 arbeitet mit der Bezeichnung Erwachsenenbildung, während sich bis zum Strukturplan für das Bildungswesen von 1970 der Oberbegriff Weiterbildung durchsetzt. Gegenwärtig scheint die Neigung zu bestehen, wieder den Begriff Erwachsenenbildung zu benutzen, zumindest aber beide Begriffe synonym zu verwenden (FISCHER 1978, S. 167). In der deutschen Übersetzung der UNESCO - Empfehlung über die "Development of adult education" wird "adult education" (als Oberbegriff für Gesamtheit organisierten Lernens) mit Weiterbildung übersetzt (vgl. UNESCO 1977, S. 1).

TIETZE 1978, S. 80; ULSHOEFER 1970, S. 21; DALLMANN 1976, S. 129). Unterschieden werden können Einführungsfortbildung (Maßnahmen zur Ergänzung der Ausbildung), Anpassungsfortbildung (Aktualisierung bzw. Erweiterung des Fachwissens) und Förderungsfortbildung (Förderung zur Wahrnehmung höherwertiger Funktionen bzw. Leitungsfunktionen) (vgl. ROßMANTH 1977, S. 104).

Den folgenden Ausführungen soll ein Arbeitsbegriff unterlegt werden, der unter Erwachsenenbildung den Oberbegriff für alle Bildungsmaßnahmen mit Erwachsenen versteht. "Fortbildung" bezieht sich dabei auf berufliche Erwachsenenbildung mit Fachkräften für den Vorschulbereich und umfaßt sowohl Einführung, Ergänzung, Anpassung an neue Aufgaben als auch Höherqualifikationen. Dieser Arbeitsbegriff entspricht auch weitgehend dem Selbstverständnis, mit dem Träger, Referenten und Teilnehmer ihre Tätigkeit bezeichnen.

Gerade im Vorschulbereich läßt sich der Ruf nach Fortbildung und die Vielfältigkeit von Fortbildungsmaßnahmen nicht übersehen. Vielfältig sind auch die Begründungen der Notwendigkeit von Fortbildung. Vier - sich teilweise überschneidende - Argumentationsstränge lassen sich unterscheiden:

- Die Ausbildung des pädagogischen Mitarbeiters im Vorschulbereich wird in der Regel als mangelhaft empfunden. Die Erzieher sind durch die viel zu kurze und schlechte Ausbildung den Problemen der Praxis wie z.B. Erwartungen der Eltern und der Öffentlichkeit oder Verhaltensschwierigkeiten der Kinder ausgeliefert und werden dadurch verunsichert (vgl. RIEMANN 1975, S. 313). Die - reglementierte - Ausbildung führt nicht zu einer Handlungsfähigkeit in der Praxis, sondern gerade zu Handlungsunfähigkeit (vgl. AG FORTBILDUNG IM KINDERZENTRUM BERLIN 1977, S. 183). Fortbildung wird hier als Möglichkeit angesehen, diese Defizite

der Ausbildung zu kompensieren<sup>2</sup>.

- Eine alleinige Begründung der Fortbildung aus den Defiziten der Ausbildung ist unbefriedigend. Auch eine qualifizierte Berufsausbildung kann nicht auf alle Situationen einer langen Berufsausübung vorbereiten (vgl. MINZ 1978, S. 77). Das berufliche Handeln von Erziehern wird durch eine Vielzahl von Faktoren bestimmt, die in der Ausbildung selbst entweder keine Relevanz haben oder durch die Struktur der Ausbildung - die ja noch kein praktisches Handeln ist - nicht berücksichtigt werden können (vgl. DILLER/SCHLAGHECKEN 1979, S. 234; ACKERMEIER u.a. 1978, S. 167f.; SEIDEL/BRÜGGEMANN 1979, S. 113). Fortbildung muß deshalb zwingenderweise die Arbeit ständig begleiten und ergänzen.

"Fortbildung hat nicht die Aufgabe, die Lücken zu schließen, die durch unzureichende Ausbildung bedingt sind. Es muß deutlich gesehen werden, daß Fortbildung eine - nicht durch Ausbildung zu ersetzende - eigene Funktion hat." (MOSKAL 1978, S. 319)

- Pädagogische Kenntnisse veralten angesichts einer sich rasch verändernden Gesellschaft und angesichts der Fortschritte in den Wissenschaften und der Weiterentwicklung von Zielen, Inhalten und Methoden der Erziehung. Es ist Aufgabe der Fortbildung auf diesen Veraltungsprozeß von Fachwissen zu reagieren und neue Kenntnisse zu vermitteln (vgl. ACKERMEIER 1978, S. 166; TIETZE 1978, S. 80; MINZ 1978, S. 77; HEYER 1974, S. 27). Pädagogisches Handeln be-

---

<sup>2</sup>Es sei hier zusätzlich auf das Problem der unvollständig oder nicht ausgebildeten Mitarbeiter im Vorschulbereich verwiesen. Fortbildung könnte hier eine Chance des Nachholens bieten. PRIESNER nimmt hier eine Einteilung in 3 historische Grundtypen der Erwachsenenbildung von SCHULLENBERG auf und bezeichnet dieses Nachholen als "kompensatorische Weiterbildung" (vgl. PRIESNER 1975, S. 50).

darf des ständigen Hinzulernens, einer "éducation permanente", eines lebenslangen Lernens<sup>3,4</sup>.

- Fortbildung von Erziehern wird auch als zentrales Moment der Bildungsreform im Vorschulbereich verstanden. Nach dem quantitativen Ausbau der Vorschulerziehung kommt ab Mitte der siebziger Jahre immer mehr die qualitative Verbesserung und Reform des Vorschulbereiches in den Blick. Innovationen und curriculare Reformen können aber nur dann gelingen, wenn die Erzieher für die neuen Aufgaben qualifiziert werden (vgl. SCHRADER 1973, S. 286f.; HEMMER 1977, S. 99f.). Fortbildung dient dann dazu, den Erziehern diese Qualifikationen zu vermitteln und sie zum Träger von Reformen zu machen (vgl. SCHMALOHR 1974, S. 12; FRUTEL u.a. 1975, S. 225; MINZ 1978, S. 79).

Der Vielfalt der Legitimationen von Fortbildung korrespondiert eine nicht minder große Vielfalt von Fortbildungsformen<sup>5</sup>, die sich vor allem unterscheiden lassen nach der zeitlichen Dauer, den Medien, der Distanz zum Arbeitsplatz/Praxisdruck, dem Ausmaß an Vorbereitung/Planung und der Zusammensetzung der Teilnehmergruppe (Teilnehmer eines

---

<sup>3</sup> In den Grundtypen von Erwachsenenbildung nach SCHULENBERG läßt sich dies als "komplementäre" Weiterbildung auffassen, obwohl sich hier - ebenso wie in der folgenden 4. Begründung für Fortbildung - Überschneidungen mit der "transitorischen Weiterbildung" ergeben; d.h. jener Weiterbildung, die - vorübergehend - Modernitätsrückstände der Erziehung aufholen will (vgl. PRIESNER 1975, S. 51).

<sup>4</sup> Zur allgemeinen Diskussion um die Veralterung, "Obsoleszenz" von Fähigkeiten oder Wissen siehe LEON 1977, bes. S. 86-91. Der Verweis auf "lebenslanges Lernen" alleine rechtfertigt noch keine Fortbildung, da lebenslanges Lernen selbst kein Lernziel sein kann. Lebenslanges Lernen kann die Notwendigkeit, die Fähigkeit oder die Möglichkeit meinen, lebenslang zu lernen. Zur Diskussion dieses Begriffes siehe BÖHME 1978.

<sup>5</sup> Die Reservierung des Fortbildungsbegriffs für strukturierte, von Institutionen angebotene Lernprozesse von HEYER 1977, S. 45f. wird hier nicht geteilt.



Kindergartens, einer Region versus zufällige Teilnehmerzusammensetzung). Ohne Anspruch auf Vollständigkeit seien einige Formen kurz angerissen:

#### - Seminare und Lehrgänge

Fortbildungsveranstaltungen in der Form von Lehrgängen beschäftigen sich mit Inhalten und Zielen der Vorschulernziehung direkt oder versuchen auch zu einer vertieften Einsicht in eigenes Verhalten zu führen. Die Veranstaltungen dauern von 1 Stunde bis zu einer Woche und werden meist an einem zentralen Ort entfernt vom Arbeitsplatz durchgeführt. Die Teilnehmer stehen in einer Distanz zu ihrer Arbeitsstätte, sie sind gewissermaßen vom direkten Praxisdruck befreit und können neue Möglichkeiten des Verhaltens ausprobieren (vgl. NACHBAUER 1978, S. 326; THÜRNUAU 1978, S. 174). Die Veranstaltung kann sorgfältig geplant und ein guter Referent ausgewählt werden. Nachteile der Distanz zum Arbeitsplatz werden in Problemen der Umsetzung des Gelernten in die Praxis und der Weitergabe der Fortbildungsinhalte an die Kollegen gesehen (vgl. SEIDEL/BRÜGGEMANN 1979, S. 113).

#### - Beratung

Fortbildung findet in diesem Fall unmittelbar an der Praxisstätte statt. Diese Form wird unter den Begriffen Fachberatung und/oder Praxisberatung diskutiert (vgl. BUDDÉ 1976; DENECKE/HOLLMANN 1977), bzw. unter dem Namen Organisationsentwicklung und Institutionsberatung (vgl. SCHWARTE 1978; MÜLLER-SCHÖLL/PRIEPKE 1976). Fachberater begleiten oft die zentralen Fortbildungsveranstaltungen eines Trägers, um so Fachberatung und zentrale Fortbildung zu verzahnen. Der Vorteil liegt hier in einer direkten Förderung berufsspezifischen Handelns. Schwierigkeiten werden darin gesehen, daß diese Beratung abhängig ist von dem Aufschlossensein einer Institution für Veränderung (vgl. THÜRNUAU 1978, S. 174) und daß im Vorschulbereich ein Fach-

berater z.T. über 100 Einrichtungen zu betreuen hat (vgl. TIETZE 1978, S. 81).

### - Fortbildungsbriefe, Fachliteratur

Material zum Selbststudium ist eine weitere Möglichkeit der Fortbildung. Beispiele wären hier die Berliner Fortbildungsbriefe, die vom SENATOR FÜR FAMILIE, JUGEND UND SPORT herausgegeben werden (vgl. ARBEITSGRUPPE FORTBILDUNG UND ERZIEHERINNEN 1976; ARBEITSGRUPPE FORTBILDUNG 1976) oder die Gedanken zur pädagogischen Arbeit und Anregung für die Praxis der Arbeiterwohlfahrt (vgl. PRAXISHEFT 14). Ebenfalls hierzu gehört das Studium von Fachzeitschriften<sup>6</sup>. Problematisch sind hier die Verbreitung und der Zugang zu diesem Material, die notwendige hohe Motivation und die praktische Umsetzung, die dem Erzieher letztlich alleine überlassen wird (vgl. TIETZE 1978, S. 82).

### - Mitarbeit in Curriculumprojekten

Ausführlich thematisiert wird die Fortbildung in der situationsorientierten Curriculumentwicklung (vgl. HEMMER 1975, 1977; SCHRADER 1973). In der Mitarbeit an der Entwicklung von didaktischen Einheiten wird hier die Qualifizierungsmöglichkeit für Erzieher gesehen. Die Grenze dieses Ansatzes liegt darin, daß nur ein geringer Teil der Erzieher in Curriculumentwicklungsprojekten direkt mitarbeiten<sup>7</sup>.

---

<sup>6</sup>Über den Einsatz von Fachzeitschriften in Fortbildungsveranstaltungen berichtet GROßE-JÄGER 1977.

<sup>7</sup>In der Lehrerfortbildung wurde ausführlich die Integration der Fortbildung in die schulnahe Curriculumentwicklung diskutiert (vgl. hierzu z.B. GSTETTNER/RATHMAYER 1978, S. 329). Diese Tendenz ist - zwar geringer - auch im Vorschulbereich vorhanden. Andererseits wird aber auch in der Lehrerfortbildung diese Integration in Frage gestellt. So formuliert LAURENZE 1978, S. 371 im Anschluß an das CELF-Projekt, daß die Verbindung mit der Curriculumforschung problematisch geworden sei und durch eine unterrichtsbezogene Fortbildung ersetzt werden sollte, die weniger die Planungs- als die Realisationsschwierigkeiten von Unterricht aufgreift (vgl. dazu auch LUTGERT/SCHÜLER 1978, S. 351ff.).

### - Regionale Arbeitskreise

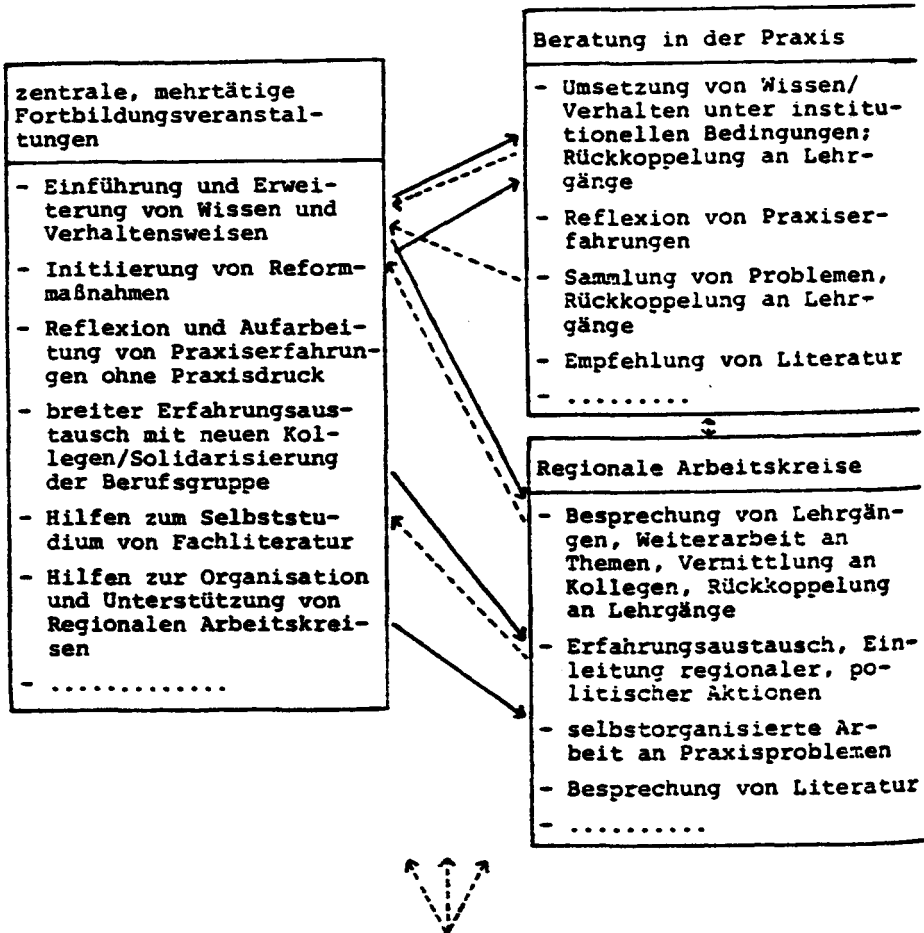
Fortbildung in regionalen Arbeitskreisen wird oft von den Erziehern benachbarter Kindergärten selbst organisiert. Schwierigkeiten werden u.a. in der Passivität der Teilnehmer, in dem Problem der Leitung und darin gesehen, daß die Inhalte oft am alten Ausbildungsstand orientiert werden (vgl. SEIDEL/BRÜGGEMANN 1979, S. 113f.). Deshalb wird Hilfe von außen nötig und das Konzept der regionalen Arbeitskreise meistens mit Fachberatung (vgl. BUDE 1976) und/oder zentralen, überregionalen Fortbildungsveranstaltungen gekoppelt, die dann die regionalen Arbeitskreise in ihrer Arbeit unterstützen sollen (vgl. WALTHER/SCHMIDT 1979; ACKERMEIER 1978, bes. S. 48f. und 184f.).

Der kurze Überblick zeigt auf, daß jede Form von Fortbildung für sich alleine Vorteile, aber auch Schwächen hat. Es gibt somit nicht die beste Organisationsform von Fortbildung, vielmehr sollte ein mehrschichtiges, mehrdimensionales System von Fortbildungsmaßnahmen aufgebaut werden, in dem Organisationsformen nach ihren Vorteilen und Funktionen ausgewählt werden und sich wechselseitig ergänzen (vgl. z.B. HEMMER 1977, S. 97; MOSKAL 1978, S. 321-325); angestrebt ist ein Verbundsystem von Lerneinheiten an zentralen und regionalen Lernorten und an dem Praxisort selbst, von längerfristigen Veranstaltungen, die sich eher auf die Verhaltensebene und -veränderungen beziehen und kurzfristige Veranstaltungen mit eher informativem Charakter (vgl. THURNAU 1978, S. 179f.; NACHBAUER 1978, S. 331)<sup>8</sup>. Abbildung 1 (S. 12) versucht modellhaft, Funktionen verschiedener Fortbildungsformen darzustellen

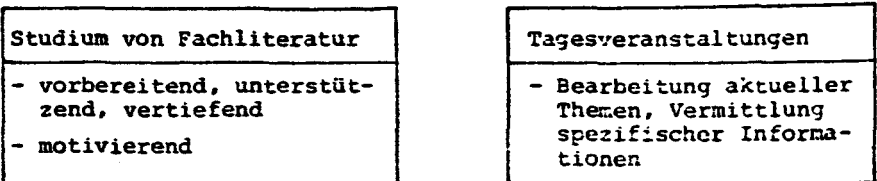
---

<sup>8</sup> Zu den verschiedenen organisatorischen Möglichkeiten von Fortbildung siehe auch die Überlegungen im Deutschen Verein für öffentliche und private Fürsorge, wo je nach Adressaten und Inhalten unterschiedliche Fortbildungsebenen angesprochen werden (vgl. EHRHARDT-PLASCHKE/IRSKENS 1977, S. 290-292; MAELICKE 1977, S. 237f.).

Abb. 1: Funktionen verschiedener Fortbildungsformen



Die Fortbildungsarbeit wird ständig begleitet von:



(Die Pfeile in der Abbildung weisen auf einige, beispielhaft ausgewählte Ergänzungen/Verbindungen der Möglichkeiten von Fortbildungsformen hin)

und einige mögliche gegenseitige Ergänzungen anzudeuten. Betrachtet man diese Abbildung, so fällt in dem skizzierten Verbundsystem die herausgehobene Stellung der zentralen, mehrtägigen Fortbildungsveranstaltungen auf. Befreit von unmittelbarem Praxisdruck, institutionell abgesichert, zeitlich längerdauernd und von meist hochqualifizierten Referenten begleitet, bietet diese Form grundlegende Möglichkeiten der Vermittlung von neuem Wissen und der Initiierung von Reformprozessen, die die anderen Fortbildungsformen nicht erreichen können. Zugleich kann dieses Potential aber erst voll entfaltet werden, wenn die mehrtägigen Lehrgänge in einem Gesamtsystem von Fortbildung eingebunden sind. Das vermittelte neue Wissen und die angebahnten neuen Verhaltensweisen müssen erst unter institutionellen Bedingungen erprobt werden. An dieser Stelle können vielfältige Beratungsprozesse die Umsetzung kontrollieren und unterstützen; zusammen mit Regionalen Arbeitskreisen werden die dabei auftretenden Probleme bearbeitet, gesammelt und an die zentralen Veranstaltungen zurückgemeldet und dort evtl. in Folgeveranstaltungen behandelt. Zum anderen - und dies ist besonders in der heutigen, oft kontrovers geführten Diskussion um die Organisation von Fortbildung wichtig - unterstützen die Lernprozesse auf zentralen Lehrgängen die anderen Formen und ermöglichen sie oft erst. Praxisberatung allein kann nur bestimmte, eher punktuelle Probleme angehen, Regionale Arbeitskreise, selbstorganisierte Fortbildungsprozesse haben zwar einen direkten Praxisbezug, beide Formen sind aber überlastet, wenn sie systematisiert neue Bereiche, neue Themen erarbeiten bzw. vermitteln sollen. Eine Ergänzung durch zentrale Lehrgänge ist nötig, wobei diesen auch die Aufgabe zufällt, zuerst die Grundlagen zu einer Selbstorganisation von Fortbildung zu vermitteln.

Aufgrund dieses zentralen konzeptuellen Stellenwertes mehrtägiger Lehrgänge für die Fortbildung der Erzieher und damit für die Verbesserung der Erziehungspraxis wird die Evaluationsfrage in diesem kaum bearbeiteten Bereich vorerst auf diesen Veranstaltungstyp zugespitzt gestellt.

## 2. Grundzüge der Evaluation von Fort- bildungsveranstaltungen

### 2.1 Evaluation als Optimierung von Fortbildung

Im "religionspädagogischen Förderprogramm für den Kindergarten" des Comenius-Instituts in Münster wird Evaluation als Bestandteil eines idealtypischen Modells von Fortbildung gesehen (vgl. ACKERMEIER u.a. 1978, S. 30f.); ebenfalls äußern andere Träger von Fortbildungsveranstaltungen ihr Interesse an Evaluation<sup>1</sup>. Auf der anderen Seite aber sind hier systematische Evaluationsbemühungen nur wenig entwickelt (vgl. ROSSBACH 1977, S. 136ff.). Im folgenden Abschnitt soll nun geklärt werden, welche Gründe es für Evaluation von Erzieherfortbildung gibt und welche Bedeutung sie spielen kann.

Evaluation allgemein beschäftigt sich mit der Sammlung und Verarbeitung von Informationen über pädagogische Programme (vgl. WULF 1975a, S. 567); ihr Ziel besteht in der Feststellung des Wertes und Nutzens des Programmes mit Hilfe wissenschaftlicher Methoden, um dadurch Entscheidungen über das Programm bzw. Programmtteile vorzubereiten (vgl. z.B. STUFFLEBEAM 1974, S. 13f.; MATUSZEK 1978, S. 346; ALKIN 1975, S. 201; MITTER/WEISHAUP 1978, S. 79). Evaluation wird

---

<sup>1</sup> In Gesprächen mit Trägervertretern von Fortbildungsinstitutionen und Referenten wurde durchgehend der Wunsch nach einer Überprüfung geäußert. In der eigenen Praxis würden zwar schon kleine Fragebögen angewendet, die aber meist in keiner Weise überprüft sind. An den Autor wurde auch der Wunsch nach der Entwicklung gesicherter Evaluationsverfahren herangetragen. Die Vorstellungen der Trägervertreter waren aber sehr allgemein und unspezifisch.

hier im größeren Rahmen des empirisch-analytischen Paradigmas<sup>2</sup> verstanden, sie unterscheidet sich aber insofern von Forschung, als es ihr nicht primär auf die Entdeckung von neuen allgemeinverbindlichen Gesetzen bzw. Quasi-Gesetzen ankommt (vgl. HAEBERLIN 1975, S. 657), sondern auf den Aufweis einer Nützlichkeit (vgl. GLASS 1972, S. 169f.; STUFFLEBEAM 1974, S. 10)<sup>3</sup>.

---

<sup>2</sup> Evaluation steht heute im Spannungsfeld zweier Paradigmen: der empirisch-analytischen Sozialforschung und der Handlungsforschung, so daß von einer empirisch-analytischen und einer handlungsorientierten Evaluation gesprochen werden kann (vgl. ähnlich WULF 1975b, S. 587-593). In dieser Arbeit wird sich orientiert an empirisch-analytischen Verfahren bzw. an anglo-amerikanischen Ansätzen. Auf die Auseinandersetzung mit Handlungsforschung wird hier erachtet, sie verlangt eine eigene Arbeit. Es wird aber - implizit - hier aufgezeigt werden, daß ein eher "klassisch" orientiertes Vorgehen Funktionen erfüllen kann, die heute direkt nur als für Handlungsforschung angemessen deklariert werden.

<sup>3</sup> Die Unterscheidung zwischen Evaluation und Forschung bedeutet also nicht, daß Evaluation aus dem empirisch-analytischen Paradigma herausfällt, nur müssen - s.w.u. - für die Evaluation die Gütekriterien der Forschung durch Kriterien ergänzt werden, die sich speziell auf diese Nützlichkeit beziehen (vgl. STUFFLEBEAM 1974, S. 6). Indem Evaluationsstudien ein pädagogisches Programm intensiv und komplex untersuchen, ähneln sie eher einer "Einzelfallstudie" und dienen so der Aufhellung der pädagogischen Wirklichkeit (vgl. SCHMIED 1977).

Zwei Beispiele für die Konsequenzen aus der Unterscheidung zwischen Evaluation und Forschung seien noch benannt:

- Empirisch-analytisch orientierte Evaluation verlangt keine Konstanthaltung von Variablen der Erziehungswirklichkeit über längere Zeit und verhindert so nicht Verbesserungen in der Praxis. Evaluation dient gerade dazu, Entscheidungen über Veränderungen der Praxis vorzubereiten.
- In Forschungsstudien besteht tendenziell die Zeit, Untersuchungen zu wiederholen, andere Variablen oder Operationalisierungen auszuprobieren, bevor ein Gesetz aufgestellt wird. Die Auswahl von zu untersuchenden Variablen ist für die Evaluation weitaus problematischer, da hier in begrenzter Zeit Informationen für - teilweise unwiderrufbare - Entscheidungen gesammelt werden müssen (vgl. STRATON 1977, S. 60).



"Thus, while the comment "even if the results of a study were not utilized, its redeeming feature is its intrinsic value and its contribution to the corpus of knowledge" is appropriate as a statement for a study designed primarily as research, it is not appropriate for an evaluation study. In short, in an evaluation it is important if someone "needs" the evaluation and someone "cares"." (ALKIN 1975, S. 201).

"Nützlich" sein kann eine Evaluation nun für verschiedene Adressaten: für z.B. die Teilnehmer von Erzieherfortbildung, die Referenten, die Veranstaltungsleiter, die Träger oder den Geldgeber; für die Kollegen der Teilnehmer, die Eltern oder die Gemeinde. Ebenso wie die Adressaten unterschiedlich sind, gibt es auch verschiedene Funktionen, die die Evaluation für diese erfüllen kann. Es seien einige aufgezählt (vgl. hierzu FASSBENDER 1974, S. 128-130; GERL 1977, S. 8; SCHINDLER 1979, S. 9f.):

- Evaluation zum Abschluß einer Veranstaltung ist ein Ritual, sie dient der Beruhigung der Teilnehmer, sie ist ein Ventil für Kritik
- Evaluation gibt einer Fortbildungsveranstaltung eine "wissenschaftliche Weihe"
- Evaluation dient der Kontrolle der Lernvoraussetzungen und der Lernwirkungen des Unterrichtes
- Evaluation dient der Vergrößerung der Selbststeuerungsmöglichkeiten und Reflexionsfähigkeiten der Lerngruppe
- Evaluation dient der Steuerung der Lernprozesse, sie will konkrete Anregungen zur Verbesserung der Fortbildung geben, sie will feststellen, welche Ziele erreicht wurden und welche nicht
- Evaluation dient der Disziplinierung der Referenten
- Evaluation dient der Legitimation des Veranstalters gegenüber der Öffentlichkeit
- Evaluation will den Nutzen einer Veranstaltung im Verhältnis zu den Kosten überprüfen, sie will eine ökonomische Mittelverwendung sicherstellen.

Es lassen sich hier zwei Hauptfunktionen identifizieren:

- Evaluation dient der Verbesserung, Modifizierung, Optimierung einer Fortbildungsveranstaltung.

- Evaluation dient der Begründung und Legitimation der eigenen Arbeit z.B. gegenüber einem Träger oder Verband (vgl. allgemein LEWY 1973, S. 317ff.; STUFFLEBEAM 1974, S. 15; WULF 1975b, S. 582-586)<sup>4</sup>.

Fortbildungsveranstaltungen für Erzieher - so lassen sich die Begründungen für Fortbildung aus dem vorherigen Kapitel zusammenfassen - sehen ihr Ziel in der Verbesserung der Erziehungspraxis. Indem Evaluation hier dieses Interesse teilt, soll ihre Hauptfunktion in der Verbesserung, Optimierung von Fortbildung bestehen, um damit indirekt Erziehungspraxis zu verändern. Evaluative Informationen sollen somit nützlich für Entscheidungen über Optimierung von Fortbildung sein. Damit wird ihr potentieller Adressatenkreis primär eingengt auf die direkt an Fortbildung Beteiligten, d.h. teilnehmende Erzieher, die Referenten und Fortbildungsleiter. Die Auswahl dieser Adressaten der Evaluation bein-

---

<sup>4</sup> Diese Unterscheidung zwischen Optimierung und Legitimation nimmt die bekannte SCRIVENSche Unterscheidung zwischen formativer und summativer Evaluation auf (vgl. SCRIVEN 1972a, S. 62), die sich nicht nur oder nicht primär auf die Frage des Zeitpunktes einer Evaluation bezieht. Die Unterscheidung kennzeichnet vielmehr Funktionen der Evaluation (vgl. WULF 1974, S. 205; STAKE 1972, S. 110). Formative Evaluation, die während der Entwicklung eines Programmes stattfindet, hat die Funktion, zur Verbesserung beizutragen, während summative Evaluation eher die Funktion hat, ein Programm (abschließend) zu bewerten und damit für seinen Einsatz in der Schule z.B. zu legitimieren. Formativ und summativ sind hier relative Begriffe. Eine (vorläufige) summative Evaluation kann zur Programmodifikation benutzt werden und damit eine formative Funktion erfüllen.

WULF führt noch eine 3. Hauptfunktion an, die Entscheidungshilfe (vgl. WULF 1975b, S. 586f.). Indem der Evaluator aber selbst nicht ein Programm verbessert, sondern nur Informationen für Entscheidungen über Verbesserungen bereitstellt, ist die Entscheidungshilfefunktion weniger eine unabhängige 3. Funktion neben Optimierung und Legitimation, sondern vielmehr eine übergeordnete Kennzeichnung der evaluativen Tätigkeit.

haltet zugleich eine Einschränkung, da die Informationen für diese Gruppen nicht immer auch relevant für andere Gruppen sind und umgekehrt. So könnten bei dem Träger oder Geldgeber evtl. die Kosten einer Veranstaltung im Vordergrund stehen, Informationen also, die für Teilnehmer, Referenten und Fortbildungsleiter erst in zweiter Linie relevant sind<sup>5</sup>.

Die Bestimmung der Evaluation als Hilfsmittel zur Vorbereitung von Entscheidungen bedeutet nun zugleich, daß der Evaluator selbst keine Entscheidungen fällt, sondern diese seinen Adressaten überläßt. Der Eingriff in die Praxis, d.h. in die Modifikation oder Konstruktion einer Veranstaltung, liegt - theoretisch - außerhalb der evaluativen Tätigkeit. Die Aufgabe des Evaluators ist beschränkter. Gemäß der Orientierung am empirisch-analytischen Paradigma werden objektive Aussagen angestrebt, d.h. solche, die intersubjektiv nachvollziehbar sind. Dies wird nicht allein durch eine Distanz des Evaluators zum Gegenstand erreicht, sondern letztlich nur dadurch, daß der Weg der Erkenntnis so offengelegt wird, daß andere Evaluatoren ihn potentiell

---

<sup>5</sup> In Gesprächen mit Fortbildungsreferenten, -leitern und Trägervertretern wurde nur einmal die potentielle Legitimationsfunktion für die eigene Arbeit angesprochen; durchgehend sahen alle anderen die Funktion der Evaluation nur in der konkreten Verbesserung der Fortbildungsveranstaltungen.

Zur allgemeinen Problematik der Identifikation der Adressaten der Evaluation und deren offene und verdeckte Interessen vgl. ALKIN 1975, S. 202-204; STRATON 1977, S. 57-59.

nachvollziehen können (vgl. PRIM/TILMANN 1975, S. 14ff.; NACKEN 1977a, S. 37)<sup>6</sup>. Eine starre Einschränkung des Evaluators nur auf solche Sätze verhindert aber seine Arbeit. Diese ist oft so komplex, bzw. geschieht unter Zeit- und Ressourcenmangel, daß er gar nicht in der Lage ist, alle seine Aussagen nach diesem Modell streng zu überprüfen. Er muß aber trotzdem den Weg, wie er zu seinen Aussagen kommt, offenlegen und ihre Absicherung relativieren. In diesem Sinne kann er Empfehlungen zur Verbesserung aussprechen, muß aber deutlich machen, daß diese nicht eine "objektive" Dignität haben<sup>7</sup>.

Evaluationsinformationen lassen sich außerdem nicht linear in Entscheidungen umsetzen, d.h. es folgt aus ihnen nicht zwingend, daß ein bestimmter Aspekt in einer bestimmten

---

<sup>6</sup> Gerade weil Evaluation es immer mit Werten und Bewertung zu tun hat, wird diese Frage in der Evaluationsliteratur sehr breit diskutiert (vgl. z.B. SCRIVEN 1972a, S. 72; GLASS 1972, S. 171ff.; STAKE 1970; MEYER 1974, S. 135f.; BERLAK 1970; ROSSERACH 1977, S. 27-36, zu dem allgemeinen Verhältnis von wissenschaftlichen Aussagen in den anwendenden Sozialwissenschaften zu politischen Entscheidungen siehe LOMPE 1972).

Intersubjektive Überprüfbarkeit bedeutet gerade hier, daß es abzulehnen ist, wenn Evaluationsberichte für Adressaten geschrieben werden und dann nicht veröffentlicht in "Schubladen" verschwinden. Auch in mehr informellen Ansätzen ist dies abzulehnen (vgl. dagegen SCHNEIDER/STAKE 1976, S. 40; HOWE 1978, S. 20). In größeren bildungspolitischen Fragen muß ein Bericht vor der Entscheidungsfällung veröffentlicht werden, damit er von anderen "Meta"-Evaluatoren kritisch diskutiert werden kann.

<sup>7</sup> Eine andere Position nimmt SAWIN ein. Er möchte Evaluation nur auf jene Arbeiten beschränken, die mit hoher Reliabilität und Objektivität zu leisten sind. Der größte Teil des komplexen Geschehens z.B. der Bewertung eines Schulversuches bliebe damit zwar ohne Untersuchung, dies könne der Evaluator aber auch gar nicht nach wissenschaftlichen Maßstäben leisten. Entsprechend möchte SAWIN den Titel "Evaluator" durch "descriptive inquiry specialist" ersetzen (vgl. SAWIN 1976, S. 44f.).

Weise verändert werden muß. Vielmehr sind die evaluativen Informationen ein Element in einem komplexen Prozeß der Verbesserung von Fortbildungsveranstaltungen. Die Erwartungen an eine Evaluation müssen also in diesem Aspekt relativiert werden<sup>8</sup>.

## 2.2 Der Ansatz der "zielfreien" Evaluation

Eine Evaluation, die der Verbesserung der Fortbildungspraxis dienen soll, braucht Erfolgskriterien, d.h. Maßstäbe für eine "gelungene" Fortbildungsveranstaltung, mit denen die gerade untersuchte Veranstaltung verglichen werden kann (vgl. NACKEN 1977b, S. 72). Soll nun ein Evaluationsinstrument entwickelt werden, welches in einer Vielzahl unterschiedlicher Veranstaltungen einsetzbar ist, so kann es nicht die Ziele einer bestimmten Veranstaltung als Erfolgskriterien setzen, sondern muß sich von ihnen lösen. Hierfür

---

<sup>8</sup> Vor allem im größeren bildungspolitischen Rahmen von Curricula und Schulversuchen wird eine enge Verbindung von Evaluation und Entscheidung als Überschätzung der Wirkung einer Evaluation gesehen. Die Evaluation ist faktisch nur ein Element unter den vielfachen Bedingungen bildungspolitischer Entscheidungen (vgl. hierzu ALKIN 1975, S. 210; KLEINBERGER 1976, speziell S. 223; für die BRD BROCKMEYER 1978, S. 121; für britische und amerikanische Verhältnisse HAWKRIDGE 1978, S. 67). Dies bedeutet natürlich nicht, daß der Evaluator der Verwendung seiner Ergebnisse blind gegenüber steht. Zu seinen Aufgaben sollte es auch gehören, den Entscheidungsprozeß kritisch zu analysieren und - öffentlich - auf jene Stellen hinzuweisen, wo - unter Beruf auf seinen Bericht - die evaluative Rationalität unterlaufen wird. Ein Beispiel hierfür wäre es, wenn aus einem Bericht ein für einen Schulversuch negativer Aspekt aus dem Zusammenhang der positiven Aspekte herausgenommen und zu einem vorgeschobenen Grund für die Ablehnung eines Schulversuches wird.

Andererseits ist es möglich, daß aus bildungspolitischen Gründen eine andere Entscheidung getroffen wird, als der Evaluationsbericht sie nahe legt. Z.B. zeigt ein Bericht negative Englischleistungen von Gesamtschülern auf, so kann es trotzdem bildungspolitische Gründe geben, die für die Einführung der Gesamtschule sprechen.

bietet sich eine Erweiterung des Konzepts der "goalfree" Evaluation von SCRIVEN an (vgl. SCRIVEN 1978)<sup>9</sup>.

Der Grundgedanke der zielfreien Evaluation besteht darin, daß ein Evaluator dann seine Objektivität verliert, wenn er seine Arbeit nur auf die Überprüfung der Zielerreichung eines pädagogischen Programms hin ausrichtet. Beeinflußt von den Zielen und Intentionen besteht zudem die Gefahr, daß der Evaluator unintendierte Nebeneffekte übersieht. Aus diesem Grunde möchte SCRIVEN unabhängig von den Projektzielen die Evaluation durchführen. Dazu soll sich der Evaluator Gedanken über mögliche Folgen eines pädagogischen Programms machen - angeleitet durch seine Erfahrungen aus der Forschung, aus eigenen Untersuchungen und aus der Evaluationsliteratur - und diese in Erfolgskriterien umsetzen, die dann - unabhängig von den konkreten Zielen der Veranstaltung - die Maßstäbe für ein pädagogisches Programm bieten<sup>10</sup>.

---

<sup>9</sup> SCRIVEN diskutiert diesen Ansatz bezogen auf die Evaluation eines Curriculums. STUFFLEBEAM betrachtet die "goal-free evaluation" als eine Alternative zu vergleichenden, experimentellen Versuchsplänen (vgl. STUFFLEBEAM 1974, S. 48).

<sup>10</sup> LARCHER/RATHMAYR sehen hier die Gefahr, daß als Kriterien nur jene der bisherigen Schulpraxis genommen werden und deshalb die Evaluation zu einer konservativen Fessel, zu einem Stolperstein für curriculare Innovationen wird (vgl. LARCHER/RATHMAYR 1975, S. 691f.). Diese Gefahr kann bestehen, betrifft aber nicht den Ansatz als solches, sondern eine unzureichende Kriterienauswahl. Innovationstendenzen müssen ebenso wie Bestehendes in dem Kriterien-satz enthalten sein.

Gerade in formativen Phasen berührt eine Zielabänderung der Curriculumkonstrukteure die zielfreie Evaluation nicht. In diesem Sinne ist der Ansatz gerade keine konservative Fessel.

Zielfreie Evaluation braucht deswegen nicht blind gegenüber den konkreten Projektzielen sein. Der Evaluator soll sicherstellen, daß diese sich auch unter seinen Kriterien befinden. Er verliert dann seine Objektivität nicht, wenn er zuerst eigene Kriterien formuliert und danach überprüft, ob sich darunter auch die Projektziele befinden (vgl. McMORRIS 1975, S. 147; LINN 1978, S. 83f.).

Ein zielfreies Evaluationsinstrument für Erzieherfortbildung verlangt somit eine ausführliche Sammlung und Begründung von Erfolgskriterien, in denen sich die Intentionen der konkreten Veranstaltungen wiederfinden lassen. Evaluation ist hier nicht völlig zielfrei, da sie sich dem übergeordneten Ziel einer Verbesserung der Erziehungspraxis verpflichtet; sie abstrahiert nur von den spezifischen Veranstaltungszielen und damit auch von den spezifischen Inhalten.

Diese Abstraktion bedeutet zugleich aber auch ein Informationsverlust:

- indem Evaluation zielfrei durchgeführt und nur an Erfolgskriterien verglichen wird, gehen Informationen über die Erreichung der spezifischen Veranstaltungsziele verloren.
- indem Evaluation von den Inhalten abstrahiert, gehen Informationen bezüglich der Behandlung spezifischer Inhalte verloren. Es läßt sich z.B. nicht mehr feststellen, wie sich der Inhalt E der Inhalte A - F bewährt hat, wohl aber, wie die Inhalte insgesamt von den Teilnehmern aufgenommen wurden.

Dieser Informationsverlust erscheint jedoch akzeptabel, falls ein Evaluationsinstrument entwickelt werden kann, welches konkrete Hilfen zur Verbesserung einer Vielzahl von Erzieherfortbildungsveranstaltungen geben kann. Die Tragweite der Abstraktion von Zielen und Inhalten kann aber nicht allein theoretisch bestimmt werden, sondern hier kommt es ebenso auf Erfahrungen beim Einsatz des Instruments an.

## 2.3 Kriterien des Erfolgs

Während im vorherigen Abschnitt die Bedeutung eines "ziel- und inhaltsfreien" Evaluationsansatzes und die Wichtigkeit von Erfolgskriterien diskutiert wurde, soll nun geklärt werden, welche Variablen und Merkmale einer Fortbildungsveranstaltung als "Erfolgskriterien" betrachtet werden können. Der Begriff des "Erfolgskriteriums" wurde in der allgemeinen Unterrichtsforschung lange im Rahmen des "Paradigmas des Unterrichtserfolgs" diskutiert.

Paradigmen sind Werkzeuge für die Erforschung des Lehrens und Unterrichtens; sie sind Denkweisen, Forschungsmuster, mit deren Hilfe Forschung vorstrukturiert wird. Ist ein Paradigma gewählt, so ist eine ausschlaggebende Entscheidung darüber getroffen, welche Arten von Variablen und Beziehungen zwischen Variablen untersucht werden sollen (vgl. FEGE/TROTSBURG 1973, S. 13; KEIL/PIONTKOWSKI 1973, S. 19f.). Im Paradigma des "Kriteriums des Erfolgs" werden zu untersuchende Variablen in zwei Kategorien eingeteilt:

Kriteriumsvariablen als abhängige und Korrelate als unabhängige Variablen. Kriteriumsvariablen sind solche, an denen sich der Erfolg des Lehrens und Unterrichtens niederschlagen soll, während Korrelate jene Variablen meint, die mit diesen Kriteriumsvariablen korrelieren. Schematisch läßt sich dies an einem Beispiel verdeutlichen:

<u>Variablen der Korrelate</u>	wirken auf	<u>Variablen des Kriteriums</u>
z.B.	→	z.B.
- autoritäre Lehrerpersönlichkeit		- kognitive Schülerleistungen
- didaktisch/method. Fähigkeiten des Lehrers		- Schülerbeteiligung

Die Aufgabe der Unterrichtsforschung besteht nun darin, die Beziehungen zwischen Korrelaten und Kriteriumsvariablen zu untersuchen. Dabei kann es verschiedene Arten von Krite-



riumsvariablen geben. Zur Vereinfachung wurde versucht, diese Kriterien in einem hierarchischen Kontinuum zu ordnen. Dies reicht von letztlich "endgültigen" Kriterien (z.B. die Leistung des Schülers im Leben) bis hin zu einfachen Kriterien (z.B. Intelligenz des Lehrers). Inwieweit ein Kriterium der unteren Hierarchiestufe valide für das endgültige Kriterium ist, wird dann allein bestimmt durch seine Korrelation mit den Kriterien der höheren Positionen (letztlich der Leistung des Schülers im Leben) (vgl. FEGER/TROTSENBURG 1973, S. 32). Hier zeigt sich, daß die Begriffe "Kriterium" und "Korrelat" nur untersuchungsspezifisch bestimmt werden können, denn eine Variable wie "Intelligenz des Lehrers" kann einmal ein Kriterium sein, indem sie nämlich ein valider Indikator für Schülerleistung ist, während sie in einer anderen Untersuchung als unabhängige Variable gesetzt werden kann.

Das Paradigma des Kriteriums des Erfolgs ist nicht zwangsläufig reduziert auf die Wirksamkeit von Lehrerverhalten. Ein Lehrer ist nicht der alleinige Faktor für einen "gelungenen Unterricht". Es erscheint ohne weiteres möglich, als Korrelate nicht nur Lehrervervariablen zu untersuchen, sondern auch komplexere Variablen von Unterricht<sup>11</sup>.

Bezieht man diesen Begriff des Kriteriums auf Fortbildungsveranstaltungen im Vorschulbereich, so lassen sich zwei

---

<sup>11</sup>Vgl. auch ALEAMONI/SPENCER 1973, S. 670; KEIL/PIONTKOWSKI 1973, S. 275; die unabhängigen Variablen sind hier sehr komplex und beziehen sich auf die Bereiche Dozent, Umgebung, Prozess und Student.

Zur weiteren Diskussion des Paradigmas des Unterrichtserfolges siehe FEGER/TROTSENBURG 1973.

Gruppen von Erfolgskriterien unterscheiden<sup>12</sup>:

- Produktkriterien

Dies sind all jene Kriterien, die sich auf die kognitiven, emotionalen und praktischen "Produkte" beziehen, die sich am Ende einer Lerneinheit, der Fortbildungsveranstaltung oder einige Zeit später in der Erzieherpraxis feststellen lassen. Darunter fällt auch ein letztlich endgültiges, hierarchisch oberstes Kriterium des Erfolges einer Fortbildungsveranstaltung, nämlich die Verbesserung der Erziehungspraxis in den vorschulischen Einrichtungen.

- Prozeßkriterien<sup>13</sup>

Dies sind all jene Kriterien, die sich auf Lernprozesse während einer Fortbildungsveranstaltung beziehen. Dies

---

<sup>12</sup> Eine ähnliche Unterscheidung ist ein Ergebnis der Untersuchung zu Struktur und Prozeß im Hochschulunterricht von KEIL/PIONTKOWSKI. Die Autoren nehmen aus verschiedenen Faktorenanalysen die Faktoren heraus, die als direkte Operationalisierung von Lerneffekten zu interpretieren sind. Aufgrund der Interkorrelationen dieser Faktoren kommen sie zu einem Zirkumplexmodell der Bewertungskriterien, die sich dabei in zwei Kategorien einteilen lassen:

- direkte Kriterien, die sich beziehen auf Reaktion, Verarbeitung von Unterricht: emotionale Stabilität, passive Rezeption, engagierte Konzentration, Anreiz, stoffebezogene Zufriedenheit.
- indirekte Kriterien, die sich beziehen auf den Stoff und den Inhalt des Hochschulunterrichts: Schwierigkeit, Theoriehaltigkeit, Relevanz, Ordnung. (Vgl. KEIL/PIONTKOWSKI 1973, S. 278-285. Vgl. auch die Unterscheidung zwischen "process of teaching" und "product of teaching" bei JAEGER/FREIJO 1974, S. 417).

<sup>13</sup> Der Begriff der Prozessevaluation wird oft unklar benutzt. Dabei meinen "Prozesse" und "Prozessevaluation" etwas unterschiedliches, obwohl sich Überschneidungen ergeben können. SCRIVEN und STUFFLEBEAM (letzterer in CIPP-Modell) verstehen unter Prozessevaluation die Bewertung der "vorläufigen Implementation" eines Curriculums, um Verbesserungen zu ermöglichen, d.h. ein Stadium im Entwicklungsprozeß (vgl. SCRIVEN 1972a, S. 71; STUFFLEBEAM 1972, S. 135f.). "Prozesse" meint dagegen nun - egal ob es sich um die Bewertung eines Curriculums im Entwick-

betrifft z.B. die ausgewählten Inhalte, die Vermittlungsmethoden oder die Arbeitsatmosphäre in der Veranstaltung. Oftmals wird eine Evaluation anhand von Produktkriterien als nicht möglich bezeichnet und sich deshalb auf Prozeßkriterien beschränkt. Prozeßkriterien sind normative Setzungen,

"die bestimmte Bedingungen des Lernprozesses negativ oder positiv mit dem nicht unmittelbar meßbaren Lernprodukt in Verbindung bringen". (HOLZAPFEL u.a. 1977, S. 6)

Es wird in der Evaluation dann zwar überprüft, ob die normativen Setzungen erfüllt waren, nicht aber, ob sie die postulierten Lernprodukte hervorbringen. In der Terminologie des Paradigmas des Erfolgs sind diese Prozeßkriterien Korrelate, die genügend hoch mit den Produktkriterien und somit letztlich mit der "Verbesserung der Erziehungspraxis im Vorschulbereich" korrelieren. Diese Indikatorfunktion läßt sie erst zu Erfolgskriterien werden.

Es soll hier aber nicht der Eindruck eines einheitlichen, hierarchisch aufgebauten und sowohl theoretisch wie empirisch begründeten Systems von Erfolgskriterien geweckt werden, denn ein solches System existiert nicht<sup>14</sup>. Vielmehr muß folgendes berücksichtigt werden:

---

noch 13 lungsprozeß oder im Endstadium handelt - jene Prozesse, die sich bei der Anwendung und Durchführung eines Curriculums ergeben (vgl. CRONBACH 1972, S. 53f.; STAKE 1972, S. 98f.). In diesem letzteren Sinne wird hier von Prozessen innerhalb von Fortbildungsveranstaltungen gesprochen.

<sup>14</sup> Für die Erwachsenenbildung allgemein sieht GERL ein Dilemma bei der Suche nach inhaltlichen Kriterien. Einerseits kann die Evaluation die richtigen Fragen nur dann stellen, wenn ein stringenter theoretischer Bezugsrahmen von Erwachsenenunterricht existiert, der aufzeigt, welches z.B. die relevanten Merkmale einer Lernsituation sind. Zum anderen liegt aber gerade dieser theoretische Bezugsrahmen nicht vor (vgl. GERL 1977, S. 10f.); zum allgemeinen Stand der Unterrichtsforschung in der Weiterbildung siehe SCHLUTZ 1979.

- Obwohl das endgültige Kriterium "Verbesserung der Erziehungspraxis" allgemein geteilt wird, wird es unterschiedlich interpretiert und es lassen sich aus ihm keine konkreten Erfolgskriterien, die Aussagen über wünschenswerte Gestaltung von Fortbildungsprozessen machen, deduktiv ableiten und damit legitimieren. Konkrete Erfolgskriterien mittlerer Ebene lassen sich zwar allgemein auf die Verbesserung der Erziehungspraxis beziehen, müssen aber selbst in eigener Argumentation inhaltlich legitimiert werden.

- Korrelationsstudien, die die Beziehungen zwischen Korrelaten und Kriterien, bzw. Erfolgskriterien "unterer" Ebene und Erfolgskriterien "höherer" Ebene untersuchen, fehlen weitgehend. In der Literatur findet sich zwar eine Vielzahl von normativen Setzungen und Postulaten für die Gestaltung von Fortbildungsprozessen, empirische Überprüfungen der postulierten Beziehung zu den gewünschten Produkten der Fortbildung sind kaum zu finden (vgl. GERL 1977, S. 11; für den Bereich der allgemeinen Unterrichtsforschung ROENSHINE/FURST 1971, S. 37-40). Aussagen über Beziehungen zwischen methodischen Arrangements und erfolgreichem Lernen in der Fortbildung im Vorschulbereich beruhen meist auf Plausibilitätsannahmen; die oben angesprochene Indikatorfunktion wurde oft nicht empirisch überprüft<sup>15</sup>.

---

<sup>15</sup> Es könnte eingewandt werden, daß vor Evaluationsstudien dann zuerst Erforschungen zumindest korrelativer Beziehungen - möglichst aber als experimentelle Studien - durchgeführt werden müßten. So legitim dieser Einwand ist, so muß doch andererseits festgehalten werden, daß dann heute Evaluationsadressaten keine Informationen erhalten würden. So muß für die vorliegende Arbeit - die hic et nunc Adressaten eine Hilfe geben will - akzeptiert werden, daß die gelieferten Informationen mit Schwächen behaftet sind. Es ist die Aufgabe des Evaluators, diese Schwächen und deren Gründe aufzuzeigen, damit die evaluativen Informationen in ihrer Tragweite richtig eingeschätzt werden können. (vgl. auch Kapitel 4.1, Anmerkung 6, S. 107f.). Außerdem bietet sich hier letztlich auch die Möglichkeit, die Evaluationsfrage in Beziehung zur Forschungsfrage zu stellen. So kann die Arbeit an den Erfolgskriterien das

Heuristisch lassen sich die Prozeßkriterien in die folgenden 4 Bereiche der Gestaltung einer Veranstaltung gliedern<sup>16</sup>:

### - Thematik

Unter Thematik werden Inhalte verstanden, die unter pädagogischer Zielsetzung für Fortbildung ausgewählt wurden (vgl. KLAPKI 1976, S. 83). Welche Inhalte werden bei welchen Zielvorstellungen für relevant gehalten? Die Zusammenfassung der Kategorien Ziele/Intentionen und Inhalte zu Thematik soll

---

noch 15 Bemühen stimulieren, verschiedene Korrelatsvariablen von Fortbildung mit dem hier entwickelten Kriterium zu korrelieren. Solche Erforschungen können dabei neben dem Erkenntnisgewinn zugleich die angebotenen Erfolgskriterien differenzieren (siehe hierzu auch den ersten Ansatz im Kapitel 6.4 über die Validität des Fragebogens S. 196ff.).

<sup>16</sup> Es wird hier nicht der Anspruch erhoben, daß hiermit ein Modell von Fortbildung vorliegt, wohl aber sollen mit dieser Strukturierung die Merkmale eines Modells angestrebt werden: das komplexe Feld Fortbildung soll auf wenige Strukturelemente reduziert werden; damit werden bestimmte Aspekte akzentuiert, was zusammen mit der Reduktion das Feld Fortbildung transparenter macht; die einseitige Hervorhebung bestimmter Merkmale erzeugt eine perspektivische Sicht des Feldes, die evtl. zur Ausarbeitung anderer Perspektiven herausfordert (vgl. POPP 1970, S. 53-56). Mit dieser Strukturierung wird "Fortbildung" nicht abgebildet, sondern eine mögliche Sichtweise selegiert, die durch andere Sichtweisen ergänzt und relativiert werden kann.

Die hier vorgelegte Strukturierung entstand in Auseinandersetzung mit folgenden anderen Strukturierungen von Unterricht allgemein bzw. Erwachsenenunterricht:

- SCHULZ 1970 gliedert die Didaktik als Theorie des Unterrichts in die Entscheidungsfelder Ziele, Inhalte, Methoden und Medien und in die Bedingungsfelder anthropologische und sozial-kulturelle Voraussetzungen.
- Davon beeinflusst beschreibt KLAFKI 1976, S. 78 vier Hauptdimensionen der Didaktik i.w.S., in denen Entscheidungen getroffen werden müssen: Ziele, Inhalte, Organisations- und Vollzugsformen von Unterricht und Medien.
- In Erweiterung dieser Vorstellungen beschreibt MEISTER 1974, S. 61f. ein 8 Komponenten- bzw. Bedingungsmodell von Hochschulunterricht. Er unterscheidet: Lehr- und Lernziele, Organisation und Umfeldbedingungen, Lehr- und Lernmethoden, Medien, Lehr- und Lerninhalte, sozialpsycho-

hier pragmatisch verstanden werden, da sich z.B. in der unten beschriebenen Literaturanalyse zeigte, daß es oft sehr schwierig ist, Ziele und Inhalte voneinander zu trennen.

---

noch 16 logische und soziokulturelle Bedingungen, personale Faktoren der Lernenden, personale Faktoren der Lehrenden.

- Ebenfalls für Hochschulunterricht gliedern KEIL/PIONT-KOWSKI 1973, S. 25 das Gesamtsystem Hochschulunterricht in vier Teilsysteme: Dozent, Student, Umgebung und Prozeß.

-DIEHL/KOHR 1977, S. 70 kommen aufgrund von Faktorenanalysen zu vier Dimensionen der Bewertung von Hochschulseminaren: Relevanz und Nützlichkeit der Inhalte, Verhalten der Dozenten gegenüber Teilnehmern, Schwierigkeiten und Umfang von Inhalten, Methodik und Aufbau der Veranstaltung.

- Zur Effizienzkontrolle von Fortbildungsmaßnahmen in der Praxis evangelischer Jugendarbeit beschreibt WEBER 1977, S. 16, vier Einflußvariablen von Fortbildung: Teamstruktur der Dozenten eines Instituts, Teilnehmerstruktur, Didaktik und Ziele.

- FEIG 1972, S. 140ff. entwirft 6 Grundvariablen für die Erwachsenenbildung, deren Beziehungen die Pädagogische Psychologie zu erforschen hat: Lehrziele, Lehrpersonen, Lehrgegenstände, Lehrmethoden, Lernpersonen und situative Bedingungen.

- LEON 1977, S. 35f. unterscheidet 4 Komponenten des pädagogischen Aktes: Kursleiter, Teilnehmergruppe, Unterrichtsgegenstand.

- GERL 1977, S. 4-6 nennt als Bereiche der Evaluation neben den Produkten: alle Prozesse der Auseinandersetzung mit dem Lerngegenstand, alle in den Lernprozeß eingeführten Materialien und Konzepte, alle fördernden und hemmenden Aspekte der Umwelt der Lerngruppe. Diese Begriffe (Produkte, Prozesse, Input und Kontext) lehnen sich an das CIPP-Modell von STUFFLEBEAM an (vgl. STUFFLEBEAM 1972, S. 132-137); GERL definiert die Bedeutung der Begriffe jedoch leicht um, indem er sie weniger auf den Prozeß der Curriculumentwicklung - so bei STUFFLEBEAM - bezieht, sondern auf die Bedingungen und Aspekte der konkreten Lernarbeit in Gruppen. Damit ähneln die Begriffe "Produkte", "Prozesse" und "Context" mehr den Evaluationsfeldern "Voraussetzungen", "Prozesse" und "Ergebnisse" bei STAKE (vgl. STAKE 1972, S. 98-100), wobei als 4. Feld die Bewertung des Materials hinzukommt. Obwohl diese Versuche teilweise verschiedene Aspekte ansprechen, lassen sich doch gemeinsame Strukturen erkennen. So werden fast durchgehend Ziele, Inhalte und Methoden angesprochen. Analytisch zu trennen vom Bereich Methode sind die Interaktionsprozesse im Unterricht, die sich

### - Methodik

Damit ist der methodische Aufbau einer Fortbildungsveranstaltung gemeint. Wie wird die Thematik den Teilnehmern vermittelt? Hierunter fallen auch die Medien der Fortbildung, da in diesem Bereich es kaum eine Eigenständigkeit von Medien gibt.

### - Interaktionsprozesse

Dies bezieht sich auf die verschiedenen Interaktionen zwischen Teilnehmern und Referenten/Leitern, bzw. der Teilnehmer untereinander. Wie läßt sich der Arbeitsstil<sup>17</sup>, das Arbeitsklima der Fortbildung charakterisieren?

### - situative und organisatorische Bedingungen

Darunter fallen alle fördernden oder hemmenden Randbedingungen, unter denen die eigentliche Fortbildungsarbeit stattfindet. Unter welchen organisatorischen und situativen Bedingungen wird gelernt?

Bei dieser Strukturierung wird nicht angenommen, daß die vier Bereiche der Gestaltung von Fortbildungsprozessen voneinander unabhängig sind. So gibt es engere Beziehungen

---

noch 16 auf die Interaktionen zwischen Lehr- und Lernpersonen beziehen. Ein weiterer wichtiger Bereich wären dann noch allgemeine Organisations- und Umweltbedingungen. Durch eine Beschränkung auf diese Aspekte gehen evtl. Informationen verloren. Trotzdem wird erhofft, alle relevanten Aspekte zu erfassen; relevant in dem Sinne, daß entsprechende Informationen unter der Evaluationsfrage zur Bewertung wichtig sind.

<sup>17</sup> TIETGENS/WEINBERG sprechen für Erwachsenenbildungsveranstaltungen lieber von Arbeitsstil als von Erziehungsstil, da es sich in der Erwachsenenbildung nicht um ein ausdrückliches Erziehungsverhältnis handelt (vgl. TIETGENS/WEINBERG 1971, S. 158f.). Für SCHULENBERG ist letzteres die Grundlinie für alle Theorie der Erwachsenenbildung (vgl. SCHULENBERG 1971, S. 65f.).

zwischen den Bereichen Methodik und Interaktionsprozesse, indem bestimmte Forderungen bezüglich der Interaktionsprozesse (z.B. "teilnehmerorientierter Arbeitsstil") Auswirkungen haben auf die Art der Vermittlungsformen (z.B. "Partizipation an der Verlaufsplanung"). Betont auch die Strukturierung - implizit - die planenden und aktiven Referenten und Fortbildungsleiter, so wird damit nicht der teilnehmende Erzieher nur als passiver Adressat der Planungsbemühungen gesehen. Fortbildungsprozesse und Fortbildungserfolg sind ein Ergebnis der Zusammenarbeit von Referenten/Leiter und Teilnehmern. Aber auch die Forderung nach einem aktiven Lerner, nach Teilnehmerorientierung oder reziproker Kommunikation weisen dem Referenten/Leiter eine herausgehobene Funktion als Initiator, Kontrolleur solcher Forderungen zu, da es unrealistisch ist anzunehmen, daß die Teilnehmer diese Leistung per se erbringen.

Prozeßkriterien des Erfolges einer Fortbildungsveranstaltung müssen in allen vier Bereichen gesucht werden. Schematisch in einem Raster zusammengefaßt bedarf eine umfassende Evaluation somit folgender Erfolgskriterien:

<u>Prozeßkriterien</u>	<u>Produktkriterien</u>
bezügl.	bezügl.
- Thematik	- kognitiven, emotionalen und
- Methodik	praktischen Lernertrag
- Interaktionsprozesse	- Veränderung der Erziehungs-
- situative und organisatorische Bedingungen <sup>18</sup>	praxis im Vorschulbereich

<sup>18</sup> Hier überschneidet sich die Evaluation der Prozesse von Fortbildung zum Teil mit der Evaluation der Voraussetzungen. Darunter versteht STAKE jene Bedingungen, die vor dem Unterricht bestehen und Einfluß auf die Ergebnisse haben können, z.B. Geldmittel, Raumausstattung (vgl. STAKE 1972, S. 98).



## 2.4 Die Funktion einer Literaturanalyse

Der im letzten Abschnitt entwickelte "Raster" gibt nur die Bereiche an, aus denen konkrete Erfolgskriterien gewonnen werden können. Der ziel- und inhaltsfreie Evaluationsansatz muß zwar die Maßstäbe seiner Adressaten berücksichtigen (vgl. S. 23), seine Hauptaufgabe besteht aber primär darin, unabhängig von den Adressaten potentielle Erfolgskriterien zu entwickeln. Dies ist hier nur durch eine umfassende und differenzierte Literaturanalyse möglich. Diese Literaturanalyse muß einmal den Produkt- und Prozeßbereich ausdifferenzieren und konkrete Erfolgskriterien beschreiben und begründen. Desweiteren - speziell angesichts oft fehlender Korrelationsstudien zwischen Prozeßmerkmalen und Ergebnissen von Fortbildung (vgl. S. 28) - muß ein besonderes Augenmerk auf plausible Begründungen für die vielfältigen Prozeßkriterien und deren Auswirkungen gelegt werden.

In der - durch den obigen Raster strukturierten - Literaturanalyse kann es im Rahmen eines ziel- und inhaltsfreien Evaluationsansatzes nicht um die Beschreibung und Begründung eines spezifischen, didaktisch-methodischen Fortbildungsmodells gehen. Zum einen steht zu erwarten, daß die alleinige Berücksichtigung eines theoretischen Konzepts der Vielfalt von Fortbildungsprozessen nicht gerecht wird. Zum anderen soll ja ein Evaluationsdesign und -instrument entwickelt werden, welches von unterschiedlichen Adressaten benutzt werden kann. Dies setzt voraus, daß die Adressaten der Evaluation die vom Evaluator vorgelegten Erfolgskriterien und normativen Setzungen teilen bzw. zumindest als mögliche Prüfsteine anerkennen können<sup>19</sup>. Die Operationalisierung eines theoretischen Ansatzes (z.B. der The-

---

<sup>19</sup> Vgl. dazu den Ansatz zur Auswertung der Lernprozesse auf dem Bildungsurlaub der "Wissenschaftlichen Begleitung des Bildungs - Urlaubs - Versuchs- und Entwicklungsprogramms (BUVEB)" bei HOLZAPFEL u.a. 1977, S. 8ff.

menzentrierten Interaktion nach COHN) erreicht dieses Ziel nicht, sondern eher der Versuch, Aussagen aus verschiedenen theoretischen Ansätzen und Konzepten zu integrieren. Dieser Pluralismus theoretischer Bezugskonzepte soll sichern, daß eine Fortbildungsveranstaltung nicht nur an einseitigen, sondern an möglichst vielfältigen Erfolgskriterien gemessen wird.

Erfolgskriterien für eine Fortbildungsveranstaltung, Postulate und Regeln für z.B. eine "gutes" Referentenverhalten sind in der Literatur oft sehr allgemein und pauschal formuliert. Dies reflektiert nicht nur den Forschungsstand bezüglich der Fortbildungstheorie, sondern primär auch den Orientierungscharakter solcher Erfolgskriterien. M.a.W. heißt dies, daß auch ein Verstoß gegen solche Kriterien (z.B. Kriterien der methodischen Gestaltung) manchmal einen theoretisch unerwarteten Lernerfolg erzielen kann (vgl. TIETGENS 1974, S. 19f.). Bei hohem Allgemeinheitsgrad der Regeln und Kriterien wird nämlich zwangsweise von den konkreten, situativen Realisierungsbedingungen der Kriterien abgesehen. Dazu gehört auch, daß Erwachsene nicht unbedingt in der Weise am besten lernen, die theoretisch als die beste gilt (z.B. "Eigenständigkeit", "kooperativer Arbeitsstil"), sondern so, wie sie es gewohnt sind. Thesen über das Verhältnis von Leitterverhalten und Lernerfolg müssen also situationsbedingt relativiert werden (vgl. TIETGENS/WEINBERG 1971, S. 55). Aufgrund der empirischen Forschung können nur Aussagen über wahrscheinliche Zusammenhänge von Lehr- und Lernverhaltensweisen gemacht werden, die im Einzelfall problematisiert werden müssen (vgl. SIEBERT 1977, S. 129).

Dies ist aber kein Argument gegen den Orientierungswert solcher Erfolgskriterien und Postulate. Bewertung und Verbesserung von Fortbildung als Funktionen der Evaluation bedeuten nicht ein blindes Abwickeln eines Kriteriums-systems und ein Bestehen auf abstrakten Maßgaben für z.B.

das Leiterverhalten. Vielmehr muß der Evaluator in seinem Bericht deutlich machen, daß die evaluativen Informationen nur durch die theoretische Brille solcher Erfolgskriterien und Orientierungsregeln gewonnen werden können. Evaluative Informationen stellen somit ein Bewertungsangebot dar, welches nicht "technologisch" verwandt werden sollte, sondern ein Angebot für ein Gespräch zwischen den Verantwortlichen ist, und dazu gehören potentiell auch die Teilnehmer.

## 2.5 Das Verhältnis der Evaluation zu ihren Adressaten

Evaluation dient der Verbesserung von Fortbildung, dazu soll sie primär Teilnehmern, Referenten und Fortbildungsleitern nützlich sein (vgl. S. 18). Dies impliziert zwar, daß jene Fragen identifiziert werden müssen, die den genannten Adressaten relevant und wichtig zur Verbesserung von Fortbildung sind, bedeutet aber nicht, daß Evaluation - quasi "adressatenblind" - nur jenes erhebt, was die Adressaten selbst als relevant bezeichnen. Der ziel- und inhaltsfreie Evaluationsansatz integriert vielmehr die Auswahl der "relevanten" Fragen in ein mehrphasiges Modell (vgl. LINN 1978, S. 83f.)<sup>20</sup>:

- Unabhängig von den potentiellen Adressaten und deren Ziele bzw. Erfolgskriterien einer Fortbildungsveranstaltung sammelt, strukturiert und begründet in der ersten Phase die oben beschriebene Literaturanalyse einen Satz von Erfolgskriterien. Diese Unabhängigkeit der Kriterienauswahl von den Adressaten - und darin besteht das Charakteristikum des ziel- und inhaltsfreien Ansatzes - besteht einmal aus Gründen der Objektivität und soll zum anderen

---

<sup>20</sup> Ähnliche mehrphasige Modelle der Auswahl der für die Adressaten/Entscheidungsträger relevanten Fragen - nicht speziell bezogen auf den zielfreien Evaluationsansatz - finden sich z.B. bei STRATON 1977, S. 59-60 und ausführlich bei MATUSZEK 1978, S. 349-352.

sicherstellen, daß unintendierte Nebeneffekte nicht übersehen werden (vgl. Kap. 2.2).

- In der zweiten Phase wird nach dieser Formulierung durch den Evaluator der Kriteriensatz auf verschiedenen Wegen potentiellen Adressaten sowie Fortbildungsexperten vorgelegt; zu denken wäre z.B. an Interviews oder Prioritätenlisten. Diese Phase erhebt eine kritische Einschätzung und überprüft zugleich, ob sich unter diesen Kriterien auch die Erfolgsmaßstäbe und kritischen Aspekte befinden, die den Evaluationsadressaten relevant erscheinen (vgl. bes. S. 23 dieser Arbeit).

- Die Stellungnahmen werden anschließend gesammelt und bei verschiedenen Revisionen des Kriteriensatzes berücksichtigt.

Dieses mehrphasige Vorgehen des ziel- und inhaltsfreien Ansatzes soll sicherstellen, daß das zu entwickelnde Evaluationsinstrument - im Endstadium - für die konkreten Adressaten wichtige und nützliche Informationen erhebt, die außerdem über eine Fortbildungsveranstaltung einschließlich unintendierter Nebeneffekte umfassend informieren und die verschiedenen Aspekte unbeeinflusst von spezifischen Zielen und Kriterien bestimmter Adressaten erfassen.

Das Verhältnis der Evaluation zu ihren Adressaten wird noch unter einem weiteren, das hier skizzierte Vorgehen durchgehend beeinflussenden Gesichtspunkt wichtig. Es handelt sich dabei um die Frage, wer die Fortbildungsveranstaltungen evaluieren soll<sup>21</sup>. Eine Antwort muß getrennt

<sup>21</sup> Allgemein wird dies diskutiert unter der Differenzierung zwischen interner und externer Evaluation. Bei der internen Evaluation ist der Evaluator Mitglied des Konstruktions-teams, bei der externen ein Außenstehender. Interne Evaluation wird oft für Zwecke der Optimierung, externe für Zwecke der Legitimierung empfohlen. Schwierigkeiten bei einer internen Evaluation werden oft in mangelnder Objektivität und Glaubwürdigkeit für bestimmte Adressaten gesehen (vgl. STUFFLEBEAM 1974, S. 9 u. S. 20; WULF 1972, S. 22f.). Eine externe Evaluation wird von den Konstrukteuren oft als eine Bedrohung wahrgenommen (vgl. z.B. HAWKRIDGE 1978, S. 65f.; SAWIN 1976, S. 42). Zu diesen Organisationsfragen der Evaluation in der Fortbildung

für das Entwicklungs-/Erprobungsstadium und für die Anwendung des - relativ - fertig entwickelten Evaluationsinstruments gegeben werden:

- Angestrebt ist eine Selbstevaluation durch die Beteiligten mit Hilfe eines erprobten Instrumentes. Das Ziel der Evaluation besteht in der Optimierung von Fortbildungsveranstaltungen, um damit indirekt einen Beitrag zur Verbesserung der Erziehungspraxis zu leisten. Letztere ist nur möglich, wenn in der Veranstaltung selbst die teilnehmenden Erzieher gleichberechtigte Partner und Subjekte ihrer Lernprozesse sind (vgl. hierzu die ausführliche Begründung in Kap. 3.1). Auch die Evaluation kann dies nun unterstützen, indem die Teilnehmer selbst ihre Lernerfahrungen reflektieren und damit ihre eigene Veranstaltung auswerten<sup>22</sup>. M.a.W. wird hier in diesem - nicht primär bildungspolitisch bestimmten - Bereich für eine Evaluation durch die Betroffenen selbst plädiert.

Eine solche Selbstevaluation gelingt aber nicht ohne Hilfsmittel von außen. Den Beteiligten fehlt oft eine Distanz und ein systematischer Überblick über die eigenen Lernprozesse. Deshalb müssen für diese Selbstevaluation Instrumente bereitgestellt werden, die eine objektivierende Distanz zu den eigenen Erfahrung herstellen. Die Beteiligten erheben mit dem Evaluationsinstrument systematisch Informationen, bereiten sie auf und leiten sie an eine gemeinsame Diskussion zwischen Teilnehmer, Referenten und Leiter weiter, in der dann zusammen über die Lernerfahrungen noch 21 tischen Versuchen in der BRD vgl. BROCKMEYER 1978, S. 120ff.; MITTER/WEISHAUPT 1978, S. 83f. und S. 92f.

<sup>22</sup> Für Grundqualifikationsseminare in der Erwachsenenbildung betont GERL diesen zentralen Evaluationszweck einer didaktischen Evaluation zur Verbesserung der Selbststeuerungsmöglichkeiten der Lerngruppe und der Selbstkontrolle des Kursleiters (vgl. GERL 1977, S. 15).

reflektiert und ggf. über Verbesserungsmöglichkeiten entschieden wird. Gerade für diesen Anwendungskontext bedarf es aber eines gut entwickelten und in seiner Qualität abschätzbaren Evaluationsinstrumentes<sup>23</sup>.

- Dieser Auswertungskontext kann während der Entwicklung und Erprobung nicht erreicht werden. Ein in seiner Qualität abschätzbares Evaluationsinstrument verlangt umfangreiche Erprobungen in einem größeren Sample von Fortbildungsveranstaltungen und ebenfalls ausführliche teststatistische Untersuchungen, die von dem außenstehenden Verfasser übernommen werden müssen. Es wird diesem "außenstehenden" Evaluator nicht möglich sein, an allen Veranstaltungen selbst teilzunehmen. Dies bedingt, daß die in den Veranstaltungen eingesetzten Instrumente dem Verfasser zugeschickt, von ihm ausgewertet und dann dem Veranstalter zurückgemeldet werden müssen. Damit sind mehrere Probleme zu erwarten.

Einmal liegen die Auswertungsinformationen erst 1 - 2 Wochen nach einer Veranstaltung vor, d.h. sie erreichen nur Träger, Fortbildungsleiter und Referenten, da die teilnehmenden Erzieher zu diesem Zeitpunkt schon abgereist sind. Ebenso findet die Auswertung nur zum Abschluß einer Veranstaltung statt. Auch wenn hier die Auswertungsinformationen rechtzeitig für alle Adressaten geliefert werden könnten, nütz-

---

<sup>23</sup> Der Einwurf, daß das Ziel der Selbstevaluation und gleichberechtigten Partnerschaft schon dadurch nicht erreicht wird, daß das Evaluationsinstrument die Teilnehmer ja zu Objekten macht, wird nicht geteilt. Der kurzfristige "objekthafte" Status während des Ausfüllens ist aufhebbar und an einen gemeinsamen Diskussionskontext gebunden, indem die Subjekthaftigkeit der Teilnehmer gerade gesichert ist. Außerdem bedeutet die Vergegenständlichung der Teilnehmer zwecks Informationserhebung ja nicht, daß auch im zwischenmenschlichen Geschehen der andere ein Objekt ist. Auf diesen Unterschied zwischen objektivierender Betrachtung und zwischenmenschlichem Handel hat schon WINNEFELD hingewiesen (vgl. WINNEFELD 1972, S. 47f.).

lich wären sie nicht primär für die Verbesserung der gerade abgeschlossenen, sondern nur für folgende Veranstaltungen. In der Entwicklungsphase ist eine Auswertung in der Mitte der Veranstaltung nicht möglich.

Es ist ebenfalls zu erwarten, daß die Fortbildungsreferenten sich durch eine Bewertung ihrer Arbeit durch Fremde bzw. Anonyme bedroht fühlen und sich gegen die Evaluation sperren<sup>24</sup>.

Diese Probleme reflektieren keine prinzipiellen Grenzen des - an empirisch-analytischer Methodologie orientierten - Evaluationsansatzes, sondern Zwänge, die durch die Notwendigkeit entstehen, für den oben beschriebenen idealen Auswertungskontext ein in seiner Qualität abgesichertes Evaluationsinstrument zu entwickeln.

## 2.6 Zur Auswahl eines Evaluationsinstruments

Im vorherigen Abschnitt wurde der Gedanke einer Selbstevaluation von Fortbildung durch die unmittelbar Beteiligten

---

<sup>24</sup> Nach der brieflichen Mitteilung eines Fortbildungsreferenten. Noch stärker wird diese Bedrohung, wenn Evaluationsergebnisse dazu führen können, daß Referenten ausgewechselt werden; siehe hierzu Tendenzen in der Auswertung von Weiterbildungskursen in der Industrie (vgl. FASSBENDER 1974, S. 10f. und S. 15). Solch weitreichende Entscheidungen setzen aber zumindest eine Qualität der Evaluation voraus, die beim momentanen Stand der Evaluationsverfahren in diesem Bereich kaum zu erreichen ist.

Ein anderes Beispiel für Ängste, Unsicherheiten und Mißverständnisse gegenüber der Evaluation sind 12 leere Fragebögen, die aus einer Fortbildungsveranstaltung in der 1. Entwicklungsphase stammen: Als die Bögen schon verteilt waren, opponierte der Referent. Alle Teilnehmer schrieben dann - anscheinend gelenkt durch den Referenten - den folgenden, identischen Text auf den Bogen:

"Ich bitte um eine Fortführung der Arbeit mit Dr....., damit tiefgreifend gearbeitet werden kann."

mit Hilfe eines Distanz schaffenden Instruments erläutert. Dieser grundsätzliche Gedanke stellt unumgängliche Ansprüche an das auszuwählende Evaluationsinstrument:

- Die Evaluation ist eingebettet in einen gemeinsamen Kommunikationsprozeß "Auswertungsgespräch" der Beteiligten, in dem die teilnehmenden Erzieher gleichberechtigte Partner der Referenten/Fortbildungsleiter und Subjekte selbstbestimmter Lernprozesse sind. Für die Teilnehmer sind zur Bewertung einer Veranstaltung primär ihre eigenen Erfahrungen und Wahrnehmung der Veranstaltung relevant und ein-sichtig. Das Evaluationsinstrument muß mithin so aufge-baut sein, daß es diese Erfahrungen und Meinungen der Teil-nehmer sammelt. Selbstevaluation von Fortbildung bedeutet dann, daß den Teilnehmern ein für sie handhabbares Instru-ment zur Verfügung gestellt wird, mit dem sie ihre eigenen Erfahrungen reflektieren, sammeln und einem gemeinsamen Auswertungsgespräch zuleiten können. Die gleichberechtigte Partnerschaft in der Auswertung der Lernprozesse impliziert auch für den Aufbau und die Einführung des Evaluationsin-strumentes ein offenes Vorgehen, in dem die teilnehmenden Erzieher keine Objekte sind, aus denen mit Tricks etwas herausgeholt wird, was sie möglichst nicht verstehen sollen.
- Zugleich muß das Evaluationsinstrument einen kontrollier-ten, objektiven Rückmeldeprozeß, eine Methodisierung der Erfahrungssammlung bieten. Die Aufmerksamkeit der Teil-nehmer soll auf bestimmte, ausgewählte und begründete Aspekte einer Fortbildungsveranstaltung - d.h. auf die oben beschriebenen Erfolgskriterien - kanalisiert werden. Das Instrument soll eine objektivierende Distanz zu den eigenen Erfahrungen schaffen und diese systematisiert sam-meln; es soll dabei die Erfahrungen und Meinungen von allen Teilnehmern - auch solchen, die im freien Gespräch gehehmt sind - erfassen. Dieser Sammlungsprozeß soll objektiv sein, d.h. die Erfahrungen und Wahrnehmungen einer Veranstaltung sollen unverzerrt durch bewußte oder unbewußte Beeinflussun-



gen, die sich z.B. im persönlichen Gespräch mit Referenten und Leitern ergeben können, erhoben werden. Dies bedeutet zudem, daß das Instrument in seiner Qualität mit empirischen Mitteln überprüft werden muß.

- Das Evaluationsinstrument soll Informationen liefern, die die Grundlage für ein gemeinsames Auswertungsgespräch bieten. Die Anwendung des Instruments und die Auswertung und Aufbereitung der Informationen darf dann nur wenig Zeit in Anspruch nehmen, damit die Informationen auch rechtzeitig vor Beginn des Gespräches vorliegen. Angestrebt ist hier ein Instrument mit guter Einsetzbarkeit bei niedrigem Aufwand an Kosten, Zeit und Energie. Erst dadurch wird eine Praxisrelevanz des beschriebenen Evaluationsansatzes und -instruments erreicht.

Abschließend ist noch eine relativierende Bemerkung nötig. Es ist nicht möglich, mit einem Evaluationsinstrument alle Aspekte einer Fortbildungsveranstaltung erschöpfend zu erfassen, vielmehr verlangt eine umfassende Evaluation die Benutzung mehrerer Instrumente, die jeweils verschiedene Aspekte aus evtl. verschiedenen Blickwinkeln betrachten und so die Informationen aus der einen Quelle durch weitere aus einer anderen Quelle validieren oder auch relativieren. Es läßt sich hier von einem "multimethodischen Ansatz" der Evaluation sprechen (vgl. LEWY 1973, S. 321f.; McMORRIS 1975, S. 148; SABAR 1975, S. 193). Evaluation kann dabei nicht alles formell untersuchen (vgl. LEWY 1973, S. 323); informelle Verfahren gehören notwendig dazu, wenn unter Zeit- und Ressourcenmangel eine Veranstaltung in allen Aspekten erfaßt und Entscheidungen vorbereitet werden sollen<sup>25</sup>. Das Evaluationsinstrument - welches in dieser

---

<sup>25</sup> In der neueren Evaluationsliteratur - z.B. in den "studies in educational evaluation (see)" - werden solche informelle Vorgehensweisen breit diskutiert. Dies umfaßt z.B. die Sammlung von Materialien in einer "locker organisier-

Arbeit entwickelt wird - ist hier nur eine mögliche Vorgehensweise, die schnell und routinemäßig Auswertungsinformationen erhebt und nicht den Anspruch auf Totalerfassung stellt. Ein solches Instrument muß z.B. durch informelle Informationssammlung ergänzt werden<sup>26</sup>.

---

noch 25 ten Mappe" (program's self-evaluation portfolio) (vgl. SCHEYER/STAKE 1976) oder die - m.E. unbefriedigende - Konzeptionierung eines "idiographischen" Evaluationsansatzes gegenüber einem angeblich "nomothetischen" (vgl. KEMMIS 1978). Besonders breit diskutiert wird der Ansatz der "responsiven evaluation" (vgl. STAKE 1976, GALLAGHER 1976), die sehr stark informell vorgeht. Dabei wird dieser Ansatz aber weniger als Alternative denn als notwendige, wünschenswerte Ergänzung eingeschätzt (vgl. RAKEL 1976, S. 35; LEWY 1977, S. 146).

Ein weiter, neuer Evaluationsansatz, der mit formellen und informellen Verfahren arbeitet, ist die "Adversary Evaluation", d.h. ein Vorgehen, in dem innerhalb der Evaluation gegensätzliche Meinungen, Blickpunkte, pro- und contra Stellungen erzeugt werden (vgl. WORTHEN/OWENS 1978).

<sup>26</sup> Neben einem Fragebogen sollte in der vorliegenden Arbeit auch ein informelles Instrument erprobt werden. Dabei wurde sich an das Konzept der "Peak Learning Experiences (PLE)" angeschlossen und nach den 3 interessantesten Aspekten der Veranstaltung gefragt (vgl. BLUM 1975; CHOPPIN/FRANKEL 1976). PLEs sind eine relativ kleine Anzahl von Schlüsselerfahrungen, die sich bei den Teilnehmern einprägen und deren Lernen entscheidend beeinflussen bzw. organisieren. Werden nun von den Teilnehmern übereinstimmend die gleichen "interessanten" Aspekte der Veranstaltung benannt, so können diese als Schlüsselerlebnisse betrachtet und für Verbesserungen benutzt werden. Die Frage nach den 3 interessantesten Aspekten stand jedoch zu Ende des Fragebogens und wurde von den Teilnehmern kaum beachtet, so daß keine Erfahrungen mit diesem informellen Instrument vorliegen.

## **2.7 Zusammenfassung: Leitlinien der Evaluation von Erzieherfortbildung**

Kapitel 2 versuchte ein Modell zur Evaluation von Fortbildungsveranstaltungen für Erzieher zu entwickeln und zu begründen. Die Grundgedanken dieses Evaluationsmodells lassen sich in den folgenden 12 Leitlinien festhalten:

1. Evaluation wird orientiert am empirisch-analytischen Paradigma; sie dient der Vorbereitung von Entscheidungen über pädagogische Programme.
2. Die Funktion der Evaluation wird in der Verbesserung von Fortbildungsveranstaltungen gesehen, um dadurch indirekt Erziehungspraxis zu verbessern.
3. Die Adressaten der Evaluation sind damit die direkt an Fortbildung Beteiligten, d.h. Evaluation muß für teilnehmende Erzieher, Referenten und Fortbildungsleiter nützliche Informationen zur Verbesserung der Fortbildungsarbeit liefern.
4. Eine Evaluation braucht Maßstäbe für eine "gelungene" Fortbildungsveranstaltung, mit denen die gerade untersuchte Veranstaltung verglichen werden kann. Aus Gründen der Objektivität des Vorgehens und der Erfassung auch unerwarteter Nebeneffekte übernimmt der "ziel- und inhaltsfreie" Evaluationsansatz nicht die Erfolgskriterien der Veranstalter oder Teilnehmer, sondern entwickelt unabhängig von ihnen - verpflichtet dem übergeordneten Ziel der Verbesserung der Erziehungspraxis - einen ausführlichen Satz potentieller Erfolgskriterien.
5. Es lassen sich zwei Gruppen von Erfolgskriterien unterscheiden: einmal direkte Produkte, Lernergebnisse des Besuchs einer Fortbildungsveranstaltung und zum anderen Merkmale der Gestaltung von Fortbildungsprozessen. Letztere werden dann zu Erfolgsmaßstäben, wenn sie genügend hoch mit Produktkriterien und damit letztlich

mit der "Verbesserung der Erziehungspraxis" korrelieren. Diese Indikatorfunktion von Prozeßkriterien beruht angesichts mangelnder Forschungsstudien oft nur auf Plausibilitätsannahmen.

6. Die Prozeßkriterien können heuristisch gegliedert werden in die 4 Bereiche Thematik, Methodik, Interaktionsprozesse und situative/organisatorische Bedingungen; zusammen mit den Produktkriterien lassen sie sich in einem Raster anordnen.
7. Zur Ausdifferenzierung dieses Rasters ist eine umfassende Literaturanalyse notwendig; diese beschreibt und begründet konkrete Erfolgskriterien und sucht zugleich nach plausiblen Begründungen für den Zusammenhang von Prozeßmerkmalen mit erwünschten Ergebnissen von Fortbildung.
8. Damit unterschiedliche Adressaten diese Erfolgskriterien als mögliche Maßstäbe anerkennen können, operationalisiert die Literaturanalyse nicht nur einen theoretischen Ansatz, sondern orientiert sich vielmehr an unterschiedlichen theoretischen Konzepten.
9. Die Literaturanalyse ist nur der erste und unbedingt notwendige Schritt in einem mehrphasigen Vorgehen. Der ziel- und inhaltsfreien Evaluation ist immanent, daß zuerst die Literaturanalyse Erfolgskriterien sammelt und danach in einer zweiten Phase überprüft wird, ob sich unter der Sammlung auch jene Maßstäbe befinden, die den Evaluationsadressaten relevant erscheinen. Anschließend wird der Kriteriensatz entsprechend modifiziert.
10. Das auf die Erfolgskriterien bezogene, fertig entwickelte Evaluationsinstrument wird als Vorbereitung eines Kommunikationsprozesses gemeinsamer Auswertungsgespräche eingesetzt. Teilnehmende Erzieher sind gleichberechtigte Partner der Veranstalter und Subjekte

ihrer Lernprozesse. Die Evaluation kann dies unterstützen, in dem die Teilnehmer selbst ihre Lernerfahrungen reflektieren und auswerten. Gerade diese Selbstevaluation durch die Beteiligten - eingebettet in gemeinsame Auswertungsgespräche - bedarf als Hilfsmittel eines ausführlich erprobten, in seiner Qualität überprüften Evaluationsinstrumentes. (Der ideale Auswertungskontext kann im Entwicklungs- und Erprobungsstadium des Instruments nicht erreicht werden.)

11. Die Einbettung des Evaluationsinstrumentes in einen gemeinsamen Kommunikationsprozeß der Beteiligten stellt unumgängliche Ansprüche an dieses Instrument: Die Teilnehmer selbst sollen mit ihm ihre Erfahrungen und Wahrnehmungen der verschiedenen Aspekte einer Fortbildungsveranstaltung reflektieren, sammeln und einem gemeinsamen Auswertungsgespräch zuleiten können. Das Instrument muß einen kontrollierten Rückmeldeprozeß erlauben, in dem die Informationen bezüglich der Erfolgskriterien systematisch und objektiv erhoben werden. Praxisrelevanz wird erst bei niedrigem Aufwand an Kosten, Zeit und Energie erreicht. Vor allem muß das Instrument so schnell anwendbar und auswertbar sein, daß die evaluativen Informationen auch rechtzeitig zu Beginn der Auswertungsgespräche vorliegen.
12. Das zu entwickelnde Evaluationsinstrument stellt nicht den Anspruch auf Totalerfassung aller Fortbildungsaspekte; es ist vielmehr als routinemäßig einsetzbares Auswertungsinstrument zu verstehen, welches ggf. durch weitere, auch informelle Verfahren (Gespräche, Protokolle) ergänzt werden muß.

Diese 12 Leitlinien geben die Grundlage für das in den folgenden Kapiteln beschriebene Vorgehen ab. Jede Evaluation selbst muß evaluiert werden (vgl. LEWY 1973, S. 323;

SCRIVEN 1972b, S. 84ff.; STUFFLEBEAM 1974, S. 1). Nach STUFFLEBEAM fragt diese Meta-Evaluation, ob in der konkreten Durchführung der Evaluationsarbeit ein bestimmter Standard, ein bestimmtes Set von allgemeinen Gütekriterien erreicht wurde. Da diese Gütekriterien in den oben beschriebenen Leitlinien schon enthalten sind, seien sie hier nur abschließend benannt. Evaluation, hier speziell auf das zu entwickelnde Evaluationsinstrument bezogen, soll den folgenden Maßstäben entsprechen (vgl. STUFFLEBEAM 1974, S. 6-11; ähnlich ALKIN 1975, S. 206-210)<sup>27</sup>:

- technische Angemessenheit

Das Evaluationsinstrument soll die Auswertungsinformationen objektiv, reliabel und valide erheben.

- Nützlichkeit

Das Evaluationsinstrument soll Informationen erheben, die für die teilnehmenden Erzieher, die Referenten und Fortbildungsleiter relevant und wichtig sind. Es soll möglichst umfassend über die Veranstaltung informieren; die Auswertungsinformationen sollen alle angestrebten Adressaten rechtzeitig für ein gemeinsames Auswertungsgespräch erreichen. Die Ergebnisse müssen für die Adressaten verständlich und in dem Sinne glaubwürdig sein, daß sie das Instrument als eine mögliche Erfahrungsquelle anerkennen und nicht apriori als unvalide und unzutreffend ablehnen.

- Kosten - Nutzen

Der Auswertungsaufwand und die Kosten sollen dem speziellen

---

<sup>27</sup> Einen anderen, aber prinzipiell ähnlichen und vergleichbaren Satz von Gütekriterien findet man bei ALEAMONI/SPENCER für die Einschätzung von Lehrern (vgl. ALEAMONI/SPENCER 1973, S. 671f.).

Die besondere Betonung des Kriteriums der Nützlichkeit, Utilität findet sich schon bei SPRUNG/SPRUNG im Rahmen einer Theorie der allgemeinen Fragebogenkonstruktion. Utilität ergänzt die klassischen Kriterien der Objektivität, Reliabilität und Validität, sie sollte vor diesen Kriterien erhoben werden (vgl. SPRUNG/SPRUNG 1978a, S. 19f.).

Nutzen für die Adressaten angemessen sein, bzw. im Ideal soll das Instrument mehr wert sein als es kostet.

In den folgenden Kapiteln sollen das Vorgehen der Evaluation und die Ergebnisse der Erprobung beschrieben werden, damit der Leser überprüfen kann, ob das entwickelte Evaluationsinstrument diese Gütekriterien erreicht.

### 3. Erfolgskriterien von Fortbildungsveranstaltungen

Anknüpfend an das mehrphasige Modell des ziel- und inhaltsfreien Evaluationsansatzes (vgl. S. 35) besteht die Aufgabe des Evaluators in der ersten Phase darin, unabhängig von den Adressaten einen Satz von Erfolgskriterien für Fortbildungsveranstaltungen mit Hilfe einer Literaturanalyse zu sammeln und zu begründen. Eine spezielle Literatur für Fortbildung für Erzieher im Vorschulbereich fehlt aber weitgehend, so daß ebenso Arbeiten aus den ähnlich gelagerten Bereichen Lehrerfortbildung, Hochschuldidaktik und Erwachsenenbildung einbezogen werden müssen<sup>1</sup>. Dies führt zwangsläufig zu relativ bereichs- und berufunspezifischen Erfolgskriterien, an denen fast jede beliebige Fortbildungsveranstaltung gemessen werden kann. Eine Zuspitzung auf berufsspezifische Fortbildungsveranstaltungen für Erzieher im Vorschulbereich läßt sich nur erreichen durch eine Betrachtung des Tätigkeitsfeldes der Erzieher.

#### 3.1 Reziproke Kommunikation als regulatives Prinzip für Fortbildung

Der Kern der beruflichen Tätigkeit von Erziehern besteht in dem Umgang mit jüngeren Kindern, in einer pädagogischen Interaktion, die durch eine starke Asymmetrie der Fähigkeiten und Möglichkeiten der Interaktionspartner gekennzeichnet ist. Damit sind zwei Fragen aufgeworfen:

---

<sup>1</sup> Eine gewisse Vergleichbarkeit der unterschiedlichen Bereiche ist dadurch gegeben, daß sie sich auf die Altersgruppe der Erwachsenen beziehen, sowie oft in einer ähnlichen Lehrgangs- und Seminarform ablaufen. Auf die Nähe zur Allgemeinen Erwachsenenbildung weisen besonders PRIESNER 1975, S. 51 und RIEMANN 1975, S. 315 hin.



- Was ist das Charakteristikum dieser Interaktion und damit das Spezifische der Vorschulerziehung insgesamt?
- Welche beruflichen Qualifikationen braucht ein in diesem Feld handelnder Erzieher?

Eingehend mit diesen Fragen beschäftigt sich URSULA PEUKERT in dem Versuch, eine handlungstheoretische Begründung von Vorschulerziehung zu leisten (vgl. PEUKERT 1979)<sup>2</sup>:

PEUKERT untersucht die Entwicklung der kognitiven, sprachlichen und interaktiven Kompetenzen, die in dem Konzept der kommunikativen Kompetenz integriert werden (vgl. ib., S. 48). Gegen Ende des Vorschulalters - in einem Übergangsfeld vom 5. bis 7. Lebensjahr - läßt sich aufgrund von entwicklungspsychologischen Forschungsergebnissen eine Zäsur in der Entwicklung des Kindes ansetzen. Nach den Theorien der kognitiven Entwicklung ist der Übergang von der präoperationalen zur Phase der konkreten Operationen zu leisten; nach Untersuchungen der interaktiven Dimension muß ein Kind eine erste Reflexivität des Rollenhandelns und damit die Grundqualifikationen Empathie, Rollendistanz und Ambiguitätstoleranz erwerben; nach psychoanalytischer Theorie ist der ödipale Konflikt auszutragen und die Alters- und Geschlechtsrolle zu übernehmen; nach Theorien der Entwicklung des moralischen Bewußtseins wird hier der Übergang von der präkonventionellen zur konventionellen Orientierung (Egozentrismus zu Soziozentrismus) gesehen; diese Leistungen gelingen nur bei einer genügend differenzierten sprachlichen Kompetenz des Kindes. All diese Entwicklungslinien konvergieren darin, /

---

<sup>2</sup>Die umfangreiche Arbeit von PEUKERT kann hier nur sehr verkürzt in einigen grundlegenden Zügen berichtet werden.

"daß das Kind im kommunikativen Handeln und darin eingebettet auch im Umgang mit sachhafter Wirklichkeit eine erste Form reflexiver Reziprozität gewinnt. (im Original gesperrt, H.G.R.)" (ib., S. 188)

Die einzelnen Lernprozesse sollten auf diese Grundqualifikation hin organisiert werden. Als ursächlich ausschlaggebender Faktor für die Entwicklung des Kindes läßt sich aufgrund der entwicklungspsychologischen Forschungen das soziale Lernen, die Entwicklung der interaktiven Kompetenz identifizieren, deren Förderung im Vorschulalter dann ein Vorrang zukommt (ib., S. 170).

Im Zentrum der Vorschulerziehung und ihrer Theorie steht somit die Interaktion zwischen Erzieher und Kind und die Frage nach der normativen Grundstruktur dieser Interaktion (vgl. ib., S. 187).

Im Rückgriff auf Untersuchungen über die Wirkung schichtspezifischer Interaktionsstrukturen und vor allem auf Ergebnisse familientherapeutischer Forschung, in der es um die Gesundheit oder Krankheit von Beziehungsstrukturen und deren Auswirkung auf die Ausbildung elementarer Kompetenzen und psychischer Instanzen im Sozialisationsprozeß geht, leitet PEUKERT für die pädagogische Interaktion im Vorschulbereich Konsequenzen ab:

Der Aufbau einer stabilen und bei krisenhaften Übergängen auf ein höheres Entwicklungsniveau belastbaren Ich-Identität sowie des Erwerbs einer flexiblen kommunikativen Kompetenz hängt davon ab, inwieweit der Erzieher eine intergenerationale Reziprozität herstellen kann; Eltern und Erzieher müssen die Kinder als prinzipiell ebenbürtig und gleichberechtigt anerkennen und behandeln, obwohl deren Handlungsmöglichkeiten noch nicht den gleichen Rang haben.

"Die normative Grundstruktur bestünde dann auch in der Beziehung zwischen den Generationen darin, den anderen und gerade auch den Schwächeren, etwa das heranwachsende Kind, als Partner eigenen Rechts von vorneherein anzuerkennen und dadurch die Entstehung pathogener, ausbeuterischer Interaktionsstrukturen zu verhindern, ..." (ib., S. 185f.)

Dieses normative Grundprinzip tritt im Vorschulbereich wegen der krassen Asymmetrie zwischen Erzieher und Kind besonders scharf hervor. Es bedarf einer besonderen Leistung des Erziehers, eines stellvertretend-antizipatorischen Handelns, welches das Kind gerade in seiner erreichten Stufe des Kompetenzerwerbes und der Identitätsausbildung unbedingt ernst nimmt und welches trotzdem handelnd die Überwindung der krisenhaften Übergänge zu einer neuen Form der Identität und Kompetenz ermöglicht. Der Erzieher muß auf die voll ausgebildeten Interaktionsweisen des Kindes vorausgreifen und mit ihm so umgehen, als ob es schon voll ausgebildete Intentionen hätte. Dieses "Fingieren" des autonom handelnden Subjektes bedeutet das Eröffnen eines Handlungsspielraumes reziproker Interaktion, in die das Kind handelnd hineinwachsen kann (vgl. ib., S. 196)<sup>3</sup>.

Die Bestimmung der intergenerationellen Reziprozität als Grundnorm für pädagogische Interaktion verlangt bestimmte berufliche Qualifikationen vom Erzieher. Er muß sensibel sein für die bisherige Biographie und Lerngeschichte des Kindes sowie die erreichte Stufe von Identität und Kompetenz; er muß durch Initiierung von Interaktionssequenzen den Übergang zu einer höheren Stufe in der Genese des Subjektes ermöglichen. Dazu bedarf es bestimmter Berufsqualifikationen:

---

<sup>3</sup> Gerade dieses stellvertretende vorgreifende Handeln kommt dem Lernen des Kindes entgegen. Das Kind lernt dialektisch, d.h. Erfahrungen, die mit dem bisher erlernten Regelsystem nicht verarbeitet werden können, führen zu einer Veränderung, Umstrukturierung des Regelsystems auf einer höheren Ebene. Dies ist ein extrem aktiver, selbständiger Prozeß des Entschlüsselns und Nachkonstruierens von generativen Regeln und ermöglicht den Erwerb einer höheren kommunikativen Kompetenz. Wegen des aktiven und autonomen Charakters ist dies zugleich mit dem Aufbau einer autonomen Ich-Instanz verbunden. (vgl. PEUKERT 1979, S. 191ff.).

- einer sozialwissenschaftlichen Kompetenz
- einer sozialen, pädagogisch orientierten Handlungsfähigkeit, die für den Erzieher ebenfalls eine bestimmte Kompetenz und Identität verlangt. Die eigenen biographischen Erfahrungen in der Ursprungsfamilie müssen so verarbeitet werden, daß der Erzieher keine extreme Verzerrung in seiner Wahrnehmung des Kindes erleidet, speziell daß er zwischen eigenen und Bedürfnissen des Kindes trennt und das Kind nicht zur Befriedigung eigener Bedürfnisse benutzt (vgl. ib., S. 201f.).

Fortbildungsveranstaltungen - und damit werden die Überlegungen zu den Erfolgskriterien wieder aufgenommen - müssen sich nun an dem Beitrag messen lassen, den sie zu der so bestimmten beruflichen Qualifikation des Erziehers leisten. Damit ist nun nicht gemeint, daß nur Fortbildungsveranstaltungen zur direkten Herstellung dieser Qualifikation angeboten und zu anderen Themen nicht mehr durchgeführt werden sollten, sondern die Überlegungen zur Struktur pädagogischer Interaktion im Vorschulbereich lassen sich als ein regulatives Prinzip für Fortbildung begreifen:

Fortbildung im Vorschulbereich darf - bei jedem Thema - nicht so organisiert werden, daß dem Erzieher bewußt oder unbewußt Verhaltensweisen vermittelt werden, die die Herstellung der intergenerationellen Reziprozität und der antizipatorisch-stellvertretenden Handlungen des Erziehers verhindern. Im "Arbeitsprozeß" Fortbildung müssen die Erzieher eine reziproke Kommunikation erfahren, in der sie gleichberechtigte Partner der Veranstalter und Subjekte selbstbestimmter Lernprozesse sind. Die maßgeblichen Berufsqualifikationen für Erzieher verlangen dieses gleichberechtigte Klima in der Fortbildung, welches als eine notwendige - aber nicht hinreichende - Bedingung, als Vorstufe für intergenerationelle Reziprozität betrachtet werden kann. Wenn Erzieher nicht in diesem Sinn als gleichberechtigte Partner erscheinen, so wirkt sich ein

"hidden curriculum" aus, welches die angestrebte pädagogische Handlungsfähigkeit verhindert<sup>4</sup>.

Für das regulative Prinzip "reziproke Kommunikation" in der Fortbildung sind folgende Aspekte festzuhalten<sup>5</sup>:

---

<sup>4</sup> Dies - und die Spezifität in der Vorschulerziehung - sei an einem konstruierten Beispiel erläutert:

Eine Fortbildungsveranstaltung für Programmierer könnte streng genommen allein an der Effektivität des Wissens über den Umgang mit technischen Materialien orientiert sein; wird gegen das Prinzip "gleichberechtigte Partner" verstoßen, so kann dies für den Teilnehmer zwar seine allgemeine Einstellung zum Lernen/seine Persönlichkeit verändern, ein effektiver Umgang mit der neuen Programmiertechnik als angestrebtem Lernertrag der Fortbildung braucht damit aber nicht ausgeschlossen sein. Anders bei Veranstaltungen mit Erziehern im Vorschulbereich. Fortbildung kann hier nicht allein orientiert werden an dem Wissenszuwachs bezüglich Verhaltensstörungen, sondern umfaßt auch eine bestimmte Befähigung zum Handeln mit Vorschulkindern. Erfahren Erzieher, daß sie während der Veranstaltung nur Objekte eines manipulatorischen Unterrichtes sind und nicht gleichberechtigte, ernst genommene Partner, so besteht die Gefahr, daß sie diese Erfahrung auf den Umgang mit Vorschulkindern ausdehnen und diese ebenfalls in ihren Interessen und Bedürfnissen nicht ernst nehmen. Die intergenerationelle Reziprozität ist so verhindert. Reziprozität bei der Fortbildungsveranstaltung für Programmierer wird zwar auch oft gefordert, eine Nichterfüllung verhindert aber nicht das effektive Anwenden einer neuen Technik, während im Vorschulbereich sich gerade die Nichterfüllung verhängnisvoll auswirken würde.

<sup>5</sup> Eine ähnliche Argumentation verfolgt HEYER in der Entwicklung einer integrativen, didaktisch-methodischen Konzeption für Fortbildung (vgl. HEYER 1977). Sie konzentriert sich auf Ansätze, die für den Erwerb der interaktiven und kommunikativen Kompetenzen relevant erscheinen. Auf dem Hintergrund des partner- und erfahrungszentrierten Modells von ROGERS, des pragmatischen Kommunikationsmodells von WATZLAWIK, BEAVIN und JACKSON sowie des themenzentrierten Interaktionssystems von COHN entwickelt sie Postulate für ein Modell der Fortbildung.

In unserem Zusammenhang ist von entscheidender Bedeutung, daß grundlegende kommunikative und interaktive Verhaltensweisen nicht nur durch spezielle, diesem Thema gewidmeten Veranstaltungen vermittelt werden können, sondern eben-

- Reziproke Kommunikation gilt als Prinzip unabhängig von dem spezifischen Inhalt und Ziel jeder Fortbildungsveranstaltung im Vorschulbereich.

- Reziprozität in der Fortbildung ist eine notwendige, aber nicht hinreichende Bedingung für intergenerationelle Reziprozität in der Erzieherpraxis. D.h. intergenerationelle Reziprozität wird nicht allein durch ein gleichberechtigtes Klima in der Fortbildung hergestellt, sondern nur in einer Vorstufe angebahnt. Verletzungen der Reziprozität in einer Fortbildungsveranstaltung für Erzieher dagegen verhindern die Ausbildung der spezifischen intergenerationellen Reziprozität.

---

noch 5 falls durch die Struktur und das Bedingungsgefüge des Lernangebotes (vgl. ib., S. 4). Soziale Kompetenz kann durch die Erfahrung von Reziprozität in der Fortbildung gefördert und Grundqualifikationen können vor allem durch ein modellhaftes Leiter-/Referentenverhalten erlernt werden (vgl. ib., S. 172). Solche Selbsterfahrung soll sich in der Kindergruppe fortentwickeln.

"Die Konzeption bezieht sich auf die Integration und Wechselwirkung didaktisch-methodischer Faktoren, die soziale Kompetenz fördern, ohne daß dies notwendig thematischer Lerngegenstand des Fortbildungskurses ist." (ib., S. 5)

"Das didaktische Basisverhalten kann prinzipiell bei jeder Thematik und Aufgabe zur Erweiterung der Sozialkompetenz beitragen, wenn es annähernd vom "Modell" verwirklicht wird." (ib., S. 181)

Problematisch ist - und da unterscheidet sich der hier vorgelegte Ansatz von HEYER -, daß nicht der Unterschied zwischen der Struktur der Interaktion in der Fortbildung und der Interaktion in der konkreten Arbeit mit dem Vorschulkind als "abhängigen Unerwachsenen" thematisiert wird. Fortbildung kann hier deshalb kein direktes Modell für die Erzieher-Kind-Interaktion sein. Zum anderen erprobt HEYER ihr Fortbildungsmodell an einem Kurs "Einführung in partnerzentriertes Kommunikations- und Erziehungsverhalten". Ihrem Anspruch nach hätte sie aber eine Thematik aussuchen müssen, die nicht direkt das gleiche Thema inhaltlich vermittelt wie die postulierten Wirkungen des integrativen Konzeptes. HEYER sieht dies zwar, betont aber dennoch die Lernwirksamkeit des didaktisch-methodischen Konzeptes, das hier durch die Übereinstimmung mit der thematischen Arbeit nur eine Steigerung erführe (vgl. ib., S. 330). Es wird aber nicht die These der prinzipiellen themenunabhängigen Wirkung des postulierten Konzeptes durch die praktische Erprobung des Modells unterstützt.

- Fortbildung kann kein direktes Muster, Modell für den Umgang mit Kindern sein, weil kommunikatives Handeln unter - prinzipiell - voll interaktionsfähigen Partnern sich von der Interaktion mit dem Vorschulkind unterscheidet. Kinder müssen erst noch die volle Handlungsfähigkeit erhalten, vorher ist die Interaktion eine unter Ungleichem.

- Reziproke Kommunikation ist zwar für jede Fortbildungsveranstaltung ein mögliches Postulat, für Fortbildung für pädagogische Fachkräfte bekommt sie aber einen besonderen Stellenwert. Vor allem für Veranstaltungen mit Erziehern im Vorschulbereich bekommt sie wegen der besonders stark ausgeprägten Asymmetrie der Interaktionspartner eine unumgängliche Qualität.

Bei der Literaturanalyse wird die Suche nach Erfolgskriterien somit gesteuert und reguliert durch das Prinzip der "reziproken Kommunikation". Erfolgskriterien, z.B. aus der Lernpsychologie von Erwachsenen, verlieren damit nicht ihre Bedeutung, werden aber jeweils daraufhin überprüft, inwieweit sie die gewünschte Form der Reziprozität ermöglichen oder behindern. Das regulative Prinzip leistet so einmal eine Gewichtung der Auswahl von Erfolgskriterien und wird zum anderen in der Auseinandersetzung mit den Erfolgskriterien konkretisiert.

Es wird sich in der Literaturanalyse zeigen, daß die Überlegungen zur gleichberechtigten Kommunikation von Teilnehmern und Referenten/Leitern von Fortbildung in der Diskussion um "Teilnehmerorientierung" wiederzufinden sind. "Teilnehmerorientierung" - die noch konkretisiert werden muß - wird hier im Vorschulbereich durch die normative Struktur des beruflichen Handelns von Erziehern ein zwingendes Gestaltungsmoment von Fortbildung und zwar in allen vier Bereichen der Fortbildungsprozesse (Thematik, Methodik, Interaktionsprozesse, situative und organisatorische Bedingungen).

### 3.2 Literaturanalyse von Kriterien für erfolgreiche Fortbildungsveranstaltungen

Die Literaturanalyse wird strukturiert durch den in Kapitel 2.3 (s. S. 32 .) entwickelten Raster von Prozeß- und Produktkriterien; sie will beide Bereiche ausdifferenzieren und konkrete Erfolgskriterien beschreiben und begründen<sup>6</sup>. Es besteht kein Anspruch auf Vollständigkeit oder Repräsentativität der einbezogenen Literatur, gleichwohl wurde versucht, die gesamte Spannbreite von Erfolgskriterien zu erfassen. Regulatives Prinzip bei der Auswahl und Gewichtung der Kriterien ist die oben beschriebene "reziproke Kommunikation".

#### 3.2.1 Erfolgskriterien im Bereich Thematik

In der analysierten Literatur wird durchgehend Fortbildung abgelehnt, die davon ausgeht, daß der wesentliche Mangel der Erzieherarbeit in einer unzureichenden Aufnahme von wissenschaftlichen Erkenntnissen in der Praxis besteht. Fortbildung müßte dann - so wird von den Autoren unterstellt - eine einseitige Weitergabe von Informationen aus der Wissenschaft an die Praktiker sein (vgl. z.B. AG FORTBILDUNG DES KINDERZENTRUMS ROENEBERGSTRASSE 1976, S. 9f.; SCHRADER 1973, S. 293f.). In einer solchen Fortbildung läge die Rollenverteilung vorher fest:

---

<sup>6</sup>In der Literaturanalyse zeigt sich, daß viele Aussagen zuerst nur für die didaktisch-methodischen Modellvorstellungen gelten, für die sie explizit formuliert wurden. Es werden im folgenden Teil nur jene Aussagen bezüglich Erfolgskriterien genannt, von denen plausibel angenommen werden kann, daß sie nicht nur modellspezifischen Charakter haben, sondern einen für verschiedene Veranstaltungen generalisierbaren Charakter haben.



der Referent ist Experte und aktiver Theorievermittler, der Erzieher nur ein passiver Rezipient<sup>7</sup>. Da die Umsetzung der Theorie in die Praxis den Erziehern alleine überlassen würde, könne Fortbildung auch keinen Beitrag zur Verbesserung der Praxis liefern. Ebenso stellen Befragungen den Wunsch von Erziehern fest, weniger mit theoretischen Erörterungen und Vermittlung allgemeiner Kenntnisse und Einsichten konfrontiert zu werden, sondern sowohl in der Methode als auch bei der Festlegung des Themas anders vorzugehen (vgl. FRÜTEL u.a. 1975, S. 231f.).

Ein solches Fortbildungsmodell - deren Verbreitung als alleinige Form von Fortbildung aber bezweifelt wird - widerspricht dem oben genannten Kriterium der reziproken Kommunikation zwischen allen Beteiligten. Demgegenüber wird heute durchweg Fortbildung unter dem Postulat der "Teilnehmerorientierung" in Bezug auf die Festlegung von Zielen und Inhalten diskutiert<sup>8</sup>:

---

<sup>7</sup> Eine ähnliche Kritik findet sich in allen Bereichen. Für die Mitarbeiter in der Jugendarbeit z.B. kritisiert ZILLEN eine Fortbildung nach dem Modell des "Lehrganges". Dahinter stünde die Auffassung, daß die Teilnehmer "weiße Flecken auf der Landkarte" hätten, die durch systematische Unterweisung zu beseitigen wären (vgl. ZILLEN 1977, S. 65f.); für die Lehrerfortbildung siehe z.B. BRUNNER 1976, S. 24; HENGARTEN/WEINREBE 1975, S. 33; für akademisches Lernen lehnen z.B. KEIL/PIONTKOWSKI eine Studentenrolle des passiven Rezipienten ab (vgl. KEIL/PIONTKOWSKI 1973, S. 15f.; oder für Vorlesungen SCHOTT 1973, S. 22).

<sup>8</sup> Es sei hier eine Unterscheidung von TIETGENS aufgenommen. Danach bezieht sich Adressatenorientierung auf die allgemeine Programmplanung und Teilnehmerorientierung auf die Planung und Durchführung eines Kurses (vgl. TIETGENS 1977, S. 283). Unterschieden werden muß Teilnehmerorientierung hier auch vom Konzept der Zielgruppenentwicklung, in der unter Adressatenbezug die Entwicklung und Herstellung von Zielgruppen bei der Erwachsenenbildung gemeint wird. Vgl. hierzu MADER/WEYMANN 1977; MADER 1979; SCHIEMANN 1978.

Fortbildung soll auf den Erfahrungen der Teilnehmer aufbauen und sich an deren Situationen, Bedürfnissen, Interessen und Gewohnheiten orientieren; erst in zweiter Linie sollen die Vorstellungen der Veranstalter ein Programm bestimmen (vgl. NESTMANN 1977, S. 300; FRÜTEL u.a. 1975, S. 232). Die Themenauswahl soll sich an den Bedürfnissen und Erwartungen der Teilnehmer orientieren (vgl. TEIGELER 1976, S. 4; PRIESNER 1975, S. 52). Fortbildung muß an den Problemen der Praxis ansetzen. Diese können den Erziehern nicht als ihre Probleme vorgesetzt werden, sondern Praxisprobleme müssen von den Erziehern selbst formuliert werden (vgl. AG FORTBILDUNG DES KINDERZENTRUMS ROENEBERG-STRASSE 1976, S. 11).

Die Orientierung an den Fortbildungswünschen der Erzieher soll gekoppelt werden mit dem Entwicklungsstand der Kleinkindforschung, der Praxisforschung, den Anforderungen der Praxis, den Ergebnissen der Wissenschaften (vgl. KÖNIG 1977, S. 123; MAELICKE 1977, S. 238; ERHARDT-PLASCHKE/IRSKENS 1977, S. 290; HEYER 1977, S. 87f.) oder dem objektiven gesellschaftlichen Bedarf an Erzieherqualifikationen, da dieser sich von den subjektiven, von Erziehern geäußerten Interessen unterscheiden kann (vgl. HEYER 1974, S. 68)<sup>10</sup>.

- <sup>9</sup> Nestmann formuliert 10 Prinzipien der Fortbildung in der sozialen Arbeit. Diese Prinzipien stammen aus der Fortbildungsarbeit des Instituts für Sozialarbeit und Sozialpädagogik in Frankfurt und beziehen sich auf Planungs- und Beratungsaufträge, bzw. eine in Forschungs- und experimentellen Praxisprojekten integrierte Fortbildung. NESTMANN weist aber daraufhin, daß die Prinzipien auch Richtlinien sein sollen für Fortbildung, die nicht dem integrierten Ansatz entspricht.
- <sup>10</sup> Die Forderung der Teilnehmerorientierung findet sich in der Erwachsenenbildung an vielen Stellen: für sozialkulturelle Erwachsenenbildung (Kennzeichen: Interpretierbarkeit und geringe Standardisierung des Lerngegenstandes) z.B. in der Hannover-Studie (vgl. SIEBERT 1975, S. 34). Teilnehmerorientierung ist das Charakteristikum einer Erwachsenenbildung, die den Bürger als Lerner mit seinen Motiven und Erwartungen unmittelbar ernst nimmt (vgl. WEINBERG 1978, S. 146f.); die Empfehlung zur Weiterbildung der UNESCO betont ebenfalls die Orientierung an den Bedürfnissen der Teilnehmer (vgl. UNESCO 1977, S. 6). Als Kriterium zur Auswertung des Bildungsurlaubsprojektes formulieren die Autoren das folgende didaktische Prinzip: "Wieweit ist gewährleistet, daß Auswahl und Struktur der Inhalte sich an den Teilnehmerinteressen, den Teilnehmerbedürfnissen und den Teilnehmer Voraussetzungen orientieren" (HOLZAPFEL u.a. 1977, S. 10).

Kenntnisse über Wünsche, Bedürfnisse, Interessen und Motive der Erzieher liegen aber kaum als Planungsdaten vor, so daß die Veranstalter vor einer Veranstaltung die Bedürfnisse der Erzieher antizipieren müßten. Dabei ergeben sich aber zwei Schwierigkeiten:

- Erzieherwirklichkeit ist so komplex, daß ohne die Erzieher selbst deren Bedürfnisse, Wünsche und Interessen gar nicht festgestellt werden, bzw. nur begrenzt antizipiert werden können (vgl. ACKERMEIER u.a. 1978, S. 179f.).

- Geht man von der Forderung nach Selbstbestimmung der Lernenden aus, so kann diese nicht nur in einer formalen Vergewisserung bestehen, ob denn nun die Veranstaltungsziele akzeptiert werden, sondern es bedarf einer inhaltlichen Diskussion mit den beteiligten Erziehern zu Beginn einer Veranstaltung über Inhalte und Ziele, d.h. einer Partizipation bei der Lernzielbestimmung (vgl. SIEBERT 1974, S. 87f.; TIETGENS/WEINBERG 1971, S. 200; BAUMANN u.a. 1976, S. 18).

Sowohl aus Gründen der Effektivität als auch der Legitimation der Auswahl von Zielen und Inhalten einer Veranstaltung verlangt hier die Orientierung an den Teilnehmern eine Partizipation bei der Bestimmung von Zielen und Inhalten. Leiter/Referenten und Teilnehmer müssen in einem gleichberechtigten Prozeß über sie entscheiden.

In der Erwachsenenbildung allgemein hält SIEBERT die Forderung nach Lernzielpartizipation für unverzichtbar, er weist aber daraufhin, daß sie nicht rigoristisch und schematisch vertreten werden sollte (vgl. SIEBERT 1974, S. 186)<sup>11</sup>. Lernzielpartizipation kostet Zeit und Energie,

<sup>11</sup> SIEBERT diskutiert hier die Relevanz der Curriculumforschung für die Erwachsenenbildung. Das Angebot und die Problemlage in der Erwachsenenbildung ist aber so differenziert, daß SIEBERT sich dagegenwehrt, nur eine Curriculumstrategie - z.B. offene Curricula, Teilnehmerorientierung - als akzeptabel zu betrachten. An mehreren Stellen betont er die Gleichberechtigung von offenen Curricula, Rahmencurricula und eher geschlossenen Curricula (vgl. SIEBERT 1974, S. 107-109, S. 152; auch WEINBERG 1978, S. 147).

sie kann evtl. Lernen blockieren, indem nämlich den Teilnehmern endlose Zieldebatten als unfruchtbar und unergiebig erscheinen (vgl. eb., S. 116). Andererseits kann für die Erwachsenenbildung auch nicht das Argument für die Schule gelten, daß nämlich ihre Adressaten noch nicht zur curricularen Mitbestimmung fähig seien. Diskursive, gleichberechtigte Verständigung über die Ziele einer Veranstaltung ist zwar als theoretisches Postulat unverzichtbar, aber sie ist oft nicht Voraussetzung, sondern erst Ergebnis von Lernprozessen (vgl. eb., S. 138)<sup>12</sup>.

Auch unter diesen Schwierigkeiten wird Teilnehmerorientierung und Partizipation bei der Lernzielbestimmung ein notwendiger Bestandteil für Fortbildungsveranstaltungen für Erzieher, die ja unter dem Kriterium der reziproken Kommunikation stattfinden sollen. Dies kann aber nicht ein totales, völliges Eingehen auf Erzieherwünsche bedeuten, wohl aber müssen die Erzieherwünsche ernst genommen werden. Alle Beteiligten müssen ihre Interessen gleichberechtigt aushandeln, wobei hier die Veranstalter eigene Interessen, ihre Sicht der Praxisprobleme sowie Ansprüche aus den Wissenschaften einbringen müssen. Dabei kann nicht verschwiegen werden, daß dieser Prozeß sehr störanfällig ist:

- Veranstalter können ihre eigenen Interessen unbewußt im Prozeß des Aushandelns in die Erzieher hineinlegen,
- eine objektive Bestimmung der Praxisprobleme steht noch aus<sup>13</sup>,

---

<sup>12</sup> Auf ein weiteres Problem weisen DILLER und KIEL hin. Wenn nämlich Erzieher Rezeptwünsche äußern würden, so kann Fortbildung nicht ganz darauf eingehen, da die Weitergabe von Rezepten allein nicht die soziale Handlungskompetenz von Erziehern verbessert. Andererseits kann Fortbildung sich aber nicht elitär über solche Wünsche hinwegsetzen (vgl. DILLER/KIEL 1977, S. 111).

<sup>13</sup> Das Rückbeziehen der Interessen der Teilnehmer und Veranstalter auf ihren Ort und ihren Platz im Ensemble der gesellschaftlichen Kräfte - d.h. auf die Praxisgruppen-

- es gibt nicht die wissenschaftliche Theorie, die einen objektiven Anspruch an die Praxis stellt.

Für diese Schwierigkeiten kann aber hier keine Lösung gegeben werden.

Von dem Kriterium der Teilnehmerorientierung ist die Forderung nach Praxisrelevanz des Themas kaum zu trennen:

Fortbildung soll relevant sein für die Förderung direkt auf das Praxisfeld bezogener Qualifikationen; sie soll die beruflichen Planungs- und Entscheidungskompetenzen erhöhen (vgl. NESTMANN 1977, S. 300; DILLER/KIEL 1977, S. 109; EHRHARDT-PLASCHKE/IRSKENS 1977, S. 288; IRSKENS u.a. 1975, S. 254; SCHMALOHR 1974, S. 174f.; ACKERMEIER u.a. 1978, S. 180ff.). Dazu soll Fortbildung an den konkreten Praxisproblemen der Teilnehmer ansetzen und zu deren Bewältigung eine Hilfe geben (vgl. RIEMANN 1975, S. 316; FRÜTEL u.a. 1975, S. 231f.; HEYER 1974, S. 62, 66; MOSKAL 1978, S. 321). Erzieherfortbildung dient primär der Praxisveränderung im Sinne einer qualitativ besseren Erziehungsarbeit; die ausgewählten Themen müssen Relevanz für grundlegende Veränderungen haben (vgl. AG FORTBILDUNG DES KINDERZENTRUMS ROENEBERGSTRASSE 1976, S. 13)<sup>14</sup>.

noch 13 bei THÜRNAU und VOLLRATH kann solange keine Legitimationsfunktion haben, wie nicht feststeht, was nun die Ansprüche der Praxisgruppen sind (vgl. THÜRNAU/VOLLRATH 1978, S. 385).

Objektive Bedarfsanalysen von Qualifikationen für den Vorschulbereich als Legitimationsinstanz fehlen. Der Versuch einer umfassenden Anforderungsanalyse von MINZLAFF/PFEIFF (1976) durch Befragungen von Erziehern, Eltern, Trägern, Mitarbeitern, Beobachtungen der Alltagspraxis konnte leider realisiert werden (persönliche Mitteilung). Zu den Schwierigkeiten solcher objektiven Bedarfsanalysen für die Erwachsenenbildung siehe auch den Versuch bei VONTOBEL (vgl. VONTOBEL 1972, S. 54ff. und bes. Anmerkung 12, S. 152).

<sup>14</sup> Die Forderung nach Praxisrelevanz der Themen wird oft gestellt; einige Beispiele: In der Diskussion curriculum-theoretischer Ansätze nennt SIEBERT praktische Verwendungssituationen vor dem Ausgehen von einer Wissenschaftsanalyse (vgl. SIEBERT 1974, S. 172). Mitarbeiterfortbildung im sozialpädagogischen Feld kann nur dann erfolgreich sein, wenn sie ihren Ausgangspunkt in der Komplexität der Praxisprobleme hat (vgl. KLAWE 1978, S. 65); in der Auswertung des Bildungsurlaubsprojektes sprechen die Autoren von dem unbestrittenen

Relevanz der Themen der Erzieherfortbildung für konkrete Verwendungssituationen in der Praxis ist mehrfach von Bedeutung:

- Fortbildungsveranstaltungen sind kein Selbstzweck. Lernerfolg und Lernmotivation werden gesteigert durch den Bedeutungsaufweis der behandelten Themen für die Praxiszwecke (vgl. z.B. VONTOBEL 1972, S. 39f.; SIEBERT 1975, S. 111f.; BRANDENBURG 1974, S. 44)<sup>15</sup>.

- Konkrete Praxisrelevanz erhalten Lernergebnisse dann, wenn sie auch in die Praxis übertragen werden können. Um einen solchen Transfer zu ermöglichen, bedarf es einer gewissen Strukturgleichheit zwischen der thematischen Lernaufgabe und dem Anwendungsbereich (vgl. LÉON 1977, S. 117; FROMMER 1978, S. 55f.). Das Kategoriale des Themas, das Exemplarische muß herausgearbeitet werden, damit Strukturverwandtheit mit den Verwendungssituationen erkannt werden kann (vgl. TIETGENS/WEINBERG 1971, S. 208-210).

Diese Bestimmung der Praxisrelevanz eines Themas für die Verwendungszwecke der Teilnehmer ist nun schwierig zu

---

noch 14 Axiom des Ausgehens von Lebenssituationen der Teilnehmer (vgl. HOLZAPFEL u.a. 1977, S. 18); in dem Konzept der Fortbildung in zentraler Trägerschaft der Arbeiterwohlfahrt wird ebenfalls der Ansatz an den Praxisproblemen der Mitarbeiter betont (vgl. THURNAU 1978, S. 177).

<sup>15</sup> Mit diesem Begriff eines Lernzweckes wendet sich SIEBERT gegen die ständige Forderung nach alleiniger intrinsischer Motivation in der Erwachsenenbildung. Lernen ist kein Selbstzweck, sondern ein Mittel zum Zweck. Die Forderung nach Reduzierung von sekundären Motiven zugunsten primärer läßt das Lernen zum Selbstzweck werden und isoliert den Lernenden aus seinen realen Lebenszusammenhängen (vgl. SIEBERT 1975, S. 111).

antizipieren. Es besteht zumindest die Gefahr, daß vom Veranstalter Probleme als praxisbezogen formuliert werden, die von den Erziehern gar nicht als solche erkannt werden. Letztlich wird die Praxisrelevanz nur dann erreicht, wenn die Erzieher an der Auswahl, Diskussion und Abstimmung der Inhalte und Ziele beteiligt sind (vgl. ACKERMEIER 1978, S. 182)<sup>16</sup>. Auf die Notwendigkeit, die Praxisrelevanz während einer Veranstaltung ständig zu verdeutlichen, wird weiter unten eingegangen werden (vgl. S. 71f.).

Erzieherfortbildung soll nicht nur streng themenbezogen arbeiten, sondern auch Raum für einen informellen Erfahrungsaustausch geben:

Fortbildung darf nicht auf Kosten des Erfahrungsaustausches zwischen Praktikern gehen. Vor und/oder neben dem eigentlichen Programm soll auch Gelegenheit zu informellen Gesprächen der Erzieher untereinander gegeben werden (vgl. NESTMANN 1977, S. 304). Entsprechend äußern sich auch Erzieher in Befragungen (vgl. HEYER 1974, S. 66; SCHMALOHR 1974, S. 174).

Ein informeller - themenunabhängiger - Erfahrungsaustausch scheint hier für die Gruppe der Erzieher wichtig. Sie lernen so nicht nur die Probleme und Lösungen anderer kennen und stellen evtl. fest, daß sie mit ihren eigenen Problemen nicht alleine stehen. Das Gespräch auf Fortbildungsveranstaltungen kann auch eine Solidarisierung einer Berufsgruppe, eine Verständigung über gemeinsame Ziele und

---

<sup>16</sup> In diesem Zusammenhang sei eine Vermutung geäußert: Aus der Teilnahme an einer bestimmten Fortbildungsveranstaltung muß nicht unbedingt auch auf Interesse am Thema und vom Erzieher erkannte Praxisrelevanz geschlossen werden. Eine Veranstaltung wird auch aus anderen Gründen ausgewählt, z.B. weil Kollegen hinfahren, weil sie in den Terminkalender paßt, weil der Erzieher aus dem Kindergarten heraus und mit anderen kommunizieren will. Verständigung über Inhalte/Ziele und über die Praxisrelevanz wird dann unumgänglich.

und Probleme ermöglichen<sup>17</sup>. Daneben sollte eine Fortbildungsveranstaltung aber auch das Motiv nach zwischenmenschlichen, nach geselligen Kontakten ernst nehmen und in der Gestaltung berücksichtigen. Dies nicht nur, weil jeder Teilnehmer in seinen Motiven ernst genommen werden muß, wenn er als gleichberechtigter Partner anerkannt werden soll, sondern auch weil die Befriedigung oder Nicht-Befriedigung des Bedürfnisses nach geselligen Kontakten Einfluß auf die direkt themenbezogene Arbeit haben kann.

Es wurde schon erwähnt - Kapitel 1 -, daß zentrale Fortbildungsveranstaltungen oft zur Unterstützung von regionalen Arbeitskreisen oder auch Formen anderer selbstorganisierter Fortbildung angeboten werden. Aber nicht nur speziell hier, sondern auch für andere Veranstaltungen wird die Forderung nach Befähigung zur Selbstorganisation von Fortbildung gestellt:

Selbstorganisierte Fortbildung kann oft von Beratern und Fortbildnern nicht genügend inhaltlich und methodisch unterstützt werden. Ein wesentliches Ziel von Fortbildung liegt dann darin, entsprechende Qualifikationen für solche Fortbildung zu fördern (vgl. EHRHARDT-PLASCHKE/IRSKENS 1977, S. 289). Fortbildung sollte weniger Fakten oder Rezepte vermitteln, sondern die Basis schaffen für die Entwicklung von Strategien und Problemlösungen in Eigeninitiative (vgl. NESTMANN 1977, S. 302). Letztlich sollen Fortbildungsveranstaltungen das Endziel haben, Praktiker zu motivieren, ihre eigene Fortbildung selbst in die Hand zu nehmen (vgl. eb., S. 304).

Fortbildungsveranstaltungen für Erzieher lassen sich demnach auch daran messen, inwieweit sie - unabhängig vom

---

<sup>17</sup> Für Mitarbeiterfortbildung in der Erwachsenenbildung spricht SIEBERT hier von einer "Multifunktionalität". Solche Intentionen und Motive sind für ihn weder zweitrangig noch minderwertig (vgl. SIEBERT 1974, S. 156).



spezifischen Thema - Anregungen und Anstöße zur selbständigen Weiterarbeit am Veranstaltungsthema oder auch zur selbständigen Erarbeitung anderer Themen geben<sup>18</sup>.

Durchgehend soll in der analysierten Literatur Erzieherfortbildung vor allem dazu dienen, interaktives und kommunikatives Verhalten von Erziehern zu verbessern. Man kann dies als Förderung der "sozialen Kompetenz" verstehen, womit die Fähigkeit zum gezielten Einsatz von interaktiven Mitteln bei komplexen sozialen Verhaltensproblemen gemeint ist. (vgl. PROSE 1973, S. 116):

Wichtigste Zielsetzung der Erzieherfortbildung ist die Erweiterung der kommunikativen Fähigkeiten der Erzieher, d.h. Erzieher sollen befähigt werden, sich ihren Partnern im Erziehungsprozeß auf sprachlicher und nicht-sprachlicher Ebene verständlich zu machen (vgl. RIEMANN 1975, S. 316). Fortbildung soll sensibilisieren für das Kommunikations- und Interaktionsgefüge, in dem der Kindergarten steht (vgl. IRSKENS u.a. 1975, S. 254). Sie soll das Verständnis und die Wahrnehmungsfähigkeit für Gruppenprozesse entwickeln und fördern; dazu sind Trainingsveranstaltungen im Kommunikationsverhalten notwendig (vgl. HEYER 1974, S. 89). Die Kluft zwischen den Erziehungsansprüchen der Erzieher einerseits und der Realisierung in der täglichen Kindergartenpraxis andererseits soll durch Trainingsseminare zur Verbesserung der Handlungskompetenz überwunden werden (vgl. UNGELENK u.a. 1978, S. 337). Neben der kognitiven Vermittlung von Erkenntnissen über Kommunikationsverhalten sollen Verhaltensweisen in Versuchs- und Realsituationen praktisch erprobt, entwickelt und trainiert werden (vgl. NESTMANN 1977, S. 302).

<sup>18</sup> Dies schließt sich an Forderungen in der allgemeinen Erwachsenenbildung an. TIETGENS und WEINBERG z.B. weisen auf negative, passive, defensive Einstellungen zum Lernen bei vielen Erwachsenen hin. Lernen wird oft zur Vermeidung von etwas Unangenehmen angestrebt. Deshalb fordern sie, neben der Sachvermittlung zugleich immer ein leistungsangemessenes Selbstbewußtsein der Lernenden zu schaffen, indem Lernen in seinem Ablauf und seiner Funktion verdeutlicht wird (vgl. TIETGENS/WEINBERG 1971, S. 91f.). Zugleich soll aktuelles Lernen sich zwar an der Lernbereitschaft anschließen, aber versuchen, ein ruhendes kognitives Potential zu mobilisieren und so die Lernfähigkeit des einzelnen zu fördern (vgl. eb., S. 129; SIEBERT 1975, S. 115f.).

Auf die Diskussion von Trainingsveranstaltungen soll hier nicht eingegangen werden<sup>19</sup>. Eine pauschale und prinzipielle Ablehnung von Trainings wird hier auch nicht geteilt<sup>20</sup>. Erzieherfortbildung findet aber nicht nur in der Form von "Trainings" in interaktiven und kommunikativen Verhaltensweisen statt. Ein solches Thema wäre auch kein Erfolgskriterium eo ipso. Wohl aber sollte Fortbildung so gestaltet werden, daß durch die Art der Kommunikation zwischen Referent/Leiter und Teilnehmern unabhängig vom spezifischen Thema - und dies meint das regulative Prinzip der reziproken Kommunikation - auf Seiten der Erzieher Verhaltenskompetenzen angebahnt werden, die Voraussetzung für den interaktiven und kommunikativen Umgang mit dem Vorschulkind sind (vgl. Kapitel 3.1). Wird Fortbildung ganz oder in größeren Teilabschnitten als Verhaltenstraining durchgeführt, so verlangt das Prinzip der reziproken Kommunikation, daß der Erzieher nicht nur "Objekt" eines Trainingsgeschehens ist. Zum anderen folgt aus der normativen Grundstruktur der intergenerationellen

---

<sup>19</sup> Vgl. hier die Begründung und Darstellung von Trainingsseminaren vor allem aus dem Bereich der Lehrerfortbildung, z.B. TEEGEN/FITTKAU 1972; FITTKAU 1972; MINSEL/MINSEL 1973; PROSE 1973; JESSEN 1975.

<sup>20</sup> Kritisiert wird z.B., daß hinter dem Training die Hoffnung stünde, durch Einstellungsänderung alleine eine Verbesserung der Praxis zu erreichen. Indem die objektive Bedingung ihrer Arbeitswirklichkeit nicht berücksichtigt wird, werden Erzieher durch Trainings nur weiter verunsichert (vgl. AG FORTBILDUNG DES KINDERZENTRUMS ROENEBERGSTRASSE 1976, S. 10). Trainings seien am technologischen Verhaltenskonzept und einem output Modell des Lernens gebunden, der Erzieher würde nur instrumental gesehen, d.h. er soll darauf trainiert werden, Kinder vom Zustand A zu Zustand B zu bringen. Dieses Trainingsgeschehen sei fremdbestimmt und widerspricht dem Ziel der Mündigkeit von Lehrenden und Lernenden (vgl. SCHRADER 1973, S. 295-297).

Reziprozität der Erzieher-Kind-Interaktion, daß dem Trainingsverfahren kein Erziehungsverständnis zugrunde liegen darf, in dem das Vorschulkind nur ein Objekt totaler Manipulation ist (vgl. im Bereich der Lehrerfortbildung hierzu BRUNNER 1973, S. 68f.).

#### **Zusammenfassung:**

Die Thematik von Fortbildungsveranstaltungen für Erzieher soll teilnehmerorientiert ausgesucht werden und praxisrelevante Themen umfassen. Beides verlangt, daß die Erzieher daran beteiligt werden. Die Veranstaltungen sollen Raum für informellen Erfahrungsaustausch bereitstellen und ebenfalls das Bedürfnis nach geselligem Kontakt ernst nehmen. Neben der themenspezifischen Arbeit soll zugleich versucht werden, die Erzieher zur Selbstorganisation ihrer Fortbildung zu befähigen. Einen besonderen Stellenwert nimmt die Förderung der interaktiven und kommunikativen Kompetenzen der Erzieher ein.

Von diesen Merkmalen der Fortbildungsprozesse wird erwartet, daß sie sich fördernd auf die Lernprodukte bei den Teilnehmern auswirken und so einen Beitrag zur Praxisveränderung und -verbesserung leisten.

In der analysierten Literatur finden sich neben diesen Erfolgskriterien noch eine Vielzahl von Forderungen nach spezifischen Inhalten und Zielen für Erzieherfortbildung. Diese werden hier nicht aufgenommen, da versucht wird, Erfolgskriterien zu sammeln, die bei vielen Fortbildungsveranstaltungen den Maßstab für Erfolg oder Mißerfolg liefern können.

### 3.2.2 Erfolgskriterien im Bereich Methodik

In der Fortbildungsveranstaltung muß der Erzieher ein gleichberechtigter Partner seiner Lernprozesse sein. Dies endet nicht mit der Abstimmung des Themas mit den Interessen der Erzieher, sondern berührt ebenso die methodische Gestaltung der Vermittlung. Die Erzieher sollen bei der Verlaufsplanung und Strukturierung der Veranstaltung partizipieren:

Die Fortbildungsteilnehmer sollen konkrete Einwirkungsmöglichkeiten auf Kursplanung und praktische Kursdurchführung haben (vgl. NESTMANN 1977, S. 300; HEYER 1977, S. 87f.; FORSCHUNGSGRUPPE KEIN 1978, S. 205). Lernen ist ein partnerschaftlicher Prozeß von Erwachsenen; diskursive Verständigung bezieht sich nicht nur auf die Auswahl der Ziele, sondern sie ist Leitlinie für den gesamten Lernprozeß (vgl. EHRHARDT-PLASCHKE/IRSKENS 1977, S. 289). Unmittelbare Orientierung an den praktischen Bedürfnissen ist einer umfassenden Vorausplanung überlegen, da dies eine praxisferne Fehlplanung verhindert (vgl. KÜCK 1976, S. 42). Beteiligung von Erziehern setzt auch Freiräume voraus, um von Referenten vorgeschlagene Programme und Problemstellungen zu verworfen und andere Wege einzuschlagen (vgl. HEMMER 1975, S. 7). Ausgangspunkt der Fortbildungsarbeit sind die Bedürfnisse der Teilnehmer auch bezüglich der Arbeitsformen. Die methodischen Vorgehensweisen werden unter weitestgehender Beteiligung der Erzieher ausgewählt (vgl. ARBEITSGRUPPE FORTBILDUNG IM KINDERZENTRUM BERLIN 1977, S. 192f.)<sup>21</sup>.

Die teilnehmenden Erzieher sollen somit ihre Lernprozesse selbst bestimmen und gleichberechtigt an Entscheidungen

---

<sup>21</sup> Das Postulat der Partizipation, Mitbestimmung bei der methodischen Gestaltung der Fortbildung findet sich in vielen Bereichen, in der Lehrerfortbildung z.B. bei BRUNNER 1976, S. 27, 37; HENCARTEN/WEINREBE 1975, S. 41-45; das Prinzip einer weitgehendsten Demokratisierung der Binnenorganisation der Fortbildung betont auch das Konzept der Fortbildung in zentraler Trägerschaft der Arbeiterwohlfaht (vgl. THÜRNAU 1978, S. 180). Partizipation an den Entscheidungen über methodische Gestaltung wird als Erfolgskriterium für die Auswertung des Bildungsurlaubsprojektes aufgestellt (vgl. HOLZAPFEL u.a. 1977, S. 11). In dem Projekt "Entwicklung und Erprobung von Qualifikations-

über die methodische Gestaltung partizipieren. Zudem ist es lernmotivierend und -aktivierend, wenn die Teilnehmer merken, daß sie verlaufsbestimmend mitwirken können (vgl. TIETGENS/WEINBERG 1971, S. 126; SIEBERT 1975, S. 111f.). Selbstbestimmung als Richtziel schließt aber systematische Strukturierung und Planung der Lernprozesse nicht aus (vgl. SIEBERT 1975, S. 26).

Es zeigt sich bei der Partizipation an der methodischen Gestaltung aber ein Problem. Gehen die Referenten und Leiter auf alle Erzieherbeiträge inhaltlicher und verlaufsbestimmender Art ein, so besteht die Gefahr, daß die Veranstaltung in eine Beliebigkeit abrutscht, da die Erzieher - die ja erst sich noch das Thema erarbeiten müssen - nicht alle Teilinhalte und Teilschritte übersehen können<sup>22</sup>. Beliebigkeit bedeutet dann aber für die

---

noch 21 seminaren für nebenberufliche pädagogische Mitarbeiter in der Erwachsenenbildung. NQ-Projekt" wird das Postulat eines selbstbestimmten Lernens vertreten (vgl. KRULL u.a. 1977, S. 5). In der allgemeinen Erwachsenenbildung ist dieses Postulat ebenfalls weit verbreitet. So untersucht die Hannover Studie vor allem Unterrichtsprozesse in der Erwachsenenbildung hinsichtlich der Partizipation der Teilnehmer (vgl. SIEBERT 1975, S. 33). Allgemein ist für Erwachsenenbildung ein Erzieher-Zögling-Verhältnis zu überwinden. Teilnehmer sind gleichberechtigte Subjekte ihrer Lernprozesse und nicht Objekte, die nur auf Impulse der Kursleiter reagieren (vgl. eb., S. 24). Den anthropologischen Hintergrund von "aktiver Partizipation" untersucht SCHERER 1978.

<sup>22</sup> TIETGENS und WEINBERG betonen zwar das Prinzip der didaktischen Selbstwahl, sehen jedoch das Problem darin, ob es sinnvoll ist, bei Wegen mitzubestimmen, die man nicht kennt und nicht übersehen kann. Didaktische Selbstwahl sei vornehmlich sinnvoll, wenn es einen Spielraum für Zielvorstellungen innerhalb eines Kurses gibt (vgl. TIETGENS/WEINBERG 1971, S. 237).

Teilnehmer, daß sie Zeit und Energie in einen Lernprozeß gesteckt haben, der sie nicht weiterbringt. "Sackgassen" und "Umwege" im Lernprozeß sind zwar bei einem selbstbestimmten Lernen nicht auszuschließen, nur nimmt ein Referent den teilnehmenden Erzieher nicht ernst, wenn er ihn nur ohne Hilfe in "Sackgassen" laufen läßt. An dieser Stelle bedarf es eines legitimierten Kriteriums, mit dem ein Fortbildungsleiter verlaufsbezogene Teilnehmerbeiträge auf das - allgemein akzeptierte - Veranstaltungsziel hin selektieren kann. TIETGENS und WEINBERG weisen besonders auf dieses Problem hin und bieten als Selektionskriterium die Sachgerechtigkeit bzw. Sachstringenz des zu Lernenden an. Das Eingehen auf methodische Teilnehmervorschläge darf auf keinen Fall in Widerspruch zu der Sachgerechtigkeit des Themas führen, denn damit ist den Teilnehmern nicht mehr gedient (vgl. TIETGENS/WEINBERG 1971, S. 127). Der Lehrende ist auch ein Treuhänder der Sache und selektiert Teilnehmerbeiträge danach - motivierende Um- und Irrwege hier ausgeschlossen -, ob sie zum Veranstaltungsziel hinführen oder nicht (eb., S. 159f.).

Es liegt hier aber nur dann ein legitimes Selektionskriterium vor, wenn objektiv feststeht, was die "Sache", die "Sachstringenz" oder das "zu Lernende" ist. Dies dürfte bei vielen Inhalten im Vorschulbereich nicht immer gegeben sein. Das "zu Lernende" bzw. die "Sache" entsteht hier oftmals erst durch die Art und Weise, wie der Referent/Leiter an ein Thema herangeht, d.h. durch seine Planungsüberlegungen. Unter dieser Perspektive steht Teilnehmerpartizipation an der Verlaufsplanung - an der unter dem regulativen Prinzip der reziproken Kommunikation natürlich festgehalten wird - letztlich unter zwei Gefahren:

- Abrutschen in eine Beliebigkeit
- verfeinertes Manipulationsinstrument für die Absichten der Veranstalter zu sein, indem diese sich auf eine

vermeintliche Sachstringenz berufen (vgl. auch dazu RUMPF 1973, S. 404f.).

Das Kriterium der Praxisrelevanz des Themas zeigt sich im methodischen Bereich als Forderung nach anwendungs- bzw. praxisbezogener Vermittlung:

Die zu vermittelnden Informationen bzw. der von den Teilnehmern zu erarbeitende Stoff soll immer auf die konkreten Situationen und Erfahrungen am Arbeitsplatz, sowie auf die dort vorhandenen Sachzwänge und Verwertungszusammenhänge gerichtet sein (vgl. NESTMANN 1977, S. 300f.; AG FORTBILDUNG DES KINDERZENTRUMS ROENEBERGSTRASSE 1976, S. 15). Die Lerngegenstände sind Praxis und Praxisprobleme; es sollen keine praxisentzogenen Lernsituationen hergestellt, sondern Praxisprobleme sollen erfahrungsbezogen besprochen werden (vgl. SCHRADER 1973, S. 309f.). Erzieher selbst wünschen sich vor theoretischem Vortrag Fallbesprechungen und Simulationsspiele (vgl. FRÜTEL u.a. 1975, S. 229). Ein integraler Bestandteil des Lernens soll ein systematischer themenbezogener Erfahrungsaustausch sein. Erfahrung sollte sich nicht nur auf Veranschaulichung abstrakter Sätze durch Geschichten aus der Praxis beschränken, sondern eine zweite Säule neben den Beiträgen des Leiters sein (vgl. PRIESNER 1975, S. 54). In der Fortbildung selbst soll die Transferproblematik, die Übertragung des Gelernten in die Praxis angesprochen werden (vgl. ARBEITSGRUPPE FORTBILDUNG IM KINDERZENTRUM BERLIN 1976, S. 195). Die Forderung der Praxisorientierung soll auch durch die didaktisch-methodische Konzeption verwirklicht werden, indem in den methodischen Elementen Praxisbezug ermöglicht und der Transfer konkret vorbereitet wird (vgl. HEYER 1977, S. 88)<sup>23</sup>.

---

<sup>23</sup> Im schon erwähnten NQ-Projekt soll die Praxiserfahrung der Adressaten zum Gegenstand von Lernprozessen gemacht werden. Deshalb wird an Fallbeispielen gearbeitet. Die Autoren sprechen von einer erfahrungsbezogenen Methode des exemplarischen Lernens (vgl. KRULL u.a. 1977, S. 25).

Lernprozesse in der Fortbildung sollen somit als Gegenstand die Praxiserfahrungen<sup>24</sup> der Teilnehmer haben, sowie deren Praxis durch Fallbeschreibungen, Rollenspiele, Videoaufnahmen u.s.w. soweit wie möglich in die Veranstaltung hereingeholt werden sollen. Die Lernbereitschaft der Teilnehmer wird durch die Arbeit an praktischen Beispielen gesteigert (vgl. BRANDENBURG 1974, S. 53); zugleich stellt Lernen an Erfahrungen ein dauerhaftes Lernen dar (vgl. MINSEL/MINSEL 1973, S. 92). Die konkrete Arbeit an eigenen Praxisproblemen soll eine spätere Übertragbarkeit sichern. Transfer stellt sich dabei nicht von alleine ein, sondern bedarf der ausdrücklichen Hinweise (vgl. VONTOBEL 1972, S. 49; TIETGENS/WEINBERG 1971, S. 209). Bei jeder Lösung sollten die Teilnehmer angeregt werden, über das Warum und Wie der Lösung nachzudenken, damit die Lösungsstrategie abstrahiert und für neue Aufgaben zur Verfügung gestellt werden kann (vgl. FROMMER 1978, S. 59). Die Übertragbarkeit in die Praxis hängt noch von der Intensität der Auseinandersetzung mit der Lernaufgabe ab (vgl. ib., S. 56f.). Gelerntes muß auch geübt werden; Übungen sind notwendiger Bestandteil jedes Lernprozesses. Üben des Gelernten an praktischen Beispielen muß zudem nicht mechanisches Üben bedeuten, sondern an neuen Beispielen kann zugleich der Einsichtsprozeß wiederholt werden. Variabilität und Abwechslungsreichheit der Übungen, d.h. Veränderungen in

---

<sup>24</sup> Anknüpfungen an die Erfahrungen der Teilnehmer als Maxime ist nach TIETGENS/WEINBERG nur dann gerechtfertigt, wenn Erfahrung nicht mit Erlebnis gleichgesetzt wird. Erfahrung ist erst dann eine Lernhilfe, wenn das sich Berufen auf sie nicht die Distanzierung von der eigenen Situation verhindert, d.h. Erfahrungen sind nicht gegen Neues immun, sondern können durch Neues verändert werden (vgl. TIETGENS/WEINBERG 1971, S. 63f.).



den Aspekten der Aufgabe oder Situation, erweitern die Übertragungsfähigkeit (vgl. TIETGENS/WEINBERG 1971, S. 226; KUYPERS 1975, S. 75-78)<sup>25</sup>.

Praxisbezogene Vermittlung schließt aber eingeschobene Theorieteile nicht aus, erfordert aber, daß deren Stellenwert für die Praxiserfahrungen der Teilnehmer deutlich sein muß.

Mit dem Kriterium der Partizipation und praxisbezogener Vermittlung im Zusammenhang steht auch die Forderung nach Aktivierung der teilnehmenden Erzieher zu einem selbständigen Lernen:

Fortbildung soll darauf achten, daß ihre Methodik nicht die Eigeninitiative der Erzieher verhindert (vgl. HEYER 1974, S. 80). Nur in begrenztem Maße soll der Leiter Informationen in die Gruppe geben; die größtmögliche Aktivität soll den Teilnehmern überlassen bleiben (vgl. NESTMANN 1977, S. 302). Rechtes Lernen besteht nicht in kritikloser Übernahme und Einprägung, sondern in einem Akt selbständiger Arbeit. Methoden sollen danach beurteilt werden, inwieweit sie eine intensive Auseinandersetzung mit einem Problem ermöglichen (vgl. PRIESNER 1975, S. 57). Lernen auf Fortbildung schließt persönliches Engagement ebenso mit ein wie die Übernahme von Verantwortung für den eigenen Lernprozeß (vgl. DILLER/SCHLAGHECKEN 1978, S. 235)<sup>26</sup>.

---

<sup>25</sup> Zu einer Rehabilitation des oft verschmähten "Übens" siehe auch DIEDERICH 1978.

<sup>26</sup> Für die Erwachsenenbildung allgemein weist TIETGENS auf eine 40jährige Diskussion der Aufforderung zur Aktivierung der Lernenden hin; Selbständigkeit ist oberstes Lernziel (vgl. TIETGENS 1974, S. 24). Vorschläge zur Förderung der Aktivität der Teilnehmer finden sich an vielen Stellen, z.B. in einer Einführung für Kursleiter an Volkshochschulen (vgl. BAUMANN u.a. 1976, S. 47ff.). Die UNESCO-Empfehlung zur Weiterbildung betont ebenfalls die aktive Teilnahme der Lernenden an allen Phasen der Bildungsprozesse (vgl. UNESCO 1977, S. 6). In der VONTOBEL-Befragung über die schweizerische Erwachsenenbildung bevorzugen die Befragten ein aktives-produktives Lernen gegenüber einem "Nürnberger Trichter" (vgl. VONTOBEL 1972, S. 120).

Viele Untersuchungen weisen daraufhin, daß Aktivität, weitgehende selbständige Erarbeitung eines Themas den Lernerfolg steigert (vgl. VONTOBEL 1972, S. 49; SIEBERT 1975, S. 102; BRANDENBURG 1974, S. 36-38). Speziell können Prinzipien, die vom Lernenden selber entdeckt werden, leichter in neue Situationen übertragen werden (vgl. LEON 1977, S. 115). Neben diesen Begründungen von der Lerneffektivität her entspricht der aktive, sich ein Thema selbständig erarbeitende Erzieher mehr dem regulativen Prinzip der gleichberechtigten Partnerschaft als ein nur passiv aufnehmender Rezipient. Die Aktivität der Erzieher in der Fortbildung schließt aber nicht passivere Phasen aus, in denen der Referent z.B. Informationen weitergibt, nur müssen solche Phasen dann immer eingebettet sein in eine prinzipielle Aktivität der Teilnehmer.

---

noch 26 Ein neuer Ansatz, der besonders stark die Aktivierung, Motivierung, Selbständigkeit und Kreativität fördern will, ist das Konzept der "Animation". Dieses Konzept bezieht sich aber überwiegend auf den Freizeitbereich und weniger auf berufliche Fortbildung (vgl. OPASCHOWSKI 1978, S. 116; JOCHUM 1978, S. 128). Vgl. hierzu die Diskussion des Animationskonzeptes in den Themenheften der Vierteljahresschrift ERWACHSENENBILDUNG 24 (1978), Heft 3 und THEORIE UND PRAXIS DER ERWACHSENENBILDUNG 10 (1977), Heft 3 und 4 oder zur Jugendarbeit LIMBOS 1977.

In der Hochschuldidaktik wird die Forderung nach Aktivierung ebenfalls aufgestellt (vgl. SCHOTT 1973, S. 77; KEIL/PIONTKOWSKI 1973, S. 184f.).

Für Schulunterricht definieren ANDERSON und SCOTT als Erfolgsvariable den hohen Grad der Schülerbeteiligung (involvement in learning). Die Autoren stellen fest, daß bei heterogenen Klassen Methoden einer Zweiweg-Kommunikation (Diskussion und Gruppenarbeit) Schülerunterschiede in der Beteiligung minimieren. Diskussionsmethoden hatten durchschnittlich die höchste Beteiligung (vgl. ANDERSON/SCOTT 1978, S. 57). Positive Beziehungen zwischen Schüleraktivität und Schülerleistung weist HECHT nach (vgl. HECHT 1978, S. 288). Den positiven Einfluß von aktivitätsförderndem Unterricht auf die Leistung und die Freude von Schülern am Unterricht (Berufsschule) weist PÄTZOLD 1977 nach.

Die Erfahrung eines aktiven Lernens in der Fortbildung kann zugleich als modellhaft für die Erzieher-Kind-Interaktion betrachtet werden, da der Erwerb von Kompetenzen für das Kind ein aktiver, selbständiger Prozeß ist (vgl. Anmerkung 3, S. 51 der Arbeit).

Zusammen mit Aktivität wird auch oft die Forderung nach Kleingruppenarbeit gestellt:

Kernstück des didaktischen Vorgehens sollte nicht ein Frontalunterricht, sondern Gruppenarbeit sein (vgl. TEIGELER 1976, S. 2; FORSCHUNGSGRUPPE KEIN 1978, S. 204). Vor allem, wenn es in der Fortbildung um Kooperation und Dialog in der Praxis geht, bekommt Gruppenarbeit einen entscheidenden Stellenwert; erfolgreiche inhaltliche Zusammenarbeit in der Gruppe stärkt zugleich das Selbstbewußtsein von Erziehern (vgl. AG FORTBILDUNG DES KINDERZENTRUMS ROENEBERG-STRASSE 1976, S. 15). Die Gruppe sollte als Lernform sowohl zur kognitiven Verarbeitung von Informationen als auch für affektiv-soziales Lernen benutzt werden (vgl. HEYER 1977, S. 88f.). In Befragungen äußern Erzieher ebenfalls ihren Wunsch nach Arbeit in kleinen Gruppen (vgl. FRÜTEL u.a. 1975, S. 229; SCHMALOHR 1974, S. 175).

Gruppenarbeit meint hier die Aufteilung eines Fortbildungsplenums in Untergruppen, mit dem Ziel, eine bestimmte Aufgabe zu bearbeiten. Danach kommen die Teilnehmer wieder im Plenum zusammen. Der Vorteil liegt in der größeren Aktivität der Teilnehmer sowie die Kleingruppe sich leichter dem individuellen Lerntempo anpassen kann. Ein Nachteil der Gruppenarbeit liegt in der hohen Vorbereitungszeit (vgl. KUYPERS 1975, S. 68). Weiterhin kann ein Gespräch in der Kleingruppe unproduktiv werden, indem die Gruppenarbeit nur noch um sich selber kreist (vgl. TIETGENS/WEINBERG 1971, S. 235). Hoher Grad an Gruppenintegration kann zwar in dem Sinne wünschenswert sein, daß für Verhaltensänderungen die Gruppe eine stärkere Autorität als ein einzelner Referent hat, dies kann aber je nach dem Lernziel auch lernhemmend wirken (vgl. VONTOBEL 1972, S. 46; BRANDENBURG 1974, S. 88f.). Arbeit in kleinen

Gruppen ist also kein Erfolgskriterium an sich, es kommt immer darauf an, unter welchen Bedingungen und unter welchen Zwecken sie stattfindet<sup>27</sup>.

Gegenüber der Kleingruppenarbeit wird die Arbeit im Plenum oft negativ beurteilt. Plenen sind aber andererseits in Fortbildungsveranstaltungen nötig, z.B. gerade wenn in Kleingruppen gearbeitet und anschließend das Ergebnis von der Gesamtgruppe zusammengefaßt wird. Die negative Einschätzung liegt wohl oft daran, daß das Plenum in seinen Möglichkeiten nicht richtig eingeschätzt wird (vgl. REICHLE 1977, S. 431)<sup>28</sup>.

Neben der Lerneffektivität bekommt Kleingruppenarbeit in der Erzieherfortbildung noch eine weitere Begründung. Reziproke Kommunikation setzt voraus, daß Teilnehmer sich untereinander kennen und ihre Ängste abbauen können. Innerhalb der Kleingruppenarbeit fällt es leichter den Partner kennenzulernen und Ängste - z.B. sich zu äußern und seine Meinung kundzutun - können leichter abgebaut werden.

---

<sup>27</sup> Gruppenarbeit ist nach TIETGENS und WEINBERG sinnvoll bei großer Teilnehmerzahl, bei Vorkenntnissen und bei instruktivem Arbeitsmaterial (vgl. TIETGENS/WEINBERG 1971, S. 235). Für BRANDENBURG sind kleine Gruppen günstig, wenn

- Diskussionsfähigkeit geübt werden soll,
- stark kontrollbedürftige Ziele erreicht werden sollen,
- die Teilnehmer aktiv werden sollen,
- Werturteile begründet werden sollen,
- allgemeine Folgerungen aus einer Diskussion gezogen werden sollen (vgl. BRANDENBURG 1974, S. 90).

<sup>28</sup> REICHLE versucht die Chancen der Plenumsarbeit aufzuzeigen. Jedes Plenum muß geplant werden, dazu gibt er 7 Punkte an: Funktion des Plenums bestimmen/Ablauf strukturieren/Plenum vorbereiten/Ablauf durchsichtig machen/nicht abschweifen/ermöglichen, daß sich alle Teilnehmer im Plenumsgeschehen wiederfinden/Plenumsverlauf anschließend auswerten (vgl. REICHLE 1977).

Die Erfolgskriterien der Verlaufspartizipation, der praxisbezogenen Vermittlung, der Aktivierung und der Arbeit in kleinen Gruppen alleine garantieren noch keine gelungene Fortbildungsveranstaltung. Diese Kriterien schließen eine differenzierte Vorplanung nicht aus, sondern verlangen sie gerade. Teilnehmerpartizipation und Selbststeuerung der Gruppe z.B. entstehen nicht dadurch, daß der Leiter auf jede Lenkung verzichtet, sondern daß er sie durch Maßnahmen unterstützt (vgl. SIEBERT 1975, S. 29). Ohne "Gegensteuerung" durch den Leiter - so formuliert in der Hannover-Studie - dominieren bei Teilnehmerpartizipation vermutlich die lerngeübten und sprachgewandten Teilnehmer (vgl. eb., S. 77). Ein Kursleiter hat also gerade die Aufgabe, ~~den~~ lernungeübten Teilnehmer Hilfen zum aktiven Einbezug zu geben (vgl. HOLZAPFEL u.a. 1977, S. 28). Zudem scheint es so, daß viele Teilnehmer eher Angst vor der Dominanz einzelner anderer Teilnehmer haben, als daß sie eine zu starke Lenkung des Leiters kritisieren (vgl. SIEBERT 1975, S. 85).

Auf Vorplanungen kann nicht verzichtet werden. Fortbildungsveranstaltungen kosten den Erziehern Zeit und Energie. Da sie kommen, um etwas zu lernen, soll jede Veranstaltung so lernintensiv wie möglich gestaltet werden. Systematische Strukturierung der Lernprozesse wird auch von den Teilnehmern gewünscht (vgl. eb., S. 90, 94). Didaktische Vorplanung ist kein Hindernis für aktives Lernen, sondern gerade eine Vorbedingung von kreativem, emanzipatorischem Lernen. Fortbildung darf aber nicht völlig verplant werden, und so jede Selbstbewegung, jedes eigenständige Suchen der Teilnehmer verhindern. Es sollte ein Mittelweg zwischen Planungslosigkeit (und völliger

Improvisation) und minutiöser Lernplanung gefunden werden (vgl. SIEBERT 1974, S. 123f.)<sup>29</sup>.

Diese Spannung zeigt sich auch in den Äußerungen von Erziehern. Diese wünschen eine abgewogene Form von informellen Möglichkeiten des Austausches und stärker strukturiierter Vermittlungshilfen zwischen Theorie und Praxis (FRÜTEL u.a. 1975, S. 232). Ihre Bedürfnisse nach systematischer Information erfüllt Gruppenarbeit und Diskussion alleine nicht; die Konsequenz sind aber nicht lange Refe-  
rate, sondern flexible Organisationsformen und Arbeitsweisen (vgl. PRIESNER 1975, S. 5ff.). Im Folgenden soll nun auf Erfolgskriterien eingegangen werden, die sich auf im engeren Sinne "methodische" Planungsgesichtspunkte beziehen. Es muß aber immer berücksichtigt werden, daß reziproke Kommunikation als Prinzip von Fortbildung die Anbindung an und Unterstützung von Partizipation, aktivem und selbständigem Lernen verlangt.

Allgemein soll die Gestaltung, Methode einer Fortbildungsveranstaltung den zu vermittelnden bzw. zu erarbeitenden Inhalten entsprechen:

Ist die Thematik "spielerisches Lernen", so soll auch in der Veranstaltung spielerisches Lernen im Vordergrund stehen. Werden Verhaltensprobleme besprochen, so muß auch die Verhaltensebene (z.B. durch Rollenspiele) einbezogen werden. Werden abstrakte und mehr

<sup>29</sup> Einen solchen Mittelweg strebt auch RAAPKE an. Lernen von Erwachsenen ist vor allem von den Motiven her bestimmt. Unter dem Begriff Teilnehmerorientierung hat sich die Diskussion auf diese Aspekte konzentriert. Geht man auf die Motive ein, so müssen die Teilnehmer ja noch nicht ihre Lernprozesse adäquat organisieren können. In offenen Situationen setzen sich zudem auch oft die Robusten mit ihren Interessen durch. Demgegenüber betont RAAPKE die Bedeutung eines systematischen und sorgfältigen Aufbaus und Planung. Das Pegel darf aber auch nicht völlig umschlagen, so daß man bei optimaler Lernorganisation blind gegenüber den Motiven der Lernenden wird. Jede Gruppe braucht lockere Phasen ohne Planung, wenn der Spielraum für Produktivität, Phantasie, intensive Kommunikation sowie Auseinandersetzung nicht verloren gehen soll. Zwischen diesen beiden Polen steht jede Veranstaltung der Erwachsenenbildung (vgl. RAAPKE 1977, insbes. S. 726).

theoretische Informationen vermittelt, so sollte dies in Form von Kurzreferaten geschehen oder z.B. in Gruppendiskussionen entwickelt werden (vgl. NESTMANN 1977, S. 301). Methoden müssen immer abgestimmt werden in bezug auf die gewählten Inhalte (vgl. ACKERMEIER u.a. 1978, S. 66).

Bei der Frage nach geeigneten Methoden kann es zwei Gefahren geben:

- Hypotasierung der Bedeutung einer bestimmten Methode, d.h. eine bestimmte methodische Anlage wie Diskussionsverfahren löst alleine Postulate wie Selbstbestimmung, Teilnehmerorientierung und -partizipation ein.
- Vernachlässigung der Methodenfrage, d.h. z.B. Teilnehmerorientierung wird allein in und durch die Inhalte hergestellt (vgl. HOLZAPFEL u.a. 1977, S. 23f.).

Eine Theorie der Verfahrenswege fehlt noch weitgehend<sup>30</sup>. Idealtypisch läßt sich zwischen Vortrag/darbietender Verfahrensweise und Diskussion/erarbeitender Verfahrensweise unterscheiden, obwohl diese in reiner Form kaum vorkommen und dann auch Lernen hemmen würden. Nur Darbietung läßt die Lernbereitschaft erlahmen; nur Erarbeitung führt leicht zu einer Verwirrung der Lernwege (vgl.

---

<sup>30</sup> WINKEL kritisiert eingehend den Wirrwarr methodischer Ansätze auf der Ebene des Begriffs, der didaktischen Richtungen und der Einteilungskriterien einzelner Methoden. Um eine Systematik der Methoden aufzubauen, lehnt er sich an die kommunikative Didaktik von SCHÄFER und SCHALLER an und diskutiert 17 Unterrichtsmethoden danach, wieviele Elemente des Beziehungsgeflechtes (Lehrer, Schüler, Gegenstand...) sie einbeziehen. Erste Ansätze zur Objektivierung einer Methodenauswahl versucht er mit fünf Fragen zu geben: Ist die Methode inhaltsbezogen und -adäquat? Trägt die Methode der Individuallage Rechnung? Ist sie päd.-didaktisch zu legitimieren? Kann ich die Methode handhaben? Ist sie unter gegenwärtigen schulischen Bedingungen realisierbar? (vgl. WINKEL 1978, insbes. S. 679-682).

TIETGENS/WEINBERG 1971, S. 123)<sup>31</sup>. Pauschal - d.h. ohne Kenntnis des Themas, der Teilnehmergruppe, des Veranstaltungsabschnittes und des Veranstaltungszieles - läßt sich die Frage, ob Diskussion oder Vortrag besser sei, nicht beantworten. In vielen Untersuchungen wurde keine eindeutige Antwort gefunden; es kommt jeweils darauf an, für welchen Zweck welche Vermittlungsform ausgewählt wird. Darbietende Verfahren scheinen Vorteile bei einführender Wissensvermittlung, bei Überblicken über ein Gebiet zu haben, während erarbeitende Verfahren dort Vorteile haben, wo es um Einstellungsänderung, Problemlösen, Abstrahieren, Definieren geht. Darbietende Methoden - die immer mit der Möglichkeit der Rückkoppelung verbunden werden müssen - erfordern relativ wenig Vorbereitung und können auch gut unter Zeitdruck benutzt werden. Erarbeitende Methoden fördern eher Individualität, Aktivität und Selbstbestimmung (vgl. BRANDENBURG 1974, S. 69-76; TIETGENS/WEINBERG 1971, S. 229-234; KUYPERS 1975, S. 55-72).

Festzuhalten bleibt, daß es nicht die Methode der Fortbildungsarbeit mit Erziehern gibt. Erarbeitende Verfahren fördern mehr die Aktivität und Selbstbestimmung und kommen damit dem Kriterium der gleichberechtigten Partnerschaft näher. Andererseits schließt dies nicht Phasen von Informationsvermittlung durch Referate aus. Werden Erzieher kurzfristig zu passiven Rezipienten eines Vortrages, so wird damit das Prinzip der gleichberechtigten Partner-

---

<sup>31</sup> BRANDENBURG äußert mit MILLER noch eine kritische Rückfrage an die Befürworter der Diskussionsmethode. Dahinter stünde oft eine fragwürdige Theorie der Diskussion, nach der demokratische Werte und Verhaltensweisen am besten in Diskussionen vermittelt werden und durch kommunikative Prozesse ein Konsens erzielt werden könnte. Dabei käme es aber oft nicht zu einer Verinnerlichung demokratischer Werte, sondern nur zur Anpassung an den Gruppenkonsens eines gesellschaftlichen Subsystems (vgl. BRANDENBURG 1974, S. 76).

FROMMER bevorzugt sinnvolles-entdeckendes Lernen, welches aber durch lenkende Eingriffe unterstützt werden soll. Mit gelenkter Entdeckung meint er: "So viel Lenkung wie notwendig - so wenig Lenkung wie möglich". (FROMMER 1978, S. 43ff.).



schaft noch nicht verletzt, wenn sie im Kontext des Vortrages Subjekte ihrer Lernprozesse sind, d.h. über die Lernschritte in - prinzipiell - gleichberechtigter Weise mit den Referenten/Leitern abstimmen.

Ebenso wie die Methode den Inhalten entsprechen soll, soll die methodische Vermittlung sich den Voraussetzungen der Teilnehmer anpassen:

Ein Kriterium für eine Fortbildungsveranstaltung ist die Frage, ob die teilnehmenden Erzieher gut folgen können oder sich unter- oder überfordert fühlen (vgl. SCHMALOHR 1974, S. 171). Die Vermittlung muß abgestimmt werden auf die jeweilige Gruppe (vgl. ACKERMEIER u.a. 1978, S. 66).

Das Verhältnis zwischen der methodischen Vermittlung und den Teilnehmervoraussetzungen wird in der Erwachsenenbildung breit diskutiert. Es seien hier - verkürzt - nur einige Aspekte angerissen:

- Ein zu vermittelndes bzw. zu erarbeitendes Thema muß an den Vorkenntnissen, dem Niveau der Teilnehmer anknüpfen (vgl. VONTOBEL 1972, S. 47f.). Dies bedeutet, daß das zu behandelnde Thema auf die teilnehmende Erziehergruppe hin aufbereitet werden muß. Einmal ist dies eine Reduktion der Stofffülle, zum anderen ein Arrangieren der Stoffelemente so, daß er dem Teilnehmer vermittelt werden kann. Für diese beiden Aufgaben wird auch das Begriffspaar "Reduktion" und "Rekonstruktion" benutzt (vgl. TIETGENS/WEINBERG 1971, S. 203-219; BAUMANN u.a. 1976, S. 40f.)<sup>32</sup>. Reduktion und Rekonstruktion verlangen hier vom Leiter bzw. Referenten, daß er die Voraussetzungen der Teilnehmer antizipieren muß.

---

<sup>32</sup> Als Prinzipien der Stoffauswahl formuliert z.B. KUYPERS: "Der Stoff muß so aufbereitet sein, daß er an die Vorkenntnisse der Lernenden anknüpft, er muß vom Leichten ausgehend zum Schweren hinführen, und er muß praxisrelevant sein. Die Beispiele müssen anschaulich und lebensnah sein." (KUYPERS 1975, S. 40).

Dies weist noch einmal auf die Notwendigkeit einer flexiblen Planung hin, damit bei Fehl-Antizipationen eine Revision der Planung möglich ist.

- Das "Tempo" der Arbeit in der Fortbildungsveranstaltung muß so sein, daß die Teilnehmer weder über- noch unterfordert werden. Beste Lernleistungen werden erzielt, wenn der Lernende sein Tempo selbst bestimmen kann (vgl. BRANDENBURG 1974, S. 27). Dies weist schon auf die Notwendigkeit hin, daß Referenten und Leiter ständig bemüht sein müssen zu kontrollieren, inwieweit das Tempo der Erarbeitung den Erziehern entspricht<sup>33</sup>.

- Sprachstil und Gebrauch von Fremdwörtern soll den Teilnehmern verständlich sein. Schwierige Sachverhalte sollen deutlich erklärt werden<sup>34</sup>. Abstraktionen und Fachwörter sind im Vorschulbereich aber notwendig und auf sie kann nicht verzichtet werden. Für Verständnisschwierigkeiten ist zudem weniger der Gebrauch von Fremdwörtern entscheidend als die Syntax der Sprache (vgl. die Diskussion der sprachlichen Verständigung bei TIETGENS/WEINBERG 1971, S. 141ff.; BAUMANN u.a. 1976, S. 42-46). Referenten und

---

<sup>33</sup> Für die Adressatengruppe der Erzieher wird keine Schwierigkeit bezüglich einer Behinderung des Lernens durch Alter der Teilnehmer erwartet. Die Bedeutung der Altersvariable für Lernen ist wohl oft überschätzt worden. SKOWRONEK weist daraufhin, daß aufgrund von Laboruntersuchungen zum kognitiven Lernen von Altersunterschieden vor dem Alter von 50 Jahren nicht zu sprechen ist (vgl. SKOWRONEK 1979, S. 303). Zu einem ähnlichen Ergebnis kommt VERRES-MUCKEL, die ausführlich Befunde zur Lern- und Leistungsfähigkeit älterer Erwachsener bespricht (vgl. VERRES-MUCKEL 1974, S. 67-69).

<sup>34</sup> Ein interessantes Ergebnis diskutieren ROSENSHINE und FURST in ihrem Sammelreferat über "Teacher Performance Criteria". In zwei von vier Studien fand sich eine positive Korrelation zwischen Wahrnehmung der Schwierigkeit durch die Schüler und der Schülerleistung (vgl. ROSENSHINE/FURST 1971, S. 54).

Leiter von Erzieherfortbildung - die oft eine akademische Ausbildung und Sprachweise haben - müssen vor allem auf Verständlichkeit ihres Satzbaues achten.

- Damit das zu vermittelnde bzw. zu erarbeitende Thema in die kognitive Struktur der Teilnehmer integriert werden kann, sollten vor dem Lernen (einmal in der Veranstaltungsankündigung, aber auch vor jedem Arbeitsschritt) Vororientierungen bzw. Ankerbegriffe (advance organizers) bereitgestellt werden, die die kognitive Struktur der Teilnehmer aktivieren (vgl. FROMMER 1978, S. 32f.; SCHOTT 1973, S. 62-64)<sup>35</sup>. Eine besondere Vororientierung ist jene über zu erwartende Schwierigkeiten bei der Arbeit (vgl. KUYPERS 1975, S. 18; BRANDENBURG 1974, S. 46).

- Das Prinzip der Passung meint eine möglichst effektive Relation zwischen Anforderungen des Lerngegenstandes und den Voraussetzungen der Lernenden (vgl. TIETGENS 1974, S. 22f.). Diese Passung bezieht sich ebenfalls auf Rollenerwartungen und Vorstellungen vom Lehren und Lernen bei Teilnehmern und Referenten/Leitern. Die Rollen beider Gruppen sollten komplementär sein (vgl. HOLZAPFEL u.a. 1977, S. 45-49). Oft besteht eine unausgesprochene Übereinkunft über die Gestaltung der Vermittlungsprozesse. Unter dem Prinzip einer reziproken Kommunikation in der Fortbildung kann dies aber nicht heißen, daß sich die Vermittlung tradierten Teilnehmererwartungen (z.B. einseitiger Transport von Wissen) anpassen soll. Das Kriterium der Aktivierung z.B. verlangt, solche internalisierten Lerngewohnheiten aufzubrechen und den Teilnehmer zu einer aktiven Partizipation und einem aktiven Lernen zu führen.

---

<sup>35</sup> Vororientierung in den Ankündigungen der Fortbildungsveranstaltungen und konkrete Arbeit müssen übereinstimmen; Diskrepanzen können zu einer Verminderung der Arbeitsmotivation führen (vgl. z.B. VONTOBEL 1972, S. 45).

Ein weiteres Kriterium ist die Strukturierung und Gliederung des zu vermittelnden bzw. zu erarbeitenden Themas<sup>36</sup>. Wird ein Thema so unübersichtlich und verworren vorgetragen oder erarbeitet, daß die Bedeutung der einzelnen Teile nicht mehr erkannt werden, so schwächt dies die Lernmotivation. Deshalb muß ein Thema in klare Teil- bzw. Lernschritte so aufgeteilt werden, daß der logische Ablauf, die didaktische Progression lückenlos nachzuvollziehen ist. Der "rote Faden" der Arbeit muß immer offen darliegen (vgl. VONTOBEL 1972, S. 47f.; TIETGENS 1974, S. 31f.; KUYPERS 1975, S. 18). Die wesentlichen Bestandteile des Themas müssen klar herausgearbeitet und u.U. schriftlich fixiert werden (vgl. BRANDENBURG 1974, S. 80). Den Teilnehmern muß zu jeder Zeit transparent sein, an welchen Problemen in welchem Zusammenhang gearbeitet wird (vgl. GERL 1975, S. 155ff.). Am Ende von Arbeits- oder Zeiteinheiten sollten die Ergebnisse kurz zusammengefaßt werden. Vor allem wenn zwischen Einheiten eine längere Pause liegt, wirkt sich eine offene Frage oder ein offenes Problem motivierend für die nächste Einheit aus (vgl. KUYPERS 1975, S. 19; TIETGENS/WEINBERG 1971, S. 229; BRANDENBURG 1974, S. 47)<sup>37</sup>.

<sup>36</sup> In SCHOTTs Untersuchungen zu Vorlesungen finden sich hier die Postulate von Schwerpunktsetzungen in der Stoffauswahl/überschaubares Angebot, Beschränkung auf das Wesentliche/Systematik und Gliederung (übersichtliche Ordnung)/Zeitökonomie (gute zeitliche Verteilung des Stoffes)/Zusammenfassungen und Wiederholungen/Erklärungsaktivitäten (Lernhilfen, Beispiele, gute Erklärungen...) (vgl. SCHOTT 1973, S. 75f.).

ROSENSHINE und FURST weisen auf positive Beziehungen hin zwischen der kognitiven Klarheit der Repräsentation eines Stoffes und der Schülerleistung sowie zwischen Statements, die einen Überblick geben über jenes, was geschehen ist und was noch geschehen wird, und der Schülerleistung (vgl. ROSENSHINE/FURST 1971, S. 44f., S. 51f.).

<sup>37</sup> KUYPERS weist noch daraufhin, daß ein leichter Zeitdruck im Lernenden eine Spannung hervorruft, die den Lernerfolg verbessert (vgl. KUYPERS 1975, S. 19; BRANDENBURG 1974, S. 47). An anderer Stelle betont BRANDENBURG aber, das Lernen unter Zeitdruck schädlich ist (ebd. S. 80)!

Ein weiterer Aspekt in der Fortbildungsveranstaltung ist die Kontrolle von Lernfortschritten:

An wichtigen Stellen im Verlauf des gesamten Lernprozesses soll unter Mitwirkung der Beteiligten das Lerngeschehen einer Kontrolle und Bewertung unterzogen werden (vgl. ACKERMEIER u.a. 1978, S. 29f.).

Kontrolle von Lernfortschritten hat in diesem Zusammenhang nichts zu tun mit "Prüfung" oder "Klassenarbeit", sondern bezieht sich auf die Vergewisserung des Standes der gemeinsamen Arbeit. Die Erkenntnis eines Lernfortschrittes wirkt motivierend für die weitere Arbeit und hilft dem Referenten/Leiter in der Planung der folgenden Arbeitsschritte (vgl. SIEBERT 1975, S. 112f.; KUYPERS 1975, S. 18; BAUMANN u.a. 1976, S. 61-64). Lernkontrolle kann hier durch Gruppendiskussion oder auch unterstützt durch kleinere feed-back Bögen geschehen.

Unter dem Prinzip der reziproken Kommunikation wird die Reflexion des Arbeitsstandes besonders wichtig. Nur wenn der Stand der Arbeit verdeutlicht und bewertet wird, wenn die Erzieher in die Lage versetzt werden, zu prüfen, was in der Veranstaltung wie erreicht wurde, können sie als gleichberechtigte Partner mit über ihre Lernprozesse entscheiden. Unklarheit über die Lernfortschritte, über den Arbeitsstand verhindern jede sinnvolle Mitbestimmung<sup>38</sup>.

Fortbildungsveranstaltungen sollen aufgelockert und abwechslungsreich gestaltet werden:

---

<sup>38</sup> Vgl. hierzu auch die Diskussion im Konzept für die Auswertung des Bildungsurlaubsprojektes. Durch Reflexionsschritte sollen die Teilnehmer Klarheit über die Lernprozesse gewinnen und überprüfen können, ob ihre eigenen Interessen sich im Lernprozeß wiederfinden, m.a.W. ob die von den Veranstaltern vorher antizipierten Teilnehmerinteressen auch die realen Interessen der Teilnehmer sind (vgl. HOLZAPFEL u.a. 1977, S. 32-34).

Grundsätzlich sollen die Arbeitsformen abgewechselt werden; spielerische, kreative Aktionen, Pausen zum Schlafen, Spaziergehen, Singen und Tanzen sind fester Bestandteil der Lernarbeit (vgl. FORSCHUNGS-GRUPPE KEIN 1978, S. 205)<sup>39</sup>

Gleichberechtigte Partnerschaft in der Fortbildungsarbeit verlangt auch, daß die Erzieher mit ihren nicht primär themenbezogenen Wünschen ernst genommen werden. In diesem Sinne ist - nach einer längeren intensiven Arbeit - der Wunsch nach einer Pause z.B. zum Spaziergehen ein legitimes Anliegen, auf das eingegangen werden sollte. Die Erfahrung, daß über die Arbeitszeit verfügt werden kann und nicht ein vorher festgelegter Stundenplan strikt "durchgezogen" werden muß, ist zudem eine Erfahrung, die für die Kindergartenarbeit direkt relevant ist. In der praktischen Arbeit sollte ebenfalls Zeit nicht rigoros verplant werden, sondern Kinder sollen auch ihre Wünsche nach Abwechslung artikulieren können. Dieses freie Verfügen über Zeit darf aber nicht - sowohl in der Kindergartenpraxis als auch in der Fortbildung - dazu führen, daß jede themenbezogene Arbeit abgebrochen wird.

Abwechslung in der direkten themenbezogenen Arbeit bezieht sich auf den Wechsel von Arbeitsformen und Methoden. Durch Methodenwechsel kann die Aktivität und Lernenergie der Teilnehmer gefördert werden, damit wird einer Monotonie entgegengewirkt, die letztlich lernhemmend ist (vgl. SIEBERT 1975, S. 103; TIETGENS/WEINBERG 1971, S. 238f.; KUYPERS 1975, S. 82-84; BIEDERMANN 1977, S. 104f.)<sup>40</sup>.

<sup>39</sup> Die Variabilität von Unterricht (Variabilität im Material und den Arbeitsformen) ist in vielen Untersuchungen positiv korreliert mit der Schülerleistung (vgl. ROSENSHINE/FURST 1971, S. 45f.).

<sup>40</sup> BIEDERMANN diskutiert hier ausführlich das Prinzip "Methodenwechsel" und die Bedeutung für das schulische Lernen. Er betrachtet Methodenwechsel als gleichberechtigtes Unterrichtsprinzip neben z.B. "Altersgemäßheit" (vgl. BIEDERMANN 1977, S. 103).

Eine bestimmte Methode fordert eine bestimmte Lernart des Teilnehmers heraus. Teilnehmer unterscheiden sich in den Fähigkeiten in den Lernarten. Methodenwechsel bietet so jedem Teilnehmer die Möglichkeit, entsprechend seinen besonderen Fähigkeiten ein Thema optimal zu erarbeiten (vgl. BAUMANN u.a. 1976, S. 27f.). Jede Methode hat zudem spezifische Vor- und Nachteile in der Fortbildungsarbeit; durch Methodenwechsel können Nachteile der einen Methode durch Vorteile der anderen ausgeglichen werden (vgl. BIEDERMANN 1977, S. 104).

Ein besonderes Augenmerk wird teilweise auf die Arbeit mit Medien gelegt:

In der praktischen Arbeit im Kindergarten zeigen Erzieher ein Defizit im Umgang mit Medien. Deshalb sollte in der Fortbildung modellhaft mit Medien gearbeitet werden (vgl. HEYER 1974, S. 73-75). Ein Medium, welches zur Einführung, Vertiefung und Problematisierung erfolgreich in mehrtägige Fortbildungsarbeit eingebracht werden kann, ist die Fachzeitschrift (vgl. GROSSE-JÄGER 1977, S. 117-119). Es wäre oft Zeitverschwendung, erste Lösungsversuche in einem Gebiet zu unternehmen, wo andere schon Erfahrungen schriftlich niedergelegt haben. Literatur ist zur thematischen Vor- und Nachbereitung wichtig (vgl. ACKERMEIER u.a. 1978, S. 95-100).

Medien können vielfältige Funktionen erfüllen, z.B. können sie Teilnehmer motivieren, sie dienen zur Veranschaulichung, sie können den Unterricht beleben, sie haben eine objektivierende, begründete Rolle, sie können Übungsobjekte sein (vgl. u.a. TIETGENS/WEINBERG 1971, S. 242-244). Medieneinsatz allein ist aber auch kein Erfolgskriterium für Erzieherfortbildung. Medien müssen immer kritisch ausgewählt und sorgfältig vorbereitet werden, ihr Gebrauch ist nur sinnvoll, wenn sie für Aufgaben und Funktionen eingesetzt werden, die sie auch erfüllen können (vgl. KUYPERS 1975,

S. 86ff.; BRANDENBURG 1974, S. 91ff.)<sup>41</sup>.

#### Zusammenfassung:

Für die methodische Ebene der Gestaltung von Fortbildung werden folgende Erfolgskriterien genannt, die zugleich die Lernprodukte auf Seiten der Teilnehmer fördern sollen: Bei Entscheidungen über den Ablauf der Veranstaltung sollen die Teilnehmer gleichberechtigt partizipieren; die Vermittlung der Themen soll anwendungsbezogen vor sich gehen<sup>42</sup>. Die Teilnehmer sollen zu einem selbständigen

<sup>41</sup> Die Verwendung von Hilfsmitteln in Vorlesungen bespricht auch SCHOTT. Er betont aber mit informations- und motivationstheoretischen Ansätzen, daß nicht der Einsatz als solches schon den Lernerfolg steigert, es kommt immer darauf an, daß die Medien (Schautafeln, Zeichnungen, Tafelbilder, Dias, Filme) auch die Stoffvermittlung intensivieren und nicht verschiedene Informationen auf zwei Kanälen auf die Teilnehmer eintreffen, bzw. ein Kanal soviel Informationen enthält, daß der andere kaum beachtet werden kann (vgl. SCHOTT 1973, S. 29-43).

<sup>42</sup> Neben den hier besprochenen methodischen Erfolgskriterien gibt es noch eine Vielzahl von spezifischen Ratschlägen zur methodischen Gestaltung, auf die hier u.a. deshalb nicht eingegangen wird, weil sie das zu entwickelnde Evaluationsinstrument zu umfangreich machen und damit die praktische Einsetzbarkeit verhindern würden.

Es sei hier nur noch auf die vielen Ratschläge hingewiesen, die innerhalb einer Veranstaltung helfen sollen, die Motivation der Teilnehmer zu erhalten und zu erhöhen:

Hilfreich für die Methoden ist ein kognitives Vakuum; der Ist-Zustand wird in Frage gestellt und es gibt eine Diskrepanz zwischen Können und Wollen. Diese Diskrepanz muß zwar stark repräsentiert sein, aber überwindbar erscheinen und die Überwindung muß den Teilnehmern bedeutsam sein. Der Teilnehmer soll während der Arbeit Erfolgserlebnisse haben. Lob und Tadel unterstützen dabei die Motivation (bei entsprechender Attribuierung: z.B. Tadel bei mangelnder Anstrengung, Lob bei realisierten Fähigkeiten), ebenso wie das Erlebnis der persönlichen Verursachung. Der Schwierigkeitsgrad muß auf die Fähigkeiten der Teilnehmer optimal passen. Schließlich unterstützt eine gute Geselligkeit auch die thematische Arbeit. Andere Interessen der Teilnehmer sollten für die Arbeit nutzbar gemacht werden, indem evtl. Beispiele aus den die Teilnehmer interessierenden Bereichen genommen werden (vgl. hierzu u.a. TIETGENS/WEINBERG 1971, S. 44ff.; VON-TOBEL 1972, S. 42f.; BRANDENBURG 1974, S. 17, 45-47; SIEBERT 1975, S. 111f.; FROMMER 1978, S. 22-30).



Lernen aktiviert werden, dazu gehört auch die Arbeit in kleinen Gruppen. Diese Forderungen schließen eine - flexible - Planung nicht aus, sondern verlangen sie gerade. Die methodische Vermittlung muß auf die Voraussetzungen der Teilnehmer und auf die Themen abgestimmt werden. Die Lernprozesse selbst müssen sorgfältig strukturiert und Fortschritte ständig kontrolliert werden. Dabei soll die Veranstaltung abwechslungsreich und aufgelockert gestaltet werden. Der Einsatz von Medien hängt davon ab, ob sie hilfreich die Arbeit unterstützen.

### 3.2.3 Erfolgskriterien im Bereich Interaktionsprozesse

Die Forderung nach reziproker Kommunikation muß sich auf allen Ebenen der Fortbildungsarbeit mit Erziehern wiederfinden. Wurde im vorherigen Abschnitt eher der methodische Aufbau und die Partizipation der Erzieher daran besprochen, so werden im folgenden Interaktionsprozesse unter dem Blickwinkel des Arbeitsstiles bzw. des sozial-emotionalen Arbeitsklimas betrachtet:

Für das Lernen mit Erziehern muß der Gruppenprozeß berücksichtigt und eine angstfreie Atmosphäre geschaffen werden. Dies verlangt den Abbau von Leiter-Zentrierung und die Herstellung von Mehrbahn-Kommunikation (vgl. DILLER/KIEL 1977, S. 113). Zur Selbstverwirklichung der Person des Erziehers sollte der Leiter ein Sprachverhalten im Sinne der Gesprächspsychotherapie zeigen, welches die Selbsterkundung und Persönlichkeitsreifung des Erziehers begünstigt (vgl. PRIESNER 1975, S. 55). Ein autoritäres oder missionarisches Leiterverhalten ist abzulehnen, allein vertretbar ist das Sachkompetenzverhältnis in partnerschaftlichem Stil. Der Erzieher ist ein prinzipiell Gleichberechtigter, Dominanz zwischen allen Beteiligten ist abzubauen (eb., S. 57). Fortbildung soll dem Erzieher ein weniger dirigistisches Verhalten vorführen, als er es in der Ausbildung erlebt hat (vgl. HEYER 1974, S. 89). Ein legitimes Ziel - wie auch Voraussetzung für die Arbeit - ist die Erfahrung einer entspannten Atmosphäre und erweiterter mitmenschlicher Begegnung (vgl. NACHBAUER 1978, S. 326).

Die interpersonalen Beziehungen der an der Fortbildung beteiligten Partner bestimmen die Arbeitsatmosphäre und beeinflussen so kognitives, emotionales und soziales Lernen (vgl. HEYER 1977, S. 175)<sup>43</sup>. Bezogen auf das Prinzip der gleichberechtigten Partnerschaft an den Lernprozessen muß hier ein Arbeitsklima vorliegen, in dem die Erzieher frei und ohne Angst ihre Meinungen, Interessen, positive und negative Gefühle äußern können. Ohne ein solches "herrschaftsfreies" Arbeitsklima ist eine reziproke Kommunikation nicht möglich und es kann nicht erwartet werden, daß die Erzieher ihre eigenen Wünsche und Bedürfnisse artikulieren. Eine entspannte und angstfreie Gruppenatmosphäre verlangt die weitgehende Abwesenheit von konkurrenzorientierten Verhaltensweisen und Einstellungen. Das Verhalten soll nicht an einem Konkurrenten sich orientieren, sondern die Teilnehmer sollen sich zu einer solidarisch lernenden Gruppe zusammenschließen, da nur so die Kapazitäten einer Gruppe entfaltet werden können (vgl. HOLZAPFEL u.a. 1977, S. 51-54). Je herrschaftsfreier das Arbeitsklima ist, desto eher erfolgt eine Motivation der Teilnehmer (vgl. SIEBERT 1975, S. 111)<sup>44</sup>.

<sup>43</sup> HEYER diskutiert sehr ausführlich die eher interaktive Seite der Erzieherfortbildung auf dem Hintergrund des partnerzentrierten Modells von ROGERS, des Kommunikationsmodells von WATZLAWIK, BEAVIN und JACKSON und der themenzentrierten Interaktion von COHN. Es führt zu weit, hier explizit auf diese Überlegungen einzugehen. Der Leser sei - vor allem bezogen auf die ausführlichen theoretischen Begründungen - auf die Arbeit von HEYER verwiesen. Die von HEYER vorgelegten Postulate zur Fortbildungsarbeit sind zudem in der hier vorgelegten Sammlung von Erfolgskriterien schon enthalten.

<sup>44</sup> Entsprechende Forderungen an ein angstfreies Klima finden sich an vielen Stellen; z.B. für Verhaltenstrainings für Erzieher bei FITTKAU, dort finden sich auch die emotional-sozialen Lernbedingungen von PINE und HORNE, die Lernen erleichtern und eine angstfreie Atmosphäre schaffen (vgl. FITTKAU 1972, S. 257f.). Die Notwendigkeit eines entspannten sozial-emotionalen Klimas, in der die Partner sich als gleichberechtigt akzeptiert fühlen, betont auch SCHOTT für die aktive Informationsverarbeitung in Vorlesungen (vgl. SCHOTT 1973, S. 77).

Eine solche Atmosphäre, in der nonkonforme Beiträge geäußert werden können und sich jeder angstfrei beteiligen kann, bedarf der Steuerung durch den Referenten/Leiter. Freigebendes Leiterverhalten garantiert keinen offenen Gesprächsverlauf. Der Leiter muß zudem versuchen, ein Gespräch für alle Beteiligten offen zu halten. Er tritt hier als Anwalt der Minderheiten auf und wirkt der Dominanz einzelner Teilnehmer entgegen (vgl. TIETGENS 1974, S. 161f.; BRANDENBURG 1974, S. 86; SIEBERT 1975, S. 26,77).

Offenes, angstfreies Klima wird also nicht durch Verzicht, sondern nur durch Wahrnehmung von Leitungsfunktionen erreicht. Bei den vielfältigen Leitungsfunktionen lassen sich zwei Grundbestandteile unterscheiden: mehr aufgabenorientierte Funktionen einerseits und mehr soziale Funktionen andererseits (vgl. GERL 1979, S. 522f.). Die erste bezieht sich auf die themenbezogene Arbeit, d.h. auf die Ausrichtung der Aktivitäten am Arbeitsziel, während die zweite die Schaffung des günstigen Arbeitsklimas meint. Untersuchungen weisen daraufhin, daß die Maximierung der sozialen Funktion, d.h. eine überwiegende Orientierung an den Beziehungsstrukturen der Gruppe für Arbeitsgruppen nicht zu einer hohen Leistung führt. Angestrebt werden sollte ein mittlerer Wert zwischen reiner Aufgaben- und reiner Beziehungsorientierung, der die Arbeitsfähigkeit einer Gruppe gewährleistet. Jede Aktivität des Fortbildungsleiters hat Auswirkungen in beiden Dimensionen:

"sie fördert oder hindert die Aufklärung der behandelten Sachfragen, sie definiert gleichzeitig die eigene Rolle oder Position in der Gruppe, sie verstärkt oder frustriert bestimmte Teilnehmer oder Teilgruppen usw." (GERL 1979, S. 524)

Angestrebt werden sollte - nach GERL - eine Verteilung der Leitungsfunktionen auf die Teilnehmer selbst.

Die sozial-emotionale Seite und die Interaktionsbeziehungen brauchen in der Fortbildungsarbeit nicht immer direkt besprochen werden. Es gibt einen Spielraum, wo diese nicht behandelte - aber immer vorhandene - Problematik latent gehalten werden kann, ohne daß sie die Arbeit an der Sache stört. Latent halten heißt aber nicht, daß die unterdrückt wird. Besprochen werden muß sie sozial-emotionale Seite dann, wenn sie die effektive Arbeit behindert, indem z.B. scheinbare Sachauseinandersetzungen nur einen Machtkampf um Führungspositionen in der Gruppe reflektieren (vgl. GERL 1975, S. 134f.)<sup>45</sup>. Der Leiter muß ein besonderes Augenmerk auf den sozial-affektiven Bereich legen und ihn u.U. thematisieren. Dies bedeutet aber nicht unbedingt, daß die Arbeit am Thema abgebrochen und nur noch über affektive Zustände geredet wird, sondern daß der Stellenwert des affektiven Bereiches im Veranstaltungsablauf deutlich gemacht und sich darüber verständigt wird, wann, wo und in welcher Weise über diese Probleme gesprochen werden soll (vgl. HOLZAPFEL u.a. 1977, S. 43).

Weitere Postulate für das Leiter- bzw. Referentenverhalten werden in der Literatur sehr breit diskutiert. Es kann hier nicht das Ziel dieser Arbeit sein, diese systematisch aufzuarbeiten, vielmehr wird sie sich auf einige Hinweise - auf dem Hintergrund des Prinzips der reziproken Kommunikation - beschränken müssen.

---

<sup>45</sup> GERL versucht hier systemtheoretische Überlegungen für die Analyse von Unterrichtsprozessen fruchtbar zu machen. An Lernsituationen sind von hier zwei Forderungen zu stellen:

- es muß eindeutig sein, welches Problem gerade besprochen wird (Problemtransparenz)
- die nicht behandelte, sozial-emotionale Thematik darf nicht übermächtig werden, d.h. es bedarf einer entspannten, unkonflikthierenden Gruppensituation oder aber diese Problematik muß bearbeitet werden (vgl. GERL 1975, S. 155).

Einige Autoren fordern - vor allem in Zusammenhang mit Trainingsveranstaltungen - ein nicht direktives, sozial-integratives Referenten- und Leiterverhalten und die Übernahme von Regeln eines eher psychotherapeutisch orientierten Verhaltens. Solches Verhalten läßt sich - auf dem Hintergrund z.B. der erziehungspsychologischen Konzeption von TAUSCH, des partnerzentrierten Modells von ROGERS und gesprächspsychotherapeutischer Modelle - kennzeichnen durch Akzeptierung der eigenen wie der anderen Person; durch Achtung, Wärme und Rücksichtnahme; Echtheit und Aufrichtigkeit; einführendes Verstehen; nicht dirigierende Aktivitäten und ein mittleres Ausmaß an Lenkung. Solches Verhalten wird nicht nur als effektiv für die Vermittlung von Wissen angesehen<sup>46</sup>, sondern das Referenten/Leiterverhalten soll ein Modell für das Erziehverhalten sein (vgl. HEYER 1977, insbes. S. 172, S. 79f.; TEEGEN/FITTKAU 1972, S. 45; FITTKAU 1972, S. 257; MINSEL/MINSEL 1973, S. 104). Wie schon erwähnt - vgl. Kapitel 3.1 - ist die Interaktion während einer Fortbildungsveranstaltung kein direktes Modell für die Erzieher-Kind-Interaktion, da für Fortbildung von prinzipiell voll handlungsfähigen Interaktionspartnern ausgegangen werden muß. Trotzdem ist das entsprechende Verhalten der Fortbildungsleiter oder -refe-

---

<sup>46</sup>Für die Schulpädagogik sei hier eine Untersuchung von JOOST genannt. Schüler, die nach dem obigen Leiterverhalten unterrichtet wurden, zeigten Unterrichtsbeiträge von höherem geistigen Niveau, verhielten sich spontaner und zeigten eine bessere Arbeitsmotivation, sie hatten weniger Angst, verhielten sich offener und hatten gegenüber den Lehrern eine emotional positivere Einstellung als eine Kontrollgruppe (vgl. JOOST 1978).

OLIVERO weist auf Untersuchungen hin, in denen wirksames Lehren gekennzeichnet ist durch häufige Anwendung von indirektem Einfluß (Akzeptieren und Erläutern von Vorstellungen) und weniger häufige Anwendung von direktem Einfluß (Anweisung, Befehle, Ablehnung, Kritik) (vgl. OLIVERO 1973, S. 28).

renten eine notwendige Erfahrung als Vorbereitung für die spezielle Erzieher-Kind-Interaktion unter der Grundnorm der intergenerationellen Reziprozität.

In der Erwachsenenbildung wird - teils mit identischen Elementen - von einem demokratischen oder hörerzentrierten Führungsstil gesprochen, der kooperative und selbständige Lernprozesse, wechselseitige Verständnisbereitschaft und Regulierung von emotionalen Spannungen ermöglichen soll (vgl. BRANDENBURG 1974, S. 84f.). Unumgänglich für eine reziproke Kommunikation in der Erzieherfortbildung ist die Forderung der Reversibilität: Ein Kursleiter soll sich den Teilnehmern gegenüber so verhalten, wie er es von ihnen sich selbst gegenüber erwartet (vgl. TIETGENS 1974, S. 30E). Prinzipielle Gleichberechtigung von Lehrenden und Lernenden entspricht am ehesten ein lernorientierter (d.h. auf den Lerner orientierter) bzw. teilnehmerorientierter Arbeitsstil, der die Produktivität der Lernenden fördert (vgl. TIETGENS/WEINBERG 1971, S. 159). Teilnehmerorientierung bedeutet hier aber nicht *laissez-faire*, der Leiter geht nicht auf jeden Wunsch der Teilnehmer ein, sondern übernimmt - als Treuhänder der Sache - explizit Führungsaufgaben. Teilnehmerorientierung meint Zurückhaltung und den Versuch, Selbststeuerungsmöglichkeiten zu schaffen (vgl. eb., S. 161f.)<sup>47</sup>. Stellen die Teilnehmer Fragen, die sich auf etwas richten, was im Hinblick auf die diskutierte

---

<sup>47</sup> Verdeutlichen läßt sich dies am Verhalten eines Diskussionsleiters. Er soll sich nicht auf eine formale Beteiligung zurückziehen, denn die Teilnehmer kommen ja, um etwas zu lernen. Der Leiter soll aber seine Steuerungsfunktion zurückhaltend wahrnehmen, sowohl Übermaß als auch Verzicht sind schlecht. Das Kriterium für die Steuerung bietet die Sachgerechtigkeit (vgl. TIETGENS/WEINBERG 1971, S. 165f.; s.a. die 10 Regeln für eine teilnehmerorientierte Gesprächsführung S.175f.). Vgl. auch die Verhaltensregel für Diskussionsleiter und dazu gehörige Spielregeln für die Gruppe bei ACKERMEIER u.a. 1978, S. 85f.

Thematik zweitrangig oder abwegig ist, so sollte das Nichteingehen auf solche Beiträge begründet werden, um hier die Verfahrensweise transparent zu machen (vgl. eb., S. 182f.).

Der Fortbildungsreferent oder -leiter verfügt zwar über einen Informationsvorsprung und Kompetenzen in der Organisation der Lernprozesse, er hat aber keine Legitimation als Meinungsführer. Nicht er allein kritisiert und bestätigt, sondern er soll die Teilnehmer gerade zur Kritikfähigkeit herausfordern (vgl. SIEBERT 1975, S. 26).

Gleichberechtigte Partnerschaft heißt nicht Harmonie und nur zufriedene Teilnehmer, der Leiter selbst sollte sachliche Auseinandersetzungen und Spannungen fördern (eb., S. 115). Voraussetzung ist aber eine unbedingte persönliche Anerkennung, die unabhängig ist von der Richtigkeit einzelner inhaltlicher Argumente (vgl. GERL 1977, S. 533).

Gerade für Erzieherfortbildungsveranstaltungen soll der Leiter und Referent nicht nur seine Veranstaltung "abwickeln", sondern sich in einem größeren Kontext der Vorschulerziehung sehen. Er soll fachlich qualifiziert sein und zudem - angesichts des Vorwurfs der Praxisferne - selbst Praxiskenntnisse besitzen (vgl. HEYER 1977, S. 177). Reziproke Kommunikation verlangt von ihm, daß er sich um eine persönliche Beziehung zum Erzieher bemüht und bereit ist, seine Vorstellungen aufgrund von berechtigter Kritik zu ändern<sup>48</sup>.

---

<sup>48</sup> ROSENSHINE und FURST berichten über positive Einflüsse auf die Schülerleistung. Darunter fallen die Lehrervariablen: begeisternder, mitreißender, stimulierender Lehrer (46f.); aufgabenorientiertes, sachliches Verhalten (47); Indirektheit (49); überwiegend negative Beziehungen zeigen sich zwischen Lehrerkritik und Schülerleistung (51) (vgl. ROSENSHINE/FURST 1971).

In der Erwachsenenbildung allgemein wird oft festgestellt, daß Veranstaltungen nicht nur aus direkt themenbezogenen Gründen besucht werden, sondern z.B. als Mittel zur Befriedigung geselliger Motive oder zur Abwechslung (vgl. VONTOBEL 1972, S. 85). Während der Veranstaltung nehmen Bedürfnisse nach Anerkennung, Sicherheit und Kontakt mit anderen einen wichtigen Stellenwert ein (vgl. BAUMANN u.a. 1976, S. 18-25; HOLZAPFEL u.a. 1977, S. 41f.). Es wird davon ausgegangen, daß diese Motive auch in der Erzieherfortbildung eine wichtige Rolle spielen<sup>49</sup>. Unter dem Prinzip der gleichberechtigten Partnerschaft müssen diese Motive ernst genommen werden. Während einer Fortbildungsveranstaltung sollte deshalb Raum und Zeit für gesellige Beziehungen und gemeinsame Spiele und Entspannung bereitgestellt werden.

#### Zusammenfassung:

Die Interaktionsprozesse in der Erzieherfortbildung sollen somit in einem angstfreien Klima stattfinden. Dies verlangt Steuerungsfunktionen vom Fortbildungsleiter und -referenten. Die nicht themenbezogene, sozial-emotionale Seite muß kontrolliert und in Balance mit der themenbezogenen Arbeit gehalten werden. Der Fortbildungsleiter soll sich partnerzentriert, nicht direktiv bzw. teilnehmerorientiert verhalten. Für die Befriedigung der Motive nach Geselligkeit soll Raum zur Verfügung gestellt werden. Werden auch diese Prozeßmerkmale postuliert, damit die Produkte, Lernergebnisse der Fortbildung auf Seiten der Teilnehmer vergrößert werden, so muß aber gerade hier auf den Eigenwert dieser

---

<sup>49</sup> So äußerten sich mehrere Erzieher in einem informellen Gespräch über den Fragebogen: "Man macht sich oft einige schöne Tage auf der Fortbildung."



Prozeßkriterien hingewiesen werden: sie werden in dieser Arbeit als unumgängliche Kriterien betrachtet, soll nicht gegen das regulative Prinzip der reziproken Kommunikation verstoßen werden.

### 3.2.4 Erfolgskriterien im Bereich situativer und organisatorischer Bedingungen

Die Idee der Teilnehmerorientierung und gleichberechtigten Partnerschaft bezieht sich auch auf den situativen und organisatorischen Kontext, unter dem die Fortbildung stattfindet, d.h. auf den äußeren Rahmen der Lernbedingungen. Sie zeigt sich hier als die Orientierung auf optimale äußere Bedingungen für die Teilnahme von Erziehern:

Aus der Diskussion von verschiedenen Problemen der Fortbildungsarbeit (z.B. familiäre Belastung, Zeitprobleme, Fragen der Vertretungskräfte, Probleme, die entstehen, wenn nur ein Teammitglied eine Veranstaltung besuchen kann, Weitergabe von Wissen, lange Anfahrtswege usw.) wird als Mindestvoraussetzung u.a. gefordert: Fortbildung muß in der Arbeitszeit liegen und innerhalb einer bestimmten Zeit alle Mitglieder eines Praxisteams umfassen (vgl. RIEMANN 1975, S. 315; NESTMANN 1977, S. 303). Nicht nur Erzieher, sondern möglichst alle Mitarbeiter einer Einrichtung sollen an der Fortbildung teilnehmen, um die Dauerhaftigkeit von Veränderungen zu sichern (vgl. MINZ 1978, S. 79f.). Fortbildung soll regelmäßig und nicht in zu weit auseinandergezogenen Zeitabständen stattfinden und sie muß Bestandteil des Arbeitsverhältnisses sein (vgl. MOSKAL 1978, S. 319). Fortbildung sollte durch Beurlaubung, Übernahme der Kosten durch den Dienstgeber gefördert werden (vgl. NACHBAUER 1978, S. 333). Erzieher selbst wünschen sich Bildungsurlaub, Vertretungskräfte und Kostenerstattung (vgl. HEYER 1974, S. 138), desweiteren: mehrtägige Veranstaltungen etwa vierteljährlich, obligatorische Kurse unter Schließung der Einrichtung oder - damit die Einrichtung nicht geschlossen werden braucht - eine schichtweise Teilnahme der Erzieher (vgl. SCHMALOHR u.a. 1974, S. 115).

Fortbildung sollte nicht nur außerhalb der bestehenden Praxisinstitutionen stattfinden, sondern auch innerhalb. Effektiv sind hier Blocklehrgänge mit eingeschobenem

Praxisphasen. Lernen außerhalb der Praxiseinrichtung ermöglicht kritisches Lernen in Distanz zu Sachzwängen und Routine; Phasen in der Praxisinstitution ermöglichen eine Überprüfung und Umsetzung des Gelernten (vgl. IRSKENS u.a. 1975, S. 256; WESTMANN 1977, S. 303). Solche diskontinuierlichen Fortbildungsprozesse können positiv gesehen werden als wechselseitig bestimmte Abfolge von Handeln und Lernen (vgl. ACKERMEIER u.a. 1978, S. 183)<sup>50</sup>.

Obwohl diese Kriterien optimaler Teilnahmebedingungen und Erfolgsaussichten von vielen Autoren aufgestellt werden, finden nicht alle Fortbildungsveranstaltungen unter ihnen statt. Die Evaluation hat hier dann weniger die Erfüllung dieser Kriterien zu überprüfen, als vielmehr die Bewertung der Veranstaltung in bezug auf solche möglichen restriktiven Bedingungen zu relativieren.

Gesicherte Aussagen zu den situativen Bedingungen des Lernens in der Fortbildungsveranstaltung finden sich kaum. Es läßt sich nur auf einige störende Drittfaktoren hinweisen: Lärm, schlechte Luft, Hitze, ungünstige Bestuhlung, schlechte Sitzverhältnisse, ungenügende Akustik, nicht erwachsenengerechte Räume behindern ein effektives Lernen (vgl. TIETGENS/WEINBERG 1971, S. 74; VONTOBEL 1972, S. 51; SCHOTT 1973, S. 76).

Erzieherfortbildung - so läßt sich zusammenfassen - sollte unter möglichst optimalen organisatorischen und situativen Bedingungen stattfinden.

---

<sup>50</sup> Auf die Möglichkeiten, neue Verhaltensweisen losgelöst von Praxisdruck auf zentralen, von der Praxis entfernten Stätten zu lernen und die Notwendigkeit, die Übertragung in die Praxis zu sichern, weisen z.B. THURNAU 1978, S. 174 und TIETGENS/WEINBERG 1971, S. 196f. hin.

### 3.2.5 Erfolgskriterien im Bereich der Produkte von Fortbildung

Erfolgskriterien, die sich auf kognitive, emotionale und praktische Lernerträge, Produkte auf Seiten der Erzieher beziehen, lassen sich ohne Berücksichtigung der Inhalte und Ziele der konkreten Fortbildungsveranstaltung kaum formulieren. In der Literatur finden sich losgelöst von den speziellen Veranstaltungen nur sehr allgemeine Forderungen. Erzieher sollen - themenbezogen - Wissen, Verhaltensweisen und Einstellungen lernen, die sie in die Alltagspraxis umsetzen können und die diese Praxis verbessern sollen. Die pädagogischen Planungs- und Entscheidungskompetenzen sowie die Fähigkeit zur Selbstorganisation von Fortbildung sollen erhöht werden. Erzieher sollen bisherige Meinungen und Haltungen im Sinne der Ziele einer Veranstaltung verändern und neue Blickpunkte und Ideen für ihre Praxis erhalten.

Die konkrete Transferierbarkeit der Produkte, des Gelernten von Fortbildung in die Praxis ist aber nicht allein der Fortbildungsveranstaltung anzulasten. Die Übertragbarkeit hängt ebenso von den konkreten Arbeitsbedingungen im Kindergarten ab, von der Struktur, in der die Inhalte erprobt werden (vgl. HEYER 1974, S. 133; KLAWE 1978, S. 65). Indem Fortbildung versuchen sollte, sich so weit wie möglich auf die Praxisbedingungen zu beziehen, verliert sie aber jede kritische Funktion, wenn sie dem Erzieher nur solche Fähigkeiten vermittelt, die ihn an die restriktiven Arbeitsbedingungen anpassen. Diesen Zusammenhang muß eine Evaluation von Fortbildung berücksichtigen.

Ein weiteres Erfolgskriterium wurde durchgehend angesprochen. Fortbildung darf dem Erzieher keine Verhaltensweise vermitteln, die der Herstellung einer intergenerationalen Reziprozität der Erzieher-Kind-Interaktion widerspricht, vielmehr sollten die Fortbildungsprozesse so

gestaltet werden, daß diese Grundnorm des Erziehverhaltens vorbereitet werden kann.

In diesem Kapitel wurde versucht, Kriterien für eine erfolgreiche Erzieherfortbildungsveranstaltung zu sammeln und zu systematisieren. Leitendes Prinzip war dabei die Vorstellung von reziproker, gleichberechtigter Kommunikation aller Fortbildungsbeteiligten. Die Analyse bezog sich nicht auf eine spezielle Veranstaltung, sondern auf eine Vielzahl von möglichen Veranstaltungen in einer gewissen Spannbreite von Zielen und Inhalten im Vorschulbereich; ebenso lehnte sie sich nicht an ein bestimmtes theoretisches Konzept an, sondern versuchte, eine Pluralität theoretischer Bezugskonzepte zu sichern.

Der zweite Schritt des mehrphasigen Modells der ziel- und inhaltsfreien Evaluation verlangt, den Kriterien-satz den potentiellen Adressaten vorzulegen und zu überprüfen, ob sich deren Ziele und Erfolgskriterien darunter befinden (vgl. S. 35f). Nun läßt sich aber die Vielzahl der angesprochenen Erfolgskriterien nicht direkt in ein Evaluationsinstrument umsetzen, welches praktikabel in begrenzt verfügbarer Zeit einsetzbar ist. Für die konkrete Evaluation einer Fortbildungsveranstaltung wird somit eine Auswahl von Erfolgskriterien nötig. Diese Auswahl wird nun - zwecks kritischer Stellungnahme - den Adressaten nicht unabhängig von dem zu entwickelnden Evaluationsinstrument vorgelegt, sondern instrumentell aufbereitet, d.h. als "Items" des Instruments. Dies ist deshalb notwendig, da letztlich die instrumentell bedingten Formulierungs- und Darbietungsschemata den Gehalt der Erfolgskriterien in der konkreten Evaluationsarbeit bedeutsam bestimmen.

Aus diesem Grund soll zuerst im folgenden Kapitel gefragt werden, wie das Evaluationsinstrument beschaffen sein muß, das die Informationen bezüglich der Erfolgskriterien erhebt.

#### 4. Evaluation durch einen Fragebogen bestehend aus Ratingskalen

Die spezifischen Arbeitsbedingungen der Erzieher verlangen von der Fortbildung eine ~~reziproke~~ Kommunikation, in der die Erzieher gleichberechtigte Partner der Veranstalter und Subjekte selbstbestimmter Lernprozesse sind (vgl. Kapitel 3.1). Evaluation soll dies unterstützen, in dem sie in einen gemeinsamen Kommunikationsprozeß "Auswertungsgespräch" eingebettet wird. Damit werden unumgängliche Ansprüche an das Evaluationsinstrument gestellt, die noch einmal kurz genannt sind (vgl. Evaluationsleitlinie 11 und ausführlich Kapitel 2.6): Mit Hilfe des Evaluationsinstruments sollen die Teilnehmer selbst ihre Erfahrungen und Wahrnehmungen der verschiedenen Aspekte einer Veranstaltung reflektieren, sammeln und einem gemeinsamen Auswertungsgespräch zuleiten können. Das Instrument muß einen kontrollierten Rückmeldeprozeß erlauben, in dem die Informationen bezüglich der Erfolgskriterien systematisch und objektiv erhoben werden; das Instrument muß schnell anwend- und auswertbar sein, damit die evaluativen Informationen rechtzeitig vor Beginn des Auswertungsgesprächs vorliegen. Diese Anforderungen sprechen für die Auswahl eines Fragebogens, dessen Items im Sinne von Ratingskalen aufgebaut sind.

##### 4.1 Vorteile von Ratingverfahren

Rating heißt übersetzt Schätzen; ein Schätzurteil ist eine systematische Schätzung des Grades, indem ein Individuum oder Objekt eine bestimmte Eigenschaft besitzt. Eine Schätzskala ist dabei ein Hilfsmittel zu dieser Beurteilung. Ratingmethoden sind somit allgemein jene Meßverfahren, die auf Einschätzung beruhen (vgl. TENT 1973, S. 13;

WEGNER 1976, S. 103). Im vorliegenden Fall wird in einem Fragebogen den teilnehmenden Erziehern eine Anzahl von Aussagen vorgelegt, die sich auf die gerade besuchte Veranstaltung beziehen. Die Teilnehmer sollen dann bezüglich jeder Aussage einschätzen, wo auf einer beigefügten Skala die Veranstaltung einzuordnen ist<sup>1</sup>.

Durch dieses Vorgehen sammeln die Teilnehmer selbst ihre subjektiven Beurteilungen der unterschiedlichen Aspekte der Fortbildungsveranstaltung. Es wird erwartet, daß dieses Sammeln der Beurteilungen mit Hilfe von Rating-skalen für die Teilnehmer ein zur Vorbereitung von Auswertungsgesprächen einsichtiges Verfahren darstellt, da Ratings dem Alltagshandeln nahe kommen. Sie schließen sich unmittelbar an die Erlebnisrealität von Menschen an. Ohne es direkt zu bemerken, "ratet" sich jeder Mensch durchs Leben, z.B. im Straßenverkehr beim Abschätzen von Entfernungsgeschwindigkeitsverhältnissen oder im Sozialverhalten bei der Abschätzung der Vertrauenswürdigkeit des Partners (vgl. LANGER/SCHULZ von THUN 1974, S. 22). Aufgabe bei der Veranstaltungsevaluation ist es aber

<sup>1</sup> Schätz- oder Ratingmethoden gehören zu den direkten Skalierungsmethoden (vgl. PETZOLD 1978, S. 31). In der Terminologie von WEGNER handelt es sich bei der Einschätzung von Aspekten der Fortbildung um die Methode des Einzelreiz-Ratings, in dem ein Objekt - die Veranstaltung - mit einer bzw. mehreren Skalen zusammen dargeboten wird (im Unterschied zu Mehrfachreizmethoden, bei denen mehrere Objekte zugleich dargeboten werden). Unterscheidet man Fremd-, Selbst- und evaluative Ratings, so ist das hier zu entwickelnde Verfahren eine Mischform, in dem die Teilnehmer das Objekt "Fortbildungsveranstaltung" (Fremd), eigene Befindlichkeiten (Selbst) einschätzen sowie zu den Objekten eine bewertende Stellung einnehmen sollen (vgl. WEGNER 1976, S. 104). Es interessieren dabei sowohl die individuellen Differenzen der Einschätzungen der Teilnehmer (subjektzentriertes Vorgehen) als auch besonders der Anteil der Variabilität, der auf die Veranstaltung, also auf die Unterschiede der Reize (reizzentriertes Vorgehen, Vpn als Replikationen) zurückgeht (vgl. ib., S. 107).

Zu einer allgemeinen Einführung in die Methodik von Ratingsverfahren siehe u.a.: LANGER/SCHULZ von THUN 1974; PETZOLD 1978, bes. S. 32-35; SCHMIDT 1965, 1966; TENT 1973, WEGNER 1976.

mögliche Verzerrungen der Ratings im Alltagshandeln zu kontrollieren bzw. zu vermindern. Ratings in einem erprobten und überprüften Fragebogen bieten einen kontrollierten Rückmeldeprozeß. Sie kanalisieren und systematisieren die Aufmerksamkeit der Teilnehmer auf bestimmte, ausgewählte und begründete Aspekte; sie methodisieren die Erfahrungssammlung, machen die Ergebnisse vergleichbar und bieten so eine ausgezeichnete Grundlage für den anschließenden gemeinsamen Kommunikationsprozeß "Auswertungsgespräch" zwischen den Beteiligten. Die Objektivität und allgemeine Güte dieser Ratings sind mit empirischen Methoden überprüfbar.

Spezifischer lassen sich die Vorteile eines Fragebogens bestehend aus Ratingskalen in den folgenden, sich teils überschneidenden Aspekten begründen:

- Ein Fragebogen mit Ratingskalen spricht die Teilnehmer als professionalisierte Fachleute ihrer Fortbildung und unmittelbar von der Veranstaltung Betroffene an. Gleichberechtigte Partnerschaft bedeutet, daß ihre Meinungen und Einschätzungen ernst genommen werden müssen und eine primäre und relevante Quelle für die Evaluation der Veranstaltung darstellen. Dies gilt auch unabhängig von dem hier angestrebten gemeinsamen und gleichberechtigten Auswertungskontext. Für den Träger von Fortbildung bieten die Teilnehmereinschätzungen immer eine relevante und ernstzunehmende Evaluationsquelle (vgl. allg. ALEMONI/SPENCER 1973, S. 669; HARARI/ZEDECK 1973, S. 265; DERRY/BRANDENBURG 1978, S. 772; OLIVERO 1973, S. 29).

- Die erhobenen subjektiven Teilnehmereinschätzungen sind eine wichtige Variable im gemeinsamen Lern- und Arbeitsprozeß. Sie bestimmen - u.U. unabhängig von dem objektiven Lernerfolg - wie zukünftiges Lernen abläuft, indem



z.B. Unzufriedenheit weitere Mitarbeit vermindert (vgl. SCHINDLER 1979, S. 39)<sup>2</sup>.

- Prinzipiell ließe sich jeder im Fragebogen angesprochene Aspekt auch im persönlichen Gespräch erheben. Wollte ein Fortbildungsleiter hier aber die Vielfalt der Aspekte ansprechen, so müßten die Auswertungsgespräche so umfangreich werden, daß die inhaltliche Arbeit gestört würde. Demgegenüber nimmt die Benutzung eines Fragebogens mit Ratingskalen weniger Zeit in Anspruch, Ratingskalen sind schnell anwend- und auswertbar; alle Teilnehmer haben die Möglichkeit zu antworten und es werden sowohl Hauptströmungen der Teilnehmereinschätzung als auch Minderheitsvoten erfaßt. Ratings liefern so eine objektive Rückmeldung über die Fortbildung, unverzerrt durch bewußte oder unbewußte Beeinflussung durch den Fortbildungsleiter im persönlichen Gespräch<sup>3</sup>.

- Steht die Erfassung der objektiven Veranstaltungsrealität im Vordergrund, so ist man ebenfalls auf Ratings und Teilnehmereinschätzungen angewiesen. Eigentlich sol-

---

<sup>2</sup> Die subjektiven Teilnehmereinschätzungen werden hier nicht als Indikator für die objektive Veranstaltungsrealität gesehen, sondern die Teilnehmermeinungen sind die zu messende, interessierende Variable selbst. Bezüglich der Validität von Ratings sprechen LANGER/SCHULZ von THUN hier von interner Ziel-Güte im Unterschied zu externer Zielgüte (vgl. LANGER/SCHULZ von THUN 1974, S. 108). Hier wird z.B. der Halo-Effekt - s.w.u. - nicht als ein verzerrender Fehler zu diskutieren sein, sondern als legitime Organisation der Teilnehmerwahrnehmung (vgl. DERRY/BRANDENBURG 1978, S. 777; TENT 1973, S. 62), die das weitere Lernen beeinflusst.

<sup>3</sup> In einer handlungsorientierten Evaluation des Modellversuchs "Soziale Studiengänge" an der Gesamthochschule Kassel verglichen BRAUNER/HERING/OELSCHLÄGEL einen informellen Gesprächsleitfaden und einen Fragebogen. Der Leitfaden war zwar leichter zu handhaben, der Dozent hatte aber einen belegbaren Einfluß auf die Ergebnisse. Demgegenüber war der Fragebogen aufwendiger und gab weniger Gelegenheit zu spontanen Reaktionen, sicherte aber gerade abweichende Minderheitsvoten im Seminar (vgl. BRAUNER/HERING/OELSCHLÄGEL 1977, S. 80).

len dann aber keine subjektiven Eindrücke erhoben werden, sondern es werden objektive Aussagen angestrebt. Mangels anderer, besserer Verfahren muß der Umweg über subjektive Beurteilungen gemacht werden, von denen man hofft bzw. aufzuzeigen hat, daß sie valide Indikatoren sind (vgl. SCHMIDT 1965, S. 90)<sup>4</sup>. Ratings werden aber auch dadurch unumgänglich, daß die Variablen einer Fortbildungsveranstaltung sehr komplex sind. Die Operationalisierung von z.B. "Strukturierung der Veranstaltung" in einzelne Indikatoren dürfte sehr schwer und nur auf dem Hintergrund eines spezifischen, methodisch-didaktischen Konzeptes möglich sein, wobei die Fortbildungsveranstaltung dann auf die Einhaltung eines solchen Konzeptes verpflichtet würde. Die Teilnehmer haben aber eine Vorstellung von "Strukturierung" und es scheint möglich, die Leistung der Aufgliederung und Erfassung dieser Variable an die Verarbeitungskapazitäten ihrer Gehirne zurückzugeben, die eine Vielzahl von Einzelindikatoren verschmelzen können und so in der Lage sind, die komplexen Merkmale einzuschätzen (vgl. LANGER/SCHULZ von THUN 1974, S. 16-22)<sup>5</sup>.

---

<sup>4</sup> Für die Bewertung von Professoren hat SCRIVEN vorgeschlagen, daß sie bewertet werden sollen durch die Ergebnisse, die sie hervorrufen, durch ihre Kollegen und durch ihre Studenten. Die Bewertung über Ergebnisse ist einmal sehr aufwendig und steht zum anderen vor methodologischen Problemen. Die Bewertung durch Kollegen ist schwierig, weil Veranstaltungen hinter verschlossenen Türen stattfinden und weil die Beurteilung wohl oft nur das Ergebnis eines Halo-Effektes ist. Aus diesem Grunde - so JAEGER/FREIJO - sei man auf die studentische Einschätzung der Kompetenz eines Professors als Lehrer angewiesen (vgl. JAEGER/FREIJO 1974, S. 416).

<sup>5</sup> Damit soll nicht der Versuch, abstrakte und komplexe Begriffe durch Indikatoren zu konkretisieren, abgebrochen werden, vielmehr soll - nach LANGER und SCHULZ von THUN - der Prozeß der Indikatorenbildung auf einer mittleren Abstraktionsebene gestoppt und menschliche Ratings eingesetzt werden (vgl. LANGER/SCHULZ von THUN 1974, S. 21).

- Ein Fragebogen aus Ratingskalen schließt sich an bereits verwendete, aber wenig formell entwickelte und abgesicherte Evaluationsinstrumente der Erzieherfortbildung an (vgl. z.B. VERON o.J.; ROSSBACH 1977, S. 136-165). Entsprechende Skalen werden u.a. auch in der Weiterbildung in der Industrie (vgl. FASSBENDER 1974) und vor allem in der Lehrerbeurteilung an Schulen und Hochschulen in den USA eingesetzt. Verschiedene Forschungsarbeiten haben allgemein aufgezeigt, daß solche Beurteilungsskalen zweckmäßige, reliable und valide Instrumente zur Erfassung der Qualität des Unterrichtes sind und auch zur Selbstkontrolle und Vervollkommenung eines Lehrers dienen können (vgl. z.B. die Sammelreferate von COSTIN u.a. 1971; TENT 1973, S. 65f.)<sup>6</sup>. Von den ausfüllenden Studenten

---

<sup>6</sup> Andererseits erheben sich auch kritische Stimmen, vor allem auf dem Hintergrund des Standes der Unterrichtsforschung. So diskutieren McNEIL /POPHAM in einem Sammelreferat über "The Assessment of Teacher Competence" ausführlich Prozeßkriterien. Sie sehen es als ein Problem an, wenn Lehrer aufgrund von Kriterien der Lernprozesse, d.h. der Verwendung bestimmter unterrichtlicher Strategien beurteilt werden, denn es sei unwahrscheinlich, daß ein bestimmtes Merkmal des Lernprozesses fest mit einem Leistungszuwachs beim Schüler verbunden sei (vgl. McNEIL/POPHAM 1973, S. 220). Es gibt zwar - im Bereich der Lehrereinschätzung - eine Überfülle von Studien, die Verbindungen zwischen Prozeßkriterien und Lernergebnissen untersuchen, jedoch wären diese zumeist als Korrelationsstudien durchgeführt worden und nicht als experimentelle Studien. Gesetzmäßige Beziehungen sind heute noch nicht entdeckt, bloß einige Hinweise oder Praktiken, die wahrscheinlicher als andere mit gewissen Ergebnissen verbunden sind (vgl. ib., S. 223f.; s.a. ROSENSHINE/FURST 1971, S. 39-42 und S. 55ff.; FRICKE 1978; zur Verbesserung der Prozeß-Produkt-Forschung durch Berücksichtigung von "Aptitude-Treatment-Interactions" siehe GARDNER 1974).

Effektives Unterrichten läßt sich so nicht durch An- oder Abwesenheit bestimmter Merkmale des Lernprozesses nachweisen, sondern nur durch den Lernzuwachs bei den Schülern. Dies gilt vor allem, wenn - wie in den USA - es um die Einstellung oder Entlassung von Lehrern geht.

werden solche Bögen als effektives Evaluationsinstrument gesehen, ihrer Wahrnehmung nach berücksichtigen viele Lehrer jedoch die Ergebnisse zu wenig und ändern evtl. Schwächen nicht ab (vgl. PENFIELD 1978, S. 21). Dieser potentiellen Gefahr, daß Ratings nur ein Ritual zum Abschluß einer Veranstaltung darstellen und ansonsten die Ergebnisse in "Schubladen" lagern, entgeht die Evaluation dann, wenn die Ergebnisse nicht nur an die Lehrer zurückgemeldet werden, sondern an ein gemeinsames Auswertungsgespräch (vgl. für die Evaluation von Hochschulveranstaltungen in der BRD DIEHL/KOHR 1977, S. 62)<sup>7</sup>.

---

noch 6 Ist man aber an der Verbesserung des Lehrens interessiert, so sind Beobachtungen und Analysen der Lernprozesse hilfreich (vgl. McNEIL/POPHAM 1973, S. 230). Hier können dann auch "student ratings" erfolgreich eingesetzt werden (vgl. ib., S. 231; MASTERS 1979, S. 222f.). Aussagen über die Anwesenheit bestimmter Prozeßmerkmale müssen also mit Vorsicht betrachtet werden. Sie haben eher - bezogen auf wünschenswerte Lernergebnisse - den Status von Hypothesen, über die im Einzelfall mit den Betroffenen diskutiert werden muß.

<sup>7</sup> DIEHL und KOHR entwickeln hier einen Fragebogen für Veranstaltungen im Fach "Psychologie". Die Übertragung dieses Bogens für erziehungswissenschaftliche Veranstaltungen scheint nur eingeschränkt möglich (KLEINE/MERKENS 1979). Besonders interessant ist auch der Versuch von PETERMANN und SOMMER, der besonders die Forderung nach Eigensteuerung des Lernens durch die Studenten aufnimmt (vgl. SOMMER/PETERMANN 1978; PETERMANN 1979). Die Studenten werden nicht nur zur Einschätzung des Ist-Zustandes bezüglich einer Liste von Aussagen aufgefordert, sondern sie sollen zugleich den Soll-Zustand angeben und die Relevanz, die sie einer Aussage bei der Bewertung einer Veranstaltung beimessen. Die Berücksichtigung aller 3 Informationen führt zu einer klareren Struktur bei der Faktorenanalyse. Indem der Befragte seine eigenen Bezugsnormen mitteilt (Relevanzurteile) wird die Befragungssituation zumindest ein Schritt in Richtung einer Subjekt-Subjekt Beziehung verändert (vgl. SOMMER/PETERMANN 1978, S. 344). Dieser Ansatz sollte weiter verfolgt werden.

## 4.2 Zur Gestaltung des Fragebogens

Die in Kapitel 3 diskutierten Erfolgskriterien aus den Bereichen Thematik, Methodik, Interaktionsprozesse und situative/organisatorische Bedingungen sollen in einen Fragebogen umgesetzt werden, der mit Hilfe von Ratings die Auswertungsinformationen erhebt. Zu der Frage des Aufbaus und der Gestaltung solcher Fragebögen liegen relativ wenig gesicherte und eindeutige Forschungsergebnisse vor; oft beruhen Gestaltungs- und Konstruktionsempfehlungen auf Erfahrungen und/oder dem "gesunden Menschenverstand". Aus diesem Grund sollen einige wichtige Gestaltungsprobleme im Hinblick auf den hier zu konstruierenden Fragebogen vorab ausführlicher diskutiert werden. Ausgangspunkt für den Aufbau des Fragebogens sind dabei wiederum die allgemeinen Überlegungen aus dem skizzierten Evaluationsmodell (vgl. Kapitel 2.6):

Die teilnehmenden Erzieher werden in der Evaluation als gleichberechtigte Partner einer gemeinsamen Arbeit angesehen. Dies impliziert für die Einführung des Fragebogens in die Veranstaltung, die Fragebogeninstruktion und die Fragebogengestaltung ein offenes Vorgehen. Die Erzieher sind keine Objekte, aus denen mit Tricks etwas herausgeholt werden soll, was sie möglichst nicht verstehen sollen. Sie sind vielmehr Experten, deren Mitarbeit zur Evaluation nötig ist; sie sollen ständig wissen, wofür und warum sie eine Frage beantworten (vgl. auch KREUTZ/TITSCHER 1974, S. 65).

### - Formulierung des Items

Die Formulierung des einzelnen Items des Fragebogens soll sich den allgemeinen Grundregeln anschließen (zur Verdeutlichung der Gestaltungsentscheidungen vgl. Anlage 1 und 16). Speziell soll eine klare, verständliche

Sprache angestrebt werden, die in der Wortwahl der Adressatengruppe entspricht; die Items sollen eindeutig und nicht einseitig formuliert werden; bevorzugt werden kürzere Items, die aufgrund der konkreten Situationen in der Fortbildungsveranstaltung beantwortbar sind; die Items sollen möglichst nach speziellen Aspekten fragen, die nicht per se in allen Veranstaltungen gleich eingeschätzt werden (vgl. allgemein FASSBENDER 1974, S. 132-137; KREUTZ/TITSCHER 1974, S. 53-61; KIRSCHHOFER-BOZENHARDT/KAPLITZA 1975, S. 97f.). Obwohl mit offenen Fragen ein sehr reiches Material an Informationen über die Veranstaltung gewonnen werden könnte, wird hier der Vorzug einem System von geschlossenen Antwortvorgaben gegeben. Damit gehen zwar spezifische inhaltliche Informationen verloren, aber eine Vielzahl von Veranstaltungen - vor allem im Erprobungsstadium - ist schnell auswertbar und einer differenzierten statistischen Analyse zugänglich. Entscheidend ist aber, daß bei geschlossenen Antwortvorgaben innerhalb einer Fortbildungsveranstaltung die Teilnehmer/Referenten/Leiter den Fragebogen um ein Vielfaches schneller auswerten können, als dies bei "offenen" Fragen möglich ist. Im fertigen Evaluationsbogen sollen die Teilnehmer zusätzlich die Möglichkeit haben - und auch dazu aufgefordert werden -, ihre Auswahl einer Antwortkategorie zu begründen; dies sollte nicht nur bei negativen Urteilen geschehen (vgl. FASSBENDER 1974, S. 137). Da aber nicht immer frei formulierte Begründungen von Teilnehmern zu erwarten sind, sichert das "Ankreuzen" einer Vorgabe dennoch eine Antwort auf die einzelnen Items.

#### - Form der Antwortskalen

In der 1. Erprobungsphase wird ein zweipoliger Skalen-  
aufbau vorgesehen. Den Teilnehmern wird dabei zu jedem  
Aspekt ein Satzanfang mit zwei möglichen, gegensätzlichen  
Ergänzungen vorgelegt. Zwischen diesen beiden Ergänzungen

befindet sich eine Zustimmungsskala von 1-6<sup>8</sup>. Tiefgreifende Arbeiten zur Konstruktion von Fragebögen fehlen noch weitgehend, viele Gestaltungsempfehlungen beruhen - wie erwähnt - auf Erfahrungen bzw. dem "gesunden Menschenverstand" (vgl. KREUTZ/TITSCHER 1974, S. 24; TENT 1973, S. 19f.). Andere Untersuchungen zeigen nur äußerst geringe Unterschiede auf<sup>9</sup>, so daß die Auswahl der zweipoligen Skalenform - in der Hoffnung, so den Teilnehmern eine Einschätzung zu erleichtern - sich in der Erprobung erst noch bewähren muß. Konzentriert werden soll sich dabei in der Erprobungsphase auf die Kontrolle von Verständnisschwierigkeiten, die in diesem Zusammenhang wichtiger ist als eine formale Kennwertanalyse (vgl. auch SPRUNG/SPRUNG 1978a, S. 14). Es wird ferner vermutet, daß eine 6er Skala der Zustimmung dem Differenzierungsvermögen der Teilnehmer entspricht, d.h. 6 Abstufungen messen so fein wie möglich, überfordern aber die Differenzierungsfähigkeit der Teilnehmer noch nicht (vgl. HOLM 1975, S. 49). Das Fehlen einer "neutralen" Position in der Mitte - wie z.B. bei einer 7er Skala die 4 - verlangt vom Teilnehmer, sich zumindest tendenziell

<sup>8</sup> Bezogen auf Klassifikationen von Ratingskalen nach der äußeren Form handelt es sich um ein numerisches Rating (vgl. SCHMIDT 1965, S. 17-28; LANGER/SCHULZ von THUN 1974, S. 103f.).

<sup>9</sup> Hierzu einige Untersuchungen: Formulierungen von Items - komplexe Sätze vs. einfache Reize; implizite Normorientierung vs. explizite Normorientierung (Vergleich mit anderen Veranstaltungen); asymmetrische (mehr positive Antwortmöglichkeiten) vs. symmetrische Skalen - bewirken keine größeren Differenzen in der Rangordnung von Professoren (vgl. JAEGER/FREIJO 1974, S. 419f.). SCHMIDT testete, ob verschiedene Skalenformen die Zuverlässigkeit - als Schätzung der Abwesenheit des Halo-Effektes - beeinflussen. Es zeigten sich nur äußerst geringe Unterschiede, ob nun schriftliche oder mündliche Instruktion, graphische oder numerische Skalen verglichen wurden; ebenso verändert der explizite Hinweis auf Normalverteilung die Einschätzung nicht (vgl. SCHMIDT 1965, S. 81-83; 1966, S. 60-69). Einfachheit, Übersichtlichkeit, "natürliche" Abstufungsfolge (z.B. bei numerischen Ratings die hohen Zahlen rechts) vorausgesetzt, haben unterschiedliche Gestaltungen nur einen geringen Einfluß auf die Güte von Ratings (vgl.

für eine Seite der Skala zu entscheiden.

Eindeutige Forschungsergebnisse fehlen auch zu der Frage, ob das positive Ende der Antwortskala - hier diejenige Satzergänzung, die eine für die Veranstaltung positive Einschätzung beinhaltet - immer auf der gleichen Seite oder nach Zufall wechselnd auf beiden Seiten stehen soll. REMMERS stellte keine Unterschiede fest (vgl. TENT 1973, S. 20f.), während LANDSHEERE daraufhinweist, daß Antwortstereotype und der Halo-Effekt, der hier - s.w.u. - als besonders verzerrend angesehen wird, unter anderem durch Skalen verringert werden können, bei denen das "gute" Ende abwechselnd auf beiden Seiten steht (vgl. LANDSHEERE 1969, S. 80f.). Um mögliche Antwortstereotypen - z.B. ein Festlegen auf einer Seite der Skala über alle Fragen - zu vermindern, soll hier ein abwechselnder Modus ausgewählt werden.

Desweiteren sollen alle Items so formuliert werden, daß die Ausprägungen linear mit dem Wertbezug kovariieren, d.h. die für eine Veranstaltung positive Ausprägung, der Soll-Wert, muß jeweils an einem Endpunkt der Skala liegen. Ein Beispiel: Je größer die Praxisrelevanz einer Veranstaltung, desto positiver ist sie einzuschätzen<sup>10</sup>.

<sup>10</sup> Auf andere Kriterienvariablen weisen KEIL/PIONTKOWSKI hin, es sind dies Variablen,  
- bei denen der Wertbezug nicht direkt und linear mit der Ausprägungsintensität kovariiert, z.B. bei mittlerem Schwierigkeitsgrad ist Unterricht am besten,  
- bei denen je nach Zeit eine positive oder negative Ausprägung einen optimalen Wert hat, z.B. emotionale Zufriedenheit von lernenden Adressaten sollte je nach Unterrichtsphase hoch oder niedrig sein (vgl. KEIL/PIONTKOWSKI 1973, S. 281f.).

Im vorliegenden Fall wird also unterstellt, daß die Soll-Werte jeweils an dem Endpunkt der Ratingskalen liegen. SOMMER und PETERMANN kritisieren dies und erheben deshalb bei den befragten Studenten neben dem Ist-Wert auch den Soll-Wert (vgl. SOMMER/PETERMANN 1978, S. 337).



### - Dramaturgie des Fragebogens und Reihenfolge der Fragen

Ein weiteres zentrales Problem der Fragebogenkonstruktion - zu dem ebenfalls sehr wenig gesichertes Wissen vorhanden ist - betrifft die Dramaturgie des Bogens und die Reihenfolge der Fragen, vor allem angesichts möglicher Ausstrahlungseffekte früherer Fragen auf spätere (vgl. allgemein KREUTZ/TITSCHER 1974, S. 40ff.; KIRSCHHOFFER-BOZENHARDT/KAPLITZA 1975, S. 93-95)<sup>11</sup>. Der Einfachheit und Schnelligkeit der Auswertung wegen sollen alle Items in einem gleichartigen Schema dargeboten werden; mögliche Ermüdungserscheinungen werden wegen der relativen Kürze des Fragebogens kaum erwartet. Da vor der Erprobung nicht eindeutig feststeht, wie sich die Items in der Wahrnehmung der beteiligten Erzieher gruppieren, wird auf eine Darbietung in thematischen Blocks verzichtet und die Items stattdessen nach dem Zufall gemischt angeordnet. Fragen mit rein psychologischen Funktionen - z.B. Puffer- oder Ablenkungsfragen -, die die Teilnehmer möglichst unbemerkt lenken sollen, werden gemäß der allgemeinen Vorstellung der Fragebogengestaltung ausgeschlossen.

### - Einführung des Fragebogens

Gleichberechtigte Partnerschaft setzt für die Evaluation voraus, daß die Teilnehmer nicht erst zu Ende einer Ver-

<sup>11</sup> Empfohlen werden hier Fragen mit psychologischen Funktionen (z.B. Einleitungs-, Spiel-, Auslöschungs-, Puffer- oder Ablenkungsfragen) und ein Wechsel von Spannung und Entspannung, von offenen und geschlossenen, monotonen und interessanten, schweren und leichten Fragen sowie ein dosierter, nicht sprunghafter Wechsel der Themen und Fragetechniken. Der hier entwickelte Fragebogen ähnelt mehr einer Fragebatterie, in der die Fragen in einem gleichartigen Schema gestellt und nach Zufall gemischt werden, sowie die negative und positive Seite wechselt (vgl. KREUTZ/TITSCHER 1974, S. 47f.); damit entspricht der Bogen den meisten in der Lehrereinschätzung verwendeten Verfahren, siehe z.B. TENT 1973; ALEAMONI/SPENCER 1973; FRENCH-LAZOVIK 1973; FREIJO 1974; DIEHL/KOHR 1977.

anstellung mit einem Fragebogen überfahren werden, sondern daß über den späteren Einsatz und den Auswertungskontext schon zu Beginn der Veranstaltung mit den Teilnehmern eine Vereinbarung getroffen wird (vgl. auch FASSBENDER 1974, S. 133). Dies gilt ebenso für das Entwicklungs- und Erprobungsstadium. In der Fragebogeneinführung und -instruktion soll noch einmal motivationsstärkend der Sinn und Zweck der Befragung erläutert und auf die Rolle der Teilnehmer als Experten, die eine Veranstaltung mit am besten einschätzen können, und die Konsequenzen, die das Ausfüllen für die Teilnehmer hat, eingegangen werden. Im Erprobungsstadium bezieht sich dies vor allem darauf, daß die möglichen Verbesserungen nicht den Teilnehmer selbst direkt, sondern eher ihren Fortbildungsnachfolgern zu gute kommen. Es ist ebenfalls darauf zu achten, daß der Aufbau des Bogens und die einzelnen Items der Erwartungshaltung entsprechen, die bei den Teilnehmern durch die Einführung und Instruktion erzeugt wird (vgl. KREUTZ/TITSCHER 1974, S. 43).

Um unbewußte Beeinflussung bei der Abgabe des eigenen Urteils über die Veranstaltung bzw. Aspekte der Veranstaltung - z.B. durch andere Teilnehmer, durch Fortbildungsleiter oder -referenten, durch Ängste bei namentlicher Kennzeichnung des Bogens - zu verhindern, sollen die Teilnehmer aufgefordert werden, den Fragebogen allein und unabhängig zu bearbeiten. Desgleichen soll ihnen Anonymität zugesichert werden.<sup>12</sup> Angestrebtes Ziel ist ein objektives, unverzerrtes Feed-back über die Veranstaltung.

<sup>12</sup> STONE u.a. belegen in einer Studie die verzerrenden Wirkungen einer namentlichen Kennzeichnung von Auswertungsbögen; in einer zweiten Studie erwies sich dieser Effekt aber als gering (vgl. STONE/RABINOWITZ/SPOOL 1977; auch ALEAMONI/SPENCER 1973, S. 680). In einer Studentenforschung über die Arbeit mit Ratingskalen äußerten 46 % der befragten Studenten, daß sie sich nicht durch die anderen Klassenmitglieder in ihrem Urteil beeinflussen lassen würden (vgl. COSTIN u.a. 1971, S. 523f.).

### 4.3 Mögliche Verzerrungen bei der Verwendung von Ratingskalen

Obwohl Ratingskalen in der pädagogischen Forschung sehr häufig und breit verwendet werden, wird ihnen eine große methodologische Skepsis entgegengebracht (vgl. z.B. LANGER/SCHULZ von THUN 1974, S. 10; SCHMIDT 1966, S. 47). Schätzurteile können mit teilweise erheblichen Fehlern und Verzerrungen behaftet sein. Diskutiert werden diese entweder nach dem teststatistischen Modell als mangelnde Objektivität, Reliabilität und Validität oder anhand von additiven Zusammenstellungen systematischer Fehler (vgl. KLEITER 1973b, S. 100). Systematische Fehler oder "constant errors" sind dabei solche, die sich in vielen Experimenten und bei vielen Gelegenheiten immer wieder beobachten lassen, d.h. Abweichungen vom "wahren" Wert, die sich nicht durch eine genügende Anzahl von Meßwiederholungen aufheben (vgl. SCHMIDT 1965, S. 28; WEGNER 1976, S. 113). Aufgrund einer letztlich mangelnden Systematik solcher Fehler und Verzerrungen<sup>13</sup> soll im folgenden eine

<sup>13</sup> Allgemeine Überblicke über Fehler finden sich bei SCHMIDT 1965, S. 28-35; COSTIN u.a. 1971, S. 512-522; TENT 1973, S. 61-63; WEGNER 1976, S. 113-117; SPRUNG/SPRUNG 1978b, S. 8f. und S. 17f. Diskutiert werden: Halo-Effekt (Hof-Effekt), logischer Fehler (logical error), error of leniency (Milde Fehler), error of central tendency (Fehler der zentralen Tendenz), contrast error (Kontrasteffekte), proximity error (Fehler der Nähe), Tendenz zur sozialen Erwünschtheit, Tendenz zur Zustimmung, Extremneigungen und differentieller Selektionseffekt. Die Unterscheidung zwischen Organisationsfehler (z.B. logischer Fehler) und Einstellungsfehler (z.B. Milde-Effekt) (vgl. SCHMIDT 1965, S. 29) systematisiert diese Aufstellung nur wenig.

Im Zusammenhang mit Fehlern beim Lehrerurteil über Persönlichkeitsmerkmale bei Schülern kritisiert KLEITER eingehend die mangelnde Systematisierung von Fehlern (vgl. KLEITER 1973a, b). Ausgehend von einer wissenschaftstheoretischen Einteilung in klassifikatorische, komparative und quantitative Begriffe unterscheidet er die Arten

- Kategorienfehler mit den Unterarten Interaktions- und Korrelationsfehlern

- Referenzfehler (ib., 1973a, S. 197-202; 1973b, S. 102-105).

Auswahl von für die Evaluation von Erzieherfortbildung besonders bedeutsamen Verzerrungen und entsprechende Kontrollmöglichkeiten diskutiert werden. Es handelt sich hier besonders um den Halo-Effekt, der als größte Fehlerquelle bei der Verwendung von Ratings betrachtet wird. Daneben werden - als Bedingungen für alle empirischen Instrumente - Reliabilität und in einem breiten Verständnis Validität angesprochen.

#### - Halo-Effekt

Ein Fragebogen liefert nur dann befriedigende Informationen für ein Auswertungsgespräch über Verbesserungsmöglichkeiten für ein Auswertungsgespräch über Verbesserungsmöglichkeiten 13 Der Begriff "Fehler" wird dabei verstanden als Abweichung der impliziten Persönlichkeitstheorie (IPT) von der expliziten Persönlichkeitstheorie (EPT) (ib. 1973b, S. 101). Diese Fehlereinteilung ließe sich prinzipiell auch auf die Veranstaltungsbeurteilung übertragen, obwohl dies dadurch erschwert würde, daß nicht ein Lehrer/ eine Lehrergruppe viele Schüler einschätzen, sondern viele Teilnehmer eine Veranstaltung und die Teilnehmer bei verschiedenen Veranstaltungen nicht die gleichen sind. In dem vorliegenden Rahmen können diese Fehler jedoch nicht untersucht werden: Zur Feststellung des Ausmaßes von Korrelations- und Interaktionsfehlern benötigt KLEITER eine konvergente, parallele Operationalisierung, d.h. für jedes Konstrukt, für jeden Aspekt einer Veranstaltung wird ein objektives Testitem und entsprechendes Schätzitem verlangt; die entstehenden Korrelationen werden in einer umfassenden Matrix organisiert (vgl. 1973a, S. 209ff.). Objektive Tests zur Evaluation von Veranstaltungsaspekten sind jedoch nicht vorhanden, so daß diese Überprüfungsmöglichkeit ausfällt.

Eine varianzanalytische Systematisierung und Überprüfbarkeit von Fehlern (vgl. 73a, S. 200f.) ist hier ebenfalls nicht möglich. Der umfangreiche 3 faktorielle Versuchsplan ohne Wiederholung verlangt, daß mehrere Lehrer mehrere (die gleichen) Schüler in mehreren Items einschätzen. Im vorliegenden Fall ist aber bestenfalls nur ein zweifaktorieller Versuchsplan ohne Wiederholung möglich, d.h. mehrere Teilnehmer schätzen eine Veranstaltung ein, wobei verschiedene Veranstaltungen auch von verschiedenen Teilnehmern besucht werden. Vernachlässigt man die Gruppierung der Teilnehmer in "Veranstaltungen" und betrachtet nur, daß viele Pbn mehrere Items einschätzen, so ließe sich mit diesem Versuchsplan nur die Reliabilität nach der Methode der Konsistenzanalyse einschätzen (vgl. LIENERT 1969, S. 228ff.).

lichkeiten, wenn er nicht nur einen globalen Eindruck von der Veranstaltung erhebt, sondern spezifische Informationen in bestimmten Bereichen (vgl. MASTERS 1979, S. 222f.). Die Kontrolle des Halo-Effektes wird in diesem Zusammenhang besonders wichtig. Der Halo-Effekt - eine der größten Fehlerquellen bei Ratings - besteht in der Tendenz der Beurteiler, sämtliche Einzelbeurteilungen einem globalen, allgemeinen und ganzheitlichen Eindruck von dem einzuschätzenden Objekt unterzuordnen (vgl. SCHMIDT 1965, S. 30). Bei der Veranstaltungsbeurteilung würden dann Teilnehmer - die z.B. einen allgemeinen guten Eindruck hatten - alle Aspekte gleichförmig positiv einschätzen und somit nicht mehr zwischen den einzelnen Aspekten unterscheiden. Läßt sich dieser Halo-Effekt zwar einerseits als legitime Organisation der Teilnehmerwahrnehmung betrachten (vgl. DERRY/TURNER-ENBURG 1978, S. 777), so bieten jedoch andererseits durch diesen Fehler verzerrte Informationen einen ungenügenden Input für ein gemeinsames Auswertungsgespräch. Im Bereich der Beurteilung von Lehrveranstaltungen gibt es Hinweise, daß Studenten zuverlässig zwischen verschiedenen Aspekten unterscheiden können und somit ein Halo-Effekt vermutlich recht niedrig ausfällt (vgl. TENT 1973, S. 66). Inwieweit sich ein solcher Halo-Effekt in dem vorliegenden Fragebogen auswirkt, wird eher auf indirektem Wege mittels der folgenden multivarianten Verfahren überprüft werden<sup>14</sup>:

Ein Halo-Effekt führt zu einer unechten Erhöhung der Item-Interkorrelation (vgl. SCHMIDT 1965, S. 30). Werden dann

<sup>14</sup> SCHMIDT schlägt eine varianzanalytische Technik zur Kontrolle des Halo-Effektes vor. Der Halo-Effekt wird demnach bestimmt über die Wechselwirkungsvarianz zwischen Beurteilern und Objekten (vgl. SCHMIDT 1965, S. 45f., sowie die Experimente S. 48-81; SCHMIDT 1966, S. 52ff.). Da diese Technik voraussetzt, daß mehrere Beurteiler mehrere (die gleichen) Objekte einschätzen, kann sie hier nicht zur Kontrolle des Ausmaßes des Halo-Effektes benutzt werden.

die Items faktorisiert<sup>15</sup>, so müßte ein allgemeiner Faktor zu extrahieren sein, auf dem alle Items höher laden. Entsprechend zeigt eine komplexe Faktorenstruktur zugleich einen geringen Halo-Effekt an (vgl. JAEGER/FREIJO 1974, S. 421), da diese nur entsteht, wenn die Item-Interkorrelationen genügend streuen.

Eine weitere Möglichkeit, den Halo-Effekt zu kontrollieren, findet sich in Zusammenhang mit Regressionsanalysen (vgl. FRENCH-LAZOVIK 1974, S. 382)<sup>16</sup>. In die Items des Fragebogens soll eine Globaleinschätzung der Veranstaltung mit aufgenommen werden. Diese wird als Kriterium betrachtet, während die anderen Items die Prädiktoren bilden. Es läßt sich dann die Squared Multiple Correlation (SMC) als Maß

---

<sup>15</sup> Fast alle Untersuchungen zu entsprechenden Fragebögen führen eine Faktorenanalyse über alle befragten Personen durch, ohne zu berücksichtigen, daß diese Personen unterschiedliche Veranstaltungen besucht haben (vgl. z.B. ALEAMONI/SPENCER 1973; DIEHL/KOHR 1977; DERRY/BRANDENBURG 1978). Veränderungen der Faktorenstruktur in Abhängigkeit von verschiedenen Veranstaltungen werden nicht erfaßt. Faktorenanalysen innerhalb jeder Veranstaltung oder über aggregierte Daten (Einheiten = Veranstaltungen) dürften aber praktisch nicht möglich sein, weil aufgrund der Forderung der linearen Unabhängigkeit der Einzelitems mathematisch mindestens 1 Pb mehr als Items, praktisch aber 2-3 mal soviel Pbn wie Items vorhanden sein müssen. Deshalb wird auch hier die Faktorenanalyse über alle befragten Teilnehmer gerechnet werden; Gruppierungseffekte werden aber dadurch versucht aufzuzeigen, daß zufällig Subsets von Veranstaltungen ausgewählt werden, über die vergleichend eine neue Faktorenanalyse gerechnet wird.

<sup>16</sup> Es ergibt sich hier das gleiche Problem wie bei der Faktorenanalyse. Betrachtet man alle Teilnehmer zusammen, so berücksichtigt man die unterschiedliche Veranstaltungszugehörigkeit nicht, aggregiert man auf Veranstaltungsebene, so müßten zu viele Veranstaltungen untersucht werden (mindestens Anzahl der Items + 1; allgemein wird man bei ca. 40-50 Prädiktoren 100-150 Einheiten fordern). Da hier nur die erste Alternative möglich sein wird, d.h. die Gruppierung der Teilnehmer nicht beachtet wird, sind alle Regressionsanalysen mit Vorsicht zu betrachten.

für den Zusammenhang zwischen Prädiktoren und Kriterium berechnen. Danach wird über eine schrittweise Regressionsanalyse ein verringerter Satz von Prädiktoren gebildet, der alleine das Kriterium fast genauso gut vorhersagt wie alle Prädiktoren zusammen. Wirkt sich nun ein Halo-Effekt aus, d.h. ist für die SMC dieses Prädiktorensatzes der Halo-Effekt verantwortlich, so müßte ein beliebig anderer, gleichumfangreicher Prädiktorensatz zu einer ähnlich hohen SMC führen. Diese letzte SMC läßt sich dann als Maximalschätzung für den Halo-Effekt betrachten.

#### - Reliabilität

Mittels Ratingskalen gewonnene Informationen sind dann für ein Auswertungsgespräch von wenig Nutzen, wenn sie unzuverlässige, ungenaue Meßdaten darstellen. Ausreichende Reliabilität ist also eine notwendige Forderung an das Evaluationsinstrument. Ratings im Bereich der Lehrer- und Hochschulveranstaltungsbeurteilung zeigen oft eine recht hohe Stabilität, split-half und Konsistenzreliabilität, die um  $r_{tt} = .8$  bis  $r_{tt} = .9$  liegt (vgl. z.B. das Sammelreferat von COSTIN u.a. 1971, S. 512f.; ALEAMONI/SPENCER 1973, S. 674; DIEHL/KOHR 1977, S. 69). Als Grundlage zur Reliabilitätschätzung wird die Split-half Methode und die interne Konsistenzschätzung benutzt werden müssen. Eine Überprüfung der Stabilität wird nicht möglich sein, da die teilnehmenden Erzieher nur zu einem Zeitpunkt befragt werden können (Abschluß der Veranstaltung).

An anderer Stelle wird die Reliabilität als Beurteilerübereinstimmung bzw. Interraterzuverlässigkeit bestimmt (vgl. z.B. LANGER/SCHULZ von THUN 1974, S. 22f.; WEGNER 1976, S. 118). Diese wird dann über varianzanalytische Versuchspläne berechnet, die voraussetzen, daß mehrere Veranstaltungen von mehreren (den gleichen) Ratern eingeschätzt

würden. Diese Möglichkeit wird im vorliegenden Fall nicht gegeben sein. Zudem führt nach SCHMIDT diese Überprüfung allenfalls zu einer Bestimmung der Objektivität von Ratings, da alle Beurteiler ein ähnliches Fehlurteil abgeben könnten (vgl. SCHMIDT 1966, S. 49f.)<sup>17</sup>.

### - Validität

Ein gemeinsames Auswertungsgespräch zwischen den an der Fortbildung Beteiligten wird nur dann sinnvoll durch einen Fragebogen vorbereitet, wenn die mit ihm erhobenen Informationen der Veranstaltungswirklichkeit entsprechen. Ein unvalides Evaluationsinstrument würde das Gespräch nur belasten und könnte für die Beteiligten keinen Nutzen bieten. In der Entwicklung und Erprobung des Fragebogens - hier vor allem in der 2. Phase - muß deshalb auf die Überprüfung der Validität ein besonderes Gewicht gelegt werden.

Validierung betrachtet die Interpretation von Informationen:

"The phrase validation of a test is a source of much misunderstanding. One validates, not a test, but an interpretation of data arising from a specified procedure." (CRONBACH 1971, S. 447)<sup>18</sup>

Es sind zwei Arten von Interpretationen möglich:

1. Fragebogeninformationen werden interpretiert, um eine Wahl zwischen Handlungsalternativen zu rechtfertigen. Da

<sup>17</sup> SCHMIDT definiert deshalb die relative Abwesenheit von systematischen Beurteilungsfehlern - vor allem des Halo-Effektes - als geeignetes Maß zur Schätzung der Reliabilität (vgl. SCHMIDT 1965, S. 46). WEGNER dagegen hält diese Definition - da hier Variationen der Meßoperationen und/oder des Reizmaterials erforderlich sind - für ein Gültigkeitsmaß (vgl. WEGNER 1976, S. 118). Zur Abschätzung des Ausmaßes des Halo-Effektes bei SCHMIDT siehe Anmerkung 14.

<sup>18</sup> Validiert wird nicht nur ein "nacktes" Instrument, sondern die gesamte Prozedur der Fragebogenanwendung einschließlich Kontext des Einsatzes, Instruktion, Einführung, antizipierte Folgen usw. Es muß festgelegt sein, was als die signifikanten Aspekte dieser Anwendungssituation gilt (vgl. CRONBACH 1971, S. 449).



solche Entscheidungen oft weitreichende Konsequenzen haben - siehe z.B. bei Einstufungstests -, setzt dies eine kriterienbezogene Validität voraus, d.h. befriedigende Korrelationen mit einem für die Entscheidung sinnvollen Außenkriterium.

2. Fragebogeninformationen werden interpretiert, um Personen bzw. Situationen beschreiben zu können; eine solche Beschreibung kann oft mehreren - teils zukünftigen - Zwecken dienen. Dieser Interpretationsmöglichkeit werden oft Content- und Konstruktvalidierung zugeordnet (vgl. ib., S. 448).

Die mit dem Fragebogen erhobenen Informationen werden in dem gemeinsamen Auswertungsgespräch weniger für Entscheidungen mit weitreichenden Konsequenzen benutzt, sondern eher zur Vergegenwärtigung der abgelaufenen Fortbildungsprozesse und seiner Stärken und Schwächen. Werden Entscheidungen getroffen, so beziehen sie sich auf Verbesserungsmöglichkeiten dieser Prozesse. Kriterienbezogene Validität wird hierfür nicht unbedingt benötigt, so daß die Beschränkung auf vorwiegend Validität für Situationsbeschreibungen - also Content- und Konstruktvalidität - legitim ist:

- Content- oder inhaltliche Validität meint, daß ein Fragebogen selbst das bestmögliche Kriterium für ein Merkmal darstellt, bzw. dies ihm in einem Expertenrating zugeschrieben wird (vgl. LIENERT 1969, S. 260). Sie bezeichnet den Grad, indem der Fragebogen theoretischen Überlegungen zur Fortbildung entspricht (vgl. CRONBACH 1971, S. 451f.), bzw. - da keine explizite Fortbildungstheorie besteht - inwieweit die Itemsauswahl repräsentativ für das Gesamt der diskutierten Postulate zur Gestaltung von Fortbildung ist<sup>19</sup>.

---

<sup>19</sup> Inhaltliche Validität ist von psychologischer Validität oder face-validity zu trennen. Diese meint die Plausibilität des Fragebogens für die Teilnehmer, die dem Bogen zugesprochen werden kann.

FALTER bezieht Inhaltsvalidität auf die Übereinstimmung der Indikatoren und der angenommenen Grundgesamtheit

Dies soll durch eine Experten- und Fortbildungsleiterbefragung überprüft werden. Ebenfalls wird hier gefragt, ob die zusammengestellten Items überhaupt für die Bewertung einer Fortbildungsveranstaltung wichtig sind. Dies wird als Importance oder Wichtigkeit bezeichnet (vgl. ib.)<sup>20</sup>.

noch 19 aller Items, die durch ein fragliches Konzept erfaßt werden. Inhaltsvalidität bezieht sich hier auf die sprachliche Ebene, d.h. darauf, wie theoretisches Konstrukt und Operationalisierung im Test übereinstimmen (vgl. FALTER 1977, S. 378). DIETERICH diskutiert auf dem Hintergrund von sprachlogischen Ansätzen (z.B. der "Logischen Propädeutik" von KAMLAH und LORENZEN) das Problem der Validität. Mit Bedeutungsumfang bezeichnet er die Gesamtheit dessen, was bei der Einführung als der dem Prädiktor entsprechende Gegenstand festgelegt wurde (vgl. DIETERICH 1973, S. 99); m.a.W.: "der Bedeutungsumfang umfaßt die an Einzelgegenständen feststellbaren relevanten Attribute, die gemäß der terminologischen Vereinbarung die Zugehörigkeit der Gegenstände zu einer Gegenstandsklasse festlegen." (ib., S. 100). Wenn ein Verfahren es erlaubt, einen Gegenstand in seinem gesamten Bedeutungsumfang zu erfassen, so kann ihm logische Validität zugesprochen werden: "Das vom Test erfaßte Merkmal ist hinsichtlich des gesamten Bedeutungsumfanges identisch mit dem, was er erfassen soll." (ib.).

Davon trennt DIETERICH die repräsentative Validität: Der vom Test erfaßte Teilgegenstand ist repräsentativ für den gemeinten Gegenstand, d.h. er erfaßt den Gegenstand nicht in seinem gesamten Bedeutungsumfang, sondern nur in einem repräsentativen Teilbereich (bzw. dies muß glaubhaft gezeigt werden).

In dem hier vorliegenden Zusammenhang wäre in diesem Sprachgebrauch also eine repräsentative Validität angezielt.

<sup>20</sup> "Content validation asks whether the test fits the developers's blueprints, and the question about importance asks whether the user would have chosen the same blueprint." (CRONBACH 1971, S. 452).

- Unter der Evaluationsfragestellung reichen inhaltliche Validität und Importance aber nicht aus. Die Informationen müssen für die konkrete Arbeit und die konkreten Zwecke der Fortbildungspraktiker - neben Auswertungsgesprächen z.B. Programmplanung, Protokollerstellung - nützlich sein. Die Experten sollen hier diese mögliche Nützlichkeit antizipieren, während die Fortbildungsleiter nach der tatsächlichen Nützlichkeit gefragt werden sollen<sup>21</sup>. Da hier aus untersuchungspraktischen Gründen nur Fortbildungsleiter und Experten befragt werden können, wird ebenfalls nach den potentiellen Auswirkungen des Ausfüllens auf die Teilnehmer zu fragen sein.

- Konstruktvalidierung ist eher theoretisch orientiert, ihr Ziel liegt in einer theoretischen Klärung dessen, was der Fragebogen mißt<sup>22</sup>. Faktoren-, Regressions- und Diskriminanzanalysen können Hinweise auf den inhaltlich-logischen Zusammenhang der Items geben.

<sup>21</sup> Ein Problem wird hier sein, daß die Frage nach dem Nutzen sich direkt nur auf den Fragebogen in dem Entwicklungsstadium bezieht und nicht auf den fertigen Bogen im idealen Auswertungskontext.

<sup>22</sup> Ausführlich diskutiert FALTER die Konstruktvalidität. Sie beruht auf einem Zweisprachenmodell: Beobachtungssprache und Theoriesprache. Die Konstrukte sind durch theoretische Netzwerke miteinander verbunden; jedes noch so rudimentär entwickelte Konstrukt ist darin eingebunden. Aus den Netzwerken werden empirisch überprüfbare Hypothesen abgeleitet und durch Verwendung geeigneter Meßverfahren getestet. Sind die Ableitungen der Hypothesen aus den Netzwerken logisch konsistent und trifft die abgeleitete Hypothese zu, so kann das Forschungsinstrument als valide angesehen werden (vgl. FALTER 1977, S. 379f.).

Logisch streng genommen ist hier nur der Umkehrschluß möglich. Leitet man aber eine möglichst große Zahl empirisch gehaltvoller Hypothesen ab und unterzieht sie einer empirischen Überprüfung, so wächst mit der Zahl der nicht falsifizierten Hypothesen die Wahrscheinlichkeit, daß Test und Konstrukt valide sind (vgl. ib., S. 382f.).

Faktorenanalysen sind wichtig zur Aufklärung der internen Struktur der Items. Diese werden zwar aufgrund der theoretischen Überlegungen zu den Erfolgskriterien ausgewählt, ob die Teilnehmer aber die Items gemäß der unterlegten Struktur auffassen, ist damit noch nicht erwiesen. Zu beachten ist, daß die Faktorenanalyse insofern nur einen relativen Hinweis auf die Struktur der Aspekte gibt, als sie ja nur jenes strukturieren kann, was an Items im Fragebogen enthalten ist.

Die Regressionsanalyse erlaubt die Verfolgung einer Frage, die für die Einschätzung der Informationen für ein Auswertungsgespräch wichtig ist. In den Fragebogen werden eine größere Anzahl von Veranstaltungsaspekten eingehen, die alle gleichgewichtet betrachtet werden. Auch die Faktorenanalyse reduziert nur die Anzahl der Items auf Dimensionen, die in ihrer Bedeutung aber prinzipiell gleichwertig sind. Für ein Auswertungsgespräch ist es aber wichtig, das Präferenzsystem der Teilnehmer zu kennen, d.h. jene Veranstaltungsaspekte, die für sie besonders relevant sind und so die Gesamtbeurteilung der Veranstaltung mehr beeinflussen bzw. "vorhersagen" als andere Aspekte. Regressionsanalytisch gesprochen sind dies die Items, die den größten Beitrag zur Vorhersage des Kriteriums "Gesamteinschätzung" leisten (vgl. FRENCH-LAZOVIK 1974)<sup>23</sup>.

<sup>23</sup> In diesem Rahmen läßt sich auch ein weiterer Vorwurf verfolgen, der vor allem an die Lehrerbeurteilung durch die Schüler herangetragen wird. Zur Gesamtbeurteilung trägt danach weniger die methodische Gestaltung und der Aufbau einer Veranstaltung usw. bei, sondern vor allem der Humor, die Popularität, das Entertainment eines Lehrers. Untersuchungen stützen diesen Vorwurf aber weniger, Schüler und Studenten sind eher kompetente Rater, die "showmanship" von gutem Unterricht trennen können (vgl. COSTIN u.a. 1971, S. 517f.; FRENCH-LAZOVIK 1974, S. 377; HARARI/ZEDECK 1973, S. 265). Konkret müßte also über die Regressionsanalyse gefragt werden, ob Items, die z.B. nach der Freizeitgestaltung fragen, besonders viel zur Gesamtbeurteilung beitragen, wobei aber festzuhalten ist, daß eine gute Freizeitgestaltung ein positives und berechtigtes Element der Erzieherfortbildung ist.

Die Güte des Fragebogens zur Vorbereitung eines Auswertungsgespräches wird weiterhin davon abhängig sein, ob der Fragebogen alle relevanten Aspekte der Veranstaltungsbeurteilung umfaßt. Ein - vorsichtiger - Ansatz zur Beantwortung dieser Frage läßt sich ebenfalls auf der Regressionsanalyse aufbauen (vgl. ib., S. 376). So läßt sich die Korrelation der Kombination aller Items mit der Globaleinschätzung betrachten. Die SMC gibt nun den Anteil an gemeinsamer Varianz an; ist dieser sehr hoch, so bedeutet dies, daß im Fragebogen alle für die Globaleinschätzung relevanten Items enthalten sind.

Schließlich sollte ein Evaluationsinstrument die vielfältigen Aspekte erfassen können, in denen sich Fortbildungsveranstaltungen für Erzieher unterscheiden. Unzureichend in der Güte ist hier ein Fragebogen, in dem sich solche Unterschiede nicht niederschlagen. Es wird deshalb über eine Diskriminanzanalyse zu fragen sein, wie sensibel die ausgewählten Fragen auf Veranstaltungsunterschiede reagieren.

Neben diesen Überprüfungen der eher inhaltlich-logischen, internen Struktur der Items soll im Rahmen der Konstruktvalidierung ein Fragebogen entwickelt werden, der nach einigen theoretisch relevanten, methodisch-didaktischen Aspekten von Erzieherfortbildung fragt und den Veranstaltern vorgelegt werden soll. Auf der Ebene der Veranstaltungen als Einheiten können dann Beziehungen zwischen Veranstaltungsunterschieden und der Teilnehmerbeurteilung untersucht werden. Als Beispiel eine Hypothese:

"Veranstaltungen, die aktive Mitarbeit ermöglichen, werden besser eingeschätzt als solche, in denen die Teilnehmer nur passiv aufnehmen können."

Zugleich sollen hier - auf Individual- und Veranstaltungsebene - die Beziehungen zwischen den Beurteilungen und einigen anderen, externen Variablen untersucht werden wie z.B. Fortbildungserfahrung der Teilnehmer, Bekanntheit der Referenten, Alter der Teilnehmer, Anzahl der Teilnehmer in einer Veranstaltung, Art des Kurses, Erfahrung der Referenten usw.<sup>24</sup>.

- Eine kriteriumsbezogene Validität des Fragebogens wird sich nur schwer feststellen lassen. Ratingskalen werden u.a. auch ausgewählt, weil andere und/oder "bessere" Verfahren und somit Kriterien nicht vorhanden sind<sup>25</sup>.

Zum anderen weisen zwar Untersuchungen im Bereich der Lehrereinschätzung auf Beziehungen zwischen Ratings und verschiedenen Kriterien hin (vgl. COSTIN u.a. 1971, S. 516-520; ALEAMONI/SPENCER 1973, S. 679-681; LICKONA/HASCH 1976, S. 452), aber sie zeigen auch, daß diese Beziehungen oft gering sind. Niedrige Korrelationen bedeuten aber nicht unbedingt zugleich niedrige Validität; sie können auch aufzeigen, daß die Teilnehmereinschätzungen einen eigenständigen Beitrag zur Veranstaltungsbeurteilung leisten (vgl. COSTIN u.a. 1971, S. 530). Unter diesen Einschränkungen kann ein rudimentärer Ansatz der Kriteriumsvalidierung dadurch versucht werden, daß die Referenten und Leiter den identischen Fragebogen ausfüllen und die aggregierten Teilnehmereinschätzungen mit diesen Veranstalterereinschätzungen verglichen werden.

<sup>24</sup> In der Veranstaltungsbeurteilung in Schule und Hochschule werden solche Beziehungen breit diskutiert, siehe hierzu z.B. ELMORE/POHLMANN 1978; COSTIN u.a. 1971, S. 520-530; TENT 1973, S. 66).

<sup>25</sup> Nach SCHMIDT sollen Ratingverfahren nur da angewendet werden, wo andere Methoden versagen bzw. noch nicht entwickelt sind. Deshalb fehlen oft Validitätskriterien; Ratings selbst sind wohl die am häufigsten verwendeten Validitätskriterien für andere Verfahren (vgl. SCHMIDT 1965, S. 91f.).

## 5. Ergebnisse der ersten Erprobungsphase

Wurden in einem ersten Schritt unabhängig von den potentiellen Adressaten in einer Literaturanalyse Erfolgskriterien zur Bewertung von Fortbildungsveranstaltungen gesammelt, so verlangte der ziel- und inhaltsfreie Evaluationsansatz, daß in einer zweiten Phase sich vergewissert wurde, ob die Erfolgs- und Beurteilungsmaßstäbe der Adressaten sich unter dieser Sammlung befinden (vgl. Kapitel 2.5, S. 35f.). Dazu wurden die Erfolgskriterien instrumentell aufbereitet, d.h. gemäß der Konstruktionsleitlinien aus Kapitel 4.2 als Items eines Fragebogens formuliert, da letztlich die instrumentelle Umsetzung der Kriterien ihren Gehalt in der praktischen Evaluationsarbeit bestimmt (vgl. S. 100). Dieser Fragebogen wurde in verschiedenen Schritten den Adressaten vorgelegt.

Diese Vorlagen bzw. Erprobungen verfolgten zugleich die Frage, ob die Qualität des Evaluationsinstruments den in Kapitel 2.6 entwickelten Ansprüchen genügt. Ausführlich sollten dabei die in Kapitel 4.3 beschriebenen Probleme und Kontrollmöglichkeiten des Halo-Effektes, der Reliabilität und der Validität untersucht werden. Die Darstellung der Ergebnisse der Erprobungsphasen geschieht in der Reihenfolge der durchgeführten Untersuchungsschritte, d.h. geordnet nach den verwendeten statistischen Verfahren, Untersuchungsschritten und Analyseebenen, um dem Leser das Verständnis zu erleichtern. Zum Abschluß der jeweiligen Erprobungsphase werden die Ergebnisse bezüglich des Halo-Effektes, der Reliabilität und der Validität kurz zusammengefaßt:

### - Vorarbeiten

Neben der Literaturanalyse wurden gleichzeitig Kontakte mit verschiedenen Fortbildungsinstitutionen aufgenommen. In teils längeren Gesprächen wurde dabei die Möglichkeit überprüft, den zu entwickelnden Fragebogen in verschiede-

nen Veranstaltungen einzusetzen. Die Gesprächspartner äußerten zwar durchgehend den Wunsch nach Evaluation und Hilfsmitteln zur Auswertung, nannten aber relativ globale und unspezifische Vorstellungen über die Evaluationsarbeit. Auch aus diesem Grunde wäre somit eine ausführliche Literaturanalyse nötig gewesen. Die Fortbildungsinstitutionen erklärten sich nur dann mit einem probeweisen Einsatz eines Fragebogens bereit, wenn dieser nur wenig der knappen Fortbildungszeit in Anspruch nehmen würde.

Anschließend an die Literaturanalyse wurden die dort gefundenen und begründeten Erfolgskriterien beispielhaft in potentielle Items umformuliert, wobei dies weniger als strenge Operationalisierung, sondern als Verdeutlichung des mit den Kriterien Gemeinten zu verstehen ist. Ein Vergleich mit ähnlichen Fragebögen ergab, daß diese Itemsammlung die dort abgefragten Aspekte durchweg umfaßt, wenn auch mit anderen Gewichten<sup>1</sup>.

Aus dieser Itemsammlung wurden 50 Items ausgewählt und in einem vorläufigen Fragebogen zusammengestellt - in der Hoffnung, damit eine zumindest ansatzweise repräsentative Auswahl zu haben (vgl. zum Repräsentanzproblem SPRUNG/

---

<sup>1</sup> Die Umformulierung der Erfolgskriterien ergab ca. 350 verschiedene Items. Diese Sammlung wurde mit den Fragebögen verglichen, die von den folgenden Autoren genannt werden:

DIEHL/KOHR 1977; FASSBENDER 1974; FRENCH-LAZOVIK 1974; JAEGER/FREIJO 1974; TENT 1973; ALEAMONI/SPENCER 1973; OLIVERO 1973; FITTKAU 1972; GERL 1975; HEYER 1977; MEISTER 1974; MINSEL/MINSEL 1973; PETERMANN 1979; PROSE 1973; SCHOTT 1973; SIEBERT 1975; TIETGENS 1974; SACKMANN 1977; VONTOBEL 1972; WEBER 1977 sowie mehrere Bögen, die von verschiedenen Trägern von Fortbildung eingesetzt werden. Es kann davon ausgegangen werden, daß die hier vorliegende Itemsammlung keine potentiell wichtigen Bereiche übersehen hat.



SPRUNG 1978b, S. 13f.). Dieser Bogen wurde Fortbildungsleitern und Trägervertretern vorgelegt und probeweise in 2 Veranstaltungen mit Erziehern eingesetzt. Diese Adressatengruppen wurden gefragt - die Leiter und Trägervertreter schriftlich, die teilnehmenden Erzieher ebenfalls schriftlich in einer Frage am Ende des Bogens sowie in einem Gespräch über den Bogen -, ob ihrer Meinung nach alle wichtigen Aspekte, Erfolgskriterien angesprochen würden. Überwiegend wurde dies attestiert, fehlende Aspekte wurden in einer anschließenden Revision ergänzt.

Die Adressaten wünschten zudem eine Verkürzung des Bogens. Zugleich stellte sich bei der Betrachtung der Antworten der teilnehmenden Erzieher heraus, daß - ähnlich wie bei dem Milde-Effekt (vgl. WEGNER 1976, S. 115) - die Antworten fast durchweg für die Veranstaltung sehr positiv ausfielen. Deshalb wurden die positiven Pole leicht "positiver" und die negativen Pole abgeschwächt "negativer" formuliert, um die Antworten zur Mitte tendieren zu lassen (vgl. KREUTZ/TITSCHER 1974, S. 61)<sup>2</sup>.

#### - Einsatz und Rücklauf

Von September bis Dezember 1978 wurde dann ein Pool von 44 Items einer ersten, ausführlichen Erprobung unterzogen. Der Fragebogen (siehe Anlage 1) wurde zum Abschluß der Fortbildungsveranstaltung den Teilnehmern mit einigen erläuternden Worten gegeben (siehe Anlage 2 und Anlage 3).

---

<sup>2</sup> Diese Antwortschiefe wird in den meisten Fragebögen zur Veranstaltungsbeurteilung festgestellt. JAEGER und FREIJO versuchten zu einer symmetrischeren Verteilung zu kommen, indem sie den globalen Skalenpunkten "below average, average usw." numerische Anker in Form von Prozentzahlen hinzufügt. Es zeigte sich zwar eine leicht symmetrischere Verteilung, die Differenzen waren aber praktisch nicht signifikant (vgl. JAEGER/FREIJO 1974, S. 419f.).

Beteiligt an der Erprobung waren verschiedene Fortbildungsträger in der BRD<sup>3</sup>. Von 27 Veranstaltungen liegen ausgefüllte Fragebögen vor; in 7 Veranstaltungen wurde der Bogen vom Verfasser selbst eingeführt, ansonsten üben die dafür instruierten Fortbildungsleiter diese Aufgabe aus. Da in der ersten Erprobungsphase hauptsächlich Hinweise auf eine Revision des Fragebogens gefunden werden sollten, wurden - angelehnt an das Konzept der klassischen Testtheorie<sup>4</sup> - hier nicht alle möglichen Probleme von Ratings untersucht. Vielmehr wurde sich auf die wichtigsten, revisionsrelevanten Aspekte beschränkt. Die engzeilig beschriebenen Ergebnisse enthalten dabei Teilaspekte, zusätzliche Analysen oder erläuternde Hinweise.

Aus den untersuchten Veranstaltungen lagen 467 ausgefüllte Bögen vor; bei 35 Bögen hatten die Teilnehmer offensicht-

<sup>3</sup> Einbezogen wurden Veranstaltungen von folgenden Trägern:

- Caritas Verband für die Diözese Freiburg
- Caritas Verband für die Diözese Hildesheim
- Caritas Verband für die Diözese Münster
- Caritas Verband für das Erzbistum Köln
- Deutscher Verein für öffentliche und private Fürsorge.
- Fortbildungswerk für sozialpädagogische Fachkräfte, Frankfurt.
- Deutsches Rotes Kreuz, Landesverband Niedersachsen, Hannover.
- Evangelischer Fachverband der Tageseinrichtungen für Kinder in Westfalen-Lippe, Münster.
- Sozialpädagogisches Fortbildungswerk des Landes Rheinland-Pfalz, Mainz.

<sup>4</sup> Fast durchweg werden alle in Anmerkung 1, Kapitel 5 genannten Veranstaltungsbeurteilungsbögen nach "klassischer" Testtheorie analysiert (vgl. z.B. DIEHL/KOHR 1977; zum Konzept LIENERT 1969; SCHWARTZ 1978; speziell für Fragebögen SPRUNG/SPRUNG 1978a, b). Alle Berechnungen wurden am Rechenzentrum der Universität Münster durchgeführt; benutzt wurde das Programmsystem SPSS (siehe NIE u.a. 1975; ab der 2. Phase wurde die dann implementierte Version 8 benutzt, vgl. HULL/NIE 1979).

lich Schwierigkeiten beim Ausfüllen. Bei 13 Bögen war dies zu korrigieren, während 22 Bögen nicht auswertbar waren. Es blieben somit 445 (95,3 %) auswertbare Bögen übrig.

Durch einen technischen Fehler wurde die Anzahl der Teilnehmer pro Veranstaltung nicht erhoben, so daß allgemein keine Aussage über den Prozentsatz der Nicht-Ausfüller gemacht werden kann. Nur bei 14 Veranstaltungen liegt die Teilnehmerzahl pro Veranstaltung vor. Diese Veranstaltungen besuchten 274 Teilnehmer, während hier nur 220 (80,3 %) Fragebögen vorliegen, von denen 216 (78,8 %) auswertbar waren. Fehlende Fragebögen sind teilweise darauf zurückzuführen, daß manche Teilnehmer zum Abschluß einer Veranstaltung - wegen Krankheit, besserer Zugverbindung bei früherer Abreise usw. - nicht mehr anwesend waren. Es ist jedoch auch nicht auszuschließen, daß die Nicht-Ausfüller eine spezielle, z.B. mit der Veranstaltung sehr unzufriedene Teilnehmergruppe waren.

Die Fragebogenanalyse berücksichtigte zuerst alle 27 Veranstaltungen.

Den größten Raum nahmen hierbei Faktorenanalysen bei unterschiedlichen Itemzahlen, unterschiedlichen Itemkodierungen und unterschiedlichen Faktorenzahlen ein. Ebenso wurden unterschiedliche Veranstaltungsarten betrachtet, speziell einmal "mehrtägige Erzieherfortbildungsveranstaltungen" und zum anderen "Tages (bzw. Informations-)veranstaltungen". Aus 7 Tagesveranstaltungen liegen 77 auswertbare Bögen vor, was schon auf die relativ hohe Zahl der "Verweigerungen" hinweist. Ein für die weiteren Analysen wichtiges Ergebnis bestand aber darin, daß die Faktorenanalysen unterschiedliche Itemkonfigurationen bei "mehrtägigen Veranstaltungen" im Unterschied zu "Tagesveranstaltungen" feststellten, wobei die Faktoren bei Tagesveranstaltungen nicht zu interpretieren waren. Aus diesem Grunde wurden die Teilnehmer von Tagesveranstaltungen aus den folgenden Analysen ausgeschlossen. Es wäre interessant, die Itemstruktur bei Tagesveranstaltungen und die Unterschiede zu mehrtägigen Veranstaltungen mit differenzierteren Verfahren - z.B. der Transformationsanalyse (vgl. TARNAI 1977) - weiterzuverfolgen.

Die folgenden Ergebnisse beziehen sich somit nur auf die Untergruppe der Teilnehmer, die mehrtägige Veranstaltungen besucht haben; es sind dies 20 Veranstaltungen mit 368 Teilnehmern.

### - Schwierigkeitsanalyse

Jedes Item wurde so gepolt, daß ein hoher Wert (also der Zahlenwert "6") für die besuchte Veranstaltung eine positive Einschätzung bedeutet; danach wurden die Klassen 1 - 3 und 4 - 6 zusammengefaßt. In der die Ergebnisse der ersten Erprobungsphase kurz zusammenfassenden Tabelle 1 (S. 141-142 ) ist unter  $p_1\%$  der Anteil der Teilnehmer aufgelistet, die in dem jeweiligen Item die Veranstaltung negativ einschätzten (bezogen auf die Anzahl aller Antworten).

Bei den Fragen nach den Referenten/Leitern - Item 38 bis 44 - konnte es maximal pro Item 4 Antworten geben; diese Werte wurden zu einem Wert zusammengefaßt (Aufaddieren und Dividieren durch die Anzahl der Referenten/Leiter, die eingeschätzt wurden).

Ein Item wurde aus der weiteren Analyse ausgeschlossen, wenn weniger/gleich 5 % der Teilnehmer eine Veranstaltung in diesem Aspekt negativ beurteilten; dies betraf die Items 2, 15, 16, 24 und 41. Desweiteren wurde Item 38 ausgeschlossen, da dieses nach Randbemerkungen von sehr vielen Teilnehmern mißverstanden wurde.

Die Fragen nach den Referenten/Leitern wurde häufig ausgelassen; dies kann an der Stellung am Ende des Bogens liegen, aber auch daran, daß man sich scheut, diese einzuschätzen. Die Auslassungen bei Item 14 (Kleingruppenarbeit), 15 (Medieneinsatz), 23 (Passung auf unterschiedliche Voraussetzungen), 40 (Zugänglichkeit des Referenten/Leiters für Kritik) und 41 (Eingriff des Referenten/Leiters in Diskussionen) wurden von den Teilnehmern dadurch begründet, daß der entsprechende Aspekt in der Veranstaltung nicht vorkam.

### - Transformationen

Die Items sind nach den Schwierigkeitsindizes sehr schief verteilt; positive Einschätzungen überwiegen. Deshalb wurde die 6er Antwortskala aufgrund der empirischen Häufigkeiten - nicht-linear - transformiert (vgl. LIENERT 1969, S. 332ff.; LIENERT 1962).

- alle Items wurden so gedreht, daß hohe Werte für die Veranstaltung positiv sind,
- daran schloß sich eine Prozentrangtransformation an,
- die Prozentränge wurden in z-Äquivalente umgewandelt; dadurch entstehen quasi normalverteilte Items mit einer Streuung von ca. 1 und einem Mittelwert von ca. 0; die 6 stufigen Items werden im Folgenden mit VAR angesprochen,
- die Items wurden dichotomisiert, d.h. alle z-Werte kleiner als Null erhalten den neuen Wert Null, alle anderen Werte den neuen Wert 1; diese transformierten, dichotomisierten Items werden mit VADI angesprochen (in Tabelle 1 stehen die Anteile der Teilnehmer mit dem neuen Wert Null, jene also, die eine Veranstaltung unterdurchschnittlich einschätzen, unter  $p_2$ ).

Aus Zeitgründen wurden die Transformationsgewichte für die Gesamtgruppe benutzt (siehe Anlage 4), sie unterscheiden sich aber nicht von den Gewichten für die Untergruppe "mehrtägige Veranstaltungen" (siehe Anlage 5).

Die neu entstehenden Items VAR und VADI sind zwar weniger schief verteilt, dafür aber relativ unanschaulich: Bedeutete vorher das Ankreuzen einer "2" ein "stimme der linken Satzergänzung ziemlich zu", so wird dies durch die Transformation zu einem "beurteile diesen Veranstaltungsaspekt unterdurchschnittlich gut mit einem bestimmten Abstand zum Median 0". VADI verkürzt diese Information dann auf unter- oder überdurchschnittlich positive Beurteilung.

Alle folgenden Analysen wurden mit den transformierten Items VAR und VADI berechnet.

#### - Trennschärfen und Reliabilitäten

Durch Aufaddieren wurde für VAR und VADI jeweils ein Gesamtwert gebildet.

Bei VAR wurden fehlende Werte mit Null, also annäherungsweise dem Mittelwert, bei VADI durch .5 ersetzt. Beide Gesamtwerte zeigen in Einwegvarianzanalysen signifikante Gruppenunterschiede auf ( $F = 8.6$  bzw.  $7.6$  bei einem kritischen Wert von  $F_{.05; 19; 348} = 1.63$  bzw.  $F_{.01; 19; 348} = 1.96$ )<sup>5</sup>.

<sup>5</sup> Teilweise ist die Voraussetzung der Varianzhomogenität nicht gesichert. Nach MITTENECKER muß in diesem Fall ein strengeres Signifikanzniveau gefordert werden, d.h. ein 1 % Niveau statt eines 5 % Niveaus (vgl. MITTENECKER 1970, S. 80f.).

Diese und die folgenden Berechnungen setzen eine Intervallskalierung der Items voraus. Für die dichotomisierten Items ist dies von ihrer Art her erfüllt (vgl. NIE

Alle Trennschärfenkoeffizienten, d.h. Korrelationen von VAR und VADI mit den Gesamtwerten, sind zumindest auf dem 0,1 % Niveau signifikant von Null verschieden und reichen von  $r_{it} = .27$  bis  $r_{it} = .75$  (siehe die Zusammenfassung der Itemstatistiken in Tabelle 1). Für die transformierten und dichotomisierten Items VADI und den entsprechenden Gesamtwert wurde als Schätzung der internen Konsistenz KUDER-RICHARDSONs Formula 8 berechnet (vgl. LIENERT 1969, S. 226). Es ergab sich eine sehr hohe Reliabilität auf der Ebene des Gesamtfragebogens von  $r_{tt} = .92$  (eine hohe Reliabilität ist auch zu erwarten, da durch die Transformationen die Schwierigkeiten für die Items an .5 angeglichen wurden; dadurch wird in Formel 8 der Bestandteil  $p \times p$  maximiert).

Für die Gruppe aller Teilnehmer wurden ähnlich hohe Reliabilitäten berechnet:

- nach K-R 8  $r_{tt} = .93$
- nach der split-half Methode  $r_{tt} = .89$ , korrigiert nach SPEARMAN-BROWN (ib., S. 221)  $r_{tt} = .94$
- für die transformierten, 6 stufigen Items VAR wurden ebenfalls durch Zufall einige Hälften gebildet und korreliert:  $r_{tt} = .9$  (korrigiert  $r_{tt} = .94$ ) und  $r_{tt} = .89$  (korrigiert  $r_{tt} = .94$ ).

Es läßt sich somit eine - auf Gesamtfragebogenebene - befriedigende Reliabilität feststellen.

noch 5 u.a. 1975, S. 5f.). Für die mehrstufigen Items VAR - die in der Sozialforschung gewohnheitsmäßig wie quantitative Items behandelt werden - werden ebenfalls "gleiche Abstände" unterstellt. Auch falls diese Forderung nicht erfüllt ist und die Items nur Ordinalskalenniveau haben, so ist nach LABOVITZ der entstehende Fehler vernachlässigbar (vgl. HOLM 1975, S. 43f.). ALLERBECK - der dieses Problem ausführlich behandelt - empfiehlt, solche strittige Skalen so zu behandeln, als ob sie intervallskaliert wären, aber zusätzlich zu überprüfen, ob sich unterschiedliche Ergebnisse bei unterschiedlichen Kodierungen ergeben (vgl. ALLERBECK 1978, sp. S. 211). Empirisch hat er verglichen, ob man bei strittigem Niveau durch Verwendung von  $r$  (setzt Intervallniveau voraus) oder TAUB (setzt Ordinalniveau voraus) zu unterschiedlichen Ergebnissen kommt. In 91 % aller Vergleiche war dies nicht der Fall (vgl. ib., S. 206-209).

## - Faktorenanalysen

Bei der Literaturanalyse wurde eine theoretische Strukturierung der Erfolgskriterien unterstellt (vgl. z.B. den Raster S. 32). Mit Hilfe von Faktorenanalysen wurde nun untersucht, wie die Teilnehmer selbst den Itemsatz intern strukturieren und gruppieren. Das Vorgehen in dieser Arbeit liegt in der Mitte zwischen explorativen und konfirmatorischen Analysen (vgl. MARRADI 1978, S. 498-500). Weder wurden die Items "blind" gesammelt, noch soll bei der Revision starr an der vorherigen theoretischen Struktur der Items festgehalten werden. In Kapitel 3.2 wurde die Idee einer Strukturierung unterstellt, die hier weiter verfolgt werden sollte. Dabei sollte man sich aber der subjektiven Komponenten dieses Verfahrens bewußt sein, z.B. bei der Bestimmung der Faktorenzahl, der Rotation, der Interpretation und bei der Auswahl der zu analysierenden Items. Zugleich sollte aufgrund der Faktorenanalyse eine erste Schätzung des Ausmaßes des Halo-Effekts versucht werden.

Bei der Bestimmung der optimalen Anzahl der Faktoren wurden verschiedene Kriterien berechnet.

Die Inspektion des Eigenwertverlaufes (vgl. FÜRNRATT 1969, S. 68f.; ÜBERLA 1977, S. 127f.) weist auf 5 Faktoren hin; 9 Eigenwerte der Korrelationsmatrix mit Einheiten in der Diagonale ( $R$ ) sind größer als 1, dies kann als Obergrenze für die Anzahl optimaler Faktoren gesehen werden (vgl. RÖHR 1977, S. 63); nach dem Varianzbeitrag-Kriterium gibt es zwei Faktoren in  $R$ , die mehr als 5 % der Gesamtvarianz erklären (vgl. ÜBERLA 1977, S. 124); nach dem Kriterium der positiven Eigenwerte aus  $R_h$  (Korrelationsmatrix mit SMC in der Diagonale) gibt es 3 Eigenwerte mit  $\lambda_i > .05 = \sum \lambda_i$  (mit  $\lambda_i > 0$ ) (vgl. RÖHR 1977, S. 73).

Die optimale Faktorenzahl liegt hiernach zwischen 2 und 9 Faktoren. Es wurden deshalb sukzessiv 3, 4, 5 und 6 Faktoren extrahiert (Modell gemeinsamer Faktoren; mit Iteration, Abbruch bei Differenzen  $< .01$ ; VARIMAX-Rotation). Nach dem FÜRNRATT-Kriterium wurde die Lösung ausgesucht, bei der jeder Faktor von mindestens 3 Items bedeutsam definiert wird (bedeutsam heißt hier  $a_{ij}^2 / \lambda_i^2 \geq .5$ ), bzw. unter sonst gleichen Bedingungen sollte die Lösung vorgezogen werden, bei der die Gesamtzahl von bedeutsamen Items am größten ist (vgl. FÜRNRATT 1969, S. 63ff.).

	bedeuts. Ladungen auf Faktor						Gesamtzahl an bedeuts. Ladungen	% an aufgeklärter Gesamtvarianz
	1	2	3	4	5	6		
3 FL	14	10	9				33	35,2
4 FL	7	10	7	2			26	37,5
5 FL	5	6	3	3	4		21	39,7
6 FL	6	4	3	4	2	2	21	41,4

Nach dem FÜRNRATT-Kriterium ist somit die Anzahl von 5 Faktoren optimal, da bei der 6-Faktorenlösung 2 Faktoren nur durch eine Doublette gekennzeichnet sind, bzw. eine 3-Faktorenlösung zu unspezifisch ist. In der Struktur der Ladungen und Kommunalitäten ergibt sich kein Unterschied zwischen den Faktoren bei den transformierten Items VAR und den transformierten, dichotomisierten Items VADI; deshalb wird hier nur über VAR berichtet.

Aufgrund des FÜRNRATT-Kriteriums wurden 5 Faktoren ausgewählt, die nun interpretiert werden sollen (Kommunalitäten und Ladungen siehe Tabelle 1, sowie ausführlicher Anlage 6).

Die 5 Faktoren erklären 39,7 % der Gesamtvarianz; auf die einzelnen Faktoren entfallen (in Klammern die % Anteile bezogen auf die aufgeklärte Varianz):

F1	F2	F3	F4	F5
10,3(25,9)	7,8(19,6)	8,3(20,8)	7,6(19,1)	5,8(14,6)

Der erste Faktor läßt sich mit Methodischer Aufbau benennen. Die bedeutsam ladenden Items beziehen sich auf die übersichtliche Strukturierung und Gliederung einer Fortbildungsveranstaltung (Item 4, 11, 36) und auf ausreichende Informationen über Lernfortschritte (19). Mit klarer Strukturierung zusammenzugehen scheint der Aufweis einer Praxisrelevanz des Themas (3).

Die nicht bedeutsamen Items, die ihre höchste Ladung aber auf dem 1. Faktor haben, unterstützen diese Interpretation; sie beziehen sich auf Praxisrelevanz (1), auf die Verteilung der zur Verfügung stehenden Zeit (6), auf das Auffangen unterschiedlichen Vorwissens (23) und auf ein angemessenes Anforderungsniveau (9).

Höhere Ladungen (über .4) haben noch 3 Items (18, 28, 37), die eher nach einer Globalbeurteilung fragen.



In diesem Faktor läßt sich - bezogen auf die Strukturierung in der Literaturanalyse - ein Teil des Bereichs "Methodik" wiedererkennen und zwar jener, der sich auf die methodischen Planungsgesichtspunkte im engeren Sinne bezieht (vgl. S. 78). Der folgende, 4. Faktor umfaßt einen anderen Bereich der Methodik; er bezieht sich auf einen speziellen Aspekt der methodischen Gestaltung, auf die Aktivität der Teilnehmer, d.h. auf ihre Möglichkeit, sich aktiv mit dem Thema auseinanderzusetzen (10, 34), wobei gerade die Arbeit in Kleingruppen dem entgegenkommt (14; wegen der niedrigen Kommunalität von  $h_1^2 = .14$  gibt es hier Interpretationsschwierigkeiten).

Aktivität scheint hier einherzugehen mit den Vorstellungen der Teilnehmer von einem sinnvollen Lernen (29, höchste Ladung auf Faktor 4, aber nicht bedeutsam) und mit abwechslungsreicher und aufgelockerter Gestaltung zusammenzugehen (7, 17), die hier eine Aufmerksamkeit aufrecht erhält und so die aktive Auseinandersetzung ermöglicht. Aktivität ist z.T. nur möglich, wenn die Erarbeitung auch auf die konkreten Praxiserfahrungen der Teilnehmer eingeht (25) und die Teilnehmer ihre Ideen und Wünsche einbringen können (5; Item 21 und 33, die ebenfalls nach der Partizipation fragen, haben mit .36 und .37 höhere Ladungen auf diesem Faktor).

Faktor 2 läßt sich am besten mit Arbeitsklima umschreiben; alle noch in der Analyse verbliebenen Items, die nach dem Referenten bzw. Leiterverhalten bezüglich der Arbeitsatmosphäre fragen, bestimmen diesen Faktor bedeutsam (39, 40, 42, 43, 44). Dazu kommt noch die Möglichkeit, abweichende Meinungen frei zu äußern (30), und die Wahrnehmung einer gleichberechtigten Partnerschaft bei der Durchführung der Veranstaltung (33; höchste Ladung auf Faktor 2, aber nicht bedeutsam). Dieser Faktor reproduziert den theoretischen Bereich der Interaktionsprozesse, d.h. Kriterien eines teilnehmerorientierten Leiterverhaltens.

Faktor 3 beschäftigt sich mit der Einschätzung des Lernertrages bzw. einer Globalbeurteilung der Veranstaltung,

d.h. inwieweit sie in den Augen der Teilnehmer nur eine "Zeitvergeudung" war oder konkret etwas für die Arbeit in der Praxis gebracht hat (18, 27, 32).

Dies läßt sich stützen durch jene Items, die ihre höchste Ladung auf diesem Faktor haben, ihn aber nicht bedeutsam definieren; es sind dies die Aspekte von befriedigenden geselligen Kontakten (8), einem klaren Praxisbezug des Themas (28), Entsprechung der Thematik und der Teilnehmerbedürfnisse (20) sowie das als Globalbeurteilung aufgenommene Item 37.

Die Einschätzung des Bereichs der Produktkriterien ist in den Augen der Teilnehmer anscheinend verbunden mit einer eher globalen Beurteilung des Gesamteindrucks einer Veranstaltung.

Der 5. Faktor schließlich umfaßt mehr den äußeren Rahmen einer Fortbildungsveranstaltung, d.h. passende Teilnehmerzahl (22), äußere Bedingungen wie z.B. unbequeme Sitze (12) oder Anfahrt, Dauer, Unterkunft (31), aber auch die Möglichkeit zu einem mehr informellen Erfahrungsaustausch untereinander (26).

Inwieweit konkrete Praxis der Teilnehmer durch Rollenspiel, Tonbandaufnahmen usw. einbezogen werden kann (13; höchste Ladung auf diesem Faktor, aber nicht bedeutsam), hängt anscheinend ebenfalls von diesen Möglichkeiten ab. Etwas unklar ist die Zuordnung der Frage nach der Teilnehmerpartizipation (21).

Gegenüber der theoretischen Strukturierung der Erfolgskriterien zeigen sich nur geringere Unterschiede. Zwar läßt sich ein eigenständiger Bereich Thematik nicht aufweisen, vielmehr verteilen sich Items, die in Bezug auf Thematik formuliert waren, auf die anderen Faktoren; andererseits lassen sich die Kriterienbereiche Produkte, Interaktionsprozesse, situative/organisatorische Bedingungen und Methodik (aufgeteilt in zwei Faktoren) wiederfinden. Da damit größtenteils die theoretischen Bereiche erhalten bleiben, wurde diese 5-Faktorenlösung der Revision des Fragebogens zugrunde gelegt.

Bei dieser Faktorenanalyse findet sich kein allgemeiner Faktor, auf dem alle Items höher laden, sondern eine spezifische, komplexe Faktorenstruktur, die anzeigt, daß die Item-Interkorrelationen genügend streuen. Damit dürfte hier ein Halo-Effekt nur niedrig ausfallen (vgl. S. 117f.).

#### - Regressionsanalyse

Bei den unterschiedlichen Verkodungen der Items VAR und VADI ergeben sich unterschiedliche Regressionsanalysen; die Ergebnisse dieser Analysen wurden somit bei der Revision nicht benutzt.

Problematisch ist, daß keine unabhängige Erfassung des Kriteriums möglich war, sondern die - vorher als Globaleinschätzung bestimmten - Items 18 und 37 im gleichen Itemsatz und in der gleichen Form erhoben wurden. Wegen der vermutlich geringen Reliabilität eines aus zwei Items additiv zusammengesetzten Kriteriums wurden kaum höhere Korrelationen erwartet. Bei den transformierten, 6 stufigen Items VAR ergab eine einfache multiple Regressionsanalyse eine Korrelation der Linearkombination der Prädiktoren mit dem Kriterium von  $R = .69$ , d.h. ein  $SMC = R^2 = .48$ ; korrigiert bezüglich der Anzahl der unabhängigen Items und der Anzahl der Versuchspersonen  $R^2 = .41$  (vgl. NIE u.a. 1975, S. 338; GAENSSLEN/SCHUBÖ 1973, S. 114); die entsprechenden Werte bei den transformierten, dichotomisierten Items VADI sind  $R = .78$ ,  $R^2 = .61$ ,  $R^2$  (korr.) = .55. Der Prädiktorensatz erklärt somit nur 48 % - 61 % der Varianz des Kriteriums; FRENCH-LAZOVIK zum Vergleich fand Werte von 89 % bzw. 96 % (vgl. FRENCH-LAZOVIK 1974, S. 376). Zur Vorhersage der Gesamteinschätzung, gebildet aus den Items 18 und 37, scheinen also nicht alle relevanten Aspekte in den Fragebogen aufgenommen. Es könnte aber auch interpretiert werden, daß nicht Items fehlen, sondern daß unabhängig von spezifischen Aspekten jede Veranstaltung im Gesamt gut bewertet wird, etwa in dem Sinne: besser "Fortbildung", "aus dem Alltag heraus", "mit Kollegen zusammensein" als "arbeiten müssen"<sup>6</sup>.

<sup>6</sup> VONTOBEL erklärt eine hohe allgemeine Kurszufriedenheit damit, daß die Konsequenzen einer Entscheidung im allgemeinen so beurteilt werden, daß die Entscheidung dadurch nachträglich gerechtfertigt wird. D.h. weil ich Zeit und Energie für einen Kurs aufgebracht habe, so kann ich meinen Aufwand nicht völlig entwerten, indem ich mit dem Kurs unzufrieden bin (vgl. VONTOBEL 1972, S. 91).

Die Struktur der Regressionskoeffizienten ist aber verschieden, je nachdem ob die Items VAR oder VADI betrachtet wurden (siehe Anlage 7 und 8; auch die Zusammenfassung in Tabelle 1). Interessant ist, daß bei VAR die Items signifikante Regressionskoeffizienten haben, die nicht zugleich ihre höchste Ladung auf dem Faktor "Lernertrag und Globaleinschätzung" haben, aus dem die Items 18 und 37 stammen, während dies bei VADI für die Items 20, 27 und 32 zutrifft.

Eine schrittweise Regressionsanalyse (vgl. GAENSSLEN/SCHUBÖ 1973, S. 112f.) ergab bei VAR einen Satz von 4 Prädiktoren, der allein schon ca. 40 % ( $R^2(\text{korr.})=.39$ ) der Varianz des Kriteriums aufklärt (bei VADI 10 Prädiktoren, die 57 % ( $R^2(\text{korr.})=.56$ ) aufklären). Man kann davon ausgehen, daß diese Auswahl von einigen Prädiktoren jeweils die vorhersagekräftigste eines bestimmten Umfanges von Prädiktoren ist (vgl. GAENSSLEN/SCHUBÖ 1973, S. 113). Zusätzlich wurden nun Prädiktorensätze gleichen Umfanges herausgesucht; sie erklären zwischen 17 % und 18 % der Varianz des Kriteriums gegenüber ca. 40 % beim optimalen Satz (bei VADI zwischen 35 % und 41 % gegenüber 57 % beim optimalen Satz). Nach FRENCH-LAZOVIK können die niedrigen Prozentzahlen als Maximalschätzung des Halo-Effektes betrachtet werden, d.h. im höchsten Fall erklärt ein Halo-Effekt ca. 17 % (bei VADI 35 %) der Varianz der Gesamteinschätzung, während ca. 23 % (bei VADI 22 %) nicht auf einen Halo-Effekt zurückzuführen wären.

#### - Zusammenfassung und Revision

In dieser Phase wurde ein erster, vorläufiger Fragebogenentwurf untersucht. In dem Mittelpunkt des Interesses stand die Suche nach Hinweisen zur Veränderung und Revision des Fragebogens, um die angestrebten Ziele besser erreichen zu können. Die wichtigsten - revisionsrelevanten - Ergebnisse seien kurz zusammengefaßt:

1. Ein Hauptaugenmerk lag auf der Frage, ob unter den ausgewählten 44 Items sich auch jene Maßstäbe der Adressaten befinden, die diesen für die Beurteilung einer Fortbildungsveranstaltung wichtig sind. Gespräche mit den Fortbildungsleitern und Trägervertretern gaben keine Hinweise auf fehlende Erfolgskriterien. Die Frage nach fehlenden Kriterien am Ende des Fragebogens beantworteten die Teilnehmer selten; auch hier zeigten sich - die meisten An-

Tabelle 1: Zusammenfassung der Statistiken aus Phase 1

Nr.	(1) (2)		n	Trennschärfe-koeff.		(5) 5-Faktorlösung VAR						Global-ein-schätz.		(8) signif. BETA	
	P <sub>1</sub> %	P <sub>2</sub> %		$r_{it}$		Ladung $a_{ij}$ auf					$r_{it}$		bei		
				(3)	(4)	$h^2_i$	F1	F2	F3	F4	F5	(6)	(7)	VAR	VADI
1	8.6	63.3	362	.56	.52	.40	.43	-	.36	-	-	.28	.46	-	+
2	4.4	37.9	356	-	-	-	-	-	-	-	-	-	-	-	-
3	16.8	49.7	364	.56	.57	.34	.44*	-	-	-	-	.21	.44	-	-
4	12.3	63.4	358	.56	.51	.50	.65*	-	-	-	-	.30	.33	-	-
5	14.9	49.6	363	.60	.56	.39	-	.30	-	.36	.31	.31	.34	-	-
6	18.5	63.3	362	.59	.49	.39	.44	-	-	-	.35	.32	.40	-	+
7	5.4	56.1	365	.65	.55	.42	.32	-	.32	.34	-	.28	.41	-	-
8	9.9	42.1	363	.48	.45	.25	-	-	.29	-	-	.29	.32	-	-
9	8.5	37.0	365	.56	.48	.33	.35	.32	-	-	-	.21	.37	-	-
10	20.7	38.1	357	.45	.34	.42	-	-	-	.61*	-	.09	.19	-	-
11	10.4	52.1	364	.57	.49	.54	.68*	-	-	-	-	.26	.29	-	-
12	24.6	51.0	357	.33	.28	.27	-	-	-	-	.49*	.17	.11	-	-
13	11.1	44.9	361	.64	.59	.46	.32	-	-	.32	.46	.25	.42	-	-
14	7.2	38.1	345	.32	.37	.14	-	-	-	.34*	-	.11	.25	-	-
15	4.4	40.4	342	-	-	-	-	-	-	-	-	-	-	-	-
16	4.8	43.4	366	-	-	-	-	-	-	-	-	-	-	-	-
17	7.2	46.8	363	.64	.55	.43	-	.32	-	.40	-	.56	.32	+	-
18	6.3	35.1	365	.73	.64	.66	-	-	.59*	-	-	-	-	-	-
19	6.9	59.6	345	.58	.58	.36	.44*	-	-	-	-	.26	.41	-	-
20	15.3	55.1	365	.61	.59	.38	.32	-	.37	.30	-	.30	.49	-	+
21	13.6	52.9	361	.60	.55	.41	-	-	-	.36	.39	.32	.38	-	-
22	21.3	46.3	360	.27	.32	.08	-	-	-	-	.22*	.14	.14	-	-
23	10.9	42.1	340	.61	.57	.36	.37	-	-	-	.28	.23	.35	-	-
24	3.6	20.0	360	-	-	-	-	-	-	-	-	-	-	-	-
25	10.9	55.2	350	.61	.60	.35	-	-	-	.30	.29	.26	.38	-	-
26	10.7	52.5	356	.52	.52	.37	-	-	-	-	.49*	.16	.33	+	-
27	7.2	41.1	360	.57	.55	.57	-	-	.68*	-	-	.32	.50	-	+
28	5.9	40.2	358	.68	.63	.53	.40	-	.45	-	-	.30	.49	-	-
29	10.5	56.4	355	.63	.58	.46	.37	-	-	.46	-	.34	.48	-	+
30	6.1	25.4	359	.56	.54	.44	.31	.51*	-	.29	-	.22	.35	+	-

Tabelle 1: Zusammenfassung (Forts.)

Nr.			n	Trennschärfekoeff.		(5) 5-Faktorlösung VAR						Global-einschätz.		(8) signif.BETA	
	(1)	(2)		$r_{it}$		Ladung $a_{ij}$ auf					$r_{it}$		bei		
	$p_1\%$	$p_2\%$		(3)	(4)	$h^2_i$	F1	F2	F3	F4	F5	(6)	(7)	VAR	VADI
31	16.1	55.7	353	.29	.31	.16	-	-	-	-	.38*	.06	.15	-	-
32	9.4	46.9	360	.59	.56	.58	-	-	.69*	-	-	.28	.52	-	+
33	14.2	43.2	358	.56	.54	.38	-	.42	-	.37	-	.29	.26	-	-
34	13.4	45.4	359	.62	.55	.63	-	.33	-	.69*	-	.30	.40	-	-
35	27.9	39.2	355	.28	.28	.07	-	.17	-	-	-	.02	.18	+	-
36	10.7	57.7	355	.62	.54	.53	.65*	-	-	-	-	.29	.41	+	-
37	5.9	44.5	357	.75	.66	.64	.43	-	.51	.32	-	-	-	-	-
38	67.2	46.6	292	-	-	-	-	-	-	-	-	-	-	-	-
39	7.9	34.3	329	.54	.54	.38	-	.48*	.30	-	-	.35	.42	-	+
40	5.1	22.5	316	.59	.52	.60	-	.70*	-	-	-	.32	.40	-	+
41	4.1	30.3	314	-	-	-	-	-	-	-	-	-	-	-	-
42	9.6	30.3	332	.45	.53	.27	-	.45*	-	-	-	.28	.31	+	-
43	9.6	43.3	332	.55	.54	.35	-	.46*	-	-	-	.35	.31	+	-
44	6.7	23.1	330	.43	.46	.25	-	.37*	-	-	-	.20	.28	-	-

- (1) % Anteil derjenigen, die in den gepolten Originalfragen eine Veranstaltung negativ beurteilen
- (2) Schwierigkeiten der transformierten und dichotomisierten Items VADI
- (3) Trennschärfen der transformierten Items VAR zum Gesamtwert
- (4) Trennschärfen der transformierten, dichotomisierten Items VADI zum Gesamtwert
- (5) aufgelistet sind alle Ladungen größer .30 (mit \* sind die nach dem FÜRTRATT-Kriterium bedeutsamen Ladungen); sind alle Ladungen kleiner .30, so ist die jeweils größte Ladung angegeben
- (6) Korrelationen von VAR mit dem Globalwert gebildet aus Item 18 und 37
- (7) Korrelationen von VADI mit dem Globalwert gebildet aus Item 18 und 37
- (8) mit + sind die signifikanten (5 % Niveau) Regressionskoeffizienten bezeichnet

merkungen bezogen sich auf äußerst spezifische Aspekte der gerade besuchten Veranstaltung - keine unzureichenden und fehlenden Erfolgsmaßstäbe.

2. In ersten Faktorenanalysen zeigten sich unterschiedliche Faktorenstrukturen bei den Veranstaltungstypen "Tages-/Informationsveranstaltung" und "mehrtägige Veranstaltung"; deshalb und wegen der unterschiedlichen Funktion dieser Typen (vgl. Kapitel 1) wurden für die weiteren Analysen nur mehrtägige Veranstaltungen ausgesucht.

3. Gespräche mit Teilnehmern, deren Randbemerkungen auf den Bögen und die nicht auswertbaren Antworten zeigten an, daß der zweipolige Aufbau der Ratingskalen nicht allen Teilnehmern deutlich war.

Teilweise äußerten Erzieher, daß der zweipolige Aufbau der Items ihnen eine Einschätzung erleichtere. Andererseits dauerte es aber jeweils längere Zeit, bis sie den Antwortmodus verstanden hatten (von durchschnittlich 23 Minuten Ausfüllzeit ca. 3-5 Minuten); andere Teilnehmer durchschauten das Schema nicht und ihre Bögen waren nicht auswertbar.

Ebenso erklärten mehrere Teilnehmer, daß 6 Antwortstufen ihnen zu viel wären und sie die unterschiedlichen Abstufungen nicht klar unterscheiden könnten. Für die zweite Erprobung wurde deshalb eine andere Darbietung der Items ausgewählt, bei der den Teilnehmern jeweils eine ganze Aussage vorgelegt wird, der sie dann auf einer Skala mit 4 Ausprägungen zustimmen können oder nicht. Nach ROHRMANN wurden folgende Stufen der Zustimmung ausgewählt (vgl. ROHRMANN 1978)<sup>7</sup>.

<sup>7</sup> ROHRMANN - dessen Arbeit während der Vorbereitung der 1. Phase noch nicht veröffentlicht war - untersucht empirisch Häufigkeits-, Intensitäts-, Wahrscheinlichkeits- und Bewertungsbegriffe und ordnet sie auf einer 9er Graduierungsskala an. Die Kombination der Intensitätskala mit dem Verb "zustimmen" scheint am vielseitigsten verwendbar. Die Endpunkte der Skala sollten relativ extrem sein; es wurden die Punkte 1,5 und 8,5 gewählt, so daß Graduierungsbegriffe zu den Skalenpunkten 1,5 - 3,8 - 6,2 - 8,5 gefunden werden mußten (nicht - ein wenig - ziemlich - völlig) (vgl. ROHRMANN 1978).

1	2	3	4
stimmt nicht	stimmt ein wenig	stimmt ziemlich	stimmt völlig

4. Die Items waren sehr schief verteilt, positive Beurteilungen aller Veranstaltungen überwogen. Es wurde vermutet, daß dies eher eine Antworttendenz widerspiegelt, als daß alle Veranstaltungen sehr positiv waren. Da die Transformationen zu inhaltlich sehr schwierig zu interpretierenden Items führten, sollte in der Revision versucht werden, die Verteilungsschiefe zu vermindern. Deshalb wurden alle Items entweder extrem positiv oder sehr abgeschwächt negativ formuliert; damit soll die Wahl der für eine Veranstaltung negativen Seite "erleichtert" werden.

5. In der Faktorenanalyse wurde kein allgemeiner Faktor extrahiert; ein Halo-Effekt fiel vermutlich niedrig aus. Der Fragebogen liefert somit spezifische Informationen die ein Auswertungsgespräch vorbereiten könnten.

Die Überprüfung des Halo-Effektes mittels der regressionsanalytischen Überlegungen sollte mit Vorsicht betrachtet werden, da je nach Transformation - VAR oder VADI - die Ergebnisse unterschiedlich ausfielen.

6. Auf Gesamtfragebogenebene ergab sich eine befriedigende Reliabilität, d.h. relativ zuverlässige und genaue Meßdaten.

7. In den Faktorenanalysen zeigte sich eine Strukturierung der Items, die inhaltlich interpretierbar ist und die weitgehend die theoretischen Überlegungen reproduziert. Aufgrund dieser faktoriellen Validität sollten bei der Revision des Fragebogens die Items auf diese 5 Faktoren zugespißt ausgewählt werden.

8. Ein Fragebogen mit 44 Items erschien den Fortbildungsleitern und -referenten noch zu lang, das Ausfüllen nahm ihrer Meinung nach zu viel Zeit in Anspruch. Da der Fragebogen ja schnell anwend- und auswertbar sein soll, wurde einer weiteren Verkürzung zugestimmt. In der Revision wurden pro Faktor 7 Items ausgewählt; 17 Items wurden positiv,



18 negativ formuliert. 7 Items mußten neu gebildet werden.

Dabei verteilen sich die neuen Items (vgl. Anlage 9) wie folgt auf die Faktoren (in Klammern stehen die Itemnummern der ersten Erprobungsphase; (-) deutet an, daß das Item neu formuliert wurde:

- I. Methodischer Aufbau: 1(3), 5(2), 10(36), 11(1.), 25(9), 27(4), 19(19)
- II. Arbeitsklima: 6(-), 7(-), 26(30), 32(39), 33(40), 34(42), 35(43)
- III. Lernertrag und Globaleinschätzung: 4(-), 9(27), 13(18), 16(32), 20(16), 23(-), 31(37)
- IV. Aktivität: 2(10), 11(17), 12(29), 17(34), 21(5), 28(-), 29(14)
- V. Äußerer Rahmen: 3(22), 8(8), 14(31), 18(12), 22(26), 24(-), 30(31).

Es wurde erwartet, daß nach der Revision der Fragebogen den Ansprüchen an ein Evaluationsinstrument in diesem Bereich besser genügt<sup>8</sup>.

Um einen Hinweis auf die zukünftigen Skalen zu bekommen, wurden jeweils die 6 "besten" Items pro Faktor ausgewählt und zu einer Skala aufaddiert. In allen Skalen - die in Richtung auf das für die Veranstaltung positive Ende schief verteilt waren - ergaben sich mindestens auf dem 1 % Niveau signifikante Gruppenunterschiede. Für die dichotomisierten Items und deren Skalen wurde KUDER-RICHARDSONs Formula 8 berechnet; die Reliabilitätsschätzungen ergaben Werte von  $r_{tt} = .72$  -  $r_{tt} = .87$ . Diese Werte sind eine leichte Überschätzung der Reliabilität, da bei den Item-Skalen-Korrelationen jedes Item selbst in der Skala durch Addition enthalten ist. Dadurch ergeben sich leicht erhöhte Korrelationen, die bei der Schätzung der Reliabilität nicht korrigiert wurden (vgl. LIENERT 1969, S. 115).

<sup>8</sup> Der revidierte Fragebogen enthielt zum Schluß noch eine offene Frage nach den 3 interessantesten Aspekten der Veranstaltung (vgl. Anmerkung 26 von Kapitel 2.6 auf S. 42). Diese informelle Frage konnte jedoch nicht systematisch ausgewertet werden, da kaum Teilnehmer sie beantworteten.

## 6. Ergebnisse der zweiten Erprobungsphase

Die Überprüfung der ersten Form des Fragebogens ergab einige Schwächen, die zu einer Veränderung führten. In einer zweiten Phase wurde der nun revidierte Bogen ausführlich überprüft. Im Mittelpunkt stand wieder die Frage, ob die Qualität des Fragebogens den entwickelten Ansprüchen an ein Evaluationsinstrument entspricht; speziell wurden Halo-Effekt, Reliabilität und Validität untersucht. Die Darstellung der Ergebnisse folgt wieder den durchgeführten Analyseschritten und entlang der verwendeten statistischen Verfahren; zuerst wird über teststatistische Kennwerte und die internen Beziehungen der Items untereinander berichtet, anschließend folgen die Ergebnisse eines Vergleichs der Teilnehmerbeurteilungen mit den Einschätzungen der Referenten und Leiter einer Veranstaltung, die Ergebnisse einer Experten- und Fortbildungsleiterbefragung sowie eine Analyse der die Teilnehmerbeurteilungen möglicherweise beeinflussenden Variablen.

Von Januar bis Oktober 1979<sup>1</sup> wurde der neue Fragebogen bestehend aus 35 Fragen<sup>2</sup> eingesetzt. Der Fragebogen (siehe Anlage 9) wurde wieder zu Ende einer Fortbildungsveranstaltung mit einigen erläuternden Worten (siehe Anlage 10) ausgeteilt. Zur Gewinnung von möglichen Außenkriterien sollten die Fortbildungsleiter und -referenten zugleich einige Fragen zu bestimmten Aspekten der Veranstaltung beantworten (siehe Anlage 11). Ebenfalls wurden sie gebeten,

<sup>1</sup> Der "harte" Winter im Frühjahr 1979 verlängerte diese Erprobungsphase, da in dieser Zeit mehrere Veranstaltungen ausfielen; anstatt bis Mai mußten deshalb Veranstaltungen bis Ende Oktober einbezogen werden.

<sup>2</sup> Obwohl die Items hier als "Aussagen" vorliegen, denen die Teilnehmer zustimmen können oder nicht, wird durchgehend einheitlich von "Fragen" gesprochen, da die Teilnehmer ja danach "gefragt" wurden, ob diese Aussagen zutreffen oder nicht.

auf dem identischen Fragebogen zur Veranstaltungsbeurteilung - den auch die Teilnehmer ausfüllten - ihre eigene Veranstaltung einzuschätzen. Letzteres sollte einem Versuch der Kriteriumsvalidierung dienen.

## 6.1 Analyse des Fragebogens und interne Beziehungen der Items

Beteiligt waren wieder verschiedene Fortbildungsträger in der BRD<sup>3</sup>.

Die ausgefüllten Fragebögen wurden dem Verfasser von den Fortbildungsleitern zugeschickt und von diesem ausgewertet. Die Ergebnisse wurden anschließend in einem Rückmeldeschema (siehe Anlage 12) den Fortbildungsleitern zurückgeschickt. In dem Schema enthalten waren die Nennungen bei der offenen Frage am Schluß des Bogens sowie die Antworten auf die allgemeinen Fragen, ferner die absoluten und relativen Häufigkeiten der 4 Zustimmungsstufen der 35 Fragen. Zusätzlich wurden in einem Profil die aggregierten Teilnehmereinschätzungen mit den entsprechenden Antworten der Referenten und Leiter auf dem gleichen Bogen verglichen; als zusammenfassenden Maß wurde noch das Q-Maß von HOFSTÄTTER mitgeteilt (vgl.

- <sup>3</sup>In der zweiten Phase wurden Veranstaltungen von folgenden Trägern evaluiert:
- Caritasverband für die Diözese Hildesheim (6 Veranstaltungen)
  - Evangelischer Fachverband der Tageseinrichtungen für Kinder in Westfalen und Lippe (6 Veranstaltungen)
  - Deutsches Rotes Kreuz, Landesverband Niedersachsen (5 Veranstaltungen; zusätzlich wurde eine Veranstaltung für ehrenamtliche Helfer in der Umsiedlerarbeit evaluiert; diese Veranstaltung wurde aber nicht bei den folgenden Analysen berücksichtigt)
  - Sozialpädagogisches Fortbildungswerk des Landes Rheinland-Pfalz (8 Veranstaltungen)

Die vier anderen Träger, die in der ersten Erprobungsphase beteiligt waren, nahmen nun nicht mehr teil, da entweder keine Veranstaltungen im Vorschulbereich angeboten wurden, die Fragebögen nicht eingesetzt werden konnten (zu spät erhalten) oder auch kein direkter Kontakt mit den Trägern mehr bestand. Die weiterhin teilnehmenden Träger waren aber diejenigen, deren Veranstaltungen überwiegend in der 1. Phase untersucht worden waren, so daß von daher keine größere Verzerrung zu erwarten ist.

HOFSTÄTTER 1971, S. 187ff.). Durchschnittlich benötigte der Verfasser pro Veranstaltung - einschließlich Betrachtungen der Randbemerkungen, Ablochen, Einlesen, Übertragen, Berechnung des Q-Wertes, Schreiben des Antwortbriefes, aber ohne Maschinen- und Wartezeit - eine Auswertungszeit von 2 3/4 bis 3 Stunden.

Zurückgeschickt wurden aus 25 Fortbildungsveranstaltungen 494 ausgefüllte Fragebögen, von denen 16 Bögen (3,2 %) nicht auswertbar waren.

Von diesen konnten 15 Bögen nicht berücksichtigt werden, da es sich versehentlich um die 1. Version handelte; d.h. bei nur einem Fragebogen wurde der Ausfüllmodus nicht verstanden. In diesen Veranstaltungen waren 557 Teilnehmer, bezogen auf diese Zahl sind somit 85,8 % der Fragebögen (=478) auswertbar. Die fehlenden Bögen erklärten die Veranstalter damit, daß Teilnehmer wegen besserer Fahrverbindungen früher abreisten. Ein absichtliches Nicht-Ausfüllen sei nicht vorgekommen, so daß die auswertbaren Bögen als repräsentativ für das Gesamt der Teilnehmer an diesen Veranstaltungen angesehen werden können.

#### 6.1.1 Itemstatistiken und Reliabilität

Jede Frage wurde so gepolt, daß hohe Werte für die Veranstaltung eine positive Einschätzung bedeuten; danach wurden die Klassen 1, 2 und 3, 4 zusammengefaßt. Tabelle 2 (unter  $p_1$  %) enthält wiederum den Anteil an Teilnehmern, die in dem jeweiligen Aspekt eine Veranstaltung negativ einschätzen (bezogen auf die Anzahl aller Antworten).

Tabelle 2 Schwierigkeiten

Nr.	P <sub>1</sub> %	P <sub>2</sub> %	Mis.	Nr.	P <sub>1</sub> %	P <sub>2</sub> %	Mis.	Nr.	P <sub>1</sub> %	P <sub>2</sub> %	Mis.
1	12,8	60,3	10	13	6,3	26,1	3	25	8,0	27,7	5
2	36,1	36,1	15	14	52,0	52,0	3	26	6,9	24,2	3
3	19,8	32,9	4	15	3,4	18,1	8	27	5,3	19,1	2
4	8,5	26,1	7	16	21,7	57,1	3	28	17,6	46,3	7
5	8,7	33,6	8	17	18,0	53,5	5	29	6,3	17,0	18
6	9,4	50,5	1	18	23,4	50,9	3	30	24,0	43,5	7
7	13,3	40,8	5	19	29,0	61,4	30	31	23,1	60,0	36
8	28,5	58,2	7	20	4,7	18,2	6	32	9,7	35,0	26
9	5,9	16,8	3	21	5,1	17,4	6	33	12,5	30,1	29
10	12,6	58,4	9	22	14,2	44,1	6	34	2,2	17,0	20
11	5,7	28,8	3	23	17,9	46,6	4	35	29,1	61,9	21
12	7,4	24,9	4	24	50,1	50,1	13				

Mis. = Auslassungen bezogen auf die Gesamtzahl von 478 Bögen

P<sub>1</sub>% = Prozentzahl derjenigen, die in der entsprechenden Frage eine Veranstaltung negativ einschätzen.

P<sub>2</sub>% = Prozentzahl derjenigen, die nach den gedrehten, transformierten und dichotomisierten Fragen TRADI - s.w.u. - eine Veranstaltung unterdurchschnittlich beurteilen.

3 Items fallen unter das 5 % Kriterium (bezogen auf p<sub>1</sub>%) und wurden aus den weiteren Analysen ausgeschlossen; es sind dies die Fragen 15, 20 und 34, m.a.W.: der rote Faden einer Veranstaltung war fast immer schon vorhanden (15), die Teilnehmer wurden fast immer sehr für eine Weiterbeschäftigung mit dem Thema motiviert (20) und der Referent/Leiter übergang fast nie Teilnehmerbeiträge (34).

Frage 20 ist die Frage 16 der 1. Version, die dort ebenfalls wegen mangelnder Schwierigkeit herausfiel. Die Fragen 15, 20 und 34 können ohne weiteres herausgenommen werden, da die negativen Beurteilungen sich nicht nur auf 1 oder 2 Veranstaltungen beziehen, sondern über alle Veranstaltungen streuen.

Wiederum wurden die Fragen nach den Referenten/Leitern relativ häufig ausgelassen. Bei Frage 19 gaben mehrere Teilnehmer an, sie hätten die Frage nicht verstanden und sie deshalb nicht beantwortet; der Grund für die höheren Auslassungen bei Frage 31 liegt darin, daß für diese Teilnehmer es die erste Fortbildungsveranstaltung war und ihnen deshalb die Vergleichsmöglichkeit fehlte.

Auch in der zweiten Erprobungsphase waren die Antworthäufigkeiten schief verteilt, positive Einschätzungen über-

wiegen weitgehend. Nach Tabelle 2 gibt es (unter  $p_1\%$ ) nur zwei Items, deren Schwierigkeiten bei 50 % liegen, alle anderen liegen zum Teil weit niedriger. Aus diesem Grunde wurden wieder die gleichen Transformationen wie in der ersten Phase angeschlossen. Durch diese Transformationen entstehen quasi normalverteilte Variablen, die oft die Voraussetzung für die Anwendung der statistischen Verfahren sind.

- alle Fragen wurden so gedreht, daß die Einschätzung "Stimmt völlig (4)" ein für die Veranstaltung positives Bild gibt. Diese 4-stufigen, gedrehten Fragen werden mit ITEM angesprochen (diese ITEMS dichotomisiert werden DITEM genannt und für sie wurden auch die Schwierigkeiten in Tabelle 2 unter  $p_1\%$  berechnet).

- anschließend wurden die ITEMS prozentrang transformiert und in z-Äquivalente umgewandelt; die entstehenden 4-stufigen, transformierten Items werden mit TRAI angesprochen.

- die Fragen TRAI wurden wieder dichotomisiert, die transformierten dichotomisierten Fragen werden TRADI genannt; deren Schwierigkeiten, d.h. % Anteile der Teilnehmer, die die Veranstaltung unterdurchschnittlich beurteilen, sind unter  $p_2\%$  in Tabelle 2 aufgelistet.

Anlage 13 enthält die neuen Transformationsgewichte; Anlage 14 u.a. einige Statistiken von ITEM, DITEM, TRAI und TRADI.

Das folgende Schema faßt die Transformationen zusammen:

	Items, gleichgepolt	
	Originalfragen	%rang- und z-trans- formiert
4-stufig	ITEM	TRAI
dichotomisiert	DITEM	TRADI

Die Transformationen erschweren aber die schnelle Auswertung und Interpretation, vor allem durch die Teilnehmer selbst. Aus diesem Grunde wurden alle Analysen auch mit den einfachen, nur gleich-gepolten Originalfragen ITEM berechnet. Falls sich hier bei verschiedenen Transformationen keine Unterschiede in den Ergebnissen zeigen, könnte in der praktischen Anwendung mit den einfacher zu interpretierenden Originalitems gearbeitet werden.

Für alle 4 verschiedenen Transformationsarten der Fragen (ITEM, DITEM, TRAI und TRADI) wurden jeweils durch Aufaddieren der Werte ein Gesamtwert gebildet und die entsprechenden Trennschärfen berechnet. Dies führte dazu, die Frage 14 (ausreichende Dauer der Veranstaltung) aus den weiteren Analysen auszuschließen.

Bei der Bildung der Gesamtwerte wurden fehlende Werte bei einzelnen Fragen durch den Mittelwert einer bestimmten Transformationsart ersetzt, z.B. ist bei ITEM  $\bar{x} = 3.4$  der Mittelwert aller Mittelwerte von ITEM 1 bis ITEM 35. Alle Trennschärfenkoeffizienten sind auf dem 0,1 % Niveau signifikant von Null verschieden, mit Ausnahme von Frage 14 ( $r_{it} = -.04$  bis  $.03$ ). Die Frage 14 wurde deshalb ausgeschlossen, da sie mit dem Gesamt der anderen Fragen anscheinend keinen Zusammenhang hat.

Anschließend wurden die Gesamtwerte neu addiert (d.h. nun neben Frage 15, 20 und 34 auch ohne Frage 14).

Berücksichtigt wurden nur die Teilnehmer, die mindestens 3/4 der Fragen (entspricht ungefähr 23 Fragen) beantwortet haben, dies sind 475 Teilnehmer. Die Gesamtwerte zeigen mindestens auf dem 1 % Niveau signifikante Gruppenunterschiede auf ( $F = 9.0$  bis  $F = 11.1$  bei  $F_{.01;24;453} = 1.84$ ).

Alle Korrelationen von ITEM, DITEM, TRAI und TRADI zu ihren entsprechenden Gesamtwerten sind mindestens auf dem 0,1 % Niveau signifikant von Null verschieden; sie reichen von  $r_{it} = .19$  bis  $r_{it} = .64$  (siehe Anlage 14).

Für diese 31 Fragen wurde auf Gesamtfragebogenebene auf verschiedenen Wegen die Reliabilität berechnet (zur Reliabilität auf Skalenebene vgl. Kapitel 6.1.6).

Eine erste Schätzung berücksichtigte nur die Teilnehmer, die alle Fragen beantwortet haben (339). Für diese wurde CRONBACHS ALPHA berechnet<sup>4</sup>. Je nach den Transformationen d.h. ob nun ausgegangen wird von ITEM, DITEM, TRAI oder TRADI ergibt sich ein  $r_{tt}$  von  $.82$  bis  $.87$ . Bildet man jeweils zwei Hälften zufällig und korreliert sie, so beträgt  $r_{tt} = .69$  bis  $.73$  (korrigiert nach SPEARMAN BROWN

<sup>4</sup> Dieses Reliabilitätsmaß ist für dichotomisierte Variablen identisch mit KUDER RICHARDSON FORMULA 8; ALPHA stand erst in der SPSS Version 8 zur Verfügung (vgl. HULL/NIE 1979, bes. S. 125f.).

$r_{tt} = .82$  bis  $.87$ ).

Berechnet man die Reliabilität aufgrund der Teilnehmer, die mindestens 23 Fragen (entspricht 74,2 %) beantwortet haben, so ergeben sich die gleichen Werte ( $\text{ALPHA} = .80$  bis  $.86$ ; nach Split-Half  $r_{tt}$  (korr.) =  $.81$  bis  $.86$ ).

Unabhängig von der Methode der Reliabilitätsschätzung ergibt sich eine hohe Reliabilität von  $r_{tt} = .82$  bis  $r_{tt} = .87$ . Mit Hilfe des Fragebogens werden somit relativ genaue und zuverlässige Informationen erhoben, die unter diesem Aspekt einen guten Input für Auswertungsgespräche bieten können. Besonders wichtig ist die Unabhängigkeit der Reliabilität von der Transformation der Items, d.h. die gute und zufriedenstellende Reliabilität gilt auch für die Originalfragen ITEM.

#### 6.1.2 Polungseffekte

Im Fragebogen waren 17 Aussagen so formuliert, daß eine Zustimmung die Veranstaltung positiv einschätzt, während bei den anderen Aussagen Zustimmung die Veranstaltung negativ beurteilt (für die Analysen wurden diese Aussagen 'gedreht'). Um zu verfolgen, ob diese Formulierung zu einer unterschiedlichen Beantwortung führt, wurde jeweils ein Gesamtwert aus den positiven und den ursprünglich negativ formulierten Fragen gebildet.

Nachdem vorher die Fragen 14, 15, 20 und 34 ausgeschlossen wurden, wurde hier - nach Zufall - noch die Frage 10 herausgenommen, um jeweils 15 positiv und negativ formulierte Fragen zu haben. Betrachtet wurden nur die Personen, die mindestens 22 Fragen (entspricht 73,3 %) beantwortet haben; dies sind 475 Teilnehmer. Bei fehlenden Werten wurde dann der Mittelwert eingesetzt. Die Summe der positiv formulierten Fragen wird POITEM, die der - ursprünglich - negativ formulierten NEITEM genannt; entsprechend: POTRAI-NETRAI, PODITEM-NEDITEM, POTRAI-NETRAI.

Ein t-Test für korrelierende Stichproben (vgl. CLAUSS/EBNER 1970, S. 217ff.) weist auf signifikante Unterschiede hin (siehe Tabelle 3).



Tabelle 3 t-Test

	$\bar{x}$	Diff.	t-Wert
POTRAI NETRAI	-0.7 -1.0	0.3	1.4
POITEM NEITEM	47.6 53.1	5.5	21.1
POTRADI NETRAI	7.7 10.5	2.8	20.4
PODITEM NEDITEM	11.8 13.4	1.6	14.5

der kritische Wert beträgt:

$$t_{.01;474} = 2.59$$

(für die einseitige Fragestellung:  $t_{.01;474} = 2.34$ )

Nur für die Vergleiche von POTRAI und NETRAI lassen sich keine signifikanten Unterschiede feststellen, in allen anderen Fällen jedoch hochsignifikante. Damit scheinen die ursprünglich negativ formulierten Aussagen im Gesamt für die Veranstaltung positiver beantwortet zu werden als die positiv formulierten. Dies läßt zwei Interpretationen zu:

- Die positiven und negativen Fragen wurden - um den Fragebogen nicht zu verdoppeln - nicht streng parallel formuliert, sondern es wurde nach Zufall die Formulierungsrichtung bestimmt. Die unterschiedliche Beantwortung ließe sich dadurch erklären, daß beide Teile unterschiedliche Aspekte einer Veranstaltung erfragen. Darauf weisen auch die niedrigen Korrelationen der jeweiligen Fragebogenhälften hin ( $r=.5$  bis  $r=.6$ ; POITEM und NEITEM haben z.B. nur ca. 33 % der Varianz gemeinsam).
- Eine negativ formulierte Aussage wird von den Teilnehmern stärker abgelehnt als einer positiven Formulierung zugestimmt wird; d.h. ein negativer Aspekt wird eher abgelehnt als ein positiver anerkannt. Dies kann als eine Differenzierung der Tendenz der Milde interpretiert werden.

Welche der beiden Erklärungsmöglichkeiten letztlich zutrifft, läßt sich bei dem hier durchgeführten Versuchsplan nicht entscheiden.

Die Reliabilitäten in den Fragebogenhälften stimmen untereinander und mit dem Gesamtbogen überein (z.B. ALPHA bei POTRAI=.79, bei NETRAI=.77; aufgewertet auf 31 Fragen ergibt sich  $r_{tt}=.89$  und  $.88$ ). In allen 8 Variablen zei-

gen sich hochsignifikante Unterschiede zwischen den 25 Veranstaltungen. Die oben beschriebene "Zustimmungstendenz" wirkt sich also nicht so aus, daß sich die Gruppen in den negativ formulierten Fragen nicht mehr unterscheiden.

Da nicht ausgeschlossen werden kann, daß sich in der Beantwortung der Items ein Polungseffekt niederschlägt, sollten bei weiteren Revisionen des Fragebogens nur gleich gepolte Aussagen aufgenommen werden; es bieten sich hier die positiven Aussagen an, da sie - im Unterschied zu den negativen - tendenziell weniger "schief" beantwortet wurden (vgl. Tabelle 3, der Mittelwert der Zusammenfassungen der positiven Aussagen liegt immer deutlich niedriger, also weniger "günstig" für eine Veranstaltung, als bei den negativen Aussagen). Für den praktischen Einsatz hätte eine gleiche Polungsrichtung zudem noch den Vorteil einer leichteren Auswertung; dies gilt vor allem für eine Auswertung auf Skalenebene - siehe hierzu weiter unten -, da dann bei einer Addition der Fragen ein "Drehen" entfallen würde.

### 6.1.3 Faktorenanalysen

Im Rahmen der Faktorenanalysen sollte einmal die interne Strukturierung der Items im Sinne einer faktoriellen Validität untersucht und die Faktorenstruktur mit der vorherigen theoretischen Strukturierung verglichen werden. Zum anderen dient die Faktorenanalyse zur Abschätzung des Ausmaßes des Halo-Effekts.

Die Faktorenanalysen wurden zuerst mit den transformierten, 4-stufigen Fragen TRAI gerechnet<sup>5</sup>.

<sup>5</sup> Diese Analyse umfaßte alle Fragen ohne 14, 15, 20 und 34. Probeweise wurden 4-Faktorenlösungen mit allen Fragen bzw. mit nur 27 Fragen (ausgeschlossen alle transformierten Fragen TRAI, in denen 20 % der Teilnehmer eine Veranstaltung negativ beurteilen) gerechnet. Die Struktur der Faktoren entspricht sich, ob nun mit 27, 31 oder 35 Fragen gerechnet wird. Über diese Analysen wird hier nicht weiter berichtet.

Die verschiedenen Kriterien - vgl. Kapitel 5 - sprechen für eine Faktorenzahl von 1 - 6 Faktoren. Sukzessiv wurden 3, 4 und 5 Faktoren extrahiert (Modell gemeinsamer Faktoren; mit Iteration, Abbruch bei Differenzen kleiner als .01) und varimaxrotiert.

	bedeuts. Ladungen auf Faktor					Gesamtzahl an bedeutsamen Ladungen	% an aufgeklärter Gesamtvarianz
	1	2	3	4	5		
3 FL	16	9	5			30	26,8 %
4 FL	8	11	6	5		30	29,4 %
5 FL	6	6	8	2	3	25	31,2 %

Bei der 5-Faktorlösung ist ein Faktor nur durch eine Doublette gekennzeichnet; 4 Faktoren sind differenzierter als 3 und klären zudem leicht mehr Varianz auf. Insgesamt ist die 4-Faktorlösung - von der technischen Seite her - sehr gut bestimmt, nur eine Frage definiert einen Faktor nicht bedeutsam.

Nach dem FÜRNRATT-Kriterium sind 4 Faktoren optimal. Diese Lösung soll nun interpretiert werden (Kommunalitäten und Ladungen siehe Tabelle 4, ausführlicher Anlage 15).

Die 4 Faktoren erklären 29,4 % an der Gesamtvarianz; auf die einzelnen Faktoren entfallen dabei (in Klammern die % Anteile bezogen auf die aufgeklärte Varianz):

F1	F2	F3	F4	Gesamt
8,9(30,3)	8,2(27,9)	7,9(26,7)	4,5(15,2)	29,5(100,1)

Die aufgeklärte Varianz liegt damit um fast 10% niedriger als in der 1. Phase.

Tabelle 4 Ladungen und Kommunalitäten

Nr.	$h_i^2$	F1	F2	F3	F4	Nr.	$h_i^2$	F1	F2	F3	F4
1	.36	-	-	.47*	-	19	.34	-	.30	.35*	-
2	.19	-	.38*	-	-	21	.25	.37*	.31	-	-
3	.10	-	-	-	.30*	22	.21	-	.38*	-	-
4	.35	.45*	-	.32	-	23	.44	-	-	.61*	-
5	.22	-	.41*	-	-	24	.38	-	-	.30	.49*
6	.20	-	.40*	-	-	25	.24	.45*	-	-	-
7	.23	-	.40*	-	-	26	.19	-	.42*	-	-
8	.35	-	-	.31	.47*	27	.41	.51*	-	.30	-
9	.29	.41*	-	.34	-	28	.36	-	.50*	-	-
10	.27	-	-	.38*	-	29	.15	.26	-	-	-
11	.44	.55*	.33	-	-	30	.34	-	-	-	.57*
12	.48	.62*	-	-	-	31	.42	.31	-	.47*	-
13	.54	.60*	-	.37	-	32	.16	-	.29*	-	-
16	.42	-	-	.60*	-	33	.17	-	.29*	-	-
17	.26	-	.44*	-	-	35	.16	-	.29*	-	-
18	.35	-	-	-	.55*						

(Aufgelistet sind alle Ladungen größer als .3; sind alle Ladungen kleiner als .3, so ist die jeweils größte angegeben. Mit \* sind alle bedeutsamen Ladungen mit  $a_{ij}^2/h_i^2 \geq .5$  gezeichnet).

Zuerst die beiden deutlichsten Faktoren:

- Faktor 4 "äußerer Rahmen"

Der 4. Faktor wird von den Fragen 3, 8, 18, 24 und 30 jeweils bedeutsam definiert; keine weitere Frage hat eine höhere Ladung (größer .3) auf diesem Faktor. Gefragt wird hier nach den Raumverhältnissen und den äußeren Bedingungen der Veranstaltung (18, 30) sowie nach der Angemessenheit der Teilnehmerzahl (3). Die Bedingungen des Hauses stecken auch zugleich den Rahmen ab, in dem eine gute Gestaltung der Freizeit möglich ist (8, 24).

Dieser Faktor reproduziert eindeutig den 5. Faktor "äußerer Rahmen" aus der 1. Phase. Die Frage nach dem informellen Erfahrungsaustausch, die vorher noch diesen Faktor bestimmte, geht nun eindeutig zu dem neuen Faktor 2 "Arbeitsklima" über, wo sie auch inhaltlich besser zu interpretieren ist.

#### - Faktor 2 "Arbeitsklima"

11 Fragen definieren diesen Faktor bedeutsam, er läßt sich eindeutig mit "Arbeitsklima" benennen. Inhaltlich wird der Faktor am stärksten durch die Möglichkeit, abweichende Meinungen frei und ohne Angst zu äußern (26), durch "Wohlfühlen" in der Gruppe (6, 7), durch eine Berücksichtigung der Vorstellungen der Teilnehmer bei Entscheidungen über die Veranstaltung (28), durch die Möglichkeit zu aktiver Mitarbeit (2, 17), durch das Eingehen auf Verständnisschwierigkeiten (5) und durch einen informellen Erfahrungsaustausch (22) bestimmt. Es liegt hier ein eindeutig identifizierbarer Faktor vor, der im Verhältnis zur 1. Phase - inhaltlich sinnvoll - neben Arbeitsklima ebenfalls Aspekte der Aktivität und einer Teilnehmerpartizipation umfaßt; zu interpretieren ist dies in dem Sinne, daß das allgemeine Arbeitsklima den Rahmen für aktive Mitarbeit absteckt, bzw. die Komponente der Aktivität das allgemeine Arbeitsklima weitgehend mitbestimmt. Es sei hier an die besondere Bedeutung dieses Bereiches erinnert. Erzieher können die intergenerationelle Reziprozität - vgl. Kapitel 3.1 - nur dann erreichen, wenn sie gleichberechtigte Partner ihrer Fortbildungsarbeit sind. Faktor 2 kann hierfür ein Indikator sein.

Bei den Faktorenanalysen in Phase 1 ließen sich zwei verschiedene Faktoren identifizieren, von denen einer Arbeitsklima und der andere den engeren Aspekt der Aktivität der Teilnehmer umfaßte. In diesem Sinne wurden auch Fragen neu formuliert: Frage 6 und 7 zu Arbeitsklima und Frage 28 zu Aktivität. In der 2. Phase ließen sich aber keine 2 Faktoren mehr identifizieren, sondern ein sehr stabiler Faktor "Arbeitsklima", der ebenfalls Aspekte der Aktivität (2, 17) umfaßt. Zugleich gehen aber auch einige Fragen (11, 12, 21), die in der 1. Phase den Faktor Aktivität schwächer bestimmten (höchste La-

dung , aber nicht bedeutsam, also recht komplexe Fragen) auf den neuen Faktor 1 über. Frage 11 und 21 (nicht zu trockene Veranstaltung und Teilnehmerpartizipation) laden aber mit .33 und .31 recht hoch auf dem neuen Faktor Arbeitsklima.

Etwas schwieriger zu interpretieren sind die beiden restlichen Faktoren:

- Faktor 3 "Lernertrag"

6 bedeutsame Ladungen definieren diesen Faktor, wobei er am stärksten markiert wird durch den fachlichen Ertrag einer Veranstaltung (16) und der damit zusammenhängenden Möglichkeit, praktische Probleme besser angehen zu können (23). Letzteres ist natürlich nur dann möglich, wenn die Auswahl der Themen auch die wichtigsten Praxisprobleme trifft (1), bzw. wenn die wichtigsten Aspekte des Themas herausgearbeitet werden (10). Eine Einschätzung des möglichen Lernertrages ist für die Teilnehmer u.a. nur dann möglich, wenn sie schon in der Veranstaltung Informationen über ihre Lernfortschritte (19) bekommen. In diesem Sinne befaßt sich der Faktor mit der Einschätzung des Lernertrages einer Veranstaltung und der dazu förderlichen methodischen Aspekte. Zugleich geht dies aber zusammen mit einer Globaleinschätzung der Veranstaltung (31), d.h. im Gesamt als gut wird eine Veranstaltung anscheinend dann eingeschätzt, wenn ihr vermuteter Lernertrag hoch ist.

Gestützt wird diese Interpretation durch jene Fragen, die den Faktor 1 bedeutsam definieren, aber noch höhere Anteile auf dem 3. Faktor haben. Es sind dies die Aspekte des Lernertrages (9), einer allgemeinen Zufriedenheit mit dem Gelernten (4) und dem Besuch der Veranstaltung (13) sowie der Übersichtlichkeit des Aufbaus der Veranstaltung (27). Letzteres läßt sich so interpretieren, daß der übersichtliche Aufbau dem Teilnehmer hilft, sich zu vergegenwärtigen, wo er im Lernprozeß steht und was er bereits gelernt hat.

### - Faktor 1 "Methodischer Aufbau/Gesamteinschätzung"

Die 8 bedeutsamen Ladungen dieses Faktors sind schwieriger zu interpretieren. Am stärksten wird er durch angemessene Anforderungen an die Teilnehmer (25) und einer Übereinstimmung zwischen der Veranstaltung und den Teilnehmervorstellungen bezüglich eines sinnvollen Lernens (12) bestimmt. Es folgen eine "nicht zu trockene" Veranstaltungsdurchführung (11), ein "lohnender" Besuch (13) und eine übersichtliche Gliederung (27); zugleich wird dieser Faktor - zwar weniger - durch die Einschätzung des Lernertrages (4, 9) bestimmt. Teilweise kommt hierzu noch ein Aspekt der Teilnehmerpartizipation (21). Beziehen sich zwar mehrere Fragen direkt auf den methodischen Aufbau der Veranstaltung (12, 25, 27), so scheint dies hier aber sehr eng verbunden mit einer eher gesamthaften Einschätzung der Veranstaltung. Der Stil, der Aufbau und die Anforderungen, der allgemein lohnende Besuch wird betrachtet, während Faktor 3 "Lernertrag" eher die Einschätzung des Ertrages einer Veranstaltung betrachtet. Auf die Vermischung der beiden Faktoren weisen aber die bedeutsamen Ladungen der Fragen nach dem Lernertrag auf Faktor 1 (4, 9) und der Gesamteinschätzung auf Faktor 3 (31) hin. Faktor 1 soll deshalb sehr vorsichtig mit "Methodischer Aufbau/Gesamteinschätzung" benannt werden.

Die Faktoren 1 "Methodischer Aufbau" und 3 "Lernertrag und Globaleinschätzung" der 1. Phase sind somit nicht eindeutig reproduzierbar; vielmehr spalten sie sich auf und verteilen sich - je zur Hälfte - auf die neuen Faktoren 1 und 3. Teilweise ist dies nicht überraschend, da viele Items der Voruntersuchung sehr komplex waren und sowohl auf dem alten Faktor 1 als auch auf Faktor 3 höhere Ladungen hatten. Die neuen Faktoren dagegen sind technisch gesehen eindeutiger, d.h. durch mehr bedeutsame Ladungen markiert.

Ebenso verteilen sich die zu "Lernertrag und Globaleinschätzung" neu formulierten Fragen 4 und 23 auf die neuen Faktoren 1 und 3.

Berücksichtigt man, daß die Revision nach der Voruntersuchung nicht Fragen mehr oder weniger verdoppelte, sondern neue Fragen auch neue - hypothetisch den Dimensionen zuordbare - Aspekte ansprechen sollten und daß die 2. Phase einen neuen Antwortmodus verwendete, so ist die Reproduzierbarkeit der Dimensionen teilweise zufriedenstellend. Zwei Faktoren - äußerer Rahmen und Arbeitsklima - lassen sich eindeutig reproduzieren; bei den beiden anderen ist der Bezug zur 1. Phase zwar erkennbar, wenngleich auch die Aspekte gegenseitig vermischt sind. Damit zeigen sich auch keine gravierende Unterschiede zu der theoretischen Strukturierung der Erfolgskriterien in der Literaturanalyse (vgl. auch S. 138 ); der direkt methodische Bereich tritt nicht mehr so klar und abgegrenzt auf, sondern scheint relativ eng mit der Gesamteinschätzung verbunden.

In dieser Struktur der 4 Faktoren befindet sich - ebenso wie in der 1. Phase - auch kein allgemeiner Faktor, auf dem alle Items höher laden, sondern eine spezifische, komplexe Faktorenstruktur. Ein Halo-Effekt dürfte demnach wiederum nur niedrig ausfallen.

Um auszuschließen, daß die Struktur der Items und die 4 Faktoren abhängig sind von bestimmten, methodischen oder modellbedingten Besonderheiten bei dem ausgewählten Vorgehen, wurden weitere Faktorenanalysen mit unterschiedlichen Akzentuierungen berechnet. Es zeigt sich, daß die 4-Faktorlösung dabei sehr stabil ist:

- Die Vermischung des Faktors 1 und 3 könnte evtl. darauf zurückgeführt werden, daß die Bedingungen der Orthogonalität der Faktoren dem Daten- und Itemsatz unangemessen ist.



Eine schiefwinkelige Rotation führt aber kaum zu einer Verbesserung der Interpretation<sup>6</sup>.

Erlaubt man den Faktoren nur eine geringe Korrelation (z.B. DELTA = -.4), so entspricht die schiefwinkelige Rotation der orthogonalen Lösung. Läßt man eine extreme Korrelation zu (z.B. DELTA = .4), so erhält man 3 klare, aber hoch korrelierende Faktoren "Arbeitsklima", "äußerer Rahmen" und "Gesamteinschätzung/Methodischer Aufbau". Diese sehr komplexen Faktoren - vor allem der letzte - rechtfertigen aber kaum die Arbeit mit korrelierenden Faktoren. Deshalb wurde im folgenden mit den orthogonalen 4 Faktoren weitergearbeitet (siehe auch zur Skalenbildung Kapitel 6.1.6).

- Die 4-Faktorlösung ist unabhängig von den Modellvorstellungen der Faktorenanalyse, d.h. ob nach dem Hauptkomponentenmodell oder nach dem Modell der gemeinsamen Faktoren vorgegangen wird.

Neben dem oben beschriebenen Vorgehen wurde eine Faktorenanalyse nach dem Hauptkomponenten-Modell (Einheiten in der Diagonalen der Korrelationsmatrix; keine Iteration) gerechnet. Diese Lösung ist in der Struktur fast völlig identisch mit der 4-Faktorlösung nach dem Modell der gemeinsamen Faktoren. Ein Unterschied besteht darin, daß nach diesem Modell ein Varianzanteil von 38,2 % aufgeklärt wird, d.h. bei der Beurteilung des Anteils an aufgeklärter Varianz muß besonders die Modellvorstellung der Faktorenanalyse berücksichtigt werden. Diese Ähnlichkeit der Faktorenanalyse bei verschiedenen Modellvorstellungen wird auch an anderer Stelle erwähnt (vgl. GAENSSLEN/SCHUBÖ 1973, S. 278f.).

---

<sup>6</sup> Im Unterschied zur relativ einfachen orthogonalen Rotation kompliziert sich die Faktorenanalyse bei einer schiefwinkelligen Rotation enorm (vgl. die ausführliche Diskussion bei ÜBERLA 1971, S. 167ff.; GAENSSLEN/SCHUBÖ 1973, S. 252-254); einmal korrelieren die Faktoren mehr oder weniger miteinander in Abhängigkeit eines vom Untersucher festgesetzten Wertes DELTA (hier bei der OBLIMIN-Rotation, der ROTATION WITH KAISER NORMALIZATION, vgl. NIE u.a. 1975, S. 486), zum anderen gibt es nun die Matrix der Faktorladungen (Pattern; Regressionsgewichte der gemeinsamen Faktoren auf die Variablen) und die Matrix der Faktor-Struktur (Korrelation der Variablen mit den Faktoren). Da SPSS vom Modell der Primärfaktoren ausgeht, ist die Einfachstruktur am besten in der Ladungsmatrix zu erkennen (vgl. ib.).

- Die 4-Faktorlösung ist unabhängig davon, ob die gedrehten Originalfragen ITEM, die 4-stufigen, transformierten Fragen TRAI oder die transformierten und dichotomisierten Fragen TRADI betrachtet werden. Diese Unabhängigkeit der Faktorenstruktur vom Auswertungsmechanismus ist für die weitere Arbeit von besonderer Bedeutung; auch wenn in der praktischen Anwendung des Fragebogens mit den Originalfragen ITEM gearbeitet wird, kann die hier beschriebene Faktorenstruktur unterstellt werden<sup>7</sup>.

Bei ITEM sind die Faktoren 1 und 4 völlig identisch, während bei den Faktoren 2 und 3 insgesamt 3 Fragen zwar noch am höchsten, aber nicht mehr bedeutsam laden (siehe Anlage 18). Betrachtet man die dichotomisierten Fragen TRADI, so geht mit der Dichotomisierung ein Informationsverlust einher, indem nur noch 25,9% der Gesamtvarianz aufgeklärt wird. Die Itemstruktur ist aber weitgehend identisch.

- Die 4-Faktorlösung ist weitgehend unabhängig von der Auswahl der in die Analysen einbezogenen Veranstaltungen; d.h. es zeigen sich keine unterschiedlichen Faktorenstrukturen, wenn verschiedene Fortbildungsveranstaltungen untersucht werden. Für die praktische Arbeit mit dem Fragebogen ist diese Stabilität insofern von Bedeutung, als die hier gefundene Struktur - mit der nötigen Vorsicht - auf die evaluierte Veranstaltung übertragen werden kann.

Die Abhängigkeit von der Auswahl der Veranstaltungen ließ sich nur andeutungsweise klären, indem die Gesamtstichprobe in Teilstichproben aufgeteilt wurde, obwohl zu erwarten ist, daß der Zufallsfehler dann vergrößert wird und eher weniger stabile Faktoren zu erwarten sind (vgl. GAENSSLEN/SCHUBÖ 1973, S. 261). Zufällig wurden 4 Sets von je 14 Veranstaltungen ausgewählt. Bei 2 Sets stimmen die 4-Faktorlösungen sehr gut zueinander und mit der Lösung der Gesamt-

<sup>7</sup> Eine Verteilungsschiefe der Items behindert eine Faktorenanalyse nicht unbedingt. Nach WITTE - der echt J-verteilte Items untersucht - kann die Annahme der Normalverteilung fallengelassen werden, wenn auf eine statistische Beurteilung verzichtet wird. Sind alle Items ähnlich schief verteilt, wie im vorliegenden Fall -, so gibt es bei der Rotation und Interpretations keine zusätzlichen Probleme (vgl. WITTE 1978).

gruppe von 25 Veranstaltungen überein, bei dem 3.Set ist die Übereinstimmung nicht so deutlich, läßt sich aber interpretativ erkennen; nur bei dem 4.Set läßt sich kaum eine gute Übereinstimmung feststellen. Die 4-Faktorlösung der Gesamtgruppe scheint also relativ unabhängig von der Auswahl der Veranstaltungen.

Als Ergebnis dieser Faktorenanalysen läßt sich festhalten, daß der Halo-Effekt vermutlich niedrig ausfällt. Der Fragebogen erhebt somit nicht nur einen globalen Eindruck, dem alle Einzeleinschätzungen untergeordnet werden, sondern eher spezifische Informationen. Zugleich zeigt der Bogen eine faktorielle Validität. Es scheint gerechtfertigt, mit diesen Faktoren weiterzuarbeiten und sie einer Skalenbildung bzw. einer weiteren Revision des Fragebogens zugrunde zu legen.

#### 6.1.4 Regressionsanalysen

Eine weitere Abschätzung des Halo-Effekts sollte mit Hilfe der Regressionsanalyse versucht werden (vgl. Kapitel 4.3, S.118f). Ebenso können hier zwei inhaltliche Fragen verfolgt werden: einmal die Frage nach dem Präferenzsystem der Teilnehmer - d.h. nach jenen Veranstaltungsaspekten, die diese für besonders relevant halten - und zum anderen die Frage, ob auch alle relevanten Veranstaltungsaspekte durch den Fragebogen erfaßt werden (s.S. 124f.).

Die Fragen 13 (lohnender Besuch) und 31 (Globaleinschätzung) wurden wieder aufaddiert und als Kriteriumsvariable betrachtet. Die Vorhersage soll dann durch alle übrigen Fragen als Prädiktoren geschehen.

Die Regressionsanalyse wurde nur für die Teilnehmer berechnet, die mindestens eine der beiden Fragen 13 oder 31 beantwortet haben; fehlt ein Wert, so wurde er durch den Mittelwert ersetzt. Frage 13 stammt aus Faktor 1 und Frage 31 aus Faktor 3; beide Fragen korrelieren mit .63 am höchsten mit dem Gesamtwert gebildet aus der Summe aller Fragen TRAI.

Berichtet wird hier nur über die Regressionsanalysen mit den transformierten, 4 stufigen Fragen TRAI (siehe Tabelle 5, S. 166), da die Analysen bei verschiedenen Transformationen weitgehend übereinstimmen.

Für die Fragen TRAI, ITEM und TRADI wurde je eine einfache multiple Regressionsanalyse gerechnet. Für ITEM wurde eine multiple Korrelation von  $R=.74$ , für TRADI von  $R=.67$  berechnet. Dies stimmt mit dem unten beschriebenen R bei TRAI überein; mit der Dichotomisierung TRADI geht ein leichter Informationsverlust einher, der zu einem niedrigeren R führt. Zu der Analyse TRAI zeigte sich bei ITEM und TRADI eine gute Übereinstimmung in den BETA-Gewichten und Regressions-Faktor-Struktur-Koeffizienten (RFSK), vor allem in den Markiervariablen der künstlichen Dimension. Eine gute Übereinstimmung fand sich zwischen TRAI, ITEM und TRADI bei der unten beschriebenen schrittweisen Regressionsanalyse.

- Bei einer einfachen multiplen Regressionsanalyse korreliert die Linearkombination aus allen Prädiktoren mit dem Kriterium mit  $R=.74$ ; dies ergibt eine  $SMC=.55$  (korrigiert  $=.52$ ). R ist signifikant von Null verschieden mit  $F=16.65$  bei einem kritischen Wert von  $F_{.01;29.393} = 1.75$ . Bei dieser - relativ niedrigen - multiplen Korrelation ist aber zu bedenken, daß bei der vermutlich niedrigen Reliabilität eines nur aus zwei Items (13 und 31) zusammengesetzten Kriteriums eine hohe Korrelation kaum erwartet werden kann. Bei  $R=.74$  erklärt die Linearkombination nur ungefähr 55 % der Varianz des Kriteriums. Es sind hier zwei Interpretationen möglich. Setzt man einmal voraus, daß das Kriterium eine gute und zufriedenstellend reliable Globalbeurteilung einer Fortbildungsveranstaltung wäre, so würden die in dem Fragebogen zusammengestellten Fragen/Erfolgskriterien nur gut die Hälfte von dem erfassen, wonach die Teilnehmer eine Veranstaltung global beurteilen. Damit wären also nicht alle für die Teilnehmer wichtigen Aspekte in dem Fragebogen aufgenommen worden. Zum anderen ist aber fraglich, ob das aus den Items 13 und 31 additiv zusammengesetzten Kriteriums eine gute Globaleinschätzung

darstellt. Es kann eher vermutet werden, daß ein solches Kriterium die Fähigkeit der Teilnehmer überfordert, viele Einzelaspekte einer Veranstaltung in einem Globalurteil zu integrieren. Dieses Globalurteil würde dann gar nicht die unterschiedlichen Bereiche wie Methodik, Arbeitsklima oder Lernprodukte erfassen bzw. nach Aspekten geschehen, die davon unabhängig sind. Ein solcher Aspekt wäre z.B. daß eine Veranstaltung danach global beurteilt wird, inwieweit sie es ermöglicht, aus dem Alltagshandeln herauszukommen und befriedigende Kontakte mit Kollegen aufzunehmen.

Das hier beschriebene regressionsanalytische Vorgehen - in Anlehnung an einen Ansatz in der Lehrerbeurteilung durch die Schüler (vgl. FRENCH-LAZOVIK 1974) - bleibt somit solange relativ unbefriedigend, wie nicht das Kriterium unabhängig von dem Itemsatz reliabel erhoben werden kann und zugleich wirklich eine "Global-/Gesamtbeurteilung" darstellt.

- Die aus den Prädiktorvariablen gebildete Linearkombination bildet eine künstliche Dimension, die maximal mit dem Kriterium korreliert. Es wurde nun gefragt, wodurch diese künstliche Dimension inhaltlich bestimmt ist, d.h. welche Veranstaltungsaspekte im einzelnen für die Teilnehmer besonders relevant zur Vorhersage der Gesamtbeurteilung sind. Hierbei wurde von den Regressions-Faktor-Struktur-Koeffizienten ausgegangen.

Es wurden nur die Fragen TRAI berücksichtigt, die einen signifikanten Beitrag zur Linearkombination leisten, konkret, deren F-Wert größer als 3.86 ist (siehe Tabelle 5). Die BETA-Gewichte sind die Gewichte der Vorhersage eines Item  $i$ , wenn vorher die Effekte von allen anderen Items aus  $i$  herauspartialisiert werden. Solche "Restitems" lassen sich aber inhaltlich kaum interpretieren. Inhaltlich interpretiert werden kann der Varianzbeitrag, den ein Item  $i$  an der künstlichen Dimension aufklärt; dies ist das Quadrat der einfachen Korrelation eines Items  $i$  mit der Linearkombination, auch Regressions-Faktor-Struktur-Koeffizient (RFSK) genannt (vgl. GAENSSLEN/SCHUBÖ 1973, S. 122f.).

Tabelle 5: Regressionsanalysen TRAI

Nr.	(1) BETA	(2) F	(3) r	(4) RFSK <sup>2</sup>	Nr.	(1) BETA	(2) F	(3) r	(4) RFSK <sup>2</sup>
1	.11	6.6	.44	.35	25	.04	1.0	.31	.18
2	.02	0.3	.23	.10	26	-.13	12.2	.10	.02
3	.00	0.0	.11	.02	27	.05	1.3	.45	.37
4	.21	26.1	.49	.44	28	.07	2.3	.34	.21
5	.04	0.9	.27	.13	29	.07	3.4	.30	.16
6	.07	3.4	.31	.18	30	.05	1.4	.08	.02
7	.09	4.6	.31	.18	32	-.03	0.5	.21	.08
8	.06	2.2	.21	.08	33	.00	0.0	.25	.11
9	.10	7.1	.38	.26	35	.04	1.2	.24	.11
10	.09	4.8	.39	.28					
11	.05	1.2	.39	.28					
12	.15	11.4	.47	.40					
16	.12	7.8	.42	.32					
17	.01	0.0	.30	.16					
18	-.04	1.2	.04	.00					
19	.03	0.7	.31	.18					
21	.02	0.2	.31	.18					
22	-.01	0.1	.25	.11					
23	.02	0.1	.39	.28					
24	.07	2.6	.24	.11					

(1) BETA-Gewichte

(2) Der kritische Wert für F beträgt auf dem 5 % Niveau 3.86

(3) Einfache Korrelation eines Prädiktors mit dem Kriterium

(4) RFSK = Regressions-Faktor-Struktur-Koeffizient; einfache Korrelation des Items j mit der aus allen Prädiktoren gebildeten Linearkombination; das Quadrat gibt die gemeinsame Varianz an

Schrittweise Regressionsanalyse

Nr.	(1) BETA	(2) F	(3) r	(4) RFSK <sup>2</sup>
4	.23	33.9	.50	.47
12	.20	25.3	.47	.42
16	.15	14.5	.42	.33
1	.12	9.7	.44	.36
24	.12	10.6	.24	.11
9	.12	10.0	.38	.27
7	.09	5.3	.31	.18
10	.11	8.8	.39	.28
26	-.13	12.4	.10	.02
28	.11	7.8	.34	.22
6	.09	5.7	.31	.18
29	.07	4.2	.30	.17

Die Fragen sind in der Reihenfolge aufgelistet, in der sie in der schrittweisen multiplen Regressionsanalyse aufgenommen wurden; als Abbruchkriterium wurde ein F von 3.8 gesetzt, d.h. die schrittweise Analyse wird dann abgebrochen, wenn das als nächstes hereinkommende Item ein BETA-Gewicht hat, welches auf dem 5 % Niveau nicht mehr signifikant ist.

(1)-(4) siehe oben

Eine gute Gesamteinschätzung läßt sich demnach vorhersagen bei Veranstaltungen,

- bei denen die Teilnehmer mit dem Gelernten zufrieden sind (Frage 4; erklärt 44 % der Varianz der Linearkombination)
- bei der der Stil der Veranstaltung mit den Vorstellungen der Teilnehmer von einem sinnvollen Lernen übereinstimmt (12; 40 %)
- bei der fachlich viel gelernt wird (16; 32 %)
- bei der die Auswahl der Themen die wichtigsten Probleme der Praxis trifft (1; 35 %)
- die viele neue Ideen und Gesichtspunkte für die Praxis bringen und bei der die wichtigsten Aspekte des Themas herausgearbeitet werden (9, 10; 26 % bzw. 28 %)
- sowie - schwächer -, bei der ein gutes Arbeitsklima besteht (7, 8).

Frage 26 - freies Äußern von abweichenden Meinungen - hat zwar mit  $BETA = -.13$  ein hohes Gewicht, klärt aber nur ca. 2 % der Varianz der Linearkombination auf, d.h. sie leistet keinen Beitrag zur inhaltlichen Bestimmung. Es handelt sich hier um eine sogenannte Suppressorvariable, die hier nicht interpretiert wird. Es sei noch einmal daran erinnert, daß die obige Interpretation in dem Sinne zu relativieren ist, daß die Linearkombination ja nicht identisch ist mit der Gesamteinschätzung; es gibt noch einen Anteil von 45 %, der nicht gemeinsam ist.

Die "beste" Vorhersage leisten somit Fragen, die bei der Faktorenanalyse auf Faktor 1 "Methodischer Aufbau/Gesamteinschätzung" (4, 9, 12) und auf Faktor 3 "Lernertrag" (1, 10, 16) ihre höchste Ladung haben. Da auch die Fragen 13 (F1) und 31 (F3) des Kriteriums ihre höchste Ladung auf diesen Faktoren haben, muß diese Interpretation mit Vorsicht betrachtet werden. Es dominieren bei den Fragen Aspekte des Lernertrags, d.h. bei einer hohen Einschätzung dessen, was die Veranstaltung "bringt", wird sie auch im

Gesamt gut eingeschätzt, während hier die Aspekte des Arbeitsklimas und des äußeren Rahmens kaum eine Rolle zu spielen scheinen. Inwieweit dies generalisierbar ist, kann hier nicht entschieden werden.

Ein Vergleich der Struktur der "vorhersagekräftigsten" Fragen zur Regressionsanalyse der 1. Phase ist kaum möglich, da sich ja dort unterschiedliche Lösungen je nach Transformation ergaben. Tendenziell ergibt sich aber eine Ähnlichkeit mit den transformierten, dichotomisierten Fragen VADI (siehe Anlage 8) in dem Sinne, daß dort jene Fragen die künstliche Dimension am meisten aufklären, die nach dem Lernertrag fragen (alt 1, 27, 31, 20).

Eine schrittweise multiple Regressionsanalyse<sup>8</sup> wählte einen Satz von 12 Fragen aus. Dieser korreliert mit dem Kriterium mit  $R = .73$  ( $R^2 = .55$ , korrigiert = .52); es ergeben sich somit die identischen Werte, d.h. die Linearkombination aus 12 Fragen sagt die Gesamteinschätzung genauso gut voraus, wie die aus allen 29 Fragen. Zu den 12 ausgewählten Fragen gehören die obigen 8 aus der einfachen Regressionsanalyse, die die neue Linearkombination in ihrer Struktur genau gleich bestimmen (vgl. Tabelle 5). Zusätzlich läßt sich eine gute Gesamteinschätzung weiter vorhersagen bei Veranstaltungen,

- die eine gute Freizeitmöglichkeit bieten (Frage 24; 27 % gemeinsame Varianz mit dem Kriterium)
- die die Vorstellungen der Teilnehmer bei Entscheidungen über den Ablauf berücksichtigen (28; 11 %)
- in denen die Teilnehmer sich wohlfühlen (6; 18 %)
- in denen die Kleingruppenarbeit effektiv ist (29; 17 %).

<sup>8</sup> Zur Strategie siehe GAENSSLEN/SCHUBÖ 1973, S. 112f.; NIE u.a. 1975, S. 345f. Nach GAENSSLEN und SCHUBÖ ist dies praktisch meistens der beste Prädiktorensatz eines bestimmten Umfangs.



Wenn auch hier etwas mehr Aspekte des Arbeitsklimas (6, 28) und des äußeren Rahmens (24) hinzukommen, so wird aber dennoch diese Linearkombination ebenfalls inhaltlich Überwiegend durch die Einschätzung des Lernertrages bestimmt.

- Eine weitere Aufgabe der Regressionsanalysen war die Abschätzung des Ausmaßes des Halo-Effektes. Hierzu wurden - anschließend an die schrittweise multiple Regressionsanalyse - einige Prädiktorensätze vom gleichen Umfang, d.h. von 12 Fragen ausgewählt.

Bei Set 1-3 wurden durch das Losverfahren aus allen 29 Fragen 12 zufällig ausgewählt; bei Set 4-6 nur aus jenen Fragen, die in der einfachen multiplen Regressionsanalyse kein signifikantes BETA haben und nicht zu den ausgewählten 12 Fragen bei der schrittweisen Regressionsanalyse gehören. Set 0 bezeichnet den optimalen 12er Satz.

Tabelle 6:

Set-Nr.	R	$R^2$	Diff. zu $R^2$ Set 0
0	.73	.55	-
1	.66	.44	.11
2	.66	.44	.11
3	.64	.41	.14
4	.61	.37	.18
5	.56	.31	.24
6	.60	.35	.20

Würde sich ein Halo-Effekt auswirken, so müßte Set 1-6 ein ähnlich hohes  $R^2$  ergeben wie bei Set 0. Durchgehend erfassen die Linearkombinationen inhaltlich Gleiches, nicht wegen eines allgemeinen Halo-Effektes, sondern weil sich von den 12 Fragen immer einige durchaus auf Ähnliches beziehen. Wählt man das niedrigste  $R^2$  als Schätzung des Halo-Effektes aus, hier  $R^2 = .31$ , so dürfte es sich hier um eine große Überschätzung handeln, d.h. im allerhöchsten Fall ließen sich 31 % der Varianz des Kriteriums durch einen Halo-Effekt erklären, während mindestens 24 %

der insgesamt aufgeklärten Varianz von 55 % nicht auf diesen zurückgeführt werden können. Wegen der inhaltlichen Gleichheit einzelner Fragen der einzelnen Sets aber, dürfte der Halo-Effekt in Wirklichkeit weitaus niedriger ausfallen.

Zusammenfassend läßt sich festhalten, daß - ähnlich wie bei der Abschätzung aufgrund der Faktorenanalyse - ein Halo-Effekt nur niedrig auszufallen scheint. Insgesamt sind die Ergebnisse der Regressionsanalysen aber mit Vorsicht zu betrachten, da kein unabhängig von den Prädiktoren erhobenes Kriterium vorliegt und die Qualität der Globalbeurteilung hier in Frage gestellt werden muß. Tendenziell tragen zur Vorhersage der Globalbeurteilung jene Fragen am meisten bei, die sich auf die Einschätzung des Lernertrages beziehen, während Aspekte des Arbeitsklimas und des äußeren Rahmens kaum eine Rolle zu spielen scheinen.

#### 6.1.5 Diskriminanzanalysen

Es wird davon ausgegangen, daß nicht alle Fortbildungsveranstaltungen im Gesamt und in den einzelnen Aspekten als qualitativ gleich gut betrachtet werden können, vielmehr gibt es vielfältige Unterschiede zwischen den Veranstaltungen. Als Anspruch an ein taugliches Evaluationsinstrument wurde deshalb formuliert (vgl. Kapitel 4.3, S. 125), daß diese Veranstaltungsunterschiede sich in dem Instrument niederschlagen müssen. Diese Fragestellung wurde mit einer Diskriminanzanalyse verfolgt, die versucht, inhaltliche Dimensionen der Unterschiedlichkeit zwischen Fortbildungsveranstaltungen zu identifizieren.

Die Diskriminanzanalyse - hier nur für die transformierten, 4-stufigen Fragen TRAI - wurde nur für die Teilnehmer berechnet, die mindestens 23 Fragen (entspricht 74,2 %) beantwortet haben; fehlende Werte

wurden durch den Mittelwert ersetzt. Maximal können bei 25 unterschiedlichen Veranstaltungen 24 orthogonale Diskriminanzfunktionen berechnet werden; die Funktionen sind dabei in absteigender Reihe sortiert, d.h. in der ersten Funktion unterscheiden sich die Veranstaltungen am meisten, in der zweiten am zweit stärksten usw. (vgl. allgemein COOLEY/LOHNES 1971, S. 243ff.). Inhaltlich bedeutsam sind aber weniger als 24 Dimensionen der Unterschiedlichkeit. Ein Maß für die relative Bedeutung einer Funktion ist der Prozentanteil des korrespondierenden Eigenwertes an der Summe aller Eigenwerte, die die gesamte diskriminierende Varianz angibt (vgl. NIE u.a. 1975, S. 442); ebenso kann die kanonische Korrelation einer Diskriminanzfunktion mit dem Set der 24 dummy-Variablen, die die 25 Gruppen definieren, betrachtet werden; dies ist ein Maß dafür, wie eng die Diskriminanzfunktion und die Gruppenvariable zusammenhängen. Diese Werte sind in Anlage 17 zusammengefaßt. In der vorliegenden Studie sollte eine Diskriminanzfunktion mindestens 5 % der diskriminierenden Varianz und die Gruppenvariable mindestens 30 % der Varianz der Diskriminanzfunktion erklären (Can.Corr. ungefähr größer als .55), bzw. die ausgewählten Diskriminanzfunktionen insgesamt ca. 2/3 der gesamten diskriminierenden Varianz umfassen (kumulative Prozentzahl um 66 %). In Frage kamen danach 5 Funktionen. Interpretiert werden - ähnlich wie bei der Regressionsanalyse - nicht die Diskriminanzkoeffizienten, sondern die Korrelationen der Fragen mit den Diskriminanzfunktionen, da hohe Diskriminanzkoeffizienten wegen möglicher Suppressor-effekte nicht unbedingt inhaltlich aussagekräftig sind. Eine Lösung mit 5 Diskriminanzfunktionen ließ sich aber nur unklar interpretieren. 4 Funktionen dagegen können gut interpretiert werden und wurden deshalb hier bevorzugt.

Tabelle 7 (S. 172) enthält die wichtigsten Werte einer Diskriminanzanalyse mit 4 Funktionen, die im folgenden interpretiert werden sollen:

#### - 1. Diskriminanzfunktion

In dieser Diskriminanzfunktion unterscheiden sich die 25 Fortbildungsveranstaltungen am meisten; sie erklärt ca. 20 % der gesamten diskriminierenden Varianz. Die höchsten Korrelationen mit der Dimension haben die Fragen nach der passenden Teilnehmerzahl (Frage 3 mit  $r = -.56$ ; die negativen Korrelationen kommen durch die Polung der

Tabelle 7: Diskriminanzkoeffizienten und Korrelationen der Fragen mit den Dimensionen, TRAI

Nr.	Funktion 1			Funktion 2			Funktion 3			Funktion 4		
	Disk. Koef.	r mit F1	r <sup>2</sup>	Disk. Koef.	r mit F2	r <sup>2</sup>	Disk. Koef.	r mit F3	r <sup>2</sup>	Disk. Koef.	r mit F4	r <sup>2</sup>
1	-.18			.19	.45	.20	-.03			-.01		
2	-.21			-.05			-.16	-.29	.08	-.10		
3	-.55	-.56	.31	-.05			-.11			-.25	-.30	.09
4	.10			-.04	.35	.12	.20			.15		
5	.06			.04	.30	.09	.00			.23		
6	-.04			.01	.20	.04	-.05			-.11		
7	.15			.09	.25	.06	.02			-.03		
8	-.20	-.32	.10	-.03			-.01			.34	.45	.12
9	-.13			.02	.32	.10	.07			-.04		
10	.03			.17	.41	.17	.07			.20		
11	.21			-.04	.34	.12	.06			.30		
12	-.01			.01	.35	.12	-.09			-.06		
13	-.03			.32	.58	.34	.31			.14		
16	.19			.21	.44	.19	.39			.02		
17	-.12			-.14	.19	.04	-.16			-.22		
18	-.48	-.54	.29	-.21			.30			-.26		
19	.05			-.07	.21	.04	.09			-.10		
21	-.12			.32	.51	.26	-.16			.07		
22	.05			.04			-.45	-.38	.14	-.23		
23	.15			.01	.38	.14	.06			.01		
24	-.20	-.34	.12	-.05			-.43	-.33	.11	.60	.64	.41
25	.05			-.09	.19	.04	.12			.09		
26	.14			-.08			-.12	-.20	.04	.16		
27	-.06			.38	.63	.40	.07			-.29		
28	-.13			.17	.43	.19	-.23			-.25		
29	-.10			-.13			-.19	-.17	.03	-.04		
30	-.35	-.45	.20	.09			.27			.19		
31	-.09			.09	.40	.16	.19			-.18		
32	.08			.23	.39	.15	-.28			-.18		
33	.11			.03	.20	.04	-.15			-.35		
35	-.06			.05	.23	.05	.17			.13		

Bei den Korrelationen sind nur jene Koeffizienten mit  $r$  größer/gleich .30 angegeben, bzw. ist keine Korrelation zu einer Dimension größer als .30, so ist jeweils die höchste Korrelation zu einer Dimension angegeben.

Die 4 Funktionen erfassen insgesamt 55 % der diskriminierenden Informationen; auf die einzelnen Funktionen entfallen in der Reihenfolge 20,3 %, 13,8 %, 11,4 % und 9,8 %.

Funktion zustande), nach den Raumverhältnissen (19;  $r = -.54$ ) und nach den äußeren Bedingungen der Veranstaltung (Anfahrtswege, Unterkunft) (30,  $r = -.45$ ). Zwei weitere höhere Korrelationen haben noch die Fragen nach Geselligkeit und Freizeitgestaltung (8 und 24 mit  $r = -.32$  und  $-.34$ ; höchste Korrelation mit der 4. Diskriminanzfunktion). Fortbildungsveranstaltungen unterscheiden sich somit am meisten in der Gestaltung jenes Bereiches, der in den Faktorenanalysen als "äußerer Rahmen" identifiziert und auch in der Literaturanalyse als abgegrenzter Bereich betrachtet worden war. Angesichts der unterschiedlichen räumlichen und organisatorischen Möglichkeiten der in dieser Phase beteiligten Fortbildungsträger (z.B. nur angemietete Räume vs. zentral gelegene, gut ausgestattete "Schule") überrascht dieses Ergebnis nicht.

Veranstaltungen, die in dieser Dimension positive Extremwerte haben, "schneiden" wegen der negativen Korrelation hier am schlechtesten ab. Von zwei Veranstaltungen liegen offene Randbemerkungen auf dem Fragebogen vor, wobei in einer Veranstaltung von 28 Äußerungen sich 16 und in der anderen von 23 sich 9 auf eine negative Bewertung des äußeren Rahmens beziehen. Dies kann als vorsichtiges Zeichen für die Güte des Fragebogens interpretiert werden, vorsichtig deshalb, weil diese offenen Antworten auf dem Fragebogen nicht von allen Veranstaltungen gleich häufig vorliegen. Durchgängig können diese Bemerkungen also nicht interpretiert werden.

## - 2. Diskriminanzfunktion

Diese Dimension wird von vielen Fragen markiert, während sich in ihr die Veranstaltungen weniger als in der 1. Dimension unterscheiden. Die höchsten Korrelationen (größer als .4) haben in absteigender Reihenfolge die Fragen 27, 13, 21, 1, 16, 28, 10 und 31 ( $r = .63$  bis  $r = .40$ ). Es sind dies überwiegend Fragen aus dem Bereich "Methodischer Aufbau/Gesamteinschätzung" (27, 13, 21) und "Lernertrag" (1, 16, 10, 31), während "Arbeitsklima" (28) nur

einen geringen Beitrag hat. Interpretativ läßt sich diese Dimension also eher auf die direkte methodische Gestaltung und den direkten Gewinn, den die Teilnehmer sich von der Veranstaltung versprechen, beziehen. In diesem inhaltlich klar bestimmten und von den anderen Dimensionen deutlich abgegrenzten Bereich unterscheiden sich die Fortbildungsveranstaltungen weniger als in der Gestaltung des äußeren Rahmens, der vermutlich oft situativ unveränderlich den Fortbildungsleitern und -referenten vorgegeben ist.

### - 3. Diskriminanzfunktion

Die 3. Dimension der Unterschiedlichkeit von Veranstaltungen ist nur recht schwach bestimmt. Höhere Korrelationen haben nur die Fragen nach dem informellen Erfahrungsaustausch (Frage 22,  $r = -.38$ ), nach den Möglichkeiten einer selbständigen Erarbeitung von Informationen (2,  $r = -.29$ ) sowie die Frage nach der Freizeitgestaltung (24 mit  $r = -.33$ , höchste Korrelation zur 4. Dimension). Es scheint hier - vorsichtig interpretiert - eher die Möglichkeit der Teilnehmer, selbst aktiv zu sein und aktiv an den Problemen zu arbeiten, auch evtl. neben der Veranstaltung, angesprochen zu sein. Hier werden aber nur ca. 11 % der Unterschiedlichkeiten von Veranstaltungen erklärt. In dem in der Faktorenanalyse identifizierten Bereich "Arbeitsklima" scheinen sich somit weniger Unterschiede zwischen Veranstaltungen zu zeigen.

### - 4. Diskriminanzfunktion

In der 4. Diskriminanzfunktion unterscheiden sich die 25 Veranstaltungen nur schwächer. Diese Dimension wird eindeutig durch die Fragen nach der Freizeitgestaltung und dem geselligen Beisammensein (Frage 24 und 8;  $r = .64$  und  $r = .45$ ) bestimmt, also durch einen speziellen Aspekt aus dem Faktor äußerer Rahmen. Obwohl sich hier nur geringere Unterschiede zeigen, so stützt diese Diskriminanzfunktion

zusammen mit der oben beschriebenen 1. Diskriminanzfunktion die Annahme, daß sich die stärksten Unterschiede zwischen Fortbildungsveranstaltungen in dem Bereich des äußeren Rahmens - d.h. in organisatorischen und situativen Bedingungen sowie in der Gestaltung der Geselligkeit und Freizeit - finden lassen.

Bestätigt wird die inhaltliche Interpretation der 4. Funktion auch durch die mehr informellen Angaben der Teilnehmer aus zwei Veranstaltungen, die in dieser Dimension extrem niedrige Werte haben, also "schlecht" abschneiden; hier wird von den Teilnehmern auf die ungenügenden Freizeitmöglichkeiten oder auf Gruppenzwänge in der Freizeit hingewiesen.

Interessant ist die negative Korrelation von Frage 3 ( $r = -.30$ ) mit dieser Dimension; d.h. tendenziell werden, wenn die Teilnehmerzahl eher als zu groß eingeschätzt wird, zugleich die Freizeitmöglichkeiten eher positiv angesehen. Es läßt sich vermuten, daß eine größere Teilnehmerzahl die Möglichkeiten zu geselligen Kontakten vergrößert.

Eine schrittweise Diskriminanzanalyse unterstützt diese Interpretation.

Ähnlich wie bei der Regressionsanalyse kann es sein, daß mit einer geringeren Anzahl an Fragen die Veranstaltungen sich genauso gut unterscheiden lassen wie mit dem Set von 31 Fragen. Dies wurde in einer schrittweisen Diskriminanzanalyse untersucht (siehe hierzu Anlage 18). Dieses Verfahren suchte 17 Fragen heraus; bei 4 Diskriminanzfunktionen erlauben 17 Fragen 31,2 % richtige Klassifikationen - s.w.u. -, während 31 Fragen mit 34,7 % nur knapp besser abschneiden. Das Set von 17 Fragen ist qualitativ fast genauso gut wie die 31 Fragen. Inhaltlich entsprechen sich die Dimensionen 1, 3 (umgekehrt gepolt) und 4 völlig. Bei Dimension 2 kommt die Interpretation als "methodischer Aufbau" und "Lernertrag" sogar noch stärker heraus. Bei dem schrittweisen Vorgehen wurden 13 Fragen ausgeschlossen, die vorher ihre höchste, aber absolut gesehen eine recht niedrige Korrelation zu dieser Dimension hatten, wobei mehrere Fragen mit den niedrigen Korrelationen sich auf das Arbeitsklima bezogen. Allgemein stimmt die schrittweise Diskriminanzanalyse also sehr gut mit der mit 31 Fragen überein.

Die vielfältigen Unterschiede zwischen Fortbildungsveranstaltungen spiegeln sich somit sensibel in dem Fragebogen wieder. Es lassen sich 4 inhaltlich gut interpretierbare Dimensionen der Unterschiedlichkeit identifizieren. Die größten Unterschiede liegen in der Gestaltung des äußeren Rahmen; es folgen der direkte methodische Aufbau und die Einschätzung des Lernertrages, die Möglichkeiten der Teilnehmer zur aktiven Mitarbeit und zum Schluß die Freizeitgestaltung. Bei diesen Dimensionen der Unterschiedlichkeit zeigen sich zwar keine Entsprechungen, aber deutliche Ähnlichkeiten zu den unter anderen Bedingungen berechneten künstlichen Dimensionen bei der Faktorenanalyse. In beiden Fällen finden sich die Bereiche des äußeren Rahmens, der methodischen Gestaltung und der Einschätzung des Lernertrages, nur der Faktor "Arbeitsklima" tritt nicht so deutlich unter den Diskriminanzfunktionen auf. Diese Ähnlichkeit weist auf die Bedeutung der Dimensionen bei der Faktorenanalyse hin. Die 4 Faktoren stellen somit nicht nur Dimensionen dar, zu denen sich die Items nach ihrer "Ähnlichkeit" strukturieren lassen, sondern zugleich sind es Bereiche, in denen sich in gewisser Weise Unterschiede zwischen den Fortbildungsveranstaltungen niederschlagen. Dies rechtfertigt noch einmal, daß bei der Skalenbildung - s.w.u. - von den 4 Faktoren ausgegangen werden soll.

Die Güte des Fragebogens läßt sich noch in einem weiteren Aspekt aufgrund der Informationen aus der Diskriminanzanalyse abschätzen:

Gesetzt den Fall, man wüßte nicht mehr, welche der 25 Fortbildungsveranstaltungen ein Teilnehmer besucht hat, möchte aber wissen, wie groß die Wahrscheinlichkeit der Zuordnung in die richtige - d.h. auch wirklich vom Teilnehmer besuchte - Veranstaltung ist. Liegen keine weiteren Informationen vor, so läßt sich die Wahrscheinlichkeit der Zugehörigkeit



zu einer Veranstaltung berechnen aus dem Verhältnis der Anzahl der Teilnehmer in der Veranstaltung zur Gesamtzahl aller Teilnehmer.

So besuchten 12 Teilnehmer die Veranstaltung 01; bei einer Gesamtzahl von 494 Teilnehmern hat ein Teilnehmer - dessen Gruppenzugehörigkeit unbekannt ist - mit einer Wahrscheinlichkeit von  $p=.02$  diese Veranstaltung besucht; m.a.W.: in 2 % aller Fälle würden wir ihn der richtigen Gruppe zuordnen. Es läßt sich nun für jede Veranstaltung die Wahrscheinlichkeit der richtigen Klassifikation berechnen.

Als "mittlere" Wahrscheinlichkeit der richtigen Klassifikation kann der Durchschnitt (hier 4 %) oder der Median (hier 4,2 %) der einzelnen Wahrscheinlichkeiten der Gruppenzugehörigkeit betrachtet werden. D.h.: in ungefähr 4 % - 4,2 % aller Fälle würde man einen Teilnehmer der richtigen - d.h. von ihm besuchten - Veranstaltung zuordnen.

Nun spiegelt sich aber die Unterschiedlichkeit der Veranstaltungen in dem Fragebogen wieder und es läßt sich fragen, ob aufgrund dieser Informationen nicht die Anzahl der richtigen Klassifikationen verbessert wird. Mit Hilfe der Diskriminanzanalyse wurden Klassifikationsfunktionen berechnet (vgl. COOLEY/LOHNES 1971, S. 262ff.), die Auskunft geben über die Prozentzahl der richtigen Klassifikationen. Für verschiedene Diskriminanzanalysen sind diese Werte in Tabelle 8 zusammengefaßt.

Tabelle 8 Prozentzahlen richtiger Klassifikationen bei verschiedenen Diskriminanzanalysen

(1)	(2)					(3)			(4)	
	DA mit 31 Fragen					DA mit 24 Fragen			DA mit 17 Fragen	
O DF	24 DF	7 DF	5 DF	4 DF		24 DF	8 DF	5 DF	5 DF	4 DF
4,2%	64,4%	49,1%	39,8%	34,7%		60,0%	49,3%	37,9%	38,5%	31,2%

- (1) Anzahl richtiger Klassifikationen ohne Diskriminanzanalyse (DA)
- (2) Diskriminanzanalyse (DA) mit allen Fragen, Methode DIRECT
- (3) DA mit 24 Fragen, Methode WILKS, PIN/POUT =.05
- (4) DA mit 17 Fragen, Methode WILKS, PIN/POUT =.001

Ist es also möglich, bei der Zuordnung eines Teilnehmers zu einer Veranstaltung auf Fragebogeninformationen zurückzugreifen, so vergrößert sich die Anzahl der richtigen Klassifikationen enorm. Beim "besten" Ergebnis (64,4 % richtige Klassifikationen) tätigt man ca. 15mal mehr richtige Klassifikationen, beim "schlechtesten" Ergebnis (31,2 %) ca. 7mal mehr richtige Klassifikationen, wenn man die Fragebogeninformationen ausnützt gegenüber dem Fall, wo diese Informationen nicht vorliegen. Dies ist ein Zeichen dafür, wie sensibel der Fragebogen auf Unterschiede zwischen den Veranstaltungen reagiert. Der vorliegende Bogen diskriminiert also recht gut zwischen Gruppen; er kann auch benutzt werden, um Gruppen zu vergleichen. Dies sollte als Zeichen seiner Qualität betrachtet werden<sup>9</sup>.

#### 6.1.6 Skalennalysen

Für jede Dimension der Faktorenanalyse wurde ein Skalenswert durch einfaches Aufaddieren der Markiervariablen berechnet. Dies wurde der exakten Bildung der Faktorscores aus den Linearkombinationen von allen Fragen vorgezogen, da ein einfaches Aufaddieren auch in der angezielten praktischen Anwendung des Bogens möglich sein dürfte.

Einfaches Aufaddieren hat noch weitere Vorteile. Einmal wird damit - ohne schiefwinkelige Rotation - die oft als künstlich empfundene Orthogonalität der Faktoren aufgehoben; zum anderen ist dies der inhaltlichen Interpretation der Faktoren gerechter, die ja anhand der Markiervariablen geschieht (vgl. GAENSSLEN/SCHUBÖ 1973, S. 255).

<sup>9</sup> Die Auswahl von Fragen nach dem Konzept normorientierter Tests, also z.B. nach Trennschärfe und Schwierigkeiten (zwecks Maximierung der interindividuellen Differenzen), führt also nicht dazu, daß der Bogen nicht mehr zwischen Gruppen differenzieren kann. Zu einem ähnlichen Ergebnis - mit weitaus größerem Aufwand an Versuchsplanung und statistischer Auswertung - kommen AIRASIAN/MADAUS. Ihr Fazit: "Critics who argue that item selection procedures utilized in norm - referenced test construction renders such items insensitive to school or program effects much reassess their argument." (AIRASIAN/MADAUS 1978, S. 6).

Pro Skala wurden jeweils die "besten" 5 Fragen ausgewählt.

Ausgegangen wurde von der Faktorenanalyse der transformierten, 4-stufigen Fragen TRAI (vgl. Tabelle 4, S. 156); zu der identischen Auswahl der 5 Fragen kommt man aber, wenn die transformierten, dichotomisierten Fragen TRADI oder die gedrehten, 4-stufigen Fragen ITEM betrachtet werden. Die auszuwählenden Fragen sollten eine bedeutsame Ladung zum entsprechenden Faktor haben ( $a_{ij}/h_i \geq .5$ ), bei den verschiedenen Akzentuierungen der Faktorenanalyse konstant auf dem identischen Faktor laden und innerhalb jeder Skala untereinander hoch korrelieren, zumindest aber untereinander höher als mit den Fragen anderer Skalen. Danach ergaben sich pro Faktor 4 - 6 "beste" Fragen. Weil alle Dimensionen gleich repräsentiert sein sollten, wurde sich für 5 Fragen pro Skala entschieden. Es sind dies: Skala 1 "Methodischer Aufbau/Gesamteinschätzung" die Fragen 11, 12, 13, 25, 27; Skala 2 "Lernertrag" die Fragen 1, 10, 16, 23, 31; Skala 3 "Arbeitsklima" die Fragen 5, 6, 17, 26, 28; Skala 4 "äußerer Rahmen" die Fragen 3, 8, 18, 24, 30. Eine neue Faktorenanalyse mit diesen ausgewählten 20 Fragen stellte sicher, daß sie die gleiche Faktorenstruktur bestimmen wie die gesamten 31 Fragen (Kommunalitäten und Ladungen siehe Anlage 19).

Zur Konstruktion der Skalen wurden fehlende Werte durch den Mittelwert ersetzt; aufaddiert wurde nur, wenn mindestens 3 Fragen pro Dimension beantwortet wurden. Gebildet wurden die Skalenwerte für die transformierten, 4-stufigen Fragen TRAI; diese wurden wieder prozentrang transformiert und in z-Äquivalente umgewandelt, um "technisch" gute Skalen zu haben (siehe hierzu weiter unten); durch diese Transformation haben die Skalen zudem in etwa den gleichen Maßstab. Angesprochen werden diese Skalenwerte unter dem Namen SCALE 1 bis SCALE 4. Wenn alle Skalenwerte vorlagen, wurde noch der Gesamtwert GESAMT berechnet. Ebenfalls wurden die Skalenwerte für die gedrehten, 4-stufigen Fragen ITEM aufaddiert (DIM 1 bis DIM 4; GEITEM als Gesamtwert). Alle folgenden Analysen wurden mit beiden Transformationsarten gerechnet; falls sich keine Unterschiede ergaben, wird nur über SCALE berichtet (in Anlage 20 sind einige Skalenstatistiken und die Fragen-Skalen Korrelationen aufgelistet).

Die Beschränkung auf 5 Fragen pro Dimension beinhaltet kaum einen Informationsverlust. Um dies zu überprüfen, wurden alle bedeutsam ladenden Fragen mit Ladungen über .35 (z.B. bei Skala 1 und 3 je 8 Fragen) zusam-

mengefaßt und mit den kürzeren, aus 5 Fragen bestehenden Skalen korreliert. Die Korrelation beträgt für SCALE 1 bis SCALE 4, GESAMT  $r=.90$  bis  $r=.96$ . Die kürzeren Skalen enthalten also fast die identischen Informationen wie die längeren.

Im folgenden werden einige Analysen der Skalenwerte berichtet:

- Die einzelnen Skalen korrelieren teils beträchtlich miteinander. Tabelle 9 enthält die Skalenkorrelationen:

Tabelle 9

	SCALE1	SCALE2	SCALE3	SCALE4	GESAMT	DIM1	DIM2	DIM3	DIM4	GEITEM
SCALE 1	-	.51	.45	.13	.71	.92	-	-	-	-
SCALE 2	-	-	.48	.21	.79	-	.98	-	-	-
SCALE 3	-	-	-	.16	.72	-	-	.96	-	-
SCALE 4	-	-	-	-	.55	-	-	-	.98	-
GESAMT	-	-	-	-	-	-	-	-	-	.96
DIM 1	-	-	-	-	-	-	.53	.45	.16	.73
DIM 2	-	-	-	-	-	-	-	.51	.22	.79
DIM 3	-	-	-	-	-	-	-	-	.17	.72
DIM 4	-	-	-	-	-	-	-	-	-	.60
GEITEM	-	-	-	-	-	-	-	-	-	-

Die Struktur der Interkorrelationen bei den Skalen SCALE und DIM stimmt völlig überein; ebenso korrelieren die entsprechenden Skalen sehr gut miteinander, z.B. SCALE 1 mit DIM 1 mit  $r=.92$  als niedrigste Korrelation. Wie schon aus den Faktorenanalysen zu erwarten war, korrelieren die Skalen 1 "Gesamteinschätzung/Methodischer Aufbau" und 2 "Lernertrag" höher miteinander, gegenseitig werden ca. 26 % der Varianz erklärt. Höhere Korrelationen haben noch die beiden Skalen zum "Arbeitsklima" (SCALE 3, 20-23 % gemeinsame Varianz), während die ersten 3 Skalen als ziemlich unabhängig zur 4. Skala "äußerer Rahmen" betrachtet werden können (2 - 4 % gemeinsame Varianz).

Diese Struktur der Skaleninterkorrelationen entspricht der schiefwinkligen Rotation, z.B. bei DELTA = 0.

Die Korrelationen der Skalen zu dem Gesamtwert sind alle relativ hoch (ca. 52 - 62 % gemeinsame Varianz), nur Skala 4 (30 % gemeinsame Varianz) bestimmt den Gesamtwert etwas schwächer.

- Für jede Skala wurde als Reliabilitätsschätzung CRONBACHs ALPHA berechnet; es ergeben sich folgende Werte für ALPHA bei SCALE 1  $r_{tt} = .77$ , bei SCALE 2  $r_{tt} = .74$ , bei SCALE 3  $r_{tt} = .60$ , bei SCALE 4  $r_{tt} = .59$ , bei GESAMT  $r_{tt} = .82$ . Berechnet man hypothetisch die aufgewerteten Reliabilitäten für eine Skalenverlängerung auf 20 Fragen (vgl. LIENERT 1969, Anmerkung 23 auf S. 119), so beträgt die Reliabilität bei SCALE 1  $r_{tt} = .93$ , SCALE 2  $r_{tt} = .92$ , bei SCALE 3  $r_{tt} = .86$ , bei SCALE 4  $r_{tt} = .85$ . Obwohl die Skalen nur aus 5 Fragen bestehen, kann ihnen - vor allem den Skalen 1 und 2 sowie dem Gesamtwert - eine relativ zufriedenstellende Reliabilität zugesprochen werden.

- In allen Skalen und dem Gesamtwert zeigen sich auf dem 0,1 % Niveau signifikante Unterschiede zwischen den Veranstaltungen.

Betrachtet man - nach dem DUNCAN-Test (vgl. NIE u.a. 1975, S. 426ff.) - Subsets von Veranstaltungen, in denen jeweils keine signifikanten Unterschiede zwischen den Veranstaltungen bestehen, so ergeben sich bei SCALE 1 und 3 je 6, bei SCALE 2 5, bei SCALE 4 10 und bei GESAMT 8 solcher Subsets. Dies zeigt noch einmal unterstützend zur Diskriminanzanalyse auf, daß sich im "äußeren Rahmen" die Veranstaltungen am meisten unterscheiden.

- Die Ähnlichkeit zwischen den Dimensionen aus der Faktorenanalyse und aus der Diskriminanzanalyse (vgl. S. 176) wies daraufhin, daß die 4 Faktoren auch jene Bereiche sind, in denen sich Unterschiede zwischen den Veranstaltungen zeigen. Genauer untersuchte dies eine Diskriminanzanalyse mit den aufgrund der Faktorenanalyse gebildeten 4 Skalen (vgl. Anlage 21). Es finden sich hier

keine bedeutsamen Unterschiede zu der Diskriminanzanalyse mit 31 bzw. 17 Fragen. D.h. die Sensibilität des Fragebogens für die Unterschiedlichkeit zwischen Veranstaltungen bleibt auf Skalenebene erhalten.

Es wurden 3 Diskriminanzfunktionen interpretiert, die zusammen ca. 95 % der diskriminierenden Varianz erklären. Die 1. Funktion wird eindeutig durch die Skala "äußerer Rahmen" bestimmt ( $r=.98$ ), während die anderen Skalen hier Nullkorrelationen haben; diese 1. Dimension ist eine Zusammenfassung der 1. und 4. Dimension bei der Analyse mit 31 Fragen. Die 2. Diskriminanzfunktion wird am stärksten bestimmt durch die Skalen 1 und 2, während Skala 3 "Arbeitsklima" einen geringeren Beitrag hat. Wiederum läßt sich diese Dimension der Unterschiedlichkeit beziehen auf die methodische Gestaltung und den eingeschätzten Lernertrag. Etwas unklar ist die neue 3. Diskriminanzfunktion bestimmt. Nur bei Berücksichtigung von DIM ist sie klarer zu interpretieren, da hier mit  $r=.61$  die Skala 3 "Arbeitsklima" die Funktion am stärksten bestimmt.

Wie die Anzahl der richtigen Klassifikationen bei einer Diskriminanzanalyse auf Skalenebene zeigt (vgl. Anlage 21), geht mit der Skalenbildung ein gewisser Informationsverlust einher, d.h. auf der Ebene der Einzelfragen sind mehr diskriminierende Informationen enthalten.

#### Zusammenfassung:

Bisher wurde überwiegend der interne Zusammenhang des Fragebogens untersucht, d.h. z.B. die Struktur der Beziehungen der Fragen untereinander, der inhaltlich - logische Zusammenhang der Fragebogenelemente (z.B. bei Faktoren-, Regressions- und Diskriminanzanalysen und bei der internen Konsistenzschätzung) sowie das Ausmaß des Halo-Effektes. Im allgemeinen ergaben diese Analysen ein für den Fragebogen positives Bild. Anschließend wurden 4 Skalen und ein Gesamtwert gebildet, die in wichtigen Aspekten den Ergebnissen der Gesamtanalyse entsprechen. Die Skalenbildung und Beschränkung auf 5 Fragen je Skala geschieht mit relativ wenig Informationsver-

lust. Sehr wichtig ist die durchgehende Feststellung, daß sich in den verschiedenen Analysen kaum Unterschiede zeigen, ob nun

- mit den gedrehten, 4-stufigen Fragen ITEM bzw. den entsprechenden Skalen DIM oder

- mit den transformierten, 4-stufigen Fragen TRAI bzw. den entsprechenden, nochmals transformierten Skalen SCALE gerechnet wird. Die aufwendigen Transformationen führen also zu keinen anderen Ergebnissen.

Die - vor allem technisch gesehen - günstigeren Merkmale haben die Skalen SCALE (Normalverteilung, gleicher Maßstab, vgl. Anlage 35), mit denen beim Vergleich mit externen Variablen auch primär weitergearbeitet werden soll. Erschwert aber schon die erste Transformation bei ITEM zu TRAI die Interpretation, so sind ebenfalls die Skalen SCALE im Unterschied zu DIM unanschaulicher.

Könnte bei ITEM eine "2" als "stimme ein wenig zu" gut interpretiert werden, so ist dies bei TRAI mit "stimme unterdurchschnittlich mit einem bestimmten Abstand zum Median 0 zu" nur schwer möglich. Das letztere gilt entsprechend für SCALE. Die Skalen DIM dagegen sind wieder anschaulicher zu interpretieren; ein Skalenwert von "5" bedeutet, daß ein Teilnehmer in allen Fragen die Veranstaltung negativ einschätzt (5mal den Wert 1 vergibt), während ein Skalenwert von "10" auf ein Gleichgewicht von positiven und negativen Einschätzungen dieses Aspektes von Fortbildung hinweist.

Werden nun im Abschnitt 6.3 bei den Untersuchungen der Beziehungen der Skalenwerte zu anderen Variablen - als vorsichtiger Ansatz einer Konstruktvalidierung - primär die technisch besseren Variablen SCALE benutzt, so werden immer die Variablen DIM ebenfalls berücksichtigt, damit bei Gleichheit der Ergebnisse im gegebenen Fall auf die einfacher zu interpretierenden Skalenwerte zurückgegriffen werden kann. Vorher soll aber noch über einen Vergleich zwischen den Teilnehmerbeurteilungen

und den Fortbildungsleiter-/Referentenbeurteilungen (6.2) und über eine Experten- und Fortbildungsleiterbefragung (6.3) berichtet werden.

## 6.2 Vergleich zwischen Teilnehmer- und Referenten-/Leiterbeurteilungen

Die Referenten und Fortbildungsleiter wurden in der zweiten Erprobungsphase gebeten, ihre eigene Veranstaltung ebenfalls auf dem hier entwickelten Fragebogen zur Veranstaltungsbeurteilung einzuschätzen. Damit sollte überprüft werden, inwieweit Referenten-/Leiterbeurteilungen einerseits und Teilnehmereinschätzungen andererseits übereinstimmen. Werden die Referenten-/Leitereinschätzungen als Kriterium für die Güte der Fortbildung betrachtet, so bietet der Vergleich mit den für jede Veranstaltung zusammengefaßten Teilnehmerbeurteilungen einen ersten Ansatz zur Bestimmung einer kriterienbezogenen Validität (vgl. Kapitel 4.3, S. 126). Als Maßzahl zur Bestimmung der Ähnlichkeit der Beurteilungen dieser beiden Gruppen wurde das Q-Maß ausgewählt; Tabelle 10 enthält die entsprechenden Werte.

Aus 16 Veranstaltungen liegen Referenten- bzw. Leitereinschätzungen vor (aus 6 Veranstaltungen liegen dabei zwei Einschätzungen vor); für diese Veranstaltungen wurden für jede Frage innerhalb jeder Veranstaltung der Mittelwert und der Median berechnet. Danach wurde für jede Veranstaltung die Meßwertreihe der Referenten/Leiter mit jener der Teilnehmergruppe verglichen. Dazu wurden die Korrelation der Meßwerte der beiden Einheiten über die verschiedenen Fragen berechnet (Q-Maß). Da dieses Maß von der Polung der Fragen abhängt, wurden alle vorher so gepolt, daß hohe Werte eine Veranstaltung positiv beurteilen (vgl. DIEHL/SCHÄFER 1975, S. 173ff.; HOFSTÄTTER 1971, S. 189f.). Es wird somit gefragt, inwieweit die Profile über alle Fragen von Teilnehmern und Referenten/Leitern übereinstimmen (zu solchen möglichen Profilen siehe auch Anlage 12).



Die Auswahl des Q-Maßes ist aber nicht unumstritten: Probleme liegen vor allem in den Voraussetzungen der Anwendung; die Fragen müssen alle einen einheitlichen Maßstab haben, intervallskaliert, gleichgepolt sein, und es werden nur lineare Beziehungen erfaßt. Eine entscheidende Schwäche liegt darin, daß bei der Korrelation ja zu vergleichende Meßwertreihen intern standardisiert werden. Dies hat zur Folge, daß Mittelwerts- und Streuungsunterschiede der Profile nicht erfaßt werden, sondern nur die Verlaufsgestalt. Für die vorliegende Fragestellung kann dies akzeptiert werden, da weniger die von möglicherweise subjektiven Vorlieben bestimmte durchschnittliche Höhe und Streuung der Beurteilungen interessiert, sondern die Validität sich eher an der Übereinstimmung im Verlauf, d.h. in den Beziehungen der Fragen, erweist.

Gegenüber dem Q-Maß werden oft D- oder Distanzmaße empfohlen (vgl. DIEHL/SCHÄFER 1975, S. 179-199; SCHLOSSER 1976, S. 91-143), die die oben genannten Schwächen nicht aufweisen. Das Q-Maß ist aber leichter zu berechnen - es konnte ein Taschenrechner TI Programmable 59 benutzt werden -, und D-Maße haben zudem den Nachteil, daß sie relativ unspezifisch sind: negativ lineare Beziehungen lassen sich nicht von mittleren Beziehungen unterscheiden; ein mittleres D-Maß läßt nicht erkennen, ob eine positive oder negative Beziehung vorliegt. Empirisch kommt man zudem oft unabhängig von den verwendeten Maßen zu vergleichbaren Ergebnissen (vgl. z.B. DIEHL/SCHÄFER 1975, S. 198f.).

Tabelle 10: Q-Maße

Veranstaltung	Ref./Leiter I		Ref./Leiter II		Häufigkeit der Korrelationskoeffizienten		
	r bei $\bar{x}$	Median	r bei $\bar{x}$	Median	r	f	% an aufgekklärter Varianz
1	.61	.64	.74	.73	.3-.39	1	9%-15%
2	ns	ns			.4-.49	4	16%-24%
3	.38	.39			.5-.59	3	25%-35%
5	.43	.42			.6-.69	7	36%-48%
6	.67	.67	.43	.45	.7-.79	5	49%-62%
7	.64	.64			.8-.89	1	64%-79%
9	.71	.74	.71	.72	durchschnittlich ca. 39,4%		
10	.49	.49	.55	.56			
11	.72	.75					
13	.61	.61					
15	.71	.76					
16	.37	ns					
17	.66	.65	.65	.71			
19	.69	.71					
20	.49	.42					
23	.57	.57	.59	.58			

Mindestens ein Referent/Leiter füllte in 16 Veranstaltungen den Fragebogen aus (siehe unter I), aus 6 Veranstaltungen liegt eine zweite Einschätzung vor (siehe unter II); aufgelistet sind nur die mindestens auf dem 5%-Niveau signifikant von Null verschiedenen Korrelationen.

Mit einer Ausnahme (Veranstaltung 16) stimmen die Korrelationen überein, wenn bei der Aggregation der Teilnehmereinschätzungen einmal der Median und zum anderen der Mittelwert berechnet wird. In den Veranstaltungen, in denen zwei Referenten/Leitereinschätzungen vorliegen, läßt sich ebenfalls eine Übereinstimmung in den Q-Maßen feststellen. Überwiegend zeigt sich ein signifikanter Zusammenhang zwischen der Referenten/Leitereinschätzung und den aggregierten Teilnehmereinschätzungen, gegenseitig sind durchschnittlich ca. 39 % der Varianz gemeinsam (vgl. Tabelle 10, rechte Spalte).

Werden die Referenten-/Leitereinschätzungen als ein mögliches Außenkriterium betrachtet, so läßt sich hier ein genauer Wert für die kriterienbezogene Validität nicht angeben. Die Ergebnisse deuten aber auf eine mittlere Übereinstimmung zwischen Teilnehmern und Referenten/Leitern hin (ca. 39 % gemeinsame Varianz), so daß - mit Vorsicht - von einer mittleren Kriteriumsvalidität gesprochen werden kann. Hiermit zeigt sich auch das in der Literatur erwähnte Bild (vgl. COSTIN u.a. 1971, S. 530): Teilnehmer- und Referenten-/Leitereinschätzungen stimmen zwar beträchtlich überein, trotzdem lassen sich die Teilnehmereinschätzungen als Informationsquelle eigener Art betrachten, die im Durchschnitt hier ca. 60 % der Varianz nicht mit den Referenten-/Leitereinschätzungen gemeinsam hat.

Ein entsprechendes Ergebnis findet sich, wenn wegen der Uneinheitlichkeit der Ähnlichkeitsmaße ganz auf entsprechende Indizes verzichtet und nur global überprüft wird, ob sich die Profile signifikant unterscheiden (vgl. KRISTOF 1958, S. 523-526).

KRISTOF schlägt einen  $\chi^2$ -Test vor, der nur auf Skalenebene berechnet werden kann, da dazu die Reliabilität nötig ist. Die Skalenwerte wurden aufgrund der oben

beschriebenen Faktorenanalysen für die Referenten/Leiter gebildet; mit den quadrierten Differenzen ( $D$ ) zwischen Teilnehmern und Referenten/Leiter in den 4 Dimensionen wurde folgender Wert berechnet:

$$\chi^2 = \frac{N}{(1+N) \sigma^2} \sum_{j=1}^4 \frac{D_j^2}{1-r_{jj}} \quad \text{mit } df=4$$

mit

$N$  = Anzahl der Teilnehmer in der Veranstaltung

$r_{jj}$  = Reliabilität der Skalen DIM 1 - DIM 4 (s. oben)

$\sigma^2$  = Streuung der Punktwerte in den Dimensionen; der Test setzt voraus, daß die Punktwerte in allen Dimensionen mit der Streuung  $\sigma^2$  für die Population verteilt sein müssen;  $\sigma^2$  wird als gemittelte Quadratsumme der Streuungen in den 4 Dimensionen geschätzt (nach den Streuungen der Teilnehmereinschätzungen).

In 4 von den 16 Fällen unterscheiden sich die Profile signifikant auf dem 5 % Niveau voneinander, d.h. in 3/4 aller Fälle aber finden sich keine Unterschiede in den Profilen der Leiter-/Referenteneinschätzungen verglichen mit den Profilen der Teilnehmer auf der Dimensionsebene.

Nicht berichtet wird hier über den Vergleich der Referenten-/Leitereinschätzungen mit den Teilnehmern in jeder Frage über alle Veranstaltungen hinweg (R-Korrelationen auf Veranstaltungsebene). Gleichen sich subjektive Vorlieben der Teilnehmer beim Ausfüllen der Ratingskalen bei der Aggregation auf Veranstaltungsebene aus, so sind die Referenten-/Leitereinschätzungen - die ja pro Veranstaltung meist nur von 1 Person vorliegen - mehr von diesen subjektiven Vorlieben verzerrt. Entsprechende Korrelationen lassen sich nicht sinnvoll berechnen und interpretieren.

Festzuhalten bleibt, daß Teilnehmer und Referenten/Leiter in der Beurteilung der Aspekte einer Fortbildungsveranstaltung - trotz eigenem Beitrages der Teilnehmer - weitgehend übereinstimmen und dies unabhängig von dem Verfahren des Vergleiches (Q-Maß;  $\chi^2$ -Test). Betrachtet man die Referenten-/Leitereinschätzungen als Außenkriterium, so läßt sich dem Fragebogen zufriedenstellende kriterienbezogene Validität nicht absprechen.

### 6.3 Experten- und Fortbildungsleiterbefragung

Da ein unvalides Evaluationsinstrument ein Auswertungsgespräch nur belasten würde, wurde auf die Überprüfung der Validität ein besonderes Gewicht gelegt. In diesem Abschnitt wird über eine Experten- und Fortbildungsleiterbefragung berichtet, deren Ziel in der Bestimmung der Content Validität, der Wichtigkeit und Nützlichkeit der Fragebogeninformationen lag (vgl. Kapitel 4.3, S. 121ff.).

Die Befragungen der Experten und der Fortbildungsleiter waren weitgehend parallel aufgebaut; die Vorlage für die Fortbildungsleiter entspricht jener für die Experten, der Leiterfragebogen war nur um 2 Fragen umfangreicher: einmal wurde der "Nutzen" jedes Aspektes erhoben und nicht nur der "Nutzen" der Fragebogeninformationen insgesamt, zum anderen wurde zusätzlich nach der möglichen Weiterverwendung des Bogens gefragt.

Für die Expertenbefragung wurden Fachleute aus den Bereichen Vorschulerziehung (Hochschule, Forschungsprojekte, Forschungsinstitute) und der Erwachsenenbildung (VHS, Hochschule, freie Erwachsenenbildungsträger) angesprochen. Neben der Bitte um Mitarbeit (vgl. Anlage 22) erhielten sie eine Vorlage zur Experteneinschätzung (vgl. Anlage 23), eine kurze Projektbeschreibung (vgl. Anlage 3), Erläuterungen zum Einsatz des revidierten Fragebogens (vgl. Anlage 10) und eine kurze Zusammenfassung einiger Ergebnisse der 1. Erprobungsphase (vgl. Anlage 24).

Angeschrieben wurden 22 Experten; auswertbare Exemplare schickten 15 zurück; ein weiterer Experte schrieb wegen Krankheit, ein zweiter wegen Zeitmangels ab; 2 Befragte schickten keine ausgefüllte Vorlage zurück, sondern ausführliche Begründungen, warum sie nicht teilnehmen wollten; 3 Experten reagierten nicht. Alle in der 2. Phase beteiligten Fortbildungsleiter bzw. die Fortbildungsinstitutionen nahmen an der Leiterbefragung teil; zurückgeschickt wurden aber nur 6 Bögen, da in zwei Fällen nicht jeder einzelne Leiter einen Bogen, sondern je das ganze Team der Institution einen Bogen zusammen ausfüllte.

Die Ergebnisse der Befragung - die Fragen wurden teils offen, teils geschlossen gestellt - werden im folgenden in der Reihenfolge der Expertenvorlage vorgestellt:

- Befragt (Frage 2) nach dem Grad, indem die Auswahl der Fragen repräsentativ für das Gesamt der diskutierten Postulate zur Gestaltung von Fortbildung ist, wurde dem Fragebogen das Prädikat "gute" bis "befriedigende" Auswahl zugeordnet.

Es lag eine Antwortskala von "1" (sehr gute Auswahl) bis "6" (sehr schlechte Auswahl) vor: im Durchschnitt antworteten die Experten mit 2,5 (bei einer Streuung s von 0,9), die Leiter schätzen einheitlich die Auswahl mit 2,0 ein.

Auch die Antworten auf die offene Frage - "welche Bereiche kommen zu kurz" - unterstützen dieses Bild einer guten Auswahl. 9 Befragte antworteten hier, bezogen sich aber auf Einzelaspekte und nicht auf das Fehlen eines ganzen Bereichs. Insgesamt kann dem Fragebogen "Inhaltsvalidität" zugesprochen werden.

Es können hier und bei den folgenden offenen Fragen nicht alle Antworten aufgelistet werden; der Verfasser hat sich aber bemüht, die wichtigsten Tendenzen herauszuarbeiten.

Einige Antworten kreisen darum, daß jenes, was mit Lern-ertrag angesprochen wird, stärker differenziert werden könnte. Eine interessante Anmerkung - einziger Leiter, der auf diese Frage eine offene Antwort gab - kritisiert, daß keine Frage gestellt wird, die nach dem Anteil der Teilnehmer am Lernerfolg, nach der eigenen Lernleistung und Eigenverantwortlichkeit für das Lernen fragt.

- Die Wichtigkeit bzw. Relevanz jedes einzelnen Aspektes (Frage 3) konnte auf einer Skala von "1" (sehr wichtig) bis "6" (völlig unwichtig) eingeschätzt werden. Im allgemeinen finden sich keine Unterschiede in der Beurteilung der Relevanz der Fragen bei Experten und Fortbildungsleitern (vgl. Anlage 25, Tabelle 1).

Mit einer sehr leichten Tendenz - bei diesem geringen N von 21 sollen keine weiteren statistischen Tests herangezogen werden - schätzen die Leiter die einzelnen Fragen als etwas wichtiger ein; größere Unterschiede finden sich nur bei 4 Fragen: Leiter schätzen das Ein-

gehen auf Verständnisschwierigkeiten (Frage 5), das gute Arbeitsklima (Frage 7) und das Eingehen auf die Teilnehmer (Frage 34) wichtiger als die Experten ein, während die Experten die Teilnehmerpartizipation bei Entscheidungen über den Ablauf der Veranstaltung (Frage 28) als wichtig, die Leiter diesen Aspekt aber nur mit "3" (eher wichtig als unwichtig) ansehen. Dies scheint die speziellen Erfahrungen mit einer bestimmten Teilnehmergruppe zu reflektieren.

Im allgemeinen werden die einzelnen Aspekte als wichtig angesehen; eine schlechtere Bewertung erhalten die Globaleinschätzungen Frage 13 und 31. Dies wurde aber auch erwartet, da mit diesen Fragen wenig spezifisches erfaßt wird. Sie wurden aber u.a. in den Fragebogen aufgenommen, um ansatzweise ein Kriterium für die Regressionsanalysen zu haben.

Faßt man nach den in Kapitel 6.1.6 beschriebenen 4 Skalen die Beurteilung der Relevanz der einzelnen Aspekte für die Skalen zusammen - und bildet so Relevanzwerte für die Skalen (vgl. Anlage 25, Tabelle 2) -, so ergibt sich eine gute Relevanz zwischen 1,8 und 2,6.

Insgesamt kann den einzelnen Fragen bzw. den nach den Analysen gebildeten Skalen zugesprochen werden, daß sie für die Gestaltung von Fortbildung wichtige Aspekte ansprechen.

- Ein grundlegender Anspruch an ein Evaluationsinstrument ist, daß es für die praktische Arbeit der Adressaten nützlich sein soll (vgl. Kapitel 2.). Speziell wurde bisher die Nützlichkeit zur Vorbereitung von gemeinsamen Auswertungsgesprächen angesprochen, eine Idealsituation, die unter den Bedingungen der Erprobungsphase nicht erreicht werden konnte. Insofern bezieht sich der Begriff des Nutzens hier leider auf eine eingeschränkte Möglichkeit des Fragebogens. Zwischen "1" (sehr nützlich) und "6" (völlig unnützlich) wurde dem Fragebogen insgesamt eine mittlere Nützlichkeit von 2,7 (Experten) und 3,2 (Leiter) zugesprochen.

Es ist zu beachten, daß die Experten hier einen hypothetischen Nutzen einschätzen, während die Fortbildungsleiter den realen Nutzen für ihre praktische Arbeit angeben. Die Leiter wurden auch nach dem Nutzen der Informationen bezüglich jedes Aspektes einzeln (vgl. Anlage 26, Tabelle 1) befragt. Nimmt man den theoretischen Mittelpunkt der Skala von 3,5 als Ort, wo die Einschätzung in "unnütz" umschlägt, so betrifft dies wieder nur die Globaleinschätzungen Frage 13 und 31; ansonsten liegt die durchschnittliche Einschätzung der Nützlichkeit mit 2,5 sogar höher als die Gesamteinschätzung der Nützlichkeit des Bogens; ähnlich das Bild, wenn die Einzeleinschätzungen wieder für die Skalen zusammengefaßt werden (vgl. Anlage 26, Tabelle 2).

Genauere Aufklärung über die hier zugesprochene nur "mittlere" Nützlichkeit läßt sich von den Antworten auf die offene Frage nach dem "Nutzen" (Frage 5b) erwarten:

Die Experten äußerten mehrfach, daß die im Bogen gestellten Fragen zu allgemein, zu unspezifisch seien; solche Aussagen seien teils beliebig interpretierbar und würden wenig Informationen in bezug auf konkrete Verbesserungsmöglichkeiten geben. Auf einige angesprochene Aspekte (z.B. des äußeren Rahmens) hätten die Fortbildungsleiter keinen Einfluß, so daß solche Informationen für sie unnütz seien. Betont wird aber auch, daß der Fragebogen die persönlichen Erfahrungen und Erkenntnisse des Referenten relativieren kann und daß der Fragebogen ergänzt werden sollte durch Gespräche mit Teilnehmern bzw. einer gemeinsamen Auswertung des Bogens.

Die Fortbildungsleiter - die den Nutzen der Informationen ja "hautnah" erlebt hatten - benutzten die Fragebogeninformationen zureigenen Reflexion der Veranstaltung, zur Vorbereitung der folgenden Veranstaltungen - aus dem Bogen erkennbare Wünsche und Anregungen sollten berücksichtigt werden - und zur gezielten Auswahl von bestimmten Aspekten für Gespräche mit den Referenten (vgl. auch VERON o.J., S. 22). Ein Problem wird aber darin gesehen,

daß der Bogen zu umfangreich und zu komplex sei. Dies verhindere einen weitaus sinnvolleren Einsatz während der Veranstaltung, z.B. an jedem Lehrgangsabend.

Ein Leiter benützte die Informationen nicht, da ihn störte, daß der Bogen Effektivität nur als abhängig von den Fortbildnern ansähe; ein anderer sah sich durch den Bogen in seinen eigenen Erfahrungen bestätigt; ein dritter freute sich über positive Bewertungen, konnte negative Äußerungen aber nicht benützen, da sie sich in diesem Fall nur auf Aspekte des äußeren Rahmens (Haus, Verpflegung, ...) bezogen, die kurzfristig nicht zu verändern seien.

Diese teilweise kritischen Äußerungen der Experten und Fortbildungsleiter zur Nützlichkeit des Fragebogens scheinen aber mit den einschränkenden Bedingungen des Entwicklungs- und Erprobungsstadiums zusammenzuhängen: Ein Evaluationsinstrument, das in den verschiedensten Veranstaltungen einsetzbar sein soll, beinhaltet fast zwangsläufig allgemeinere und damit weniger situationsspezifische Fragen. Diese Unspezifität läßt sich dann aufheben, wenn die Informationen den Input für ein gemeinsames Auswertungsgespräch zwischen Teilnehmern, Referenten und Leitern bieten - ein Kontext, der in der Erprobungsphase aus untersuchungspraktischen Gründen nicht erreicht werden konnte. Der Fragebogen - als Routineinstrument - bekommt dann die Funktion, die Informationssammlung zu verobjektivieren. Minderheitsvoten und auch Meinungen von im offenen Gespräch gehehmten Teilnehmern zu sammeln. Die Nützlichkeit kann - wegen der schnelleren Auswertung und Rückkoppelung in die Veranstaltung - noch gesteigert werden, wenn der Fragebogen weiter verkürzt und präzisiert wird. In diesem Fall kann er auch als Tagesreflexionsinstrument eingesetzt und die Informationen direkt zur Verbesserung der laufenden Veranstaltung benutzt werden.



- Die Auswirkungen des Fragebogenausfüllens auf die Teilnehmer (Frage 6) wurden unterschiedlich gesehen. Überwiegend vermuten Experten und Fortbildungsleiter, daß durch das Ausfüllen bei den Teilnehmern Reflexionsprozesse in Gang gesetzt werden, die die eigenen, oft diffusen Erfahrungen strukturieren und differenzieren helfen. Es können Maßstäbe für die Beurteilung gewonnen werden, die zudem eine zukünftige kritischere Teilnahme an Fortbildung anbahnen könnten. Der Fragebogen kann auch als Anregung zur Auswertung von Elternversammlungen und Mitarbeiterbesprechungen betrachtet werden. Die Teilnehmer fühlen sich durch das Ausfüllen ernst genommen; eine weitere Auswirkung besteht auch darin, daß der Teilnehmer "einiges loswerden" kann.

Kritische Anmerkungen beziehen sich darauf, daß die Ergebnisse nicht mit den Teilnehmern diskutiert werden konnten und deshalb die Teilnehmer meinen, daß die Informationen nur für die Veranstalter von Nutzen seien; zudem könnten nach der Veranstaltung Mißverständnisse und Unstimmigkeiten nicht aufgeklärt werden. Der Fragebogen sei recht lang und teilweise hätten die Teilnehmer Schwierigkeiten beim Ausfüllen gehabt; dieses sei eine größere Arbeitsbelastung, die nicht selten als reine Pflichterfüllung betrachtet würde.

Trotz dieser kritischen Anmerkungen tendieren die Experten und Fortbildungsleiter dahin, daß die Evaluation - hier das Ausfüllen des Bogens - für die Teilnehmer eine Hilfe zur Strukturierung eigener Erfahrungen bietet, mithin dem Teilnehmer "etwas bringt" und ihn nicht nur als "Objekt" beläßt.

- Danach befragt, ob sie den Fragebogen weiterverwenden würden, äußerten sich die Fortbildungsleiter zwar skeptisch, aber teilweise zustimmend (zwei Leiter dagegen

zogen rein informelle Reflexionen vor). In der Erprobungsphase erhielten die Leiter die Auswertungsinformationen ja erst einige Tage nach Abschluß der Veranstaltung. Dadurch wären gemeinsame Gespräche mit den Teilnehmern nicht möglich gewesen und die Fragebogeninformationen hätten somit direkt wenig Nutzen gehabt. Andererseits sollte - eingebettet in mündliche Auswertungsgespräche - der Fragebogen prinzipiell weiter verwendet, aber gestrafft und reduziert werden, damit er auch zwischendurch nutzbringend angewandt werden kann. Reduzierung des Umfangs und Einbettung in Auswertungsgespräche wären somit Bedingungen der Weiterverwendung.

- Stichwortartige Anmerkungen zu Ende der Vorlagen bezogen sich teils auf schon Besprochenes. Es seien hier nur - sehr verkürzt - einige besonders kritische Anmerkungen wiedergegeben:

Zwei Befragte äußerten eine Skepsis gegenüber dieser Effektivitätskontrolle, bei der gerade durch den Fragebogen die Teilnehmer zu Objekten par excellence würden. Demgegenüber wird der Ansatz der Aktionsforschung vertreten und "bloße" Fragebogenaktionen abgelehnt. Soll Aktionsforschung aber Hilfen für Kursleiter bereitstellen, um Kurse "rollend" auswerten zu können, und versteht man darunter auch eine "methodisch" angeleitete Auswertung, so wird kein Widerspruch zu dem hier dargelegten Versuch gesehen, ein Evaluationsinstrument zu entwickeln.

Eine andere Anmerkung - die von "ärgerlicher Mittelschichtorientierung" spricht - ähnelt eher einem Klischee. Einmal wird hier diese "Mittelschichtorientierung" nicht gesehen und zum anderen entstanden die Formulierungen mit Erziehern zusammen, konkreten Fortbildungsteilnehmern also und wenn deren Sprache "mittelschichtorientiert" ist, so entspricht der Bogen nur dem Adressatenkreis.

Weiterhin wurde kritisiert, daß dieser Fragebogen nicht auf heute diskutierte, innovative Formen von Fortbildung - Aqs, Blocklehrgänge usw. - passen würde. Aus diesem Grunde wurde der Fragebogen abgelehnt. Diese Ablehnung übersieht aber zweierlei: ein sehr großer Teil von Fortbildung findet in der Form von mehrtägigen, zentralen

Lehrgängen statt und zum anderen hat diese Fortbildungsform - vgl. Kapitel 1 - in den meisten innovativen Konzepten einen besonderen Stellenwert.

Faßt man die Ergebnisse der Experten- und Fortbildungsleiterbefragung kurz zusammen, so läßt sich - trotz teilweise kritischer Stimmen - dem Fragebogen zusprechen, daß die Fragen eine gute bis befriedigende Auswahl aus dem Gesamt der zur Zeit diskutierten Postulate zur Gestaltung von Fortbildung darstellen, mithin er also inhaltsvalide ist. Die erfragten Aspekte können weiterhin als wichtig für die Gestaltung von Fortbildung angesehen werden. Die mittels Fragebogen erhobenen Informationen haben eine mittlere Nützlichkeit für die praktische Arbeit. Durchweg äußern aber die Experten und Fortbildungsleiter - und hier stimmen sie mit den in Kapitel 2.5 dargelegten Grundgedanken des Evaluationsmodells überein -, daß diese Nützlichkeit gesteigert werden kann, wenn die Fragebogeninformationen in ein gemeinsames Auswertungsgespräch zwischen allen Beteiligten eingebracht werden. Dies setzt voraus, daß der Fragebogen gestrafft und reduziert wird, damit er schnell auswertbar ist und die Ergebnisse mit den Teilnehmern - nicht nur am Ende einer Veranstaltung, sondern evtl. auch in der Mitte - besprochen werden können. Für diese Teilnehmer kann das Ausfüllen des Fragebogens als eine mögliche Reflexion und Strukturierung eigener Erfahrungen betrachtet werden. Evaluation "bringt" dann zugleich etwas für die Teilnehmer die nicht nur als Objekte angesehen werden. Unter den Bedingungen der Kürzung des Fragebogens und der Einbettung in gemeinsame Auswertungsgespräche mit den Teilnehmern sprechen sich die Fortbildungsleiter für die Weiterverwendung eines solchen Bogens aus - es sei hier aber noch einmal daran erinnert, daß diese Aussage mit Vorsicht betrachtet werden sollte, da in der Fortbildungsleiterbefragung nur 6 auswertbare Bögen vorliegen.

#### 6.4 Abhängigkeit der Teilnehmerbeurteilungen von externen Einflußgrößen

Um die Güte der Fragebogeninformationen zur Vorbereitung gemeinsamer Auswertungsgespräche abschätzen zu können, ist es wichtig zu wissen, von welchen anderen Variablen die Teilnehmerratings abhängig sind (vgl. Kapitel 4.3, S. 125f.). Die Beziehungen der Teilnehmereinschätzungen zu diesen Einflußvariablen lassen sich auf zwei Ebenen untersuchen. So können einmal individuelle Teilnehmerdaten - z.B. Alter, Dauer der Berufstätigkeit - mit den individuellen Veranstaltungsbeurteilungen verglichen werden. Von einem "guten" Evaluationsinstrument sollte hier erwartet werden, daß diese individuellen Teilnehmervariablen möglichst wenig Einfluß auf die Einschätzungen der Veranstaltungsaspekte haben. Zum anderen können die Teilnehmerbeurteilungen auf Veranstaltungsebene aggregiert und mit einigen als didaktisch-methodisch für relevant erachteten Veranstaltungsvariablen verglichen werden. Letztere lassen sich bei den Referenten und Fortbildungsleitern erheben. Unterschiede in der didaktisch-methodischen Gestaltung sollten - im Sinne der Konstruktvalidität - bei einem guten Evaluationsinstrument auch ihren Niederschlag in der Einschätzung bestimmter Aspekte der Veranstaltung durch die Teilnehmer finden. Diese beiden Fragen sollen nun verfolgt werden.

Bevor aber Ergebnisse berichtet werden, ist noch eine Vorbemerkung nötig. Die untersuchten Fortbildungsveranstaltungen fanden bei unterschiedlichen Trägern statt. In den Analysen zeigte sich - vgl. weiter unten S. 202f-, daß die Variable "Trägerzugehörigkeit" eine Vielzahl von Unterschieden hinsichtlich des Einzugsbereiches, des Alters der Teilnehmer, des Referententeams, der organisa-

torischen Möglichkeiten u.v.a. umfaßte und somit indirekt Veranstaltungen und damit ihre Beurteilungen beeinflusste. Da hier aber nicht die Abhängigkeit der Beurteilungen von der Trägerzugehörigkeit interessiert, wurde versucht, die möglichen Einflüsse der Trägerzugehörigkeit soweit wie möglich zu kontrollieren.

Technisch wurde bei Variablen, die als "intervallskaliert" interpretiert werden können, neben der Produkt-Moment-Korrelation die partielle Korrelation mit der Kontrollvariable "Träger" als dummy-Variable berechnet. Liegt eine Variable nominal - und die andere intervallskaliert vor, so wurden Einwegvarianzanalysen durch Kovarianzanalysen mit der Kovariate "Träger" ergänzt. Sind beide Variablen als nominal zu betrachten, wurden Kreuztabellierungen mit  $\chi^2$ -Test ausgewählt und Trägereffekte mittels 3-weg Kreuztabellen kontrolliert. An vielen Stellen zeigte sich in den Analysen der Effekt dieser Kontrollmaßnahme darin, daß sich Beziehungsstrukturen nach Einführung der Kontrollvariable veränderten (vgl. z.B. S. 199; über solche Veränderungen wird aber im folgenden nicht berichtet).

Allgemein wurde eine Beziehung wenn möglich mit verschiedenen statistischen Verfahren überprüft, um Abhängigkeiten vom Analyseverfahren zu verhindern (vgl. ROSENSHINE/FURST 1971, S. 63).

Im Text sind gekennzeichnet mit

\* auf dem 5 % Niveau signifikante Koeffizienten

\*\* auf dem 1 % Niveau signifikante Koeffizienten

+++ auf dem 0,1 % Niveau signifikante Koeffizienten.

Für nicht signifikante, aber tendenzielle Beziehungen wird das exakte Signifikanzniveau angegeben (z.B. durch  $p = 8 \%$ ).

#### 6.4.1 Individualebene

Neben den 35 Beurteilungsaspekten wurde im Bogen noch nach dem Alter, der Dauer der Berufstätigkeit, der Anzahl der in den letzten beiden Jahren besuchten Fortbildungsveranstaltungen, den Vorkenntnissen bezüglich des Themas, nach Schwierigkeiten beim Besuch der Veranstaltung und nach der Bekanntheit der Referenten/Leiter gefragt.

Durch einen Fehler in der Vorlage wurde nicht die Bekanntheit des Leiters getrennt von der des Referenten erhoben; in einigen Veranstaltungen bekam der oft sehr bekannte Leiter das Kennzeichen A, in anderen B; dies war aber nachträglich nicht mehr zu erkennen. Aus diesem Grunde kann in den Analysen nur eine durchschnittliche Bekanntheit des Teams von Referenten und Leitern betrachtet werden. Anlage 27 enthält einige Statistiken dieser Variablen, die Interkorrelationen und die Korrelationen dieser Variablen zu den Beurteilungen der 4 Skalen gebildet nach der Faktorenanalyse.

Bei den Interkorrelationen der Variablen (Anlage 27, Tabelle 2 und 3) gibt es neben der hohen, aber trivialen Korrelation zwischen Alter der Teilnehmer und der Dauer der Berufstätigkeit zwar einige hochsignifikante Beziehungen, die aber absolut gesehen recht niedrig ausfallen (solche Signifikanzniveaus sind bei  $N = \text{ca. } 475$  nicht überraschend); am höchsten sind noch die - ebenfalls trivialen - Beziehungen zwischen dem Alter, der Dauer der Berufstätigkeit, Anzahl der in den letzten beiden Jahren besuchten Fortbildungsveranstaltungen und der durchschnittlichen Bekanntheit des Referenten/Leiterteams. Dies ist vor allem auf einen "bekannten" Leiter zurückzuführen, da Fortbildung in einem regionalen Kontext oft von einem bestimmten Leiter durchgeführt wird, der zugleich Praxisberater ist. Die anderen signifikanten Korrelationen sind so niedrig (maximal 2 - 3 % gemeinsame Varianz), daß auf eine Interpretation verzichtet werden soll.

Das entsprechende Bild ergibt sich, wenn man die Individualebene verläßt und die Beziehungen zwischen den aggregierten Daten betrachtet (Anlage 27, Tabelle 4); z.B. zwischen der durchschnittlichen Dauer der Berufstätigkeit der Teilnehmer einer Veranstaltung und der Anzahl der von den Teilnehmern in den letzten beiden Jahren besuchten Fortbildungsveranstaltungen. Im letzten Fall gibt es eine positive Tendenz ( $r = .41$ ) derart, daß in Veranstaltungen mit berufserfahrenen Teilnehmern auch jene sind, die häufiger Fortbildungsveranstaltungen besuchen.

Es lassen sich zwar einige hochsignifikante Korrelationen zwischen diesen individuellen Variablen und der Veranstaltungsbeurteilung finden, die Korrelationen fallen mit  $r = .11$  bis  $r = .26$  aber recht niedrig aus (vgl. Anlage 27, Tabelle 5 und 6), d.h. es sind nur zwischen 1 - 7 % der Varianz gemeinsam. Faktisch gesehen kann somit von

einer inhaltlichen Unabhängigkeit der Veranstaltungsbeurteilung von diesen Variablen gesprochen werden.

Eine durchgehende, sehr leichte Tendenz findet sich nur bei dem Alter der Teilnehmer und der damit korrelierenden Dauer der Berufstätigkeit: Je älter die Teilnehmer sind, je länger sie in ihrem Beruf tätig sind, desto positiver werden die einzelnen Aspekte der Veranstaltung beurteilt (partielle Korrelation von  $r=.11$  bis  $r=.26$ ). Dies braucht keine Tendenz eines "allgemeinen Wohlwollens Älterer" zu sein, der die "kritischere Jugend" gegenübersteht. Vielmehr läßt sich eine solche Tendenz auch dahingehend interpretieren, daß der ältere Teilnehmer aufgrund seiner Erfahrung mehr in der Lage ist, das für ihn Nützliche aus einer Veranstaltung herauszuziehen. Diese Beziehung ist aber faktisch sehr gering; die durch den Fragebogen erfaßten Informationen würden also kaum in dem Sinne verzerrt, daß ältere, berufserfahrene Teilnehmer eine Veranstaltung per se positiver beurteilen.

Allgemein läßt sich festhalten, daß die Beurteilung einer Fortbildungsveranstaltung im Gesamt wie in den einzelnen Skalen unabhängig ist von den Erfahrungen der Teilnehmer mit Fortbildungsveranstaltungen, dem Vorwissen bezüglich des Themas, den Schwierigkeiten, die Veranstaltung zu besuchen und der durchschnittlichen Bekanntheit des Referenten-/Leiterteams.

Bei der durchschnittlichen Bekanntheit des Referenten-/Leiterteams zeigt sich z.B. auch der Effekt der Kontrolle der Trägervariable. Ohne Berücksichtigung der Tatsache, daß Veranstaltungen von verschiedenen Trägern durchgeführt werden, zeigt sich durchweg eine positive Beziehung zur Beurteilung. Diese verschwindet in der partiellen Korrelation; es gibt nur noch eine leichte Tendenz derart, daß mit der Bekanntheit der Referenten und Leiter die positive Beurteilung des Arbeitsklimas steigt (Skala 3,  $r=.08$ , aber nur 1 % gemeinsame Varianz).

Die individuellen Einschätzungen der unterschiedlichen Aspekte einer Fortbildungsveranstaltung werden somit nicht bzw. nur kaum von den angesprochenen individuellen Teilnehmervariablen beeinflusst. Dies bedeutet für ein Auswertungsgespräch, daß z.B. eine positive Beurteilung des Arbeitsklimas nicht allein darauf zurückzuführen ist, daß der Referent/Leiter bei den Teilnehmern sehr bekannt ist oder daß ein älterer Teilnehmer per se alles "positiver" sieht. Vielmehr beurteilen die Teilnehmer anscheinend unabhängig von solchen Einflußgrößen die Qualität einer Veranstaltung. Dies kann als Hinweis auf die Güte der Evaluationsquelle "Teilnehmerratings" betrachtet werden.

Größtenteils entsprechende Ergebnisse erhält man, wenn von der Individualebene auf Veranstaltungsebene übergegangen wird, d.h. z.B. nicht mehr das individuelle Teilnehmeralter mit der individuellen Beurteilung verglichen wird, sondern das Durchschnittsalter einer Veranstaltung mit der durchschnittlichen Teilnehmereinschätzung in den 4 Skalen (vgl. Anlage 27, Tabelle 7).

Beim Durchschnittsalter und der durchschnittlichen Berufserfahrung gibt es zwar nur noch eine leichte, aber spezifische Tendenz zu "Gesamteinschätzung/Methodischer Aufbau" (Skala 1) und "Lernertrag" (Skala 2), die ebenfalls in dem Sinne interpretiert werden kann, daß ältere und berufserfahrenere Teilnehmer aufgrund ihrer Erfahrung tendenziell mehr aus einer Veranstaltung herausnehmen können ( $r=.28$  bis  $r=.36$ ;  $p=6-10\%$ ). Mit durchschnittlicher Bekanntheit des Referenten-/Leiterteams steigt wieder in der Tendenz die positive Beurteilung des Arbeitsklimas ( $r=.33$ ;  $p=7\%$ ); unklar ist, daß je größer der Prozentsatz der Teilnehmer ist, die die Inhalte vorher nicht kannten, desto positiver der äußere Rahmen beurteilt wird ( $r=.37^+$ ).

Eine neue Beziehung ist, daß mit dem Ansteigen des Prozentsatzes der Teilnehmer, die Schwierigkeiten hatten die Veranstaltung zu besuchen, zugleich durchschnittlich der Gesamteindruck/Methodischer Aufbau und der Lernertrag schlechter beurteilt wird ( $r=-.41^+$  und  $r=-.18^+$ ). Es scheint so zu sein, daß gehäufte Probleme die Aufmerksamkeit behindern, was sich vor allem in der Einschätzung des direkten Lernertrages zeigt und nicht in der Einschätzung des Arbeitsklimas. Auf der individuellen Ebene fand sich diese Beziehung nicht.



#### 6.4.2 Veranstaltungsebene

Im Sinne der Konstruktvalidität wird von einem guten Evaluationsinstrument erwartet, daß Unterschiede in der didaktisch-methodischen Gestaltung einer Veranstaltung auch einhergehen mit unterschiedlichen Beurteilungen entsprechender Veranstaltungsaspekte durch die Teilnehmer. Um diese Frage zu verfolgen, wurde den Referenten und Leitern am Ende der Veranstaltung ein kurzer Bogen (vgl. "Fragebogen zur Veranstaltung", Anlage 11) vorgelegt, der nach entsprechenden Charakterisierungen der Veranstaltung fragte. Diese - auf der Ebene von Veranstaltungen - vorliegenden Informationen sollen nun mit den pro Veranstaltung aggregierten Teilnehmereinschätzungen in den 4 Skalen verglichen werden (vgl. Anlage 28).

Zur Aggregation wurden die Mittelwerte pro Veranstaltung in den Skalen SCALE und DIM (vgl. Kapitel 6.1.6) herangezogen. Vor den Vergleichen sind aber einige relativierende Hinweise nötig:

Eine Aggregation auf Veranstaltungswerte beschränkt sich auf die Arbeit mit Durchschnittswerten einer Veranstaltung; damit geht natürlich eine Vielzahl von Informationen verloren, z.B. die Streuung von Werten in einer Veranstaltung in Abhängigkeit von didaktisch-methodischen Variablen. Dies ist problematisch angesichts der Möglichkeit, daß eine bestimmte Variable auf unterschiedliche Teilnehmer unterschiedlich einwirkt (vgl. z.B. BRACHT 1970). Vor einer Verallgemeinerung der Vergleiche muß zudem gewarnt werden, da nur 25 Veranstaltungen als Einheiten vorliegen.

Die Interpretation der Ergebnisse wird ebenfalls noch dadurch erschwert, daß bei Untersuchung einer so großen Anzahl von Beziehungen - z.B. unter Kontrolle der Trägereffekte wurden einschließlich unterschiedlicher Kodierungen 340 Beziehungen zu den Skalen betrachtet - auf dem 5<sup>ten</sup> Niveau schon ca. 17 signifikant erscheinende Koeffizienten zu erwarten sind, obwohl in Wirklichkeit eine Null-Beziehung besteht. Dies muß bei der Interpretation der 33 signifikanten Koeffizienten berücksichtigt werden, die nur dann als "Beziehungen" interpretiert werden sollten, wenn sie sich konsistent in begründbare Erwartungen einordnen lassen.

Desweiteren ist die Erhebung der die Fortbildung charakterisierenden didaktisch-methodischen Variablen mit Vorsicht zu betrachten. Um die an der Untersuchung be-

teiligten Träger, Referenten und Leiter nicht zu sehr belasten, konnte der Bogen nicht vorgetestet werden und fragte zudem nach jedem Aspekt nur mit einer Frage. Mögliche unbewußte Antwortverzerrungen konnten nicht überprüft werden. Es stellte sich heraus, daß dieser Fragebogen für die Veranstalter teils doppeldeutig und ein in der Validität vermutlich eingeschränktes Instrument war. Einige Fragen, d.h. Operationalisierungen von didaktisch-methodischen Aspekten, erwiesen sich als so unglücklich, daß sie hier nicht herangezogen werden können. Dies betrifft: die Frage nach dem größeren Rahmen (5), in dem die Veranstaltung eingebettet war; die Frage nach der Zusammensetzung der Teilnehmergruppe (6), die fast ausschließlich mit "einheitlichem Berufsfeld" umschrieben wurde, während teils wegen anderer Trägerzusammensetzung die Teilnehmer in der 1. Phase aus verschiedenartigen Bereichen kamen; die Frage nach den verwendeten Medien (7), der Planung von Geselligkeit und Freizeit (11) und der Kontrolle von Lernfortschritten (12) wurden schließlich so unterschiedlich verstanden, d.h. sie bezogen sich inkonsistent auf die unterschiedlichsten Aspekte der Veranstaltung, daß eine Auswertung nicht möglich war.

Die internen Beziehungen der erhobenen didaktisch-methodischen Variablen untereinander werden hier nicht durchgehend und systematisch berichtet, da unter der Fragestellung der Validität des Fragebogens die Beziehungen zu den Beurteilungen in den 4 Skalen interessieren. Hinweise auf die Beziehungen untereinander werden nur an einigen, vom Verfasser als besonders interessant erachteten Stellen gegeben.

Die 25 Fortbildungsveranstaltungen fanden bei insgesamt 4 Trägern statt. Unterschiedliche Trägerzugehörigkeit führt aber nicht nur zu signifikant unterschiedlichen Ausprägungen und Konstellationen der anderen didaktisch-methodischen Charakterisierungen einer Veranstaltung, sondern ebenso zu signifikant unterschiedlicher Beurteilung der Skalen. Da unter der Validitätsfragestellung nicht die Abhängigkeit der Beurteilung von der Trägerzugehörigkeit interessiert, sondern von den anderen als didaktisch-methodisch für relevant erachteten Variablen, wurden bei allen folgenden Beziehungen jeweils die Effekte der Trägerzugehörigkeit kontrolliert (vgl. zu Technik S. 197).

Veranstaltungen verschiedener Träger dauern unterschiedlich lang, die vorherrschenden Sozialformen sind verschieden, das durchschnittliche Teilnehmeralter und die durchschnittliche Dauer der Berufstätigkeit differieren ebenso wie die Anzahl der durchschnittlich in den letzten beiden Jahren besuchten Fortbildungsveranstaltungen und die Bekanntheit des Referenten-/Leiterteams; ebenso sind die Vorkenntnisse bezüglich des Themas bei den verschiedenen Trägern unterschiedlich. Neben nicht signifikanten Tendenzen unterscheidet sich vor allem die Beurteilung des "äußeren Rahmens" je nach Träger. Durch die starke Beziehung der anderen didaktisch-methodischen Aspekte zu der Trägerzugehörigkeit wird eine indirekte Beeinflussung der Beziehung dieser Variablen zur Veranstaltungsbewertung erwartet, die eine Kontrolle der Trägereffekte nötig machte. Nach Einführung der Kontrollvariable änderten sich auch viele Beziehungen.

- Die Fortbildungsveranstaltungen wurden in 5 inhaltliche Kategorien eingeteilt; es zeigen sich aber keine Unterschiede in der Beurteilung der 4 Bereiche der Veranstaltungen. Eine Erwartung, daß z.B. Veranstaltungen aus dem musischen/praktischen Bereich im Arbeitsklima oder in dem - hier eher handwerklichen - Lernertrag besser eingeschätzt werden, läßt sich nicht bestätigen.

Es ist hier aber der begrenzte Bereich der Inhalte zu berücksichtigen. So wurden 5 Kategorien gebildet: Veranstaltungen, die sich mit Schwierigkeiten auf Seiten der Kinder beschäftigen/Veranstaltungen, die sich mit der Förderung des musischen Bereiches befassen/Veranstaltungen mit planungs- bzw. curriculumsorientierter Thematik/Veranstaltungen, die sich eher auf die Persönlichkeitsbildung des Erziehers selbst beziehen/sonstige.

Eine interessante Beziehung gibt es zur Fortbildungserfahrung der Teilnehmer (durchschnittliche Anzahl der von den Teilnehmern in den letzten beiden Jahren besuchten Fortbildungsveranstaltungen): Curriculum- und planungsorientierte Veranstaltungen haben Teilnehmer mit der geringsten Fortbildungserfahrung (1,3 besuchte Veranstaltungen), während Veranstaltungen, die sich auf die Förderung des Erziehers selber beziehen, Teilnehmer mit der größten Fortbildungserfahrung haben (2,5 besuchte Veranstaltungen) ( $F = 2,7$ ; knapp nicht signifikant; wird die Trägerzugehörigkeit kontrolliert, so wird die Beziehung mit  $F = 4,6++$  in einer Kovarianzanalyse hochsignifikant). Haben Teilnehmer somit wenig Möglichkeiten, Fortbildungen zu besuchen, so wählen sie Veranstaltungen aus, die sich auf die Planung der Arbeit direkt beziehen; haben sie mehr Möglichkeiten,

so wählen sie eher persönlichkeitsfördernde Veranstaltungen; in der Mitte liegen Veranstaltungen mit der Thematik Problemkinder und musischer Bereich.

- Die durchschnittliche Teilnehmerzahl pro Veranstaltung betrug 22 - 23 ( $s = 5,4$ ; Spannbreite von 11 - 19 Teilnehmern; Frage 3 des Veranstalterbogens). Signifikante Beziehungen zwischen Teilnehmerzahl und Beurteilung gibt es keine, nur zwei leichte Tendenzen: Je mehr Teilnehmer eine Veranstaltung besuchen, desto positiver wird der Lernertrag der Veranstaltung eingeschätzt ( $r = .3$ ,  $p = 9\%$ ). Eine größere Teilnehmerzahl wirkt sich somit nicht unbedingt negativ auf jenes aus, was die Teilnehmer aus der Veranstaltung mitnehmen; "viele Teilnehmer" kann hier auch heißen, daß viele Erfahrungen eingebracht werden, aus denen gelernt werden kann. Andererseits wird der äußere Rahmen einer Veranstaltung schlechter eingeschätzt, je größer die Teilnehmerzahl ist ( $r = -.31$ ,  $p = 8\%$  bei SCALE 4;  $r = -.36$  beim DIM 4,  $p = 5,1\%$ ). (Betrachtet man nicht auf Skalenebene, sondern auf Itemebene die Beziehungen zu den einzelnen Fragen, so finden sich Korrelationen zu der Einschätzung einer passenden Teilnehmerzahl (Frage 3 mit  $r = -.56^+$ ) und eine Tendenz zu der Einschätzung der Raumverhältnisse als störend (Frage 8 mit  $r = -.33$ ,  $p = 7\%$ ).) Die Skala "äußerer Rahmen" reagiert also relativ empfindlich auf die Einschränkungen in der Veranstaltung, die eine größere Teilnehmerzahl mit sich bringen.

Einige Beziehungen zu den anderen didaktisch-methodischen Variablen: Je mehr Teilnehmer in einer Veranstaltung sind, desto eher werden nur Informationen vermittelt ( $r = -.53^{++}$ ), eine darbietende Methode gewählt ( $r = -.46$ ) und desto geringer ist die Mitbestimmung bei Entscheidungen über den Ablauf einer Veranstaltung ( $r = -.38$ ) und über die Themenauswahl ( $r = .43^+$ ).

- Die Veranstaltungsdauer betrug 3 - 5 Tage (Frage 4 des Veranstalterbogens); die Beurteilung der verschiedenen Aspekte der Veranstaltung ist unabhängig von der Dauer; eine Tendenz, daß in längeren Veranstaltungen es eher möglich ist, ein gutes Arbeitsklima herzustellen, läßt sich nicht bestätigen.
- Den größten Zeitraum in der Veranstaltung (Frage 8 des Veranstalterbogens) nahmen entweder die Arbeit in der Kleingruppe oder in der Gesamtgruppe ein (je 8 Nennungen) oder in Klein- und Gesamtgruppe gleichermaßen (9 Nennungen). Eine fast hochsignifikante Beziehung gibt es nur zu dem äußeren Rahmen (SCALE 4  $F = 4,9^+$ ; DIM 4  $F = 4,7^+$ ): Überdurchschnittlich wird dieser Bereich eingeschätzt, wenn in der Klein- und Gesamtgruppe gleichermaßen am meisten gearbeitet wurde, während eine überwiegende Zeitverteilung auf entweder Gesamtgruppe oder Kleingruppe zu einer unterdurchschnittlichen Bewertung des äußeren Rahmens führt. Es wurden hier zwar Beziehungen in Abhängigkeit von der verwendeten Sozialform erwartet, aber zugegebenermaßen am wenigsten im Bereich äußerer Rahmen. (Diese signifikante Beziehung kommt vor allem - betrachtet man die Items - durch die Fragen nach den Freizeitmöglichkeiten (Frage 24,  $F = 5,3^+$ ) und nach dem geselligen Beisammensein zustande (Frage 8,  $F = 2,9$ ,  $p = 8 \%$ ).)

Eine weitere interessante, signifikante Beziehung auf Itemebene gibt es zu Frage 22 ( $F = 3,5^+$ ): Die Möglichkeiten zu einem informellen Erfahrungsaustausch über Praxisprobleme wurden in den Veranstaltungen am besten beurteilt, in denen die Veranstalterangaben, überwiegend in der Kleingruppe zu arbeiten. Dies ist leicht zu interpretieren, da in der Kleingruppe es viel stärker zu lockeren Seitengesprächen kommen kann als in der Gesamtgruppe.

Keinerlei Beziehungen zeigen sich zu den Nennungen jener Sozialformen, denen in bezug auf die Veranstaltungsziele die größte Bedeutung beigemessen wurde (Frage 9 des Veranstalterbogens).

- Durchschnittlich wurden von den Referenten/Leitern in den letzten beiden Jahren 10 - 12 Fortbildungsveranstaltungen durchgeführt (Frage 10 des Veranstalterbogens). Dies kann als Hinweis auf die Erfahrung der Referenten/Leiter betrachtet werden.

Auch hier konnte durch einen Fehler in der Vorlage nachher nicht mehr unterschieden werden, welche der Kennzeichnungen A-D sich auf den Referenten und welche sich auf den meist erfahreneren Leiter bezogen. Es konnte wieder nur ein Durchschnittswert der Erfahrung des Teams gebildet werden. Eine mittlere Erfahrung kann nun aber auch dadurch zustande kommen, daß das Team aus einem sehr erfahrenen Leiter besteht, der mit einem unerfahrenen Referenten zusammenarbeitet. Solche Unterschiede bzw. mögliche Auswirkungen auf die Veranstaltungsbeurteilung konnten nicht verfolgt werden.

Die Erfahrung des Referenten-/Leiterteams wirkt sich unterschiedlich in den Veranstaltungsbeurteilungen durch die Teilnehmer aus. Es gibt zwar keine signifikanten Beziehungen, aber eine interessante Tendenz zur Skala 2 "Lernertrag": Je erfahrener das Referenten-/Leiterteam ist, desto negativer wird dieser Bereich eingeschätzt ( $r = -.3$  bei SCALE 2,  $p = 9\%$ ;  $r = -.33$  bei DIM 2,  $p = 7\%$ ). Um diese Beziehung genauer zu verfolgen, wurde auf die Itemebene der einzelnen Fragen zurückgegangen.

Bei höherer Referenten-/Leitererfahrung werden negativ von den Teilnehmern eingeschätzt: die treffende Auswahl des Themas ( $r = -.48^+$ , Frage 1), der Lernertrag (Frage 3,  $r = -.47^+$ , Frage 23,  $r = -.51^+$ ) sowie die Möglichkeit zu geselligem Beisammensein (Frage 8,  $r = -.37^+$ ; knapp nicht signifikant mit  $r = -.32$  Frage 24 nach den Möglichkeiten zur Freizeitgestaltung). Positiv werden bei hoher Erfahrung des Teams von den Teilnehmern beurteilt: die Anregungen zu aktiver Mitarbeit (Frage 17 mit  $r = .37^+$ ) und die Förderung einer effektiven Kleingruppenarbeit (Frage 29 mit  $r = .48^+$ ).

Während bei einem weniger erfahrenen Team die Teilnehmer also meinen, mehr zu lernen bzw. die Geselligkeit besser einschätzen - hier schlägt sich vielleicht das Engagement eines jüngeren Teams durch oder bestimmte mit der Erfah-

rung sich ändernde methodische Vorlieben der Referenten und Leiter -, so scheint sich die größere Erfahrung eher in dem engeren methodischen Bereich einer Aktivierung (z.B. auch durch effektive Kleingruppenarbeit) niederzuschlagen. Diese Tendenz sollte in weiteren Untersuchungen als Hypothese verfolgt werden.

Es scheint sich aber bei der hier angedeuteten Beziehung um ein sehr komplexes Muster zu handeln, das hier wegen der geringen Fallzahl nicht weiter verfolgt werden kann. Mit der Erfahrung des Fortbildungsteams werden nämlich tendenziell eher gruppendynamische Übungen eingebaut bzw. wird eher erarbeitend vorgegangen ( $r = .3$ ,  $p = 8 \%$ ). Zugleich können die Teilnehmer eher bei Entscheidungen über die Themenauswahl ( $r = .54^{++}$ ) und beim Ablauf der Veranstaltung ( $r = .38^{+}$ ) mitbestimmen. Diese Variablen wiederum wirken sich - s.w.u. - differenziert auf Lernertrag und Arbeitsklima aus. Anlage 29 versucht dieses in der vorliegenden Stichprobe gefundene Beziehungsmuster bzw. einen Ausschnitt in eine schematische Darstellung zu bringen.

- Die Fortbildungsveranstaltungen ließen sich auf einem Kontinuum zwischen "reiner Informationsveranstaltung" und "reine gruppendynamische Veranstaltung" einordnen (Frage 13 des Veranstalterbogens). Korreliert man dieses mit der Teilnehmerbeurteilung, so findet sich eine erwartungsge-  
mäßige positive Korrelation zu der Skala Arbeitsklima (SCALE 3  $r = .39^{+}$ ; DIM 3  $r = .42^{+}$ ): Gibt der Veranstalter an, daß er eher gruppendynamische Übungen mit in die Veranstaltung aufnimmt, so steigt die Bewertung des Arbeitsklimas an.

An dieser positiven Beziehung sind einzeln auf Itemebene fast alle die Fragen beteiligt, die die Skala Arbeitsklima bilden. Berechnet man - auf Skalenebene - nicht eine partielle Korrelation, sondern eine Kovarianzanalyse (Kontrolle des Trägereffektes) mit einer in 3 Stufen transformierten Frage 13 des Veranstalterbogens, so zeigen sich keine signifikanten Beziehungen.

Allgemein zeigen sich bei Frage 13 bis 16 des Veranstalterbogens keine Unterschiede, wenn anstatt der 5 stufigen Antwortvorgaben die partiellen Korrelationen mit einer Kodierung in 3 Antwortstufen berechnet werden.

Interessant ist aber eine weitere leichte Tendenz zum Lernertrag (SCALE 2  $r = -.22$ ,  $p = 17 \%$ ; DIM 2  $r = -.24$ ,  $p = 15 \%$ ): Je mehr gruppensdynamische Übungen einbezogen werden, desto tendenziell negativer wird der Lernertrag eingeschätzt (dies geht - auf Itemebene - auf die signifikanten bis hochsignifikanten Korrelationen zu den entsprechenden Items 16 und 23 zurück;  $r = -.55^{++}$  und  $r = -.4^{+}$ ). Unterstellt man bei "Lernertrag" Vermittlung von "handfesten" Informationen und Ideen, so ist diese negative Beziehung nicht unerwartet. Der Fragebogen spiegelt die Beziehung somit relativ sensibel wieder.

- Abstrakt werden in der methodischen Literatur als zwei Pole oft darbietende und erarbeitende Vorgehensweisen genannt; darin ließen sich auch die hier untersuchten Fortbildungsveranstaltungen unterscheiden. An erarbeitende Verfahren wird vor allem der Anspruch erhoben, daß damit Aktivität und Partizipation der Teilnehmer gesteigert werden können, also der Teilnehmer mehr ein aktiver und gleichberechtigter Lerner ist. In diesem Sinne findet sich auch eine signifikante Beziehung zum Arbeitsklima (SCALE 3  $r = .48^{+}$ ; DIM 3  $r = .53^{++}$ ): Je eher der Veranstalter angibt, erarbeitend vorzugehen, desto positiver schätzen die Teilnehmer in dieser Veranstaltung das Arbeitsklima ein (beim DIM 3 zeigt sich dies auch als Tendenz in der entsprechenden Kovarianzanalyse mit  $F = 2,8$ ,  $p = 9 \%$ ).

Neben den völlig entsprechenden Beziehungen auf Itemebene gibt es noch zwei interessante, signifikante Beziehungen: Je eher erarbeitend vorgegangen wird, desto eher entspricht der Stil der Veranstaltung den Vorstellungen der Teilnehmer von einem sinnvollen Lernen (Frage 12,  $r = .45^{+}$ ) und desto eher wird die Kleingruppenarbeit als effektiv angesehen (Frage 29,  $r = .68^{++}$ ). Eine sehr leichte negative Tendenz zeigt sich wiederum zum Lernertrag: je eher erarbeitend vorgegangen wird, desto weniger meinen die Teilnehmer fachlich gelernt zu haben (Frage 16,  $r = -.3$ ,  $p = 8 \%$ ).



- Eine weitere Dimension der Methode - in der sich die Veranstaltungen unterscheiden - ist das Kontinuum zwischen "exakte Vorausplanung" und "keine Vorausplanung" (Frage 15 des Veranstalterbogens). Hier ergeben sich keinerlei lineare Beziehungen zu den Teilnehmereinschätzungen der einzelnen Aspekte. Die Variable "Vorausplanung" wurde deshalb in drei Stufen umkodiert und eine Kovarianzanalyse berechnet, in der Hoffnung, kurvilineare Beziehungen aufzudecken. Ein signifikanter oder tendenziell signifikanter Effekt ließ sich jedoch nicht aufweisen.

Ohne Kontrolle der Trägereffekte zeigt sich zwar ein fast signifikanter Effekt (SCALE 2  $F=3,1$ ,  $p=6,7\%$ ; DIM 2  $F=2,9$ ,  $p=7\%$ ), jedoch nicht in der erwarteten Bevorzugung eines mittleren Ausmaßes an Vorausplanung. Am besten werden Veranstaltungen mit situativem Eingehen auf Teilnehmerwünsche beurteilt, gefolgt von "starker Vorausplanung", und am schlechtesten dagegen war die Einschätzung des Lernertrages bei einem mittleren Ausmaß an Vorausplanung. Bei Kontrolle der Trägereffekte bekommt diese Beziehung aber ein  $F$  von 1,5 mit einem Signifikanzniveau von 24 %.

Die Variablen "gruppendynamische Übungen vs. Informationsvermittlung" und "erarbeitend vs. darbietend Vorgehensweise" korrelieren sehr stark miteinander ( $r=.63^*$ ); eine Beziehung zur "Vorausplanung" gibt es nicht. Es gibt aber eine interessante Beziehung der Variable "Vorausplanung" zum tatsächlichen Ausmaß an Mitbestimmung in der Veranstaltung (Frage 16 des Veranstalterbogens). Je eher es in der Veranstaltung ein starkes Ausmaß an exakter Vorausplanung gibt, desto weniger Mitbestimmungsmöglichkeiten haben die Teilnehmer bei den Lernzielen ( $r=-.39^*$ ) und - schwächer, nicht signifikant - bei der Themenauswahl ( $r=-.24$ ). Diese erwartungsgemäße Beziehung weist auf einen in diesem Aspekt gültigen Veranstalterfragebogen hin.

- Mit einer letzten Frage wurde das tatsächliche Ausmaß der Mitbestimmung der Teilnehmer an Entscheidungen über Themenauswahl, Lernziele und Ablauf erhoben (Frage 16 des Veranstalterbogens).

Die Antworten streuen gleichmäßig mit  $s=0,9$  bis 1,3 über alle drei fünfstufigen Antwortvorgaben. Im Durchschnitt gab es bei der Partizipation bei der Themenauswahl und bei den Lernzielen eine mittlere Mitbestim-

mung (beidemal  $\bar{x} = 3,0$ ), während in einer leichten Tendenz die Mitbestimmung an Entscheidungen über den Ablauf mit  $\bar{x} = 2,6$  etwas stärker war. Am stärksten korreliert die Mitbestimmung bei der Themenauswahl mit der Mitbestimmung bei den Lernzielen ( $r = .72^{***}$ ), während die Mitbestimmung bei Entscheidungen über den Ablauf schwächer mit der Partizipation an Themenauswahl und Lernzielen zusammenhängt ( $r = .45^+$  und  $r = .50^+$ ). Es gibt noch eine interessante - erwartungsgemäße - Beziehung zu den Vorkenntnissen der Teilnehmer ( $r = .38^+$ ): Je mehr die Teilnehmer einer Veranstaltung vorher schon die Inhalte kannten, desto größer ist auch das Ausmaß an Mitbestimmung bei Entscheidungen über Lernziele, da - so läßt sich interpretieren - sie dann auch schon eher den Überblick über mögliche Aspekte des Themas habe.

Mitbestimmung durch die Teilnehmer bei Entscheidungen über die Themenauswahl scheint ein breiter Indikator für das Eingehen auf die Teilnehmerwünsche und - erwartungsgemäß - für das Bemühen um ein gutes Arbeitsklima zu sein. Geben die Veranstalter an, daß sie diese Partizipation den Teilnehmern eher zugestehen, so beurteilen die Teilnehmer das Arbeitsklima dieser Veranstaltung besser (SCALE 3 und DIM 3  $r = -.46^+$ ). Bei der Mitbestimmung an Entscheidungen über Lernziele und den Ablauf der Veranstaltung dagegen zeigen sich auf der Skalenebene keine signifikanten Beziehungen zur Beurteilung des Arbeitsklimas bzw. der anderen Aspekte der Veranstaltung.

Neben den klaren Beziehungen der Mitbestimmung bei der Themenauswahl zu den Fragen, die sich auf das Arbeitsklima beziehen lassen (Frage 6, 17, 21, 22, 28, 29, 33, 3 und 12,  $r = -.36^+$  bis  $r = -.57^+$ ; die negativen Korrelationskoeffizienten kommen durch die Polung der Fragen zustande), zeigen sich auf Itemebene keine klaren Beziehungen und Tendenzen zu den Skalen Lernertrag und Gesamteinschätzung/Methodischer Aufbau, in denen sich eine Mitbestimmung also nicht auswirkt. Eine überraschende positive Korrelation gibt es zur Einschätzung des geselligen Beisammenseins (Frage 8,  $r = .38^+$ ) und der Einschätzung der äußeren Bedingungen (Frage 30,  $r = .4^+$ ), deren Beurteilung mit wachsender Mitbestimmung bei der Themenauswahl sinkt. Diese Beziehung ist aber unklar zu interpretieren.

Auf Itemebene gibt es zur Mitbestimmung bei den Lernzielen kaum Beziehungen zu den Fragen, etwas mehr zu der Mitbestimmung beim Ablauf der Veranstaltung (in der Tendenz wieder zu Aspekten des Arbeitsklimas). Je mehr Mitbestimmung den Teilnehmern zugestanden wird, desto besser schätzen diese ein: den informellen Erfahrungsaustausch (Frage 22,  $r = -.47^+$ ), die Selbständigkeit (Frage 2,  $r = -.33$ ,  $p = 8\%$ ) sowie die effektive Kleingruppenarbeit (Frage 29,  $r = -.32$ ,  $p = 8\%$ ) und die passende Teilnehmerzahl (Frage 3,  $r = -.59^{++}$ ). Für die Güte des Fragebogens spricht auch folgende Tendenz (Frage 19,  $r = -.35$ ,  $p = 6\%$ ): Je eher der Veranstalter ein hohes Ausmaß an Partizipation im Ablauf angab, desto positiver schätzen die Teilnehmer die Informationen über Lernfortschritte ein. Nur wenn es solche Informationen gibt - so läßt sich interpretieren -, können die Teilnehmer wissen, wo sie stehen und am Ablauf mitbestimmen.

Das tatsächliche Ausmaß an Mitbestimmung - vor allen bei der Themenauswahl - wirkt sich somit in einem besseren Arbeitsklima, Wohlfühlen der Teilnehmer und mehr Aktivität dieser aus. Gerade für Erzieherfortbildung, die als regulatives Prinzip "gleichberechtigte Partnerschaft" in der Fortbildung anstrebt, zeigt sich, daß die Forderung nach mehr Mitbestimmung sich auf Seiten der Teilnehmer in der gewünschten Weise auswirkt. Auch wenn sich keine Beziehungen der Mitbestimmung zum Lernertrag ergeben, ist somit dennoch an der Forderung nach Mitbestimmung festzuhalten.

In diesem Abschnitt wurde die Frage verfolgt, wie sich Unterschiede in der didaktisch-methodischen Gestaltung von Fortbildung in den auf Veranstaltungsebene aggregierten Teilnehmerbeurteilungen niederschlagen. Zusammenfassend läßt sich festhalten, daß zwar es einerseits einige nicht erwartete Beziehungen gibt bzw. dort keine Beziehungen auftreten, wo größere Zusammenhänge erwartet worden waren. Andererseits zeigen sich jedoch eine Vielzahl von erwarteten Beziehungen, die sich in Theorieansätze einordnen lassen: Leiter/Referenten von Fortbildungsver-

anstellungen schätzen ihre eigenen Veranstaltungen in einigen didaktisch-methodischen Aspekten unterschiedlich ein. Diese Unterschiede in den Fortbildungsveranstaltungen gehen einher mit erwarteten Unterschieden in der Einschätzung von verschiedenen Veranstaltungsaspekten durch die Teilnehmer. Von besonderer Bedeutung sind nach den vorliegenden Ergebnissen die Beziehungen zum Arbeitsklima, welches verbessert werden kann, wenn gruppendynamische Übungen einbezogen werden, eher erarbeitend vorgegangen wird und den Teilnehmern eine stärkere Mitbestimmung - vor allem im Bereich der Themenauswahl - zugestanden wird. Indem der Fragebogen hier Beziehungsmuster aufzeigt, die sich konsistent in theoretische Überlegungen einordnen lassen, kann dem Fragebogen Konstruktvalidität zugesprochen werden. Mithin wird davon ausgegangen, daß er Veranstaltungswirklichkeit erfaßt und so mit seinen Informationen auch gemeinsame Auswertungsgespräche vorbereiten kann.

## 6.5 Zusammenfassung der Ergebnisse

Zusammenfassend sollen nun die Ergebnisse der zweiten Erprobungsphase auf die beschriebenen Ansprüche (vgl. Kapitel 2 und 4) an ein Instrument zur Evaluation von Erzieherfortbildungsveranstaltungen bezogen werden:

1. Eine entscheidende Aufgabe dieser Erprobungsphase bestand darin sicherzustellen, daß sich in dem hier dargelegten, in Fragebogenitems umgesetzten und nach den Ergebnissen der ersten Phase revidierten Satz von Erfolgskriterien auch jene Maßstäbe finden, die den Adressaten zur Beurteilung einer Veranstaltung wichtig sind (vgl. Kapitel 2.5, S.35f.). Gespräche mit Referenten und Fortbildungsleitern sowie die offene Frage zum Schluß des Teilnehmerfragebogens geben keine Hinweise auf fehlende

Aspekte. Die Experten- und Fortbildungsleiterbefragung zeigt ebenfalls keine fehlenden Bereiche auf; den Fragen bzw. den nach den Analysen gebildeten Skalen wird zugesprochen, daß sie für die Gestaltung von Fortbildung wichtige Aspekte erfassen (vgl. S. 189 f.).

2. Grundlegender Anspruch an ein Evaluationsinstrument ist, daß es für die praktische Arbeit der Adressaten nützliche Informationen erhebt. Durch die Zwänge des Erprobungsstadiums konnte der Fragebogen nicht wie angestrebt unmittelbar zur Vorbereitung gemeinsamer Auswertungsgespräche verwendet werden. Gerade dies kritisieren auch die Experten und Fortbildungsleiter (vgl. S. 190-192 ). Sie sprechen zwar dem Fragebogen eine Nützlichkeit zu, die aber sich erst dann voll zur Verbesserung von Fortbildung entfalten kann, wenn die Fragebogeninformationen in gemeinsame Auswertungsgespräche mit allen Beteiligten eingebettet werden können. Die Befragten wünschen auch eine weitere Straffung und Reduzierung des Umfangs, damit ein Einsatz in der Mitte einer Veranstaltung möglich wird und die Informationen direkt zur Verbesserung der laufenden Arbeit benutzt werden. Unter diesen Bedingungen - die genau dem angestrebten Ziel für den Einsatz des "fertigen" Fragebogens entsprechen - wollen die Fortbildungsleiter den Fragebogen auch weiter verwenden. Einen speziellen Nutzen kann der Fragebogen auch nach Meinung der Experten und Leiter für die Teilnehmer darin haben, daß der Prozeß des Ausfüllens des Bogens eine mögliche Reflexionshilfe und Strukturierung der eigenen Erfahrungen darstellt.

Die folgenden 4 Ergebnisse beziehen sich vorwiegend auf die in Kapitel 4.2 besprochenen Gestaltungsprobleme des Fragebogens. Es sei noch einmal darin erinnert, daß sich zu diesen Fragen in der Literatur wenig eindeutige Aussagen finden ließen:

3. Die Darbietung der Fragebogenitems, in der den Teilnehmern Aussagen vorgelegt werden, denen sie auf einer 4-er Skala zustimmen können, ist den Teilnehmern sehr verständlich; nur 1 von 479 Fragebögen aus der zweiten Phase war nicht auswertbar (vgl. S. 148). Diese Darbietung sollte deshalb für weitere Revisionen des Fragebogens beibehalten werden.

4. Aufgrund der Schiefe der Antwortverteilung in der 1. Erprobungsphase wurden für die 2. Phase alle Items entweder extrem positiv oder sehr abgeschwächt negativ formuliert, um die Wahl der für eine Veranstaltung negativen Seite zu erleichtern (vgl. S. 144). Dies war jedoch nicht erfolgreich; auch in der 2. Phase überwiegen positive Einschätzungen der verschiedenen Aspekte einer Fortbildungsveranstaltung (vgl. S. 150f). Da andererseits die Partikel "haargenau", "teilweise", "zu jeder Zeit sehr wohl" ... oft kritisiert werden, sollte auf sie verzichtet werden; eine Antwortverzerrung zum positiven Ende hin bauen sie nicht ab. Für die statistischen Analysen selbst wurde versucht, die Verteilungsschiefe durch einen umfangreichen, die Interpretation komplizierenden Transformationsprozeß zu vermindern. Bei allen Analysen wurden dann verschiedene Transformationen der Teilnehmerantworten verwendet, vor allem die Originalfragen ITEM bzw. die entsprechenden Skalenergebnisse DIM sowie die prozent-rangtransformierten und in z-Äquivalente umgewandelten Fragen TRAI und SCALE. Bei keiner Analyse zeigen sich aber unterschiedliche Ergebnisse (vgl. z.B. S. 162). Somit kann - und dies ist für die praktische Anwendung des Fragebogens wichtig - immer mit den Originalfragen gearbeitet und auf den komplizierten, wenig anschaulichen Transformationsprozeß verzichtet werden (vgl. auch S. 183).

5. Negativ formulierte Aussagen des Fragebogens lehnen die Teilnehmer in der Tendenz stärker ab als sie positiv formulierten Aussagen zustimmen (vgl. S. 152 - 154). Es kann

nicht ausgeschlossen werden, daß hier ein Polungseffekt vorliegt. Bei einer weiteren Revision des Fragebogens sollten deshalb alle Aussagen gleichgepolt werden. Positive Aussagen sind hier günstiger, da sie in der Tendenz von den Teilnehmern leicht weniger "schief" beantwortet werden. Gleiche Polungsrichtung hat außerdem Vorteile für eine schnelle Auswertung, vor allem auf Skalenebene, da ein "Drehen" einzelner Fragen bei der Addition zu den Skalenwerten nicht mehr nötig ist.

6. Nach den Ergebnissen der Faktorenanalysen wurde für jeden Faktor eine Skala aus je 5 Items durch einfaches Aufaddieren der Markiervariablen gebildet (vgl. S. 178-182); das einfache Aufaddieren ermöglicht für die praktische Anwendung auch eine schnelle, wenig Zeit beanspruchende Skalenbildung. Gegenüber den Einzelitems fassen Skalenwerte verschiedene Informationen zusammen und bieten - angesichts der geringeren Reliabilität eines Tests aus einem Item - einen zuverlässigeren Überblick über einen Bereich. Die hier gebildeten Skalen korrelieren - erwartungsgemäß - teils höher miteinander. Die Reduzierung von 31 auf nun 20 Fragen führt zu keiner anderen Faktorenstruktur und auch nur zu einem geringen, zu vernachlässigenden Informationsverlust auf Skalenebene. Die Reliabilität der Skalen - s.w.u. - wird als zufriedenstellend betrachtet. Bei einer weiteren Revision des Fragebogens können die Fragen nun nach Skalen sortiert dargeboten werden (vgl. auch S. 113, ein verzerrender Effekt wird nicht erwartet. Die Gruppierung nach Skalen ermöglicht auch eine schnelle Bildung der Skalenwerte.

Die folgenden Ergebnisse beziehen sich auf die in Kapitel 4.3 diskutierten möglichen Verzerrungen bei der Verwendung von Ratingskalen und Fragebögen:

7. Als eine besonders wichtige, ein Auswertungsgespräch behindernde Fehlerquelle wird der Halo-Effekt betrachtet,

d.h. die Tendenz, sämtliche Einzelbeurteilungen einem allgemeinen und ganzheitlichen Eindruck von der Veranstaltung unterzuordnen. In den Faktorenanalysen zeigt sich aber kein allgemeiner Faktor, auf dem alle Items höher laden, sondern eine spezifische, komplexe Faktorenstruktur (vgl. S. 160), mithin dürfte der Halo-Effekt nur niedrig ausfallen. Dies wird unterstützt durch die Ergebnisse der Regressionsanalysen (vgl. S. 169f.). Der Fragebogen erhebt somit spezifische Informationen bezüglich der Einzelaspekte einer Fortbildungsveranstaltung.

8. Der Fragebogen erhebt die Information mit einer zufriedenstellenden Zuverlässigkeit: auf Gesamtfragebogenebene beträgt die Reliabilität  $r_{tt} = .85$  (vgl. S. 151f) und auf Skalenebene  $r_{tt} = .59$  -  $r_{tt} = .77$  (vgl. S. 181). Die niedrigere Skalenreliabilität ist aber noch akzeptabel, da die Skalen nur aus je 5 Fragen bestehen; würde man die Skalen auf z.B. 20 Fragen aufwerten, so wäre eine Reliabilität um  $r_{tt} = .85$  bis  $r_{tt} = .93$  zu erwarten.

9. Den größten Raum nahm die Untersuchung der Validität des Fragebogens ein, da Fragebogeninformationen ein Auswertungsgespräch nur dann sinnvoll vorbereiten, wenn sie der Veranstaltungswirklichkeit entsprechen. Die Validität wurde auf verschiedenen Wegen überprüft:

- Die Faktorenanalysen weisen auf eine faktorielle Validität hin; die untersuchten Items lassen eine inhaltlich deutlich zu interpretierende Struktur erkennen (vgl. S. 156-160). Während sich die Faktoren "Methodischer Aufbau" und "Lernertrag/Gesamteinschätzung" der 1. Erprobungsphase in der zweiten etwas vermischen - es entstehen die Faktoren "Lernertrag" und "Gesamteinschätzung/Methodischer Aufbau", die hoch miteinander korrelieren -, lassen sich die Bereiche "äußerer Rahmen" und "Arbeitsklima" sehr gut reproduzieren. Von besonderer Bedeutung ist hier der gut identifizier- und interpretierbare Faktor "Arbeitsklima", der als Indikator für das regulative Prin-



zip "gleichberechtigte Partnerschaft" betrachtet werden kann. In der zweiten Erprobungsphase sind die 4 Faktoren von der technischen Seite her sehr gut bestimmt und lassen sich bei unterschiedlichen methodischen Akzentuierungen der Analysen stabil wiederfinden. Aus diesem Grunde wurden die 4 Faktoren auch der Skalenbildung zugrunde gelegt.

- Den inhaltlichen Zusammenhang der Fragebogenitems untersuchte auch die Regressionsanalyse. Es wurde hier die Frage nach dem Präferenzsystem der Teilnehmer verfolgt, d.h. nach jenen Aspekten, die für die Teilnehmer besonders wichtig sind und so eine Gesamteinschätzung vorhersagen können. Tendenziell tragen jene Fragen am meisten zur Vorhersage bei, die sich auf die Einschätzung des Lernertrages beziehen; d.h. bei hoher Einschätzung dessen, was die Veranstaltung für die Praxis "bringt", läßt sich eine gute Gesamteinschätzung vorhersagen. Aspekte des Arbeitsklimas und des äußeren Rahmens scheinen dabei kaum eine Rolle zu spielen. Die Ergebnisse der Regressionsanalyse sollten aber mit Vorsicht betrachtet werden, da kein unabhängig von den Prädiktoren erhobenes Kriterium vorliegt. Der Zusammenhang der Prädiktoren mit dem aus zwei Items additiv gebildeten Kriterium ist nicht sehr ausgeprägt ( $R = .74$ , die Linearkombination erklärt nur ca. 55 % der Varianz des Kriteriums). Dies könnte darauf hinweisen, daß nicht alle für die Teilnehmer relevanten Aspekte in den Fragebogen aufgenommen wurden - betrachtet man das hier benutzte Kriterium als eine gute und reliable Gesamteinschätzung. Es scheint aber wahrscheinlicher, daß dieses Kriterium nur eine unzuverlässige und undifferenzierte Gesamtbeurteilung darstellt, welche die Fähigkeit der Teilnehmer, verschiedene Einzelaspekte in einem Globalurteil zu integrieren, überfordert. Solange nicht unabhängig von dem Itemsatz ein Kriterium reliable erhoben werden kann, welches eine wirkliche Gesamteinschätzung umfaßt, bleibt dieses regressionsanalytische Vorgehen unbefriedigend.

- Die vielfältigen Unterschiede in der Qualität der Erzieherfortbildungsveranstaltungen spiegeln sich sensibel in dem Fragebogen wider (vgl. S. 170-176). Eine Diskriminanzanalyse identifizierte 4 inhaltliche Dimensionen der Unterschiedlichkeit: am stärksten unterscheiden sich die Veranstaltungen in dem äußeren Rahmen; es folgt der Bereich der Einschätzung des Lernertrages und des methodischen Aufbaus; schon schwächer lassen sich Unterschiede in den Aktivitätsmöglichkeiten der Teilnehmer und schließlich in der Freizeitgestaltung finden. Diese Diskriminanzfunktionen zeigen Ähnlichkeiten zu den 4 Dimensionen bei der Faktorenanalyse. Es läßt sich schließen, daß die 4 Faktoren somit auch die Bereiche darstellen, in denen sich die Unterschiede zwischen den Fortbildungsveranstaltungen niederschlagen. Dies rechtfertigt die Verwendung der 4 Faktoren bei der Skalenbildung und bei einer möglichen weiteren Revision des Fragebogens. Auf die Güte des Fragebogens weist auch noch hin, daß die Anzahl der richtigen Klassifikationen eines Teilnehmers in die von ihm besuchte Veranstaltung enorm verbessert werden kann, wenn für die Zuordnung Fragebogeninformationen herangezogen werden können.

- Die kriterienbezogene Validität des Fragebogens ließ sich nur schwer überprüfen, da andere Instrumente/Kriterien nicht entwickelt und erprobt sind. Betrachtet man aber die Referenten-/Leitereinschätzungen auf dem identischen Fragebogen als ein mögliches Außenkriterium, so läßt sich dem Fragebogen eine ausreichende Validität nicht absprechen. Unabhängig vom Verfahren des Vergleichs (Q-Maß;  $\chi^2$ -Test) stimmen - trotz eigenem Beitrag der Teilnehmer - die Referenten-/Leitereinschätzungen weitgehend mit den Einschätzungen der Teilnehmer überein (S. 184 - 187).

- Die Experten- und Fortbildungsleiterbefragung weist eine befriedigende Inhaltsvalidität und Relevanz der erhobenen Informationen aus (vgl. S. 188ff). Die Fragen

stellen eine gute bis befriedigende Auswahl aus dem Gesamt der zur Zeit diskutierten Postulate zur Gestaltung von Fortbildung dar; die erfragten Aspekte werden weitgehend als wichtig zur Bewertung von Fortbildung angesehen.

- Vergleiche mit einigen individuellen Teilnehmervariablen ergaben, daß die Beurteilungen der einzelnen Aspekte einer Fortbildungsveranstaltung nicht bzw. nur kaum beeinflusst werden durch das Alter der Teilnehmer, der Dauer ihrer Berufstätigkeit, ihr Vorwissen bezüglich des Themas, ihren Fortbildungserfahrungen, der Bekanntheit des Referenten/Leiters und den Schwierigkeiten, die die Teilnehmer hatten, die Veranstaltung zu besuchen (vgl. S. 197 - 200). Für ein Auswertungsgespräch bedeutet diese Unabhängigkeit, daß eine Teilnehmerbeurteilung nicht auf diese Variablen, sondern auf die Veranstaltung selbst zurückgeführt werden muß. Teilnehmerratings sind somit eine gute, nicht von individuellen Daten verzerrte Evaluationsquelle.

- Die Fragebogeninformationen wurden auf Veranstaltungsebene zusammengefaßt und mit einigen didaktisch-methodischen Merkmalen - die bei den Referenten/Leitern erhoben wurden - verglichen (vgl. S. 201- 212). Unterschiede in der didaktisch-methodischen Gestaltung von Fortbildung gehen einher mit erwarteten, theoretisch begründbaren Unterschieden in der Beurteilung der Veranstaltungsaspekte durch die Teilnehmer. Die Beurteilung des hier als besonders relevant erachteten Arbeitsklimas steigt an, wenn eher gruppendynamische Übungen einbezogen bzw. erarbeitend vorgegangen und den Teilnehmern eine Partizipation bei Entscheidungen über die Veranstaltung ermöglicht wird. Diese Ergebnisse weisen auf eine zufriedenstellende Konstruktvalidität hin.

Die verschiedenen Ansätze zur Überprüfung der Güte der Fragebogeninformationen zeigen alle eine gute und ausreichende Validität auf. Da die Fragebogeninformationen somit Veranstaltungsrealität widerspiegeln, können sie ein Auswertungsgespräch sinnvoll vorbereiten helfen.

## 7. Evaluation von Erzieherfortbildungsveranstaltungen: Zusammenfassung und Ausblick

Die Funktion der Evaluation wird in der vorliegenden Arbeit in der Verbesserung und Optimierung der Erzieherfortbildung gesehen; untersucht wurde dabei, ob ein Fragebogen, der mittels Ratingskalen eine Einschätzung der Veranstaltung durch die Teilnehmer selbst verlangt, ein für diese Funktion nützliches und brauchbares Evaluationsinstrument darstellt. Eine Zusammenschau des Ganges der Arbeit führt nun auf die in der Einleitung aufgeworfenen Frage einer zu empfehlenden Weiterverwendung des Fragebogens zurück: soll dieses Verfahren empfohlen oder abgelehnt werden oder ist eine Empfehlung nur bei bestimmten Modifikationen angemessen?

Erzieherfortbildungen (Kapitel 1) - als berufliche Erwachsenenbildung mit Fachkräften für den Vorschulbereich - wird als unumgängliche Maßnahme betrachtet, eine mangelhafte Ausbildung zu kompensieren, der Veralterung des Fachwissens entgegenzuwirken, berufsbegleitend spezifische Qualifikationen zu entwickeln und speziell jene Qualifikationen zu fördern, die Erzieher als potentielle Träger von Reformen benötigen. Findet Fortbildung auch in unterschiedlichen Organisationsformen statt und gibt es nicht die beste Form, sondern ein sich wechselseitig ergänzendes Verbundsystem, so werden die "klassischen", zentralen, mehrtägigen Lehrgänge nicht überflüssig, sondern nehmen auch in innovativen Konzepten einen ausgewiesenen Stellenwert ein, indem sie andere Formen unterstützten sollen.

Evaluation (Kapitel 2) läßt sich anbinden an das allgemeine Interesse der Verbesserung der Erziehungspraxis qua Optimierung von Fortbildung, für die sie - als ein Hilfsmittel - Informationen bereitstellt. Die Adressaten

der Evaluation sind damit primär die direkt an Fortbildung beteiligten Teilnehmer, Referenten und Fortbildungsleiter.

Jede Evaluation braucht Maßstäbe für eine "gelungene" Fortbildungsveranstaltung, mit denen die gerade untersuchte Veranstaltung verglichen werden kann. In dem ziel- und inhaltsfreien Evaluationsansatz werden diese Erfolgskriterien zuerst unabhängig von den Adressaten entwickelt und begründet. Damit soll verhindert werden, daß der Evaluator - in Kenntnis der Ziele der Veranstalter - seine Objektivität verliert und blind gegenüber nicht-intendierten Nebeneffekten wird. Erfolgskriterien lassen sich in zwei Gruppen einteilen: einmal kann die Güte einer Veranstaltung durch die von ihr erzeugten Produkte und Lernergebnisse auf Seiten der Teilnehmer abgelesen werden, zum anderen werden Gestaltungs- und Prozeßmerkmale von Fortbildung dadurch zu Erfolgsmaßstäben, daß sie befriedigend hoch mit den letztlich angestrebten Produkten korrelieren. Die Systematisierung und Begründung eines Satzes von Erfolgskriterien soll Aufgabe einer Literaturanalyse sein, wobei der Bereich der Fortbildungsprozesse heuristisch in Thematik, Methodik, Interaktionsprozesse und situative/organisatorische Bedingungen gegliedert werden kann. Diese Literaturanalyse soll nicht einen theoretischen Ansatz von Fortbildung operationalisieren, sondern sich an verschiedenen Konzepten orientieren, um so für unterschiedlichste Adressaten einsehbare Erfolgsprüfsteine zu entwickeln.

Die Sammlung von Erfolgskriterien bildet aber nur die erste Phase des Evaluationsmodells. Anschließend an eine Literaturanalyse ist sicherzustellen, daß sich unter den gesammelten Kriterien auch jene Maßstäbe befinden, die den Evaluationsadressaten relevant erscheinen.

Das auf die Erfolgskriterien bezogene, fertig entwickelte Evaluationsinstrument soll in einen Kontext gemeinsamer

Auswertungsgespräche zwischen allen Betroffenen eingebettet werden. Teilnehmende Erzieher sind gleichberechtigte Partner der Veranstalter, die ihre Lernerfahrungen selbst reflektieren und auswerten sollen. In dem hier vorliegenden, nicht primär bildungspolitisch bestimmten Bereich wird somit für eine Selbstevaluation durch die Betroffenen plädiert. Gerade eine solche Selbstevaluation bedarf aber eines ausführlich erprobten und qualitativ guten Evaluationsinstruments, welches einen kontrollierten Rückmeldeprozeß erlaubt, in dem die Informationen bezüglich der Erfolgskriterien systematisch und objektiv erhoben werden. Das Evaluationsinstrument soll eine objektivierende Distanz zu den eigenen Erfahrungen herstellen, die Aufmerksamkeit der Teilnehmer kanalisieren, Minderheitsvoten sichern und so die gemeinsamen Auswertungsgespräche vorbereiten helfen. Dies verlangt, daß mit dem Instrument die Informationen schnell und rechtzeitig vor Beginn der Auswertungsgespräche erhoben werden können. Das Instrument stellt dabei nicht den Anspruch auf Totalerfassung aller Informationen, sondern soll ein schnell einsetzbares Routineinstrument sein, das z.B. durch informelle Auswertungsgespräche ergänzt werden muß.

Der für das zu entwickelnde Evaluationsinstrument zentrale Kontext von gemeinsamen Auswertungsgesprächen kann in der Entwicklungs- und Erprobungsphase aber nicht erreicht werden. Vor diesem idealen praktischen Einsatz steht die Notwendigkeit einer ausführlichen empirischen Überprüfung der Qualität des Instruments.

Die bis hier beschriebenen Gedanken zu einem Fortbildungsmodell legten die Leitlinien für die weiteren Kapitel fest. Die ausführliche Literaturanalyse (Kapitel 3) entwickelte einen umfangreichen Satz von Erfolgskriterien. Diese Sammlung wurde auf Fortbildung für Erzieher im Vorschulbereich spezifiziert durch das folgende, oberste Prinzip: in der Fortbildung müssen Erzieher gleichberech-

tigte Partner der Veranstalter und Subjekte selbstbestimmter Lernprozesse sein. Dieses unabhängig von Ziel und Inhalt einer Veranstaltung geltende, alle anderen Erfolgskriterien kontrollierende Postulat ist zwar keine hinreichende, aber eine notwendige Bedingung dafür, daß Erzieher jene spezifische normative Grundgestalt der Interaktion mit dem Kind - die intergenerationelle Reziprozität - herstellen können, die für einen durch eine krasse asymmetrische Interaktionsstruktur gekennzeichneten Arbeitsbereich nötig ist. Die entwickelten Ansprüche an das Evaluationsinstrument - Selbstevaluation unterstützt durch ein Distanz schaffendes Instrument, kontrollierter und objektiver Rückmeldeprozeß, schnelle Auswertung - führte zur Auswahl eines Fragebogens, dessen Items im Sinne von Ratingskalen aufgebaut sind (Kapitel 4). Dieses Verfahren schließt sich der Erlebnisrealität der Menschen an, die - ohne es direkt zu bemerken - sich durchs Leben "raten". Der Evaluationskontext bedingte für die Gestaltung des Fragebogens ein offenes Vorgehen, in dem die Erzieher keine Objekte sind, aus denen mit "Tricks" etwas herausgeholt werden soll, sondern Experten und Fachleute, die ständig wissen, warum sie eine Frage beantworten sollen. Die Güte und Qualität eines solchen Fragebogens kann gemindert werden durch einige Verzerrungen - vor allem durch den Halo-Effekt, mangelnde Reliabilität und Validität. Diese möglichen Verzerrungen wurden in den folgenden Kapiteln umfassend überprüft.

In einer ersten Phase (Kapitel 5) wurde ein Bogen mit 44 Fragen erprobt, der anschließend revidiert und auf 35 Items reduziert wurde. Eine zweite, ausführlichere Erprobung (Kapitel 6) zeigt, daß sich unter den 35 Fragen auch jene Erfolgsmaßstäbe befinden, die den Adressaten zur Bewertung von Fortbildung relevant erscheinen. Experten und Fortbildungsleiter sprechen dem Fragebogen zu, daß er

nützliche Informationen für die Adressaten erhebt - bezogen auf den oben beschriebenen Einsatz des Instruments im Rahmen von gemeinsamen Auswertungsgesprächen. Der Fragebogen erhebt spezifische Informationen, die nur wenig durch einen Halo-Effekt verzerrt werden. Reliabilität und Validität - die von verschiedensten Seiten untersucht wurden - können als zufriedenstellend betrachtet werden (vgl. bes. die Zusammenfassung S. 212-219).

Betrachtet man zusammenfassend die Ansprüche an ein Evaluationsinstrument, so sprechen die dargelegten Ergebnisse für eine Weiterverwendung des Fragebogens in der Alltagspraxis der Auswertung von Fortbildungsveranstaltungen für Erzieher. Es liegt hier ein objektives, reliables und valides Instrument vor, dessen einzelne Fragen für die Adressaten relevante, wichtige und die Veranstaltung umfassende Informationen erheben. Der Fragebogen ist hier als ein bewußt reaktives Instrument zu verstehen, dessen Anwendung Veranstaltungswirklichkeit verändern bzw. verbessern soll und kann. Es wird erwartet, daß der für die Verbesserung der Fortbildungsarbeit entstehende Nutzen den Aufwand an Zeit und Energie überwiegt, den der Einsatz des Fragebogens verlangt. Die praktische Verwendung des Fragebogens in der Auswertung von Fortbildungsveranstaltungen kann somit hier nur empfohlen werden.

Für diesen praktischen Einsatz stellt Anlage 30 ein auf 20 Fragen reduziertes Instrument zur Verfügung. Der Fragebogen beinhaltet die Fragen aus den 4 untersuchten Skalen, für die die Ergebnisse aus Kapitel 6 gelten. Zusätzlich zu den schon oben beschriebenen Revisionen (vgl. S. 213-215) ist für jede Frage eine Begründungszeile vorgesehen, auf der die Teilnehmer - fakultativ - stichwortartige Anmerkungen geben bzw. ihre Einschätzung begründen können (vgl. auch S.109f). Werden auch nicht alle Teilnehmer diese Möglichkeit ausnutzen, so sichert doch der beibehaltene Antwortmodus einmal eine Antwort auf alle Fragen und zum anderen eine schnelle Auswertung. Die Reduzierung auf 20 Fragen beinhaltet einmal kaum eine Informationsverminderung auf Skalenebene (vgl. 179f), ermöglicht aber einen schnell-



leren Einsatz in der Praxis. Anlage 31 enthält für die Auswertung ein Schema, das auch eine Zusammenfassung der Informationen auf Skalenebene erlaubt.

Der Fragebogen sollte in der Auswertungsarbeit zu verschiedenen Zeitpunkten eingesetzt werden. In der Mitte einer Veranstaltung erlaubt er z.B. eine Verbesserung der laufenden Fortbildungsprozesse, zum Abschluß dient er eher einer kritischen Reflexion der Gesamtprozesse und somit evtl. einer Optimierung weiterer Veranstaltungen. Inwieweit aber der Fragebogen den angestrebten Nutzen erreichen kann, hängt nicht zuletzt von den betroffenen Referenten und Fortbildungsleitern und deren Gewöhnung an den Umgang mit einem solchen Instrument ab. Es wird hier die Hoffnung vertreten, daß die vorliegende Arbeit und ihre Ergebnisse zu einer - zumindest probeweisen - praktischen Anwendung motivieren können.

Für eine weitere, zum heutigen Zeitpunkt erst in den Anfängen stehende Erforschung der in Fortbildungsveranstaltungen stattfindenden Prozesse bietet der Fragebogen ebenfalls einige Möglichkeiten. Diese Forschungsarbeit kann durchaus mit dem angestrebten Einsatz des Fragebogens in gemeinsamen Auswertungsgesprächen einhergehen, indem nach diesen Gesprächen die ausgefüllten Fragebögen einer zentralen Auswertung zur Verfügung gestellt werden. Untersucht werden sollte hier z.B. durch Faktorenanalysen die Struktur der Veranstaltungsaspekte in der Wahrnehmung der Teilnehmer. Dabei sollte auch der Ansatz von SOMMER und PETERMANN aufgenommen werden, die bezüglich der einzelnen Aussagen nicht nur eine Einschätzung des Ist-Zustandes verlangen, sondern zugleich auch eine Einschätzung der Relevanz erheben, die die Teilnehmer den entsprechenden Aspekten bei der Bewertung einer Veranstaltung beimessen; anschließend können relevanz-gewichtete Items faktorisiert werden (vgl. SOMMER/PETERMANN 1978; PETERMANN 1979).

Gelingt es desweiteren, eine reliable und valide Gesamt-/Globalbeurteilung zu entwickeln, so kann die Frage der Regressionsanalyse nach dem Präferenzsystem der Teilnehmer neu aufgenommen werden: Welche Aspekte einer Fortbildungsveranstaltung sind in den Augen der Teilnehmer - bzw. auch der Referenten und Fortbildungsleiter - von besonderer Wichtigkeit?

Die größte Bedeutung kann der hier entwickelte Fragebogen aber in der empirischen Erforschung der Wirkungen von unterschiedlichen didaktisch-methodischen Veranstaltungsvariablen haben. Der Fragebogen bietet für die Abschätzung der Effekte dieser Variablen ein mögliches Kriterium, da sich unterschiedliche Gestaltungselemente von Fortbildung auch in unterschiedlichen Beurteilungen der verschiedenen Veranstaltungsaspekte durch die Teilnehmer niederschlagen.

# LITERATURVERZEICHNIS

- ACKERMEIER, H.G. u.a.: Fortbildung an der Basis (Förderprogramm für den Kindergarten. Heft 5.). Münster: Comenius Institut 1978.
- AG FORTBILDUNG DES KINDERZENTRUMS ROENEBERGSTRASSE: Situation und praxisnah. Zur Kritik und Konzeption der Erzieherfortbildung in Berlin (West). In: Neuer Rundbrief. Information über Familie, Jugend und Sport o.J. (1976), 1, S. 9-16.
- AIRASIAN, P.W./MADAUS, G.E.: A Post hoc Technique for Identifying Between Program Differences in Achievement. In: see 4 (1978), 1, S. 1-8.
- ALEAMONI, L.M./SPENCER, R.E.: The Illinois Course Evaluation Questionnaire. A Description of its Development and a Report of some of its Results. In: Educational and Psychological Measurement 33 (1973), S. 669-684.
- ALKIN, M.C.: Evaluation: Who Needs It? Who Cares? In: see 1 (1975), 3, S. 201-212.
- ALLERBECK, K.R.: Meßniveau und Analyseverfahren - Das Problem "strittiger Intervallskalen". In: Zeitschrift für Soziologie 7 (1978), 3, S. 199-214.
- ANDERSON, L.W./SCOTT, C.C.: The Relationships Among Teaching Methods, Student Characteristics, and Student Involvement in Learning. In: Journal of Teacher Education 29 (1978), 3, S. 52-57.
- ARBEITSGRUPPE FORTBILDUNG: Elternarbeit. Teil 1. Anregungen für die Zusammenarbeit von Eltern und Erzieherinnen (Fortbildungsbriefe für Mitarbeiter in Kindertagesstätten. Hrsg. vom SENATOR FÜR FAMILIE, JUGEND UND SPORT). Berlin (West) 1976.
- ARBEITSGRUPPE FORTBILDUNG IM KINDERZENTRUM BERLIN: Erzieherausbildung - Ausbildung zur Handlungsunfähigkeit und was dagegen zu tun ist. In: KOCH, R./ROCHOLL, G. (Hrsg.) Kleinkinderziehung als Privatsache? Köln 1977, S. 181-207.
- ARBEITSGRUPPE FORTBILDUNG UND ERZIEHERINNEN: Bausteine für die Planung der pädagogischen Arbeit in Kindertagesstätten (Fortbildungsbriefe für Mitarbeiter in Kindertagesstätten. Hrsg. vom SENATOR FÜR FAMILIE, JUGEND UND SPORT). Berlin (West) 1976.

- BAUMANN, W./SCHLUTZ, E./SENZKY, K./TIETGENS, H.: Einführung für Kursleiter an Volkshochschulen. Bonn 1976.
- BERLAK, H.: Values, Goals, Public Policy and Educational Evaluation. In: Review of Educational Research 40 (1970), 2, S. 261-278.
- BIEDERMANN, A.: Das Unterrichtsprinzip "Methodenwechsel" und seine Bedeutung für das schulische Lernen. In: Wirtschaft und Erziehung 29 (1977), S. 101-105.
- BLUM, A.: Peak Learning Experience in the Context of curriculum Evaluation. In: see 1 (1975), 1, S. 55-57.
- BÖHME, G.: Über den Begriff des lebenslangen Lernens und seine Folgen. In: Hessische Blätter für Volksbildung 28 (1978), 2, S. 93-101.
- BRACHT, G.H.: Experimental factors related to aptitude treatment interactions. In: Review of Educational Research 40 (1970), S. 627-645.
- BRANDENBURG, A.G.: Der Lernerfolg im Erwachsenenalter. Ergebnisse psychologischer, soziologischer und didaktischer Forschung (Kommission für wirtschaftlichen und sozialen Wandel. Bd. 7). Göttingen 1974.
- BRAUER, T./HERING, S./OELSCHLÄGEL, D.: Evaluation im Modellversuch "Soziale Studiengänge". In: BÜLOW, M. (Hrsg.): Evaluation I. Verfahren - Methoden - Erfahrungen zur Überprüfung universitärer Ausbildung (Hochschuldidaktische Arbeitspapiere. Bd. 8). Hamburg 1977, S. 71-83.
- BROCKMEYER, R.: Wissenschaftliche Begleitung - Zusammenarbeit von Verwaltung und Wissenschaft. In: Unterrichtswissenschaft 6 (1978), 2, S. 113-125.
- BRUNNER, J.: Das Projekt als didaktische Konzeption in der Lehrerfortbildung. In: AREGGER, K. (Hrsg.): Lehrerfortbildung. Projektorientierte Konzepte und neue Bereiche. Weinheim/Basel 1976, S. 16-56.
- BRUNNER, R.: Anmerkungen zur Praxis und Theorie des Trainingsverfahrens "Micro-Teaching". In: OLIVERO, J.L./BRUNNER, R.: Micro-Teaching - ein neues Verfahren zum Training des Lehrerverhaltens. München/Basel 1973, S. 53-73.
- BUDDE, J.: Fachberatung im Kindergarten. In: Welt des Kindes 54 (1976), S. 193-197.

- CHOPPIN, B./FRANKEL, R.: "The Three Most Interesting Things". In: see 2 (1976), 1, S. 57-61.
- CLAUSS, G./EBNER, H.: Grundlagen der Statistik für Psychologen, Pädagogen und Soziologen. Frankfurt a.M./Zürich 1970.
- COOLEY, W.W./LOHNES, P.R.: Multivariate Data Analysis. New York/London/Sydney/Toronto: John Wiley & Sons, Inc. 1971.
- COSTIN, F./GREENOUGH, W.T./MENGENS, R.J.: Student ratings of college teaching: reliability, validity, and usefulness. In: Review of Educational Research 41 (1971), S. 511-535.
- CRONBACH, L.J.: Test Validation. In: THORNDIKE, R.L. (ed.): Educational Measurement, Second Edition. Washington: American Council of Education 1971, S. 443ff.
- Evaluation zur Verbesserung von Curricula. In: WULF, CH. (Hrsg.): Evaluation. Beschreibung und Bewertung von Unterricht, Curricula und Schulversuchen. München 1972, S. 41-59.
- DALLMANN, G.: Curriculumorientierte Lehrerfortbildung für den Sachunterricht (COLFS-Projekt) am Pädagogischen Zentrum in Berlin. In: AREGGER, K. (Hrsg.): Lehrerfortbildung. Projektorientierte Konzepte und neue Bereiche. Weinheim/Basel 1976, S. 119-147.
- DENEKE, K./HOLLMANN, E.: Zur Entwicklung einer Konzeption für die Fort- und Weiterbildung von Fachkräften in der Fach- und Praxisberatung im Kindertagesstättenbereich. In: Nachrichtendienst des Deutschen Vereins für öffentliche und private Fürsorge 57 (1977), 9, S. 284-288.
- DERRY, S./BRANDENBURG, D.C.: Student's Ratings of Academic Programs: A Study of Structural and Discriminant Validity. In: Journal of Educational Psychology 70 (1978), 5, S. 772-778.
- DIEDERICH, J.: Das Üben planen? In: Westermanns Pädagogische Beiträge 30 (1978), 10, S. 389-393.
- DIEHL, B./SCHÄFER, B.: Techniken der Datenanalyse beim Eindrucksdifferential. In: BERGLER, R. (Hrsg.): Das Eindrucksdifferential: Theorie und Technik. Bern/Stuttgart/Wien 1975, S. 157-211.

- DIEHL, J.M./KOHR, H.U.: Entwicklung eines Fragebogens zur Beurteilung von Hochschulveranstaltungen im Fach Psychologie. In: Psychologie in Erziehung und Unterricht 24 (1977), S. 61-75.
- DIETERICH, R.: Psychodiagnostik. Grundlagen und Probleme. München 1973.
- DILLER, A./KIEL, U.: Wie kann Fortbildung Praxis verändern? In: Welt des Kindes 55 (1977), S. 109-114.
- DILLER, A./SCHLAGHECKEN, H.: Fortbildung kein Irrgarten. Vorschläge zur Fortbildung von Leiterinnen. In: Welt des Kindes 57 (1979), 3, S. 232-239.
- EHRHARDT-PLASCHKE, A./IRSKENS, B.: Überlegungen zur Weiterentwicklung der Fortbildung im Elementarbereich. In: Nachrichtendienst des Deutschen Vereins für öffentliche und private Fürsorge 57 (1977), 9, S. 288-292.
- ELMORE, P.B./POHLMANN, J.T.: Effect of Teacher, Student, and Class Characteristics on the Evaluation of College Instructors. In: Journal of Educational Psychology 70 (1978), 2, S. 187-192.
- FALTER, J.W.: Zur Validierung theoretischer Konstrukte - Wissenschaftstheoretische Aspekte des Validierungskonzeptes. In: Zeitschrift für Soziologie 6 (1977), 4, S. 370-385.
- FASSBENDER, S.: Die Beurteilung von Weiterbildungskursen durch die Teilnehmer (Führungskräfte fördern. Eine Reihe des Wuppertaler Kreises. Bd. 1). Köln/Bonn 1974.
- FEGER, H./TROTSENBURG, E. van: Paradigmen für die Unterrichtsforschung. Deutsche Bearbeitung des Kapitels 3: Paradigms for Research on Teaching, von N.L. GAGE. In: INGENKAMP, K. (Hrsg.): Strategien der Unterrichtsforschung. Teilausgabe des Handbuches der Unterrichtsforschung. Weinheim/Basel 1973, S. 11-59.
- FEIG, R.: Psychologische Forschung und Pädagogik des Erwachsenenalters. In: Zef 3 (1972), S. 137-149.
- FISCHER, F.: Erwachsenenbildung - Weiterbildung. Ein Diskussionsbeitrag zu Begriff und Inhalt. In: Erwachsenenbildung 24 (1978), S. 164-168.
- FITTKAU, B.: Kommunikations- und Verhaltenstraining für Erzieher. In: Gruppendynamik (1972), 3, S. 252-274.
- FORSCHUNGSGRUPPE KEIN: Kindergärtnerinnen - Qualifikation und Selbstbild. Eine empirische Untersuchung. Weinheim/Basel 1978.

- FRENCH-LAZOVIK, G.: Predictability of students' evaluation of college teachers from component rations. In: Journal of Educational Psychology 66 (1974), 3, S. 373-385.
- FRICKE, R.: Zusammenhänge zwischen Lehrerverhalten und Schülerleistung. In: Zef 12 (1978), 2, S. 85-94.
- FROMMER, H.: Lernpsychologie für die Praxis der Erwachsenenbildung (Schriften zur Erwachsenenbildung in Baden-Württemberg. Nr. 7). Villingen-Schwenningen 1978.
- FRÜTEL, B./HÜLSTER, M./SPREY, T.: Sozialpädagogische Fortbildung. In: Welt des Kindes 53 (1975), S. 225-232.
- FURNTRATT, E.: Zur Bestimmung der Anzahl interpretierbarer gemeinsamer Faktoren in Faktorenanalysen psychologischer Daten. In: diagnostica 15 (1969), 2, S. 62-75.
- GAENSSLEN, H./SCHUBÖ, W.: Einfache und komplexe statistische Analyse. Eine Darstellung der multivariaten Verfahren für Sozialwissenschaftler und Mediziner. München/Basel 1973.
- GALLAGHER, R.E.: Responsive Evaluation: A Personal Reaction. In: see 2 (1976), 1, S. 33-34.
- GARDNER, P.L.: Research on Teacher Effects: Critique of a Traditional Paradigm. In: British Journal of Educational Psychology 44 (1974), S. 123-130.
- GERL, H.: Analyse von Lernsituationen in der Erwachsenenbildung. In: SIEBERT, H./GERL, H.: Lehr- und Lernverhalten bei Erwachsenen (Theorie und Praxis der Erwachsenenbildung. Bd. 49). Braunschweig 1975, S. 125-178.
- Zur Systematik der Evaluation in der Erwachsenenbildung und ihre Anwendung in Grundqualifikationsseminaren. In: a u e Informationen. Sonderheft 21. Hannover 1977, S. 2-37.
- Soziale Interaktion in Lerngruppen Erwachsener. In: SIEBERT, H. (Hrsg.): Taschenbuch der Weiterbildungsforschung. Baltmannsweiler 1979, S. 515-536.
- GLASS, G.V.: Die Entwicklung einer Methodologie der Evaluation. In: WULF, CH. (Hrsg.): Evaluation. Beschreibung und Bewertung von Unterricht, Curricula und Schulversuchen. München 1972, S. 166-206.
- GROSSE-JÄGER, H.: Zeitschriften lesen lohnt sich. In: Welt des Kindes 55 (1977), S. 115-199.

- GSTETTNER, P./RATHMAYR, B.: Lehrerbeteiligung als Innovationsstrategie. Schulnahe Curriculumentwicklung zwischen Basisanspruch und staatlichem Planungsdilemma. In: Z.f.Päd. 24 (1978), 3, S. 329-349.
- HAEBERLIN, U.: Empirische Analyse und pädagogische Handlungsforschung. In: Z.f.Päd. 21 (1975), 5, S. 653-676.
- HARARI, O./ZEDECK, S.: Development of behaviorally anchored scales for the evaluation of faculty teaching. In: Journal of Applied Psychology 58 (1973), 2, S. 261-265.
- HAWKRIDGE, D.G.: British and American Approaches to Evaluation Studies in Education. In: see 4 (1978), 1, S. 55-70.
- HECHT, L.W.: Measuring Student Behavior during Group Instruction. In: Journal of Educational Research 71 (1978), 5, S. 283-290.
- HEMMER, K.P.: Erzieherfortbildung und Curriculumentwicklung. Problemfelder für Erzieher. In: Welt des Kindes 53 (1975), S. 4-14.
- Was ist Fortbildung? Fortbildung und Fachlichkeit. In: Welt des Kindes 55 (1977), S. 94-100.
- HENGARTNER, E./WEINREBE, H.: Lehrerfortbildung als handlungsorientierte Curriculumentwicklung. Ein Werkstattbericht aus einem Projekt zur Grundschulmathematik. In: Z.f.Päd. 21 (1975), 1, S. 33-53.
- HEYER, A.: Probleme, die sich bei der beruflichen Weiterbildung von Erzieherinnen (Kindergärtnerinnen) abzeichnen: eine Pilotstudie. Unveröffentlichte Diplomarbeit. Institut für Erziehungswissenschaften der Universität Münster 1974.
- HEYER, J.: Fortbildungsveranstaltungen für Erzieher: Konzeption eines integrativen, didaktisch-methodischen Modells. Dissertation, Universität Münster 1977.
- HOFSTÄTTER, P.R.: Gruppendynamik. Kritik der Massenpsychologie. Hamburg 1971.
- HOLM, K.: Die Frage. In: HOLM, K. (Hrsg.): Die Befragung 1. Der Fragebogen - Die Stichprobe. München 1975, S. 32-91.
- HOLZAPFEL, G. u.a.: Konzeption zur Auswertung des Lernprozesses auf dem Bildungsurlaub. Heidelberg 1977.



- HOWE, C.Z.: A Responsive Evaluation of a Graduate Seminar. In: see 4 (1978), 1, S. 19-23.
- HULL, C.H./NIE, N.H.: SPSS UPDATE. New Procedures and Facilities for Releases 7 and 8. New York: McGraw-Hill Book Company 1979.
- IRSKENS, B./MAELICKE, B./PLASCHKE, A.: Fortbildungswerk für Sozialpädagogen - Stand der Entwicklung. In: Nachrichtenendienst des Deutschen Vereins für öffentliche und private Fürsorge 55 (1975), 9, S. 252-257.
- JAEGER, R.M./FREIJO, T.D.: Some psychometric questions in the evaluation of professors. In: Journal of Educational Psychology 66 (1974), 3, S. 416-423.
- JESSEN, K.: Evaluation of Teacher Training on Student-centered Behavior. In: see 1 (1975), 2, S. 95-108.
- JOCHUM, M.: Animation - ein neues Berufsfeld. In: Erwachsenenbildung 24 (1978), 3, S. 127-129.
- JOOST, H.: Förderliche Dimensionen des Lehrerverhaltens im Zusammenhang mit emotionalen und kognitiven Prozessen bei Schülern. In: Psychologie in Erziehung und Unterricht 25 (1978), S. 69-74.
- KEIL, W./PIONTKOWSKI, U.: Strukturen und Prozesse im Hochschulunterricht. Weinheim/Basel 1973.
- KEMMIS, S.: Nomothetic and Idiographic Approaches to the Evaluation of Learning. In: Journal of Curriculum Studies 10 (1978), 1, S. 45-59.
- KIRCHHOFER-BOZENHARDT, A. von/KAPLITZA, G.: Der Fragebogen. In: HOLM, K. (Hrsg.): Die Befragung 1. Der Fragebogen - Die Stichprobe. München 1975, S. 92-126.
- KLAFKI, W.: Zum Verhältnis von Didaktik und Methodik. In: Z.f.Päd. 22 (1976), 1, S. 77-94.
- KLAWE, W.: Determinanten der Fortbildung von Mitarbeitern im sozialpädagogischen Feld. In: Theorie und Praxis der sozialen Arbeit 29 (1978), 2, S. 64-69.
- KLEINBERGER, A.F.: Research and Policy-Making in Education. In: see 2 (1976), 3, S. 215-226.
- KLEINE, D./MERKENS, H.: Überprüfung eines Fragebogens zur Beurteilung von Lehrveranstaltungen. In: Psychologie in Erziehung und Unterricht 26 (1979), S. 149-153.

- KLEITER, E.: Über Theorie und Modell kategorialer Fehler des Lehrerurteils. In: Psychologische Beiträge 15 (1973), S. 185-229. (a)
- Über Referenz-, Interaktions- und Korrelationsfehler im Lehrerurteil. Taxonomie und Sammelreferat von Kategorienfehler. In: Bildung und Erziehung 26 (1973), S. 100-117. (b)
- KÖCK, P.: Verhaltensgestörte Kinder im Elementarbereich (I) und (II). In: Sozialpädagogische Blätter 27 (1976), 1, S. 9-11 und 2, S. 42-44.
- KÖNIG, K.H.: Fortbildungsinstitut in kirchlicher Trägerschaft. In: Welt des Kindes 55 (1977), S. 120-125.
- KOSSOLAPOW, L.: Ausbildung pädagogischer Kompetenz des Erziehers unter persönlichkeitsorientiertem Aspekt. In: Sozialpädagogische Blätter (1978), 2, S. 33-39.
- KREUTZ, H./TITSCHER, S.: Die Konstruktion von Fragebögen. In: KOOLWIJK, J. van/WIEKEN-MAYSER, M.: Techniken der empirischen Sozialforschung. Bd. 4. Erhebungsmethoden: Die Befragung. München/Wien 1974, S. 24-82.
- KRISTOF, W.: Statistische Prüfverfahren zur Beurteilung von Testprofilen. In: Zeitschrift für experimentelle und angewandte Psychologie 5 (1958), S. 520-533.
- KRULL, M./LOEBER, H.D./RIEKEN, H.: Materialien zur nebenberuflichen Qualifikation. O. Einführung. Oldenburg 1977.
- KUYPERS, H.W.: Unterricht mit Erwachsenen. Planung und Durchführung. Stuttgart 1975.
- LANDSHEERE, G.D.: Einführung in die pädagogische Forschung. Weinheim/Berlin/Basel 1969.
- LANGER, I./SCHULZ von THUN, F.: Messung komplexer Merkmale in Psychologie und Pädagogik - Ratingverfahren. München/Basel 1974.
- LARCHER, D./RATHMAYR, B.: "Zielbezogene" versus "zielfreie" Evaluation von Curricula und Unterricht - am Beispiel eines sprachwissenschaftlichen Curriculums. In: FREY, K. (Hrsg.): Curriculumhandbuch. Bd. 2. München/Zürich 1975, S. 688-700.
- LAURENZE, A.: Curriculumentwicklung als Lehrerfortbildung. Perspektiven für eine Didaktik der Lehrerfortbildung? In: Z.f.Päd. 24 (1978), 3, S. 363-372.

- LEON, A.: Psychologie der Erwachsenenbildung. Stuttgart 1977.
- LEWY, A.: Evaluationsmodelle: Theorie und Praxis. In: EDELSTEIN, W./HOPF, D. (Hrsg.): Bedingungen des Bildungsprozesses. Psychologische und pädagogische Forschungen zum Lehren und Lernen in der Schule. Stuttgart 1973, S. 298-325.
- LEWY, A.: Responsive Evaluation: An Interpretation. In: see 3 (1977), 2, S. 143-148.
- LICKONA, T./HASCH, P.: Research on Teacher Centers. In: Educational Leadership 33 (1976), 6, S. 450-455.
- LIENERT, G.A.: Über die Anwendung von Variablen-Transformationen in der Psychologie. In: Biometrische Zeitung 4 (1962), 3, S. 145-181.
- Testaufbau und Testanalyse. Weinheim/Berlin/Basel <sup>3</sup>1969.
- LIMBOS, E.: Der animateur. Seine Aufgaben und seine Position in der Gruppe. In: Deutsche Jugend 25 (1977), 11, S. 504-510.
- LINN, M.C.: Evaluation of the Science Enrichment Activities (SEA) Program: A Decision-Oriented Model. In: see 4 (1978), 2, S. 83-90.
- LOMPE, K.: Wissenschaftliche Beratung der Politik. Ein Beitrag zur Theorie anwendender Sozialwissenschaften (Wissenschaft und Gesellschaft. Bd. 2). Göttingen <sup>2</sup>1972.
- LÜTGERT, W./SCHÜLER, H.: Curriculum - Entwicklung und Lehrer - Fortbildung. In: Z.f.Päd. 24 (1978), 3, S. 351-362.
- MCMORRIS, R.F.: Product Evaluation: A Public Speaking Film. In: see 1 (1975), 3, S. 147-164.
- MONEIL, J.D./POPHAM, W.J.: The Assessment of Teacher Competence. In: TRAVERS, R.M.W. (ed.): Second Handbook of Research on Teaching. Chicago: Rand McNally & Company 1973, S. 218-244.
- MADER, W.: Adressatenbezug: Formel für einen notwendigen Konflikt. In: Volkshochschule im Westen 31 (1979), 3, S. 125-128.
- MADER, W./WEYMANN, A.: Zielgruppenentwicklung, Teilnehmerorientierung und Adressatenforschung. Manuscript. Oktober 1977.

- MAELICKE, B.: Praxisforschung und Praxisveränderung als kooperativer Prozeß zwischen Praxis, Ausbildung und Fortbildung. In: Nachrichtendienst des Deutschen Vereins für öffentliche und private Fürsorge 57 (1977), 9, S. 233-238.
- MARRADI, A.: Die Faktorenanalyse und ihre Rolle in der Entwicklung und Verfeinerung empirisch nützlicher Konzepte. In: Kölner Zeitschrift für Soziologie und Sozialpsychologie 30 (1978), S. 488-513.
- MASTERS, J.R.: High School Student Ratings of Teachers and Teaching Methods. In: Journal of Educational Research 72 (1979), 4, S. 219-224.
- MATUSZEK, P.: Designing Internally-Based Program Evaluations in School Setting. In: Journal of School Psychology 16 (1978), 4, S. 346-355.
- MEISTER, H.: Lehrmethoden, Lernerfolge und Lernvoraussetzungen bei Studenten. Experimentelle Bedingungsanalyse eines hochschuldidaktischen Projekts zur problemorientierten Einführung in die Pädagogische Psychologie (Studien zur Lehrforschung. Bd. 7). Düsseldorf 1974.
- MEYER, H.L.: Einführung in die Curriculum-Methodologie. München 1974.
- MINSEL, W.-R./MINSEL, B.: Teachertraining in "Sozialintegrativen Verhaltensweisen" (Aufbau und Analyse von Verhaltenstrainingskursen mit Lehrern an Gesamtschulen in Schleswig-Holstein). In: ROYL, W./MINSEL, W.-R. u.a.: Teachertraining. Hannover/Darmstadt/Dortmund/Berlin 1973, S. 91-115.
- MINZ, H.: Fortbildung für Erzieher. Funktionen, Ziele, Inhalte. In: HIELSCHER, H. (Hrsg.): Früherziehung in Kindergärten, Vorklassen und Familien. Hannover/Dortmund/Darmstadt/Berlin 1978, S. 77-80.
- MINZLAFF, T./PFEIFF, E.: Organisationsplan zur Anforderungsanalyse. Unveröffentlichtes Arbeitspapier. 1976.
- MITTENECKER, E.: Planung und statistische Auswertung von Experimenten. Eine Einführung für Psychologen, Biologen und Mediziner. Wien 1970.
- MITTER, W./WEISHAUPT, H.: Pädagogische Begleitforschung - eine Entscheidungshilfe? Erste Ergebnisse einer empirischen Untersuchung. In: Mitteilungen und Nachrichten des Deutschen Instituts für Internationale Pädagogische Forschung (1978), 90/91, S. 77-95.

- MOSKAL, E.: Notwendigkeit permanenter Fortbildung. Begründung, Aufgaben und Möglichkeiten. In: MÖRSBERGER, H. (Hrsg.): Der Kindergarten. Bd. 1. Der Kindergarten in der Gesellschaft. Freiburg/Basel/Wien 1978, S. 315-325.
- MÜLLER-SCHÖLL, A./PRIEPKE, M.: Fortbildung - Organisations- und Institutionsberatung. In: Sozialpädagogik 18 (1976), S. 171-177.
- NACHBAUER, K.: Notwendigkeit permanenter Fortbildung. Fortbildungsangebote und Förderung. In: MÖRSBERGER, H. (Hrsg.): Der Kindergarten. Bd. 1. Der Kindergarten in der Gesellschaft. Freiburg/Basel/Wien 1978, S. 326-334.
- NACKEN, W.: Erfolgsbeurteilung von Bildungsprojekten (I). Teil I: Grundlagen der Erfolgsbeurteilung als zweckkritischer Beratung. In: Materialien zur Politischen Bildung (1977), 1, S. 84-92 (a).
- Erfolgsbeurteilung von Bildungsprojekten (II). Strategien der Zielbestimmung und -kritik. In: Materialien zur Politischen Bildung (1977), 2, S. 72-78 (b).
- NESTMANN, F.: Prinzipien der Fortbildung in der sozialen Arbeit. In: Theorie und Praxis der sozialen Arbeit 28 (1977), 8, S. 299-304.
- NIE, N.H. u.a.: SPSS. Statistical Package for the Social Sciences. Second Edition. New York: McGraw-Hill Book Company 1975.
- OLIVERO, J.L.: Besseres Unterrichten durch Micro-Teaching. In: OLIVERO, J.L./BRUNNER, R.: Micro-Teaching - ein neues Verfahren zum Training des Lehrerverhaltens. München/Basel 1973, S. 7-52.
- OPASCHOWSKI, H.W.: Animation. Eine freizeitpädagogische Methode als neue Handlungskompetenz der Jugendarbeit, Erwachsenenbildung und Sozialpädagogik. In: Erwachsenenbildung 24 (1978), 3, S. 113-120.
- PÄTZOLD, G.: Der Einfluß von aktivitätsfördernden Unterrichtssituationen auf die Lernleistung und Lerneinstellung von Berufsschülern und Fachoberschülern. In: Die berufsbildende Schule 29 (1977), S. 559-568.
- PENFIELD, D.A.: Student Ratings of College Teaching. Rating the Utility of Rating Forms. In: Journal of Educational Research 72 (1978), S. 19-22.

- PETERMANN, F.: Ein Fragebogen zur Beurteilung von Ist-Soll-Bedingungen von Lehrveranstaltungen. In: Psychologie in Erziehung und Unterricht 26 (1979), S. 227-229.
- PETZOLD, P.: Direkte Skalierungsmethoden. In: Probleme und Ergebnisse der Psychologie (1978), 66, S. 29-47.
- PEUKERT, U.: Interaktive Kompetenz und Identität. Zum Vorrang sozialen Lernens im Vorschulalter. Düsseldorf 1979.
- POPP, W.: Die Funktion von Modellen in der didaktischen Theorie. In: DOHMEN, G./MAURER, F./POPP, W.: Unterrichtsforschung und didaktische Theorie. München 1970, S. 49-60.
- PRAXISHEFT 14: Der Kindergarten der Arbeiterwohlfahrt (1). Gedanken zur pädagogischen Arbeit und Anregungen für die Praxis. Bonn: Arbeiterwohlfahrt Bundesverband e.V. 1975.
- PRIESNER, P.: Überlegungen zur Weiterbildung der Erzieherin. In: Kindergarten heute 3 (1975), S. 51-60.
- PRIM, R./TILMANN, H.: Grundlagen einer kritisch-rationalen Sozialwissenschaft. Studienbuch zur Wissenschaftstheorie. Heidelberg 1975.
- PROSE, F.: Teachertraining in "Sozial-integrativen Verhaltensweisen" (Gruppendynamisches Training für Lehrer an Gesamtschulen). In: ROYL, W./MINSEL, W.-R. u.a.: Teachertraining. Hannover/Darmstadt/Dortmund/Berlin 1973, S. 116-136.
- RAAPKE, H.D.: Weiterbildung. In: WULF, Ch. (Hrsg.): Wörterbuch der Erziehung. München/Zürich 1974, S. 625-630.
- Lernen und Didaktik in der Erwachsenenbildung. In: Z.f.Päd. 23 (1977), 5, S. 719-727.
- RAKEL, R.E.: A Summary: Responsive Evaluation and Family Practice. In: see 2 (1976), 1, S. 35-36.
- REICHLE, K. u.a.: Ein Plenum muß weder langweilig noch anstrengend sein - Zur Methodik der Strukturierung von Plenumsitzungen. In: Gruppendynamik 8 (1977), 6, S. 431-441.
- RIEMANN, I.: Erzieherfortbildung als Aufgabe der Volkshochschulen. In: Hessische Blätter für Volksbildung 25 (1975), 4, S. 312-317.
- RÖHR, M.: Zur Interpretation faktorenanalytischer Studien in der Psychologie. In: Probleme und Ergebnisse der Psychologie (1977), 62, S. 61-77.

- ROHRMANN, B.:** Empirische Studien zur Entwicklung von Antwortskalen für die sozialwissenschaftliche Forschung. In: Zeitschrift für Sozialpsychologie 9 (1978), 3, S. 222-245.
- ROSENSHINE, B./FURST, N.:** Research on Teacher Performance Criteria. In: SMITH, B.O. (ed.): Research in Teacher Education. A Symposium. Englewood Cliff, N.J.: Prentice-Hall Inc. 1971, S. 37-72.
- ROSSBACH, H.G.:** Vorüberlegungen zur Evaluation von Fortbildungsveranstaltungen im Vorschulbereich. Unveröffentlichte Diplomarbeit. Institut für Erziehungswissenschaft der Universität Münster 1977.
- ROSSMANITH, W.:** Fortbildung - freiwillig oder Pflicht? In: Welt des Kindes 55 (1977), S. 101-104.
- RUMPF, H.:** Divergierende Unterrichtsmuster in der Curriculumentwicklung. In: Z.f.Päd. 19 (1973), 3, S. 391-416.
- SABAR, N.:** The Utility of Various Kinds of Evaluation Data in the Process of Developing a New Biology Science Curriculum. In: see 1 (1975), 3, S. 193-200.
- SACKMANN, G.W.:** "Kriterien" zur Beurteilung einer Unterrichtsstunde. In: Wirtschaft und Erziehung 29 (1977), S. 126-130.
- SAW'N, E.I.:** Curriculum Evaluation or Descriptive Inquiry? In: see 2 (1976), 1, S. 41-51.
- SCHERER, G.:** Methode-Gruppe-Kommunikation. In: Erwachsenenbildung 23 (1978), 3, S. 121-129.
- SCHEYER, P.T./STAKE, R.E.:** A Program's Self-Evaluation Portfolio. In: see 2 (1976), 1, S. 37-40.
- SCHIEMANN, K.K.:** Zielgruppenarbeit an Volkshochschulen. Anspruch auf Offenheit. In: b:e 11 (1978), 2, S. 50-54.
- SCHINDLER, K.:** Wirkung und Erfolg der Weiterbildung. Zu Fragen der Effizienzmessung. Köln 1979.
- SCHLOSSER, O.:** Einführung in die sozialwissenschaftliche Zusammenhangsanalyse. Reinbeck b. Hamburg 1976.
- SCHLUTZ, E.:** Unterricht - ein für die Weiterbildungsfor-schung relevantes Feld? In: SIEBERT, H. (Hrsg.): Taschenbuch der Weiterbildungsfor-schung. Baltmannsweiler 1979, S. 482-514.

- SCHMALOHR, E.: Fortbildung für Kindergartenerzieher im Modellversuch von Nordrhein-Westfalen 1970 und 1971 (Beiträge zur Vorschulerziehung). Hannover/Dortmund/Darmstadt/Berlin 1974.
- SCHMALOHR, E./DOLLASE, R./HOLLÄNDER, A./SCHMERKOTTE, H./WINKELMANN, W.: Vorklasse und Kindergarten aus der Sicht der Erzieher. Ergebnisse einer Erzieherbefragung im Modellversuch von Nordrhein-Westfalen (Beiträge zur Vorschulerziehung). Hannover/Dortmund/Darmstadt/Berlin 1974.
- SCHMIDT, H.D.: Die Beurteilung menschlichen Verhaltens durch Rating-Skalen. Dissertation. Bonn 1965.
- Über die Zuverlässigkeit von Verhaltensbeschreibungen durch Rating-Skalen. In: Archiv für die gesamte Psychologie 118 (1966), S. 47-72.
- SCHMIED, G.: Alternative Einzelfallstudie (Zur Methodologie eines in Vergessenheit geratenen Untersuchungstyps). In: Zef 11 (1977), 3, S. 127-139.
- SCHOTT, E.: Zur empirischen und theoretischen Grundlegung eines Bewertungsinstrumentes für Vorlesungen. Zwischenbericht über ein Forschungsprojekt (Blickpunkt Hochschuldidaktik. Heft 28). Hamburg 1973.
- SCHRADER, H.: Professionalisierung der Erzieher im Rahmen vorschulischer Curriculumentwicklung. In: ZIMMER, J. (Hrsg.): Curriculumentwicklung im Vorschulbereich. Bd. 1 (Erziehung in Wissenschaft und Praxis. Bd. 21/1). München 1973, S. 285-318.
- SCHULENBERG, W.: Erwachsenenbildung. In: GROOTHOFF, H.H. (Hrsg.): Pädagogik (Fischer Lexikon. Bd. 36). Frankfurt a.M. 1971, S. 65-74.
- SCHULZ, W.: Unterricht-Analyse und Planung. In: HEIMANN, P./OTTO, G./SCHULZ, W.: Unterricht. Analyse und Planung. Hannover/Berlin/Darmstadt/Dortmund 1970, S. 13-47.
- SCHWARTE, N.: Fortbildung und Organisationsentwicklung. In: Sozialpädagogik 20 (1978), S. 221-229.
- SCHWARTZ, S.A.: A Comprehensive System For Item Analysis In Psychological Scale Construction. In: Journal of Educational Measurement 15 (1978), 2, S. 117-123.
- SCRIVEN, M.: Die Methodologie der Evaluation. In: WULF, Ch. (Hrsg.): Evaluation. Beschreibung und Bewertung von Unterricht, Curricula und Schulversuchen. München 1972, S. 60-91 (a).



- SCRIVEN, M.: An Introduction to Meta-Evaluation. In: TAYLOR, P.A./COWLEY, D.M. (eds.): Readings in Curriculum Evaluation. Dubuque, Iowa: Brown Company Publishers 1972, S. 84-86 (b).
- Goal-Free Evaluation. In: HOUSE, E.R. (ed.): School Evaluation: The Politics & Process. Berkeley, California: McCutchan Publishing Corporation 1973, S. 319-328.
- SEIDEL, I./BRÜGGEMANN, F.: Erzieherfortbildung auf dem Lande. In: Kindergarten heute 9 (1979), 3, S. 113-119.
- SIEBERT, H.: Curricula für die Erwachsenenbildung (Theorie und Praxis der Erwachsenenbildung. Bd. 27). Braunschweig 1974.
- Probleme, Ergebnisse und Konsequenzen einer empirischen Untersuchung. In: SIEBERT, H./GERL, H.: Lehr- und Lernverhalten bei Erwachsenen (Theorie und Praxis der Erwachsenenbildung. Bd. 49). Braunschweig 1975, S. 13-123.
- Thesen und Materialien zur Didaktik und Methodik. In: SIEBERT, H. (Hrsg.): Praxis und Forschung in der Erwachsenenbildung. Opladen 1977, S. 129-144.
- SKOWRONEK, H.: Lernpsychologische Forschung zum Erwachsenenalter. In: SIEBERT, H. (Hrsg.): Taschenbuch der Weiterbildungsforschung. Baltmannsweiler 1979, S. 286-307.
- SOMMER, J./PETERMANN, F.: Descriptive und präskriptive Aussagen in einem Fragebogen zur Beurteilung akademischer Lehrveranstaltungen. In: Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie 10 (1978), 4, S. 336-346.
- SPRUNG, L./SPRUNG, H.: Zur Theorie und Methodik der Fragebogenkonstruktion-Anspruch, Leistung und Dilemma einer Standardmethodik. Teil I: Wesen und Konstruktionstechniken von Fragebogenmethodiken. In: Probleme und Ergebnisse der Psychologie (1978), 66, S. 5-27 (a).
- Zur Theorie und Methodik der Fragebogenkonstruktion - Anspruch, Leistung und Dilemma einer Standardmethodik. Teil II: Das Dilemma der Fragebogenmethodik und einige Ansatzpunkte sowie Verfahrenstechniken für seine Überwindung. In: Probleme und Ergebnisse der Psychologie (1978), 67, S. 5-24 (b).
- STAKE, R.E.: Objectives, Priorities and other Judgement Data. In: Review of Educational Research 40 (1970), 2, S. 181-212.

- STAKE, R.E.: Verschiedene Aspekte pädagogischer Evaluation. In: WULF, Ch. (Hrsg.): Evaluation. Beschreibung und Bewertung von Unterricht, Curricula und Schulversuchen. München 1972, S. 92-112.
- A Theoretical Statement of Responsive Evaluation. In: see 2 (1976), 1, S. 19-22.
- STONE, E.E./RABINOWITZ, S./SPOOL, M.D.: Effect of Anonymity on Student Evaluations of Faculty Performance. In: Journal of Educational Psychology 69 (1977), 3, S. 274-280.
- STRATON, R.G.: Ethical Issues in Evaluating Educational Programs. In: see 3 (1977), 1, S. 57-66.
- STUFFLEBEAM, D.L.: Evaluation als Entscheidungshilfe. In: WULF, Ch. (Hrsg.): Evaluation. Beschreibung und Bewertung von Unterricht, Curricula und Schulversuchen. München 1972, S. 113-145.
- Towards a Technology for Evaluating Evaluations (Paper presented at the American Educational Research Association Annual Meeting .(Chicago, Illinois, April 15-19, 1974)). Chicago, Illinois 1974.
- TARNAI, Ch.: Methoden der Transformationsanalyse und ihre Anwendung im Rahmen der Faktorenanalyse (D.I.P.-Methodologische Arbeitsberichte). Münster 1977.
- TEEGEN, F./FITTKAU, B.: Erziehung der Erzieher. Erfahrungen aus einem hochschuldidaktischen Experiment. In: Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie 4 (1972), 1, S. 43-50.
- TEIGELER, U.: Fortbildung für Mitarbeiter in Kindertagesstätten. In: Neuer Rundbrief. Informationen über Familie, Jugend und Sport o.J. (1976), 1, S. 2-8.
- TENT, L.: Schätzverfahren in der Unterrichtsforschung. Deutsche Bearbeitung des Kapitel 7: Rating Methods in Research on Teaching, von H.H. REMMERS. In: INGENKAMP, K. (Hrsg.): Schätzen und Messen in der Unterrichtsforschung. Teilausgabe des Handbuches der Unterrichtsforschung. Weinheim/Basel 1973, S. 11-84.
- THÜRNUAU, F.: Veränderungen der sozialen Praxis durch Veränderung der Fortbildung - Zum Fortbildungskonzept einer Fortbildung in zentraler Trägerschaft der Arbeiterwohlfaht. In: Theorie und Praxis der sozialen Arbeit 29 (1978), 5, S. 172-181.

- THURNAU, F./VOLLRATH, D.: Teilnehmerorientierte Fortbildung? - Versuch einer Antwort. In: Theorie und Praxis der sozialen Arbeit 29 (1978), S. 383-391.
- TIETGENS, H.: Zur Beobachtung von Weiterbildungsprozessen (Theorie und Praxis der Erwachsenenbildung. Bd. 28). Braunschweig 1974.
- Adressatenorientierung der Erwachsenenbildung. In: Hessische Blätter für Volksbildung 27 (1977), 4, S. 283-289.
- TIETGENS, H./WEINBERG, J.: Erwachsene im Feld des Lehrens und Lernens (Theorie und Praxis der Erwachsenenbildung. Bd. 88). Braunschweig 1971.
- TIETZE, W.: Fortbildung für Erzieher. Organisationsformen und Institutionen. In: HIELSCHER, H. (Hrsg.): Früherziehung in Kindergärten, Vorklassen und Familien. Hannover/Dortmund/Darmstadt/Berlin 1978, S. 80-84.
- UBERLA, K.: Faktorenanalyse. Eine systematische Einführung für Psychologen, Mediziner, Wirtschafts- und Sozialwissenschaftler. Berlin/Heidelberg/New York 1977.
- ULSHOEFER, H.: Praxisprobleme der Kindergärtnerin. München 1970.
- UNESCO: UNESCO-Empfehlung über die Fortentwicklung der Weiterbildung angenommen von der Generalkonferenz auf ihrer neunzehnten Tagung. Nairobi, den 26.11.1976. In: Deutscher Bundestag 8. Wahlperiode. Unterrichtung durch die Bundesregierung. Drucksache 8/1130 vom 7.11.1977.
- UNGELENK, B./WOLLSCHLÄGER, P./SCHMIDT-DENTER, U./NICKEL, H.: Konzeption und Erprobung eines Verfahrens zur Verbesserung der erzieherischen Handlungskompetenz im Vorschulbereich (Erzieherinnentraining). In: Psychologie in Erziehung und Unterricht 25 (1978), S. 337-344.
- VERON, K.H.: Aspekte der Evaluation von Fortbildungsveranstaltungen im Sozialpädagogischen Fortbildungszentrum des Landes Rheinland-Pfalz. Unveröffentlichte Hausarbeit gem. § 3 Abs. 2 SoArG. o.J.
- VERRES-MUCKEL, M.: Lernprobleme Erwachsener. Befunde und Konzepte für die praktische Arbeit. Stuttgart/Berlin/Köln/Mainz 1974.
- VONTOBEL, J.: Über den Erfolg in der Erwachsenenbildung: Empirische Basisstudie zum Problem der Erfassung des Bildungserfolges (Theorie und Praxis der Erwachsenenbildung. Bd. 4). Braunschweig 1972.

- WALTHER, H./SCHMIDT, U.: Fortbildung für Teilnehmer aus regionalen Arbeitsgruppen pädagogischer Mitarbeiter in Kindertagesstätten. In: Nachrichtendienst des Deutschen Vereins für öffentliche und private Fürsorge 59 (1979), 1, S. 30-33.
- WEBER, G. u.a.: Effizienzkontrolle von Fortbildungsmaßnahmen in der Praxis evangelischer Jugendarbeit. Forschungsprojekt der AEJ im Deutschen Jugendinstitut. Großprienzenau/München 1977.
- WEGNER, R.: Ratingmethoden. In: KOOLWIJK, J. van /WIEKENMAYSER, M. (Hrsg.): Techniken der empirischen Sozialforschung. Bd. 5. Testen und Messen. München/Wien 1976, S. 103-130.
- WEINBERG, J.: Veranstaltungsformen und Arbeitsweisen der Volkshochschulen. In: Volkshochschule im Westen 30 (1978), 4, S. 146-148.
- WINKEL, R.: Zur Theorie und Praxis der Unterrichtsmethoden. In: Die Deutsche Schule 70 (1978), 11, S. 669-683.
- WINNEFELD, F.: Zur Gesetzmäßigkeit pädagogischen Geschehens. In: DOHMEN, G./MAURER, F./POPP, W.: Unterrichtsforschung und didaktische Theorie. München 1972, S. 40-48.
- WITTE, E.H.: Zur Faktorenanalyse echt J-verteilter Items. In: Zeitschrift für experimentelle und angewandte Psychologie 25 (1978), 4, S. 694-701.
- WORTHEN, B.R./OWENS, T.R.: Adversary Evaluation And The School Psychologist. In: Journal of School Psychology 16 (1978), 4, S. 334-345.
- WULF, Ch.: Curriculumevaluation. In: WULF, Ch. (Hrsg.): Evaluation. Beschreibung und Bewertung von Unterricht, Curricula und Schulversuchen. München 1972, S. 15-37.
- Evaluation. In: WULF, Ch. (Hrsg.): Wörterbuch der Erziehung. München/Zürich 1974, S. 204-207.
- Planung und Durchführung der Evaluation von Curricula und Unterricht (Kapitelkonzeption). In: FREY, K. (Hrsg.): Curriculumhandbuch. Bd. 2. München/Zürich 1975, S. 567-579 (a).
- Funktionen und Paradigmen der Evaluation. In: FREY, K. (Hrsg.): Curriculumhandbuch. Bd. 2. München/Zürich 1975, S. 580-600 (b).

ZILLEN, E.: Thesen zu einer Didaktik der Fortbildung für Gruppenleiter in der Jugendarbeit. In: Deutsche Jugend 25 (1977), S. 63-68.

verwendete Abkürzungen:

b:e	betrifft erziehung
see	Studies in Educational Evaluation
Zef	Zeitschrift für erziehungswissenschaftliche Forschung
Z.f.Päd.	Zeitschrift für Pädagogik

## ANHANG

### Verzeichnis der Anlagen

<u>Nr.</u>	<u>Seite</u>
1. Fragebogen zur Beurteilung von Fortbildungsveranstaltungen; Phase 1	1- 6
2. Zum Einsatz des Fragebogens; Phase 1	7- 8
3. Projektbeschreibung	9
4. Transformationsgewichte Gesamtgruppe	10
5. Transformationsgewichte "mehrtägige Veranstaltungen"	11
6. Faktorenanalyse Untergruppe "mehrtägige Veranstaltungen"; Kommunalitäten und Ladungen, 5 Faktoren	12
7. Regressionskoeffizienten, einfache multiple Regressionsanalyse VAR	13
8. Regressionskoeffizienten, einfache multiple Regressionsanalyse VADI	14
9. Fragebogen zur Beurteilung von Fortbildungsveranstaltungen; Phase 2	15-19
10. Zum Einsatz des revidierten Fragebogens; Phase 2	20
11. Veranstalterfragebogen; Phase 2	21-22
12. Rückmeldeschema; Phase 2	23-27
13. Transformationsgewichte	28
14. Zusammenfassung einige Itemstatistiken	29
15. Faktorenanalyse; Kommunalitäten und Ladungen, 4 Faktoren TRAI	30
16. Faktorenanalyse; Kommunalitäten und Ladungen, 4 Faktoren ITEM	31
17. Diskriminanzanalyse, DIRECT; Varianzanteile und kanonische Korrelationen	32
18. Diskriminanzanalyse, WILKS; Diskriminanzkoeffizienten und Korrelationen	33
19. Faktorenanalyse, reduzierter Itemsatz von 20 Fragen; Kommunalitäten und Ladungen, 4 Faktoren TRAI	34
20. Skalen-Statistiken und Fragen-Skalen Korrelationen	35

<u>Nr.</u>	<u>Seite</u>
21. Diskriminanzanalyse auf Skalenebene	36
22. Anschreiben der Expertenbefragung	37-38
23. Vorlage zur Experteneinschätzung	39-41
24. Erste Ergebnisse der Fragebogenentwicklung; Phase 1	42
25. Experten- und Fortbildungsleiterbefragung, Wichtigkeit	43
26. Leiterbefragung, Nützlichkeit	44
27. Vergleiche mit externen Einflußgrößen, Individualebene	45-46
28. Vergleiche mit externen Einflußgrößen, Veranstaltungsebene	47-48
29. Beziehungsmuster auf Veranstaltungsebene	49
30. Fragebogen zur Beurteilung von Fortbildungs- veranstaltungen (Schlußfassung)	50-53
31. Auswertungsschema	54-56





Fragebogen zur Beurteilung von Fortbildungsveranstaltungen

Wir möchten Sie mit der Vorlage dieses Fragebogens um Ihre Mitarbeit zur Unterstützung eines Forschungsvorhabens am Institut für Erziehungswissenschaften der Universität Münster bitten. Es geht dabei um die Entwicklung eines Fragebogens zur Bewertung von Fortbildungsveranstaltungen, deren Verlauf und Zweckmäßigkeit Sie als Teilnehmer am besten einschätzen können. Die gesammelten Informationen sollen dazu benutzt werden, Fortbildungsveranstaltungen zu verbessern. Indem Sie diesen Fragebogen ausfüllen, tragen Sie dazu bei, daß Ihnen verbesserte Fortbildungsveranstaltungen angeboten werden können.

Ein Fragebogen, der diese Aufgaben erfüllen soll, bedarf einer gründlichen Erprobung. Aus diesem Grunde ist am Ende des Fragebogens auch Raum für Ihre Anmerkungen gegeben, die bei der endgültigen Fassung des Fragebogens berücksichtigt werden sollen. Gehen Sie bitte die einzelnen Fragen zügig durch und beantworten Sie bitte alle. Arbeiten Sie unabhängig von Ihren Nachbarn. Die völlige Anonymität der Auswertung wird Ihnen garantiert.

Zu Beginn einige allgemeine Fragen:

1. Alter..... ☐ ☐
  2. Geschlecht..... männlich ☐  
weiblich ☐
  3. Welchen Ausbildungsabschluß (z.B. Fachschulabschluß) haben Sie erreicht?..... ☐ ☐
  4. Welche Berufstätigkeit (bitte genaue Berufsbezeichnung) üben Sie zur Zeit aus?..... ☐ ☐
  5. Wieviel Jahre sind Sie in diesem Beruf tätig?..... ☐ ☐
  6. An wieviel Fortbildungsveranstaltungen (sowohl mehr- tägige Veranstaltungen während einer Woche oder über mehrere Wochen verteilt als auch einzelne Tagesveranstaltungen, aber ohne selbstorganisierte Arbeitskreise) haben Sie ungefähr bereits teilgenommen?..... ☐ ☐
  7. Kannten Sie die Inhalte der gerade besuchten Fort- bildungsveranstaltung bereits vorher (z.B. durch die Ausbildung, Fachzeitschriften oder andere Fort- bildungsveranstaltungen)?..... nein ☐  
ja, ein wenig ☐  
ja, gut ☐
  8. Gab es für Sie Schwierigkeiten, diese Fortbildungsver- anstaltung beruhigt zu besuchen (z.B. weil keine be- friedigende Vertretung da war; weil dringende Pro- bleme in der Praxis anstanden; wegen Krankheit...)?... nein ☐  
ja ☐
- Wenn ja, welche Schwierigkeiten waren dies?..... ☐ ☐

Im Folgenden finden Sie einige Aussagen, die auf die gerade besuchte Fortbildungsveranstaltung zutreffen können. Dabei wird Ihnen jeweils ein Satzanfang mit zwei möglichen, gegensätzlichen Ergänzungen vorgelegt. Kreuzen Sie bitte jeweils zwischen 1 bis 6 an, wo Ihrem persönlichen Eindruck nach die Veranstaltung einzuordnen ist.

b.w.

## Ein Beispiel:

0. In der Fortbildungsveranstaltung konnte ich...  
 ...aktiv mitarbeiten 1 2 3 4 5 6 ...nur passiv aufnehmen.

Wenn Sie die 1 ankreuzen, so sind Sie der Meinung, daß die linke Satzergänzung völlig zutrifft; kreuzen Sie die 6 an, so sind Sie der Meinung, daß die rechte Satzergänzung völlig zutrifft. Wenn Sie nur zu einem Teil einer der beiden Ergänzungen zustimmen, so kreuzen Sie bitte eine Zahl zwischen 1 und 6 an.

Es bedeuten somit im Beispiel:

- 1 : Sie konnten aktiv mitarbeiten.  
 2 : Sie konnten ziemlich aktiv mitarbeiten.  
 3 : Sie konnten eher aktiv mitarbeiten als nur passiv aufnehmen.  
 4 : Sie konnten eher nur passiv aufnehmen als aktiv mitarbeiten.  
 5 : Sie konnten nur ziemlich passiv aufnehmen.  
 6 : Sie konnten nur passiv aufnehmen.

Schauen Sie sich also jeweils zuerst die rechte und die linke Satzergänzung an und entscheiden Sie sich dann nach Ihrem persönlichen Eindruck für eine der 6 Möglichkeiten.

Bitte versuchen Sie alle Aussagen zu bearbeiten. Für den Fall, daß dies Ihnen bei einer Aussage nicht möglich erscheint, versuchen Sie bitte trotzdem, noch so gut wie möglich zu antworten. Kreuzen Sie diese Aussage aber bitte an und beschreiben Sie kurz am Rand, warum Ihnen eine Bearbeitung nicht möglich erscheint.

1. Das in der Veranstaltung Erarbeitete werde ich wahrscheinlich...  
 ...sehr gut in meiner praktischen Arbeit umsetzen können. 1 2 3 4 5 6 ...weniger gut in meiner praktischen Arbeit umsetzen können.
2. Schwierige und komplizierte Sachverhalte, Fremd- oder Fachwörter wurden...  
 ...immer verständlich erklärt. 1 2 3 4 5 6 ...manchmal nicht oder unverständlich erklärt.
3. Die Auswahl des Themas der Veranstaltung...  
 ...ging an den wichtigen Problemen der Praxis vorbei. 1 2 3 4 5 6 ...traf sehr genau die wichtigen Probleme der Praxis.
4. Der Aufbau der Fortbildungsveranstaltung war...  
 ...klar und übersichtlich gegliedert. 1 2 3 4 5 6 ...oftmals unklar und unübersichtlich gegliedert.
5. Ideen, Wünsche und Vorschläge der Teilnehmer bezüglich der Veranstaltungsdurchführung wurden...  
 ...kaum berücksichtigt. 1 2 3 4 5 6 ...intensiv berücksichtigt.
6. Die zur Verfügung stehende Zeit wurde auf die einzelnen Veranstaltungsabschnitte...  
 ...unangemessen verteilt (eine andere Zeitaufteilung wäre besser gewesen). 1 2 3 4 5 6 ...optimal verteilt (man hätte es nicht besser machen können).
7. Die Fortbildungsveranstaltung wurde...  
 ...abwechslungsreich und fesselnd aufgebaut. 1 2 3 4 5 6 ...eintönig und langweilig aufgebaut.

8. Während der Fortbildungsveranstaltung gab es...  
 ...kaum befriedigende Möglichkeiten zu geselligen Kontakten. 1 2 3 4 5 6 ...befriedigende Möglichkeiten zu geselligen Kontakten.
- 
9. Die Anforderungen an die Teilnehmer in dieser Veranstaltung waren mir...  
 ...oft unangemessen (d.h. zu hoch oder zu niedrig). 1 2 3 4 5 6 ...angemessen (d.h. weder zu hoch noch zu niedrig).
- 
10. In dieser Veranstaltung konnte ich größtenteils...  
 ...mir selbstständig Informationen erarbeiten. 1 2 3 4 5 6 ...nur passiv Informationen aufnehmen.
- 
11. Die Beziehung eines Veranstaltungsabschnittes / Programmpunktes zur Gesamthematik war...  
 ...immer klar zu erkennen. 1 2 3 4 5 6 ...oftmals nicht deutlich zu erkennen.
- 
12. Die Fortbildungsveranstaltung wurde durch äußere Bedingungen (wie z.B. unbequeme Sitze, schlechte Raum- bzw. Sichtverhältnisse, Lärm von draußen, akustische Verständlichkeit im Raum)...  
 ...oftmals gestört. 1 2 3 4 5 6 ...nicht gestört.
- 
13. In dieser Veranstaltung wurde die konkrete Praxis der Teilnehmer (z.B. durch Rollenspiele, Tonbandaufnahmen, Fallberichte, Beispiele usw.)...  
 ...unzureichend einbezogen. 1 2 3 4 5 6 ...sehr gut einbezogen.
- 
14. Die Arbeit in Kleingruppen war in dieser Veranstaltung...  
 ...sehr effektiv. 1 2 3 4 5 6 ...oftmals nicht effektiv.
- 
15. Der Einsatz von Medien (z.B. Arbeitspapiere, -unterlagen, Filme, Dias, Tafelbilder usw.)...  
 ...behinderte oftmals die Arbeit am Thema. 1 2 3 4 5 6 ...unterstützte immer die Arbeit am Thema.
- 
16. Für eine selbstständige Weiterbeschäftigung mit dem Thema der Veranstaltung erhielt ich...  
 ...sehr viele Anregungen. 1 2 3 4 5 6 ...wenige Anregungen.
- 
17. Die Fortbildungsveranstaltung wurde...  
 ...öfters mit etwas Humor aufgelockert. 1 2 3 4 5 6 ...selten mit etwas Humor aufgelockert.
- 
18. Der Besuch dieser Fortbildungsveranstaltung war...  
 ...teilweise eine Zeitvergeudung. 1 2 3 4 5 6 ...eine sehr produktiv verbrachte Zeit.
- 
19. Über Fortschritte hinsichtlich der Erarbeitung des Themas erhielt ich während der Veranstaltung...  
 ...ausreichende Informationen. 1 2 3 4 5 6 ...unzureichende Informationen.
- 
20. Die Thematik dieser Fortbildungsveranstaltung entsprach meinen Bedürfnissen...  
 ...nicht völlig. 1 2 3 4 5 6 ...völlig.

21. Die Teilnehmer waren an der Strukturierung und Planung des Ablaufes der Fortbildungsveranstaltung...  
 ...unzureichend be-teiligt. 1 2 3 4 5 6 ...ausreichend be-teiligt.
- 
22. Die Teilnehmerzahl war für diese Veranstaltung...  
 ...passend. 1 2 3 4 5 6 ...unpassend.
- 
23. Unterschiedliche Voraussetzungen der Teilnehmer an Wissen, Können oder Erfahrungen...  
 ...störten den Ablauf der Fortbildungsveranstaltung. 1 2 3 4 5 6 ...wurden in der Fort-bildungsveranstaltung sinnvoll berücksichtigt.
- 
24. Bei Problemen und Verständigungsschwierigkeiten konnten die Teilnehmer...  
 ...jederzeit unter-brechen und nachfragen. 1 2 3 4 5 6 ...nicht immer unter-brechen und nachfragen.
- 
25. Die Erarbeitung des Themas war...  
 ...immer auf die kon-kreten Praxiserfahrungen der Teilnehmer bezogen. 1 2 3 4 5 6 ...oftmals von den kon-kreten Praxiserfahrungen der Teilnehmer abgehoben.
- 
26. Die Möglichkeiten zu einem informellen Erfahrungsaustausch über Praxisprobleme waren in dieser Fortbildungsveranstaltung...  
 ...unzureichend. 1 2 3 4 5 6 ...sehr gut.
- 
27. Für meine praktische Arbeit hat die Fortbildungsveranstaltung...  
 ...kaum neue Blickpunkte und Ideen gebracht. 1 2 3 4 5 6 ...viele neue Blickpunkte und Ideen gebracht.
- 
28. Die Bedeutung des Themas der Fortbildungsveranstaltung für die Praxis der Teilnehmer ...  
 ...ging oftmals ver-loren. 1 2 3 4 5 6 ...war immer deutlich zu erkennen.
- 
29. Der methodische Aufbau und der Arbeitsstil der Veranstaltung...  
 ...entsprach meinen Vor-stellungen von sinn-vollem Lernen. 1 2 3 4 5 6 ...traf meine Vorstel-lungen von sinnvolle-m Lernen nicht.
- 
30. Abweichende Meinungen konnten...  
 ...frei geäußert werden. 1 2 3 4 5 6 ...manchmal nicht frei ge-äußert werden.
- 
31. Der äußere Rahmen der Fortbildungsveranstaltung (wie z.B. Anfahrt, Dauer, evt. Unterkunft) gefiel...  
 ...mir sehr gut. 1 2 3 4 5 6 ...mir teilweise nicht gut.
- 
32. Durch den Besuch dieser Fortbildungsveranstaltung habe ich...  
 ...fachlich wenig ge-lernt. 1 2 3 4 5 6 ...fachlich viel ge-lernt.
- 
33. Bei der Durchführung dieser Veranstaltung...  
 ...waren Teilnehmer und Referenten/Leiter gleich-berechtigte Partner. 1 2 3 4 5 6 ...bestimmte meistens der Referent/Leiter allein, was gemacht wurde.
- 
34. Durch die Gestaltung der Fortbildungsveranstaltung wurde ich...  
 ...zur aktiven Mit-arbeit angeregt. 1 2 3 4 5 6 ...kaum zur aktiven Mit-arbeit angeregt.

35. In der Veranstaltung habe ich mich den anderen Teilnehmern und den Referenten/Leitern gegenüber...

...oftmals gehemmt gefühlt. 1 2 3 4 5 6 ...immer ungehemmt gefühlt.

36. In dieser Fortbildungsveranstaltung ...

...verwirrte die Behandlung von vielen Einzelheiten. 1 2 3 4 5 6 ...wurden gezielt die wichtigsten Aspekte herausgearbeitet.

37. Wenn man die gesamte Fortbildungsveranstaltung betrachtet (auch im Vergleich zu anderen Veranstaltungen), so war...

...sie sehr gut und ich bin sehr zufrieden mit ihr. 1 2 3 4 5 6 ...nicht so gut und ich bin eher unzufrieden mit ihr.

Betrachten und beurteilen Sie nun bitte jeden Referenten bzw. Leiter der Fortbildungsveranstaltung einzeln. Jeder Referent bzw. Leiter hat dafür einen Buchstaben von A bis D zugeteilt bekommen, der Ihnen vor dem Ausfüllen des Fragebogens bekannt gegeben wurde. Schätzen Sie jeweils nach jedem Aussagenpaar die Referenten bzw. Leiter ein.

Wieviel Veranstaltungen haben Sie bei dem jeweiligen Referenten/Leiter bisher (außer dieser) schon mitgemacht?.....

Frau/Herr A  
Frau/Herr B  
Frau/Herr C  
Frau/Herr D


- verhinderte die Dominanz einzelner Teilnehmer.

Frau/Herr A: 1 2 3 4 5 6  
Frau/Herr B: 1 2 3 4 5 6  
Frau/Herr C: 1 2 3 4 5 6  
Frau/Herr D: 1 2 3 4 5 6

verhinderte die Dominanz einzelner Teilnehmer nicht.

- lenkte den Verlauf der Veranstaltung unangemessen (zu viel oder zu wenig).

Frau/Herr A: 1 2 3 4 5 6  
Frau/Herr B: 1 2 3 4 5 6  
Frau/Herr C: 1 2 3 4 5 6  
Frau/Herr D: 1 2 3 4 5 6

lenkte den Verlauf der Veranstaltung angemessen (weder zu viel noch zu wenig).

- war oftmals unzugänglich für Kritik am Ablauf der Veranstaltung.

Frau/Herr A: 1 2 3 4 5 6  
Frau/Herr B: 1 2 3 4 5 6  
Frau/Herr C: 1 2 3 4 5 6  
Frau/Herr D: 1 2 3 4 5 6

war zugänglich für Kritik am Ablauf der Veranstaltung.

- griff ständig in Diskussionen ein.

Frau/Herr A: 1 2 3 4 5 6  
Frau/Herr B: 1 2 3 4 5 6  
Frau/Herr C: 1 2 3 4 5 6  
Frau/Herr D: 1 2 3 4 5 6

hielt sich in Diskussionen angemessen zurück.

- ging auf Teilnehmerbeiträge angemessen ein.

Frau/Herr A: 1 2 3 4 5 6  
 Frau/Herr B: 1 2 3 4 5 6  
 Frau/Herr C: 1 2 3 4 5 6  
 Frau/Herr D: 1 2 3 4 5 6

Übergang oftmals Teil-  
nehmerbeiträge.

- war oftmals den Teilnehmern gegenüber verschlossen.

Frau/Herr A: 1 2 3 4 5 6  
 Frau/Herr B: 1 2 3 4 5 6  
 Frau/Herr C: 1 2 3 4 5 6  
 Frau/Herr D: 1 2 3 4 5 6

versuchte eine persönliche Beziehung zu den Teilnehmern aufzubauen.

- war stark am Gegenstand der Veranstaltung interessiert.

Frau/Herr A:	1	2	3	4	5	6
Frau/Herr B:	1	2	3	4	5	6
Frau/Herr C:	1	2	3	4	5	6
Frau/Herr D:	1	2	3	4	5	6

schien oftmals wenig Interesse an Gegenstand der Veranstaltung zu haben.

Und zum Abschluß eine letzte Bitte:

Es ist möglich, daß einige für Sie wichtige Aspekte der gerade besuchten Fortbildungsveranstaltung durch den Fragebogen nicht ganz erfaßt wurden oder nicht deutlich herauskamen. Wir möchten Sie deshalb bitten, solche Punkte im Folgenden zu formulieren:

- An dieser Fortbildungsveranstaltung fiel mir auf:.....

Wir danken Ihnen für Ihre Mitarbeit.

Zum Einsatz des Fragebogens

Um die Informationen zwischen verschiedenen Fortbildungsveranstaltungen vergleichbar zu machen, soll die Durchführung der Fragebogenaktion in folgenden Rahmen stattfinden:

1. Damit die Teilnehmer sich auf den Fragebogen einstellen können, bzw. damit sie am Ende der Veranstaltung nicht überfahren werden, soll schon zu Beginn der Fortbildungsveranstaltung (1. Sitzung, z.B. bei der Programmvorstellung) darauf hingewiesen werden, daß zum Abschluß der Veranstaltung ein Fragebogen eingesetzt werden soll.
2. Vor bzw. während des Austeilens soll kurz mündlich der Sinn und Zweck der Befragung erläutert werden. Dabei sollen in lockerer Form folgende Aspekte angesprochen werden:
  - Die Befragung steht in Zusammenhang mit einem Forschungsvorhaben am Institut für Erziehungswissenschaften der Universität Münster. Bei verschiedenen Fortbildungsträgern in der BRD wird dabei diese Befragung durchgeführt.
  - Der Fragebogen fragt nach der Bewertung und Beurteilung der gerade besuchten Fortbildungsveranstaltung aus der Sicht der Teilnehmer heraus. Damit sollen Informationen über die Veranstaltung gesammelt werden, die Hinweise auf Stärken oder Schwächen geben können. Für weitere Veranstaltungen liegen dann Informationen vor, um evtl. die Fortbildungsarbeit zu verbessern. Indirekt profitieren die Teilnehmer also auch von dem Ausfüllen des Fragebogens, indem evtl. zukünftig verbesserte Veranstaltungen angeboten werden können.
  - Die Veranstaltungsteilnehmer sind dabei in ihrer Rolle als Fachleute angesprochen, denn sie haben die Fortbildungsveranstaltung ja hauteng erlebt (die Teilnehmer sind so nicht Objekte eines fremden Forschungszweckes, sondern eher Partner bei einer gemeinsamen Arbeit, nämlich der Verbesserung der Fortbildungsarbeit). Die Teilnehmer sollen nach ihrem Eindruck über die Veranstaltung befragt werden. Dabei kann durch den Fragebogen nicht die Vielfalt aller Eindrücke völlig erfaßt werden und es gehen vielleicht Aspekte verloren. Aber durch die Ausfüllung des Fragebogens können die Meinungen der verschiedenen Teilnehmer schnell erhoben und verglichen werden. Falls ganz wichtige Aspekte nicht oder nur unzureichend erfaßt werden, so steht dafür am Ende des Fragebogens eine offene Frage bereit.
  - Der Fragebogen ist nicht endgültig fertig, er steht vielmehr in einem Erprobungsstadium. So ist der vorliegende Bogen etwas länger, als er in der endgültigen Fassung sein soll; es sollen zu diesem Zeitpunkt Erfahrungen gesammelt werden darüber, welches die "besten" Fragen sind und welche Fragen nicht so wichtig sind; mit a.W.: es sollen verschiedene Fragen erprobt werden. Die Teilnehmer werden gebeten, alle Fragen zu beantworten. Sollte ihnen dies einmal nicht möglich erscheinen, so sollen sie dies am Rande anmerken und kurz den Grund beschreiben. Nur so können dann solche Fragen verbessert werden.
3. Im Fragebogen werden mit 7 Fragen auch die Referenten/Leiter (die mehr oder weniger ständig anwesend waren) eingeschätzt.

Damit dies für mich anonym bleibt, sollen die Referenten/Leiter mit den Buchstaben A bis D bezeichnet werden (evt. an einer Tafel).

Ich möchte Sie noch bitten -um die Beurteilungen der Veranstaltungsteilnehmer besser einschätzen zu können-, mir einige allgemeine Fragen zu der Fortbildungsveranstaltung zu beantworten:

1. Wieviel Teilnehmer waren in der Veranstaltung?
2. Wie kann man die Veranstaltung charakterisieren; war es mehr nur die Vermittlung von spezifischen Kenntnissen, eher ein Verhaltenstrainingskurs,....? In welcher Form fand die Veranstaltung statt?
3. Benutzen Sie bitte nun die Kennzeichnungen A bis D der Referenten/Leiter:
  - a. Geschlecht des jeweiligen Referenten/Leiters
  - b. Anzahl der Fortbildungsveranstaltungen, die der Referent/Leiter gehalten hat.
  - c. Hauptberufliche Tätigkeit des Referenten



# Projektbeschreibung "Instrumente zur Evaluation von Fortbildungsveranstaltungen für Fachkräfte im Vorschulbereich"

Fortbildungsveranstaltungen für Fachkräfte im Vorschulbereich sind wichtige Aufgaben zur Verbesserung des Kindergartenalltags. Was aber ein "Erfolg" solcher Veranstaltungen bedeutet und wie er festgestellt werden kann, dies sind meist offene Fragen. In dem Projekt "Instrumente zur Evaluation von Fortbildungsveranstaltungen für Fachkräfte im Vorschulbereich" soll diesen Fragen nachgegangen werden.

Evaluation bedeutet hier das Bereitstellen von Informationen, mit denen man eine Fortbildungsveranstaltung bewerten und -falls erforderlich- verbessern kann. Solche Informationen sollen Aufschluß über folgende Aspekte der Fortbildungsarbeit geben:

- wie schätzen die Teilnehmer (insb. Erzieher) die Veranstaltung in ihrer Didaktik und Methodik ein?
- haben die Teilnehmer etwas gelernt, was sie in ihrer praktischen Arbeit anwenden können?
- sind die Ziele und Inhalte der Veranstaltung für die Bedürfnisse der Praxis relevant?

Sowohl für die Referenten als auch für die Träger der Fortbildungsveranstaltungen können solche Informationen dokumentieren, was in den einzelnen Veranstaltungen geschehen ist, ob sie erfolgreich oder weniger erfolgreich war. Gerade letztere Informationen bieten dann die Möglichkeit, die tägliche Fortbildungsarbeit pragmatisch zu verbessern.

Diese Informationen sollen durch einen Fragebogen gewonnen werden, den jeder Teilnehmer der Fortbildung am Ende der Veranstaltung ausfüllen soll (Zeitaufwand ca. 30 Minuten). Das Ziel unseres Projekts ist die Entwicklung eines solchen Fragebogens, der nach Abschluß des Projekts den Trägern und Referenten der Fortbildungsveranstaltungen zur Verfügung gestellt werden soll. Mit Hilfe des Bogens können dann eine Vielzahl von Fortbildungsveranstaltungen unterschiedlicher Inhalte bewertet, bzw. Hinweise auf Verbesserungsmöglichkeiten gefunden, sowie die Ergebnisse bei der Planung des folgenden Jahresprogramms einbezogen werden.

Zur Entwicklung dieses Fragebogens sind folgende Schritte nötig:

- eine Befragung der beteiligten Referenten und Träger um festzustellen, welche Fragen über Fortbildungsveranstaltungen sie beantwortet haben wollen.
- eine Fragebogenaktion in mehreren Fortbildungsveranstaltungen, speziell in solchen für Fachkräfte aus dem Vorschulbereich mit pädagogischen Inhalten.
- eine Befragung der beteiligten Referenten und Träger um festzustellen, wie die Informationen des Projekts gebraucht werden konnten, ob sie für nützlich gehalten werden oder nicht.

Es ist übrigens selbstverständlich, daß den beteiligten Referenten und Trägern Anonymität zugesichert wird. Informationen über den Erfolg der eigenen Veranstaltung werden nur den Referenten bzw. Trägern selbst zugestellt. Im Projektbericht sollen die Daten nur so aufbereitet werden, daß eine Identifikation von Daten und einzelnen Kursen nicht möglich ist.

Transformationsgewichte Gesamtgruppe

In der folgenden Tabelle sind jeweils die neuen Werte von VAR und VADI nach den Transformationen angegeben. Bei VAR wurden fehlende Werte mit -9, bei VADI mit -1 kodiert.

Item-Nr.	VAR						VADI					
	1	2	3	4	5	6	1	2	3	4	5	6
1	1.0	0.0	-0.7	-1.5	-2.0	-2.5	1	0	0	0	0	0
2	0.5	-0.6	-1.2	-1.6	-2.1	-2.5	1	0	0	0	0	0
3	-2.3	-1.6	-1.1	-0.8	-0.3	0.7	0	0	0	0	0	1
4	0.9	-0.1	-0.9	-1.4	-1.8	-2.2	1	0	0	0	0	0
5	-2.0	-1.4	-1.1	-0.9	-0.3	0.7	0	0	0	0	0	1
6	-2.4	-1.7	-1.1	-0.6	0.1	1.1	0	0	0	0	1	1
7	0.8	-0.2	-1.0	-1.6	-2.0	-2.7	1	0	0	0	0	0
8	-2.2	-1.7	-1.3	-0.9	-0.4	0.7	0	0	0	0	0	1
9	-2.3	-1.8	-1.4	-1.1	-0.6	0.5	0	0	0	0	0	1
10	1.1	0.2	-0.5	-0.9	-1.3	-1.9	1	1	0	0	0	0
11	0.7	-0.3	-1.0	-1.4	-1.9	-2.4	1	0	0	0	0	0
12	-2.1	-1.4	-0.9	-0.6	-0.3	0.6	0	0	0	0	0	1
13	-2.3	-1.7	-1.3	-0.9	-0.4	0.7	0	0	0	0	0	1
14	0.6	-0.4	-1.0	-1.4	-1.7	-2.2	1	0	0	0	0	0
15	-2.5	-2.1	-1.8	-1.3	-0.6	0.6	0	0	0	0	0	1
16	0.6	-0.4	-1.1	-1.6	-1.3	-2.4	1	0	0	0	0	0
17	0.7	-0.4	-1.0	-1.4	-1.7	-2.1	1	0	0	0	0	0
18	-2.3	-1.9	-1.6	-1.2	-0.6	0.5	0	0	0	0	0	1
19	0.9	-0.1	-0.9	-1.6	-2.0	-2.6	1	0	0	0	0	0
20	-2.1	-1.5	-1.1	-0.8	-0.2	0.8	0	0	0	0	0	1
21	-2.1	-1.6	-1.1	-0.8	-0.2	0.8	0	0	0	0	0	1
22	0.6	-0.4	-0.8	-1.0	-1.3	-1.8	1	0	0	0	0	0
23	-2.5	-1.9	-1.4	-1.0	-0.4	0.6	0	0	0	0	0	1
24	0.3	-1.1	-1.7	-1.9	-2.1	-2.6	1	0	0	0	0	0
25	0.8	-0.3	-0.9	-1.4	-1.9	-2.6	1	0	0	0	0	0
26	-2.3	-1.7	-1.3	-0.9	-0.3	0.8	0	0	0	0	0	1
27	-2.2	-1.8	-1.5	-1.1	-0.4	0.6	0	0	0	0	0	1
28	-2.8	-2.1	-1.6	-1.2	-0.5	0.6	0	0	0	0	0	1
29	0.8	-0.2	-0.9	-1.3	-1.7	-2.1	1	0	0	0	0	0
30	0.3	-0.9	-1.4	-1.7	-2.0	-2.5	1	0	0	0	0	0
31	0.8	-0.1	-0.5	-1.1	-1.4	-2.0	1	0	0	0	0	0
32	-2.2	-1.7	-1.4	-1.1	-0.3	0.7	0	0	0	0	0	1
33	0.6	-0.5	-0.9	-1.2	-1.5	-2.1	1	0	0	0	0	0
34	0.7	-0.4	-0.9	-1.2	-1.6	-2.1	1	0	0	0	0	0
35	-2.1	-1.3	-0.8	-0.4	0.2	1.1	0	0	0	0	1	1
36	-2.3	-1.8	-1.4	-0.9	-0.2	0.8	0	0	0	0	0	1
37	0.6	-0.4	-1.1	-1.6	-1.9	-2.3	1	0	0	0	0	0
38	1.6	1.0	0.6	0.3	0.1	-0.7	1	1	1	1	1	0
39	-2.2	-1.7	-1.5	-1.1	-0.7	0.4	0	0	0	0	0	1
40	-2.7	-2.2	-1.8	-1.4	-0.9	0.3	0	0	0	0	0	1
41	-2.5	-2.1	-1.8	-1.4	-0.8	0.4	0	0	0	0	0	1
42	0.5	-0.7	-1.1	-1.4	-1.6	-2.0	1	0	0	0	0	0
43	-2.4	-1.9	-1.5	-1.1	-0.5	0.6	0	0	0	0	0	1
44	0.3	-0.9	-1.3	-1.5	-1.6	-1.9	1	0	0	0	0	0

Transformationsgewichte "mehrtägige Veranstaltungen"

In der folgenden Tabelle sind jeweils die neuen Werte von VAR und VADI nach den Transformationen angegeben. Bei VAR wurden fehlende Werte mit -9, bei VADI mit -1 kodiert.

Item-Nr.	VAR						VADI					
	1	2	3	4	5	6	1	2	3	4	5	6
1	0.9	-0.1	-0.9	-1.6	-2.0	-2.5	1	0	0	0	0	0
2	0.5	-0.7	-1.4	-1.9	-2.3	-2.6	1	0	0	0	0	0
3	-2.2	-1.5	-1.1	-0.8	-0.3	0.7	0	0	0	0	0	1
4	0.9	-0.1	-0.9	-1.4	-1.8	-2.3	1	0	0	0	0	0
5	-1.9	-1.4	-1.1	-0.9	-0.4	0.7	1	0	0	0	0	0
6	-2.7	-1.8	-1.1	-0.6	0.1	1.1	0	0	0	0	1	1
7	0.8	-0.3	-1.2	-1.7	-2.0	-2.6	1	0	0	0	0	0
8	-2.4	-1.9	-1.5	-1.1	-0.5	0.6	0	0	0	0	0	1
9	-2.3	-1.8	-1.5	-1.2	-0.6	0.5	0	0	0	0	0	0
10	1.0	0.1	-0.5	-1.0	-1.3	-1.9	1	1	0	0	0	0
11	0.7	-0.3	-1.0	-1.4	-1.8	-2.4	1	0	0	0	0	0
12	-2.1	-1.4	-0.9	-0.6	-0.2	0.7	0	0	0	0	0	1
13	-2.4	-1.7	-1.3	-1.0	-0.5	0.6	0	0	0	0	0	1
14	0.5	-0.5	-1.2	-1.6	-1.9	-2.3	1	0	0	0	0	0
15	-2.5	-2.1	-1.8	-1.3	-0.6	0.5	0	0	0	0	0	1
16	0.6	-0.6	-1.3	-1.8	-2.0	-2.4	1	0	0	0	0	0
17	0.6	-0.4	-1.1	-1.7	-1.8	-2.3	1	0	0	0	0	0
18	-2.3	-1.9	-1.7	-1.3	-0.7	0.5	0	0	0	0	0	1
19	0.8	-0.2	-1.0	-1.7	-2.2	-3.0	1	0	0	0	0	0
20	-2.1	-1.5	-1.1	-0.9	-0.3	0.8	0	0	0	0	0	1
21	-2.2	-1.7	-1.1	-0.9	-0.3	0.7	0	0	0	0	0	1
22	0.6	-0.3	-0.7	-0.9	-1.2	-1.8	1	0	0	0	0	0
23	-2.4	-1.9	-1.4	-1.1	-0.5	0.6	0	0	0	0	0	1
24	0.3	-1.1	-1.8	-1.9	-2.1	-2.6	1	0	0	0	0	0
25	0.8	-0.3	-1.0	-1.4	-1.9	-2.7	1	0	0	0	0	0
26	-2.5	-1.9	-1.4	-1.0	-0.3	0.7	0	0	0	0	0	1
27	-2.2	-1.8	-1.6	-1.2	-0.6	0.5	0	0	0	0	0	1
28	-3.0	-2.2	-1.7	-1.3	-0.6	0.5	0	0	0	0	0	1
29	0.8	-0.3	-1.0	-1.4	-1.8	-2.2	1	0	0	0	0	0
30	0.3	-0.9	-1.4	-1.7	-2.0	-2.5	1	0	0	0	0	0
31	0.8	-0.2	-0.7	-1.1	-1.5	-2.0	1	0	0	0	0	0
32	-2.2	-1.7	-1.5	-1.1	-0.3	0.6	0	0	0	0	0	1
33	0.6	-0.5	-1.0	-1.2	-1.5	-2.1	1	0	0	0	0	0
34	0.6	-0.4	-0.9	-1.3	-1.7	-2.2	1	0	0	0	0	0
35	-2.1	-1.3	-0.8	-0.4	0.2	1.1	0	0	0	0	0	1
36	-2.3	-1.8	-1.4	-1.0	-0.2	0.8	0	0	0	0	0	1
37	0.6	-0.5	-1.2	-1.7	-2.1	-2.5	1	0	0	0	0	0
38	1.6	1.0	0.6	0.3	0.1	-0.7	1	1	1	1	1	0
39	-2.2	-1.8	-1.5	-1.2	-0.7	0.4	0	0	0	0	0	1
40	-2.6	-2.1	-1.8	-1.5	-1.0	0.3	0	0	0	0	0	1
41	-2.5	-2.1	-1.9	-1.5	-0.8	0.4	0	0	0	0	0	1
42	0.4	-0.7	-1.1	-1.4	-1.6	-2.0	1	0	0	0	0	0
43	-2.3	-1.8	-1.5	-1.1	-0.5	0.6	0	0	0	0	0	1
44	0.3	-0.9	-1.3	-1.5	-1.6	-1.9	1	0	0	0	0	0

Faktorenanalyse Untergruppe "mehrtägige Veranstaltungen"  
VAR, 5 Faktoren Kommunalitäten und Ladungen

Nr.	L a d u n g e n					Kommunalität
	Faktor 1	Faktor 2	Faktor 3	Faktor 4	Faktor 5	
1	.43	.11	.36	.27	-.01	.40
3	.44	.11	.23	.20	.21	.34
4	.65	.18	.18	.09	.06	.50
5	.23	.30	.16	.36	.31	.39
6	.44	.17	.15	.18	.35	.39
7	.32	.20	.32	.34	.24	.42
8	.13	.26	.29	.14	.25	.25
9	.35	.32	.27	.09	.15	.33
10	.05	.07	.05	.61	.20	.42
11	.68	.21	.19	.02	.09	.54
12	.07	.17	.06	-.05	.48	.27
13	.32	.07	.25	.32	.46	.49
14	.10	.05	.10	.34	.07	.14
17	.25	.32	.19	.40	.26	.43
18	.40	.28	.59	.23	.12	.66
19	.44	.17	.18	.25	.21	.36
20	.32	.10	.37	.30	.23	.38
21	.12	.22	.27	.36	.39	.41
22	.13	.04	.02	.11	.22	.08
23	.37	.26	.15	.24	.28	.36
25	.26	.24	.23	.30	.29	.35
26	.07	.19	.19	.23	.49	.37
27	.19	.15	.69	.09	.15	.56
28	.40	.28	.45	.11	.26	.52
29	.37	.17	.28	.46	.08	.46
30	.31	.51	.04	.29	.04	.44
31	.04	.05	.04	.10	.38	.16
32	.22	.18	.69	.11	.09	.58
33	-.08	.42	.11	.37	.22	.38
34	.21	.33	.08	.69	.06	.63
35	.16	.17	.02	.14	.02	.08
36	.65	.14	.20	.14	.17	.53
37	.43	.24	.51	.32	.18	.64
39	.14	.48	.30	.09	.20	.38
40	.17	.70	.24	.11	.09	.60
42	.18	.45	.06	.13	.13	.27
43	.15	.46	.21	.19	.19	.35
44	.20	.37	.12	-.02	.25	.25

Regressionsanalyse VAR

(Kriterium besteht aus Item 18 und 37; Prädiktoren sind alle Items ohne 2, 15, 16, 18, 24, 37, 38, 41)

MULTIPLE R 0.69  
R SQUARE 0.48  
ADJUSTED R SQUARE 0.41

ANALYSIS OF VARIANCE	DF	SUM OF SQUARES	MEAN SQUARE	F
REGRESSION	36	756.459	21.013	6.621
RESIDUAL	257	815.581	3.174	

Nr.	(1) BETA	(2) r	(3) r	(4) RFSK <sup>2</sup>	Nr.	BETA	r	r	RFSK <sup>2</sup>
1	-0.01	0.0	0.28		26	-0.12	4.5	0.16	0.06
3	0.00	0.0	0.21		27	0.08	1.5	0.32	
4	0.01	0.0	0.30		28	-0.04	0.3	0.30	
5	-0.01	0.0	0.31		29	0.08	1.8	0.34	
6	0.04	0.5	0.32		30	-0.12	4.2	0.22	0.10
7	-0.10	2.9	0.28		31	-0.07	1.7	0.06	
8	0.08	2.1	0.25		32	0.02	0.1	0.28	
9	-0.07	1.4	0.21		33	0.06	1.1	0.29	
10	-0.08	1.7	0.09		34	0.01	0.0	0.30	
11	0.06	0.9	0.26		35	-0.12	5.8	0.02	0.00
12	0.01	0.0	0.17		36	0.13	4.0	0.29	0.17
13	0.00	0.0	0.26		39	0.04	0.5	0.35	
14	0.02	0.2	0.11		40	0.00	0.0	0.32	
17	0.54	69.2	0.59	0.72	42	0.14	6.9	0.28	0.17
19	-0.10	2.8	0.26		43	0.13	4.4	0.35	0.25
20	0.06	1.2	0.30		44	0.10	3.2	0.20	
21	0.11	2.9	0.32						
22	0.03	0.4	0.14						
23	-0.06	1.2	0.23						
25	-0.08	1.8	0.26						

- (1) BETA-Gewichte  
(2) Der kritische Wert für F beträgt auf dem 5 % Niveau 3.88  
(3) Einfache Korrelation eines Prädiktors mit dem Kriterium  
(4) RFSK = Regressions-Faktor-Struktur-Koeffizient; einfache Korrelation des Items j mit der aus allen Prädiktoren gebildeten Linearkombination; das Quadrat gibt die gemeinsame Varianz an; RFSK wurde hier nur für die Items berechnet, die ein signifikantes F haben.

Regressionsanalyse VADI

( Kriterium besteht aus Item VADI 18 und 37; Prädiktoren sind alle Items ohne 2, 15, 16, 18, 24, 37, 38, 41)

MULTIPLE R 0.78  
R SQUARE 0.60  
ADJUSTED R SQUARE 0.57

ANALYSIS OF VARIANCE	DF	SUM OF SQUARES	MEAN SQUARE	F
REGRESSION	36	125.624	3.490	10.774
RESIDUAL	257	83.244	0.324	

Nr.	(1) BETA	(2) $r_{ij}$	(3) $r_{ij}$	(4) RFSK <sup>2</sup>	Nr.	BETA	$r_{ij}$	$r_{ij}$	RFSK <sup>2</sup>
1	0.12	5.5	0.46	0.36	26	0.03	0.3	0.33	
3	0.06	1.3	0.44		27	0.13	5.9	0.50	0.42
4	-0.01	0.0	0.33		28	0.05	0.9	0.49	
5	-0.04	0.6	0.34		29	0.13	6.0	0.48	0.38
6	0.12	6.1	0.39	0.26	30	0.08	2.0	0.35	
7	0.04	0.7	0.41		31	-0.03	0.4	0.15	
8	0.03	0.5	0.32		32	0.15	7.8	0.52	0.46
9	0.08	2.8	0.37		33	0.00	0.0	0.26	
10	-0.03	0.4	0.19		34	0.06	1.2	0.37	
11	-0.04	0.5	0.29		35	-0.01	0.0	0.18	
12	-0.01	0.0	0.11		36	0.04	0.6	0.41	
13	0.05	0.8	0.42		39	0.12	5.2	0.42	0.30
14	0.03	0.3	0.25		40	0.13	6.5	0.40	0.26
17	-0.04	0.7	0.32		42	-0.07	1.6	0.31	
19	0.05	0.9	0.41		43	-0.08	2.7	0.31	
20	0.11	4.6	0.49	0.40	44	-0.02	0.2	0.28	
21	0.06	1.3	0.38						
22	-0.04	0.6	0.14						
23	-0.02	0.1	0.35						
25	0.02	0.2	0.38						

- (1) BETA-Gewichte
- (2) Der kritische Wert für F beträgt auf dem 5 % Niveau 3.88
- (3) Einfache Korrelation eines Prädiktors mit dem Kriterium
- (4) RFSK = Regressions-Faktor-Struktur-Koeffizient; einfache Korrelation des Items j mit der aus allen Prädiktoren gebildeten Linearkombination; das Quadrat gibt die gemeinsame Varianz an; RFSK wurde hier nur für die Items berechnet, die ein signifikantes F haben.

Fragebogen zur Beurteilung von Fortbildungsveranstaltungen

Wir möchten Sie mit der Vorlage dieses Fragebogens um Ihre Mitarbeit zur Unterstützung eines Forschungsvorhabens am Institut für Erziehungswissenschaften der Universität Münster bitten. Es geht dabei um die Entwicklung eines Fragebogens zur Bewertung von Fortbildungsveranstaltungen, deren Verlauf und Zweckmäßigkeit Sie als Teilnehmer am besten einschätzen können. Die gesammelten Informationen sollen dazu benutzt werden, Fortbildungsveranstaltungen zu verbessern. Indem Sie diesen Fragebogen ausfüllen, tragen Sie dazu bei, daß Ihnen verbesserte Fortbildungsveranstaltungen angeboten werden können.

Ein Fragebogen, der diese Aufgaben erfüllen soll, bedarf einer gründlichen Erprobung. Aus diesem Grunde ist am Ende des Fragebogens auch in einer anderen Form Raum für Ihre Anmerkungen gegeben, die bei der endgültigen Fassung des Fragebogens berücksichtigt werden sollen. Gehen Sie bitte die einzelnen Fragen zügig durch und beantworten Sie bitte alle. Arbeiten Sie unabhängig von Ihren Nachbarn. Eine völlige Anonymität der Auswertung wird Ihnen garantiert.

Zu Beginn einige allgemeinen Fragen:

1. Alter..... ☐
  2. Geschlecht..... männlich ☐  
weiblich ☐
  3. Welchen Ausbildungsabschluß (z.B. Fachschulabschluß) haben Sie erreicht?..... ☐
  4. Welche Berufstätigkeit (bitte genaue Berufsbezeichnung) üben Sie zur Zeit aus?..... ☐
  5. Wieviel Jahre sind Sie in diesem Beruf tätig?..... ☐
  6. An wieviel mehrtätigen Fortbildungsveranstaltungen haben Sie in den letzten zwei Jahren teilgenommen?..... ☐
  7. Kannten Sie die Inhalte der gerade besuchten Fortbildungsveranstaltung bereits vorher (z.B. durch die Ausbildung, Fachzeitschriften oder andere Fortbildungsveranstaltungen)?..... nein ☐  
ja, ein wenig ☐  
ja, gut ☐
  8. Gab es für Sie Schwierigkeiten, diese Fortbildungsveranstaltung beruhigt zu besuchen (z.B. weil keine befriedigende Vertretung da war; weil dringende Probleme in der Praxis anstanden; wegen Krankheit.....)?..... nein ☐  
ja ☐
- Wenn ja, welche Schwierigkeiten waren dies?..... ☐

Im Folgenden finden Sie einige Aussagen, die auf die gerade besuchte Fortbildungsveranstaltung zutreffen können. Wir möchten Sie bitten jeweils nach Ihrem Eindruck zu entscheiden, inwieweit die Aussagen für die gerade besuchte Veranstaltung zutreffend sind. Ein Beispiel finden Sie auf der Rückseite:

b.w.

-2-

Ein Beispiel:

	stimmt nicht	stimmt ein wenig	stimmt ziemlich	stimmt völlig
0. In dieser Fortbildungsveranstaltung konnte ich aktiv mitarbeiten.	(1)	(2)	(3)	(4)

Kreuzen Sie bitte eine der Zahlen von 1 bis 4 an. Dabei bedeuten:

- 1 : Die Aussage stimmt nicht für diese Veranstaltung.  
 2 : Die Aussage stimmt ein wenig für diese Veranstaltung.  
 3 : Die Aussage stimmt ziemlich für diese Veranstaltung.  
 4 : Die Aussage stimmt völlig für diese Veranstaltung.

Bitte bearbeiten Sie alle Aussagen. Wenn Ihrer Meinung nach eine der angegebenen Antwortmöglichkeiten nicht zutrifft, so wählen Sie bitte jene aus, die Ihnen noch am ehesten als zutreffend erscheint. Kreuzen Sie diese Aussage aber bitte an und beschreiben Sie kurz, warum es Ihnen schwerfällt, die Aussage zu bearbeiten.

	stimmt nicht	stimmt ein wenig	stimmt ziemlich	stimmt völlig
1. Die Auswahl des Themas der Veranstaltung traf haargenau die wichtigsten Probleme der Praxis.	(1)	(2)	(3)	(4)
2. In dieser Veranstaltung konnte ich mir überwiegend selbständig Informationen erarbeiten.	(1)	(2)	(3)	(4)
3. Die Teilnehmerzahl war für diese Veranstaltung ein wenig zu groß.	(1)	(2)	(3)	(4)
4. Ich bin mit dem, was ich in dieser Veranstaltung gelernt habe, <del>teil</del> weise unzufrieden.	(1)	(2)	(3)	(4)
5. Auf Verständnisschwierigkeiten von Teilnehmern wurde immer sofort klärend eingegangen.	(1)	(2)	(3)	(4)
6. Während der Veranstaltung habe ich mich in der Gruppe zu jeder Zeit sehr wohl gefühlt.	(1)	(2)	(3)	(4)
7. Es gab einzelne Situationen in dieser Veranstaltung, in denen mir das Arbeitsklima nicht so gut gefiel.	(1)	(2)	(3)	(4)
8. Durch den äußeren Rahmen der Fortbildungsveranstaltung gab es ausgezeichnete Möglichkeiten zu geistlichem Beisammensein.	(1)	(2)	(3)	(4)
9. Für meine praktische Arbeit hat diese Veranstaltung relativ wenig neue Gesichtspunkte und Ideen gebracht.	(1)	(2)	(3)	(4)
10. In dieser Veranstaltung wurden gezielt ausnahmslos alle wichtigsten Aspekte des Themas herausgearbeitet.	(1)	(2)	(3)	(4)



-3-

	stimmt nicht	stimmt ein wenig	stimmt ziemlich	stimmt völlig
11. Manchmal wurde die Veranstaltung ein wenig zu "trocken" durchgeführt.	(1)	(2)	(3)	(4)
12. Der Aufbau und Stil der Veranstaltung entsprach nicht immer meinen Vorstellungen von sinnvollem Lernen.	(1)	(2)	(3)	(4)
13. Es gab Momente, da war der Besuch dieser Veranstaltung weniger lohnend.	(1)	(2)	(3)	(4)
14. Die Dauer dieser Veranstaltung war völlig ausreichend.	(1)	(2)	(3)	(4)
15. Die Beziehung eines Veranstaltungsabschnittes / Programmpunktes zur Gesamthematik (der "rote Faden") war manchmal etwas schwer zu erkennen.	(1)	(2)	(3)	(4)
16. Durch den Besuch dieser Fortbildungsveranstaltung habe ich fachlich außerordentlich viel gelernt.	(1)	(2)	(3)	(4)
17. Durch die Gestaltung der Fortbildungsveranstaltung wurde ich ausnahmslos zur aktiven Mitarbeit angeregt.	(1)	(2)	(3)	(4)
18. Teilweise störten die Raumverhältnisse (z.B. unbequeme Sitze, schlechte Sichtverhältnisse, Lärm von draußen) die effektive Arbeit.	(1)	(2)	(3)	(4)
19. Die Teilnehmer erhielten über ihre Lernfortschritte in der Veranstaltung treffende Informationen. <i>Teilnehmer</i>	(1)	(2)	(3)	(4)
20. Für eine selbständige Weiterbeschäftigung mit dem Thema der Veranstaltung wurde ich nur wenig motiviert.	(1)	(2)	(3)	(4)
21. Ideen, Wünsche und Vorschläge der Teilnehmer bezüglich der Veranstaltungsdurchführung wurden manchmal zu wenig berücksichtigt.	(1)	(2)	(3)	(4)
22. In dieser Veranstaltung hätte es mehr Möglichkeiten zu einem informellen Erfahrungsaustausch über Praxisprobleme der Teilnehmer geben können.	(1)	(2)	(3)	(4)
23. Ich glaube, daß ich nach dem Besuch dieser Veranstaltung viele praktische Probleme entschieden besser angehen kann.	(1)	(2)	(3)	(4)
24. Neben der Veranstaltung gab es hervorragende Möglichkeiten zur Gestaltung der Freizeit.	(1)	(2)	(3)	(4)

b.w

	stimmt nicht	stimmt ein wenig	stimmt ziemlich	stimmt völlig
25. Teilweise waren die Anforderungen in der Veranstaltung speziell für mich unangemessen (d.h. zu hoch oder zu niedrig).	(1)	(2)	(3)	(4)
26. Abweichende Meinungen konnten in der Veranstaltung frei und ohne Angst geäußert werden.	(1)	(2)	(3)	(4)
27. Der Aufbau der Fortbildungsveranstaltung war manchmal etwas unübersichtlich gegliedert.	(1)	(2)	(3)	(4)
28. Bei Entscheidungen über den Ablauf der Veranstaltung wurden die Vorstellungen der Teilnehmer immer voll berücksichtigt.	(1)	(2)	(3)	(4)
29. Die Arbeit in Kleingruppen war in dieser Veranstaltung teilweise nicht sehr effektiv.	(1)	(2)	(3)	(4)
30. Die äußeren Bedingungen der Veranstaltung (z.B. Anfahrtsweg, Unterkunft, Veranstaltungsort) hätten günstiger sein können.	(1)	(2)	(3)	(4)
31. Wenn ich diese Veranstaltung im Ganzen betrachte, so war sie eine der besten Fortbildungsveranstaltungen, die ich je besucht habe.	(1)	(2)	(3)	(4)

Betrachten und beurteilen Sie nun bitte jeden Referenten bzw. Leiter der Fortbildungsveranstaltung einzeln. Jeder Referent bzw. Leiter hat dafür einen Buchstaben von A bis D zugeteilt bekommen, der Ihnen vor dem Ausfüllen des Fragebogens bekannt gegeben wurde. Schätzen Sie jeweils nach jeder Aussage die Referenten bzw. Leiter ein:

32. Der Referent / Leiter lenkte manchmal den Verlauf der Veranstaltung unangemessen stark.				
Frau/Herr A	(1)	(2)	(3)	(4)
Frau/Herr B	(1)	(2)	(3)	(4)
Frau/Herr C	(1)	(2)	(3)	(4)
Frau/Herr D	(1)	(2)	(3)	(4)
33. Der Referent / Leiter schien immer zugänglich für Kritik an der Veranstaltung zu sein.				
Frau/Herr A	(1)	(2)	(3)	(4)
Frau/Herr B	(1)	(2)	(3)	(4)
Frau/Herr C	(1)	(2)	(3)	(4)
Frau/Herr D	(1)	(2)	(3)	(4)

-5-

34. Der Referent / Leiter Übergang  
samt Teilnehmerteilnahme.

	stimmt nicht	stimmt ein wenig	stimmt ziemlich	stimmt völlig
Frau/Herr A	(1)	(2)	(3)	(4)
Frau/Herr B	(1)	(2)	(3)	(4)
Frau/Herr C	(1)	(2)	(3)	(4)
Frau/Herr D	(1)	(2)	(3)	(4)

35. Der Referent / Leiter versuchte sehr  
intensiv, eine persönliche Beziehung  
zu den Teilnehmern aufzubauen.

	(1)	(2)	(3)	(4)
Frau/Herr A	(1)	(2)	(3)	(4)
Frau/Herr B	(1)	(2)	(3)	(4)
Frau/Herr C	(1)	(2)	(3)	(4)
Frau/Herr D	(1)	(2)	(3)	(4)

Wieviel Veranstaltungen haben Sie bei dem jeweiligen  
Referenten / Leiter bisher (außer dieser) schon mit-  
gemacht?.....

Frau/Herr A  
Frau/Herr B  
Frau/Herr C  
Frau/Herr D


Und zum Abschluß noch eine letzte Bitte:

Es ist möglich, daß durch den Fragebogen einige für Sie besonders  
wichtige Aspekte nicht erfaßt wurden oder nicht deutlich herauskamen.  
Wir möchten Sie bitten, solche Aspekte im Folgenden zu formulieren.  
Schreiben Sie bitte stichwortartig die 3 auffallendsten Aspekte auf,  
die Sie während dieser Fortbildungsveranstaltung bemerkt haben. Es  
kann sich dabei um positive und/oder negative Aspekte handeln; die  
Aspekte können sich direkt auf die Veranstaltung beziehen oder auf  
sonst etwas, was für Sie in dieser Zeit besonders auffallend war.

Besonders auffallend war für mich während dieser Fortbildungsver-  
anstaltung:

- .....  
.....  
.....
  - .....  
.....  
.....
  - .....  
.....  
.....
- (ggf. Rückseite benutzen)

Wir danken Ihnen für Ihre Mitarbeit.

Zum Einsatz des revidierten Fragebogens

Um die Informationen zwischen verschiedenen Fortbildungsveranstaltungen vergleichbar zu machen, soll die Durchführung der Fragebogenaktion in folgendem Rahmen stattfinden:

1. Damit die Teilnehmer sich auf den Fragebogen einstellen können, bzw. damit sie am Ende der Veranstaltung nicht überfahren werden, soll schon zu Beginn der Fortbildungsveranstaltung (1. Sitzung, z.B. bei der Vorstellung des Programms) darauf hingewiesen werden, daß zum Abschluß der Veranstaltung ein Fragebogen eingesetzt werden soll.
2. Vor bzw. während des Austeilens soll kurz mündlich der Sinn und Zweck der Befragung erläutert werden. Dabei sollen in lockerer Form folgende Aspekte angesprochen werden:
  - Die Befragung steht in Zusammenhang mit einem Forschungsvorhaben am Institut für Erziehungswissenschaften der Universität Münster. Bei verschiedenen Fortbildungsträgern in der BRD wird dabei diese Befragung durchgeführt.
  - Der Fragebogen fragt nach der Bewertung und Beurteilung der gerade besuchten Fortbildungsveranstaltung aus der Sicht der Teilnehmer heraus. Damit sollen Informationen über die Veranstaltung gesammelt werden, die Hinweise auf Schwächen oder Stärken geben können. Für weitere Veranstaltungen liegen dann Informationen vor, um evt. die Fortbildungsarbeit zu verbessern. Indirekt profitieren die Teilnehmer also auch von dem Ausfüllen des Fragebogens, indem evt. zukünftig verbesserte Veranstaltungen angeboten werden können.
  - Die Veranstaltungsteilnehmer sind dabei in ihrer Rolle als Fachleute angesprochen, denn sie haben die Fortbildungsveranstaltung ja hauteng erlebt (die Teilnehmer sind so nicht Objekte eines fremden Forschungszwecks, sondern eher Partner bei einer gemeinsamen Arbeit, nämlich der Verbesserung der Fortbildungsarbeit). Die Teilnehmer sollen nach ihrem Eindruck über die Veranstaltung befragt werden. Dabei kann durch den Fragebogen nicht die Vielfalt aller Eindrücke völlig erfasst werden, und es gehen vielleicht einige Aspekte verloren. Aber durch die Ausfüllung eines Fragebogens können die Meinungen der verschiedenen Teilnehmer schnell erhoben und verglichen werden. (Auch um evt. Verlorenes wieder aufzufangen, ist am Ende des Fragebogens noch einmal in einer anderen, offeneren Form Platz gegeben für die Meinungsäußerungen der Teilnehmer.)
  - Der Fragebogen ist nicht endgültig fertig, er steht vielmehr in einem Erprobungsstadium. Die Teilnehmer werden gebeten, alle Fragen zu beantworten. Sollte ihnen dies einmal nicht möglich erscheinen, so sollen sie dies kurz am Rand anmerken und den Grund beschreiben. Nur so können solche Fragen dann verbessert werden.
3. Im Fragebogen werden mit 4 Fragen auch die Referenten/Leiter speziell eingeschätzt. Damit dies für mich anonym bleibt, sollen die Referenten/Leiter mit den Buchstaben von A bis D bezeichnet werden (z.B. an der Tafel).

Fragen zur Veranstaltung

Um die Beurteilung der Fortbildungsteilnehmer besser einschätzen zu können und auch um mehr über die Veranstaltung selber zu erfahren, möchte ich Sie bitten, mir einige allgemeine Fragen zu der Veranstaltung zu beantworten:

1. Unter welchem Titel wurde die Veranstaltung angekündigt?.....
2. Charakterisieren Sie bitte kurz den inhaltlichen Schwerpunkt der Veranstaltung:.....
3. Wieviel Teilnehmer besuchten diese Veranstaltung?.....
4. Wieviel Tage dauerte die Veranstaltung?.....
5. In welchem größeren Rahmen wurde diese Veranstaltung durchgeführt?
  - ☐ als einmalige Fortbildungsveranstaltung
  - ☐ als Teil einer Veranstaltungsserie
  - ☐ als Blockseminar mit eingeschobenen Praxisphasen
  - ☐ als einmalige Veranstaltung; es finden Nachbesprechungen oder andere Praxisbegleitungen der Teilnehmer statt
  - ☐ sonstiges, und zwar.....
6. Wenn Sie aus Ihren Erfahrungen die Teilnehmergruppe einschätzen, kamen die Teilnehmer dann
  - ☐ aus einem einheitlichen Berufsfeld (z.B. nur Erzieher oder nur Sozialarbeiter in Heimen....) oder
  - ☐ aus verschiedenartigen Berufsfeldern (z.B. Kinderpfleger aus Kindergärten und Sozialarbeiter in der Altenbetreuung gleichzeitig)?
7. Welche Medien wurden hauptsächlich in der Veranstaltung eingesetzt?.....
8. Wenn sie die verschiedenartigen Sozialformen betrachten, welche nahm dann den größten Zeitraum in der Veranstaltung ein?
  - ☐ Arbeit in der Gesamtgruppe
  - ☐ Arbeit in Kleingruppen
  - ☐ Einzelarbeit
  - ☐ sonstiges, und zwar.....
9. Welche Sozialform hatte in bezug auf die Veranstaltungsziele die größte Bedeutung (dies muß nicht mit Frage 8 identisch sein)?
  - ☐ Arbeit in der Gesamtgruppe
  - ☐ Arbeit in Kleingruppen
  - ☐ Einzelarbeit
  - ☐ sonstiges, und zwar.....
10. Wieviel Fortbildungsveranstaltungen haben die Referenten/Leiter bisher in den beiden letzten Jahren gehalten? (bitte unterstreichen)
 

Frau/Herr A:	keine	1-3	4-6	7-9	10-12	über 12
Frau/Herr B:	keine	1-3	4-6	7-9	10-12	über 12
Frau/Herr C:	keine	1-3	4-6	7-9	10-12	über 12
Frau/Herr D:	keine	1-3	4-6	7-9	10-12	über 12
11. Wurde in der Fortbildungsveranstaltung Geselligkeit/Freizeit vorher geplant?
  - ☐ nein
  - ☐ ja, und zwar.....

12. Wurden in der Fortbildungsveranstaltung Verfahren angewandt, um Lernfortschritte von Teilnehmern zu erfassen?

☐ nein  
☐ ja, und zwar.....

In der Literatur zur Erwachsenenbildung/Fortbildungsarbeit lassen sich verschiedene Schwerpunkte auffinden. Hierzu einige kurze Fragen:

13. In manchen Veranstaltungen überwiegt die reine Vermittlung von Informationen, in anderen ein rein gruppendynamisches Training von Kommunikationsverhalten. Ordnen Sie bitte Ihre Veranstaltung in diese Spannbreite ein:

☐ es war mehr eine reine Informationsveranstaltung  
☐ es überwog der Aspekt der Informationsvermittlung  
☐ ausgeglichen  
☐ es wurden mehr gruppendynamische Übungen durchgeführt  
☐ es war eine reine gruppendynamische Veranstaltung  
☐ sonstiges, und zwar .....

14. In der methodischen Gestaltung einer Veranstaltung wird oft -abstrakt- unterschieden zwischen einer darbietenden (z.B. Referate, schriftliche Materialien) und erarbeitenden (z.B. Diskussionen) Vorgehensweise. Kreuzen Sie bitte an, wo sich Ihre Veranstaltung überwiegend einordnen läßt:

☐ rein darbietende Vorgehensweise  
☐ mehr darbietende als erarbeitende Vorgehensweise  
☐ ausgeglichen  
☐ mehr erarbeitende als darbietende Vorgehensweise  
☐ rein erarbeitende Vorgehensweise  
☐ sonstiges, und zwar .....

15. Fortbildungsveranstaltungen werden in einem unterschiedlichen Ausmaß vorgeplant (z.B. abhängig vom Thema, von der Adressatengruppe..) und unterscheiden sich auch darin, wie stark die Durchführung der Veranstaltung an der Vorausplanung orientiert bleibt. Bitte stufen Sie Ihre Veranstaltung nach diesem Gesichtspunkt ein:

☐ starkes Ausmaß von exakter Vorausplanung; Planung wurde überwiegend in der Veranstaltung eingehalten.  
☐ starkes Ausmaß von exakter Vorausplanung; Planung wurde aber teilweise nach Teilnehmerreaktionen modifiziert.  
☐ mittleres Ausmaß von Vorausplanung; flexible Anpassung auf Teilnehmerreaktionen.  
☐ nur Grobplanung vorher, Feinplanung geschieht mit den Teilnehmern zusammen.  
☐ vorwiegend keine Vorplanung im Einzelnen; Eingehen auf situative Gegebenheiten.  
☐ sonstiges, und zwar .....

16. Beteiligung, Mitwirkung und Mitbestimmung von Teilnehmern an Entscheidungen über Themenauswahl, Lernziele und Veranstaltungsablauf variieren von Veranstaltung zu Veranstaltung (dies ist abhängig z.B. von der Art des Themas, von den äußeren Bedingungen, von der Interessiertheit der Teilnehmer, von der Teilnehmergruppe....). Beurteilen Sie bitte das tatsächliche Ausmaß von Beteiligung und Mitbestimmung der Teilnehmer an Entscheidungen in folgenden Bereichen der Veranstaltung:

	starke Mitbestimmung			keine Mitbestimmung		
Themenauswahl	1	2	3	4	5	
Lernziele	1		3	4	5	
Ablauf	1	2	3	4	5	

Vielen DANK für Ihre Mitarbeit.









-4-

	stimmt nicht		stimmt ein wenig		stimmt ziemlich		stimmt völlig		a
	n	b	n	b	n	b	n	b	
34. Der Referent / Leiter Übergang manchmal Teilnehmerbeiträge. Frau/Herr A Frau/Herr B Frau/Herr C Frau/Herr D									
35. Der Referent / Leiter versuchte sehr intensiv, eine persönliche Beziehung zu den Teilnehmern aufzubauen. Frau/Herr A Frau/Herr B Frau/Herr C Frau/Herr D									

In den folgenden Abbildungen (Profile) werden die durchschnittlichen Antworten der Teilnehmer verglichen mit den Antworten, die der Referent / Leiter A-D auf dem gleichen Fragebogen abgegeben hat.

Dazu wurden einige Fragen "umgedreht", d.h.: In den Fragebogen waren einige Fragen positiv formuliert (z.B. Frage 1 und Frage 2: eine 4 ist für die Veranstaltung positiv), andere waren negativ formuliert (z.B. Frage 7 und Frage 9; eine 1 ist für die Veranstaltung positiv). Für die Abbildungen auf den folgenden Seiten wurden nun alle die Fragen, die negativ formuliert waren, quasi umformuliert, sodaß nun ebenfalls hohe Werte (z.B. eine 4) für die Veranstaltung eine positive Einschätzung bedeuten. Dadurch werden die Abbildungen übersichtlicher.

Ein Beispiel:

Frage 11 lautete: "Manchmal wurde die Veranstaltung ein wenig zu trocken durchgeführt". Ein Teilnehmer, der nicht dieser Meinung ist, kreuzt die 1 (stimmt nicht) an; ein Teilnehmer, der mit dieser Aussage übereinstimmt, kreuzt die 4 (stimmt völlig) an. (4 ist also ein für die Veranstaltung negatives Urteil, 1 ein positives). Nun wird die Frage "umgedreht", d.h. der Teilnehmer, der der Aussage nicht zustimmte, erhält nun eine 4 und derjenige, der dieser Aussage zustimmte, der also die Veranstaltung für "trocken" hielt, erhält nun eine 1.

Durch dieses "Umdrehen" wird erreicht, daß z.B. eine 4 immer ein Wert ist, der eine für die Fortbildungsveranstaltung positive Beurteilung bedeutet. Von diesem "Umdrehen" sind betroffen die Fragen: 3, 4, 7, 9, 11, 12, 13, 15, 18, 20, 21, 22, 25, 27, 29, 30, 32, 34 (In den Abbildungen werden diese Fragen mit einem \* gekennzeichnet).

Bei diesen Vergleichen sind sowohl die Aspekte der Fortbildungsveranstaltung interessant, in denen die Einschätzungen der Teilnehmer mit denen der Referenten / Leiter übereinstimmen, als auch die Aspekte, in denen sie sich unterscheiden (dabei können Unterschiede von weniger als 1 Punkt vernachlässigt werden).

Das Gesamtausmaß der Ähnlichkeit der Einschätzungen läßt sich auch durch eine Zahl kennzeichnen (Q-Maß von Hofstätter). Dieses Maß kann von -1 bis +1 variieren; ein Wert von +1 bedeutet, daß es eine sehr gute, totale Übereinstimmung gibt; ein Wert von -1 sagt aus, daß die Einschätzungen genau gegensätzlich sind; ein Wert um 0 weist auf keinen Zusammenhang der Teilnehmereinschätzungen mit denen der Referenten / Leiter hin. Hier die Werte:

- Q-Wert für Ref./Leiter A =
- Q-Wert für Ref./Leiter B =
- Q-Wert für Ref./Leiter C =
- Q-Wert für Ref./Leiter D =



Transformationsgewichte

In der folgenden Tabelle sind jeweils die neuen Werte von TRAI und TRADI nach den Transformationen angegeben. Bei TRAI wurden fehlende Werte vorerst mit -9, bei TRADI mit -1 kodiert.

Item-Nr.	TRAI				TRADI			
	1	2	3	4	1	2	3	4
1	-2.2	-1.4	-0.3	0.8	0	0	0	1
2	-1.5	-0.7	0.2	1.4	0	0	1	1
3	-1.6	-1.0	-0.6	0.4	0	0	0	1
4	-2.0	-1.5	-0.9	0.3	0	0	0	1
5	-2.3	-1.6	-0.8	0.4	0	0	0	1
6	-2.0	-1.5	-0.5	0.7	0	0	0	1
7	-2.0	-1.3	-0.6	0.5	0	0	0	1
8	-1.7	-0.9	-0.2	0.8	0	0	0	1
9	-2.2	-1.7	-1.2	0.2	0	0	0	1
10	-2.0	-1.4	-0.4	0.8	0	0	0	1
11	-2.2	-1.7	-0.9	0.4	0	0	0	1
12	-2.1	-1.6	-1.0	0.3	0	0	0	1
13	-2.1	-1.7	-1.0	0.3	0	0	0	1
14	-0.9	-0.1	0.3	1.1	0	0	1	1
15	-2.6	-2.0	-1.2	0.2	0	0	0	1
16	-2.0	-1.1	-0.3	0.8	0	0	0	1
17	-1.9	-1.2	-0.4	0.7	0	0	0	1
18	-1.4	-0.9	-0.3	0.7	0	0	0	1
19	-1.5	-0.8	-0.1	0.9	0	0	0	1
20	-2.4	-1.9	-1.4	0.1	0	0	0	1
21	-2.3	-1.8	-1.2	0.2	0	0	0	1
22	-2.0	-1.3	-0.5	0.6	0	0	0	1
23	-2.1	-1.2	-0.5	0.6	0	0	0	1
24	-1.1	-0.3	0.3	1.1	0	0	1	1
25	-2.1	-1.6	-0.9	0.4	0	0	0	1
26	-2.4	-1.7	-1.0	0.3	0	0	0	1
27	-2.2	-1.8	-1.2	0.2	0	0	0	1
28	-1.8	-1.2	-0.5	0.6	0	0	0	1
29	-2.2	-1.7	-1.2	0.2	0	0	0	1
30	-1.5	-0.9	-0.4	0.6	0	0	0	1
31	-1.7	-1.0	-0.2	0.8	0	0	0	1
32	-2.3	-1.6	-0.8	0.5	0	0	0	1
33	-1.8	-1.3	-0.8	0.4	0	0	0	1
34	-3.0	-2.3	-1.3	0.2	0	0	0	1
35	-1.6	-0.8	-0.1	0.9	0	0	0	1

In dieser Tabelle sind die Transformationsgewichte für die schon gleichgepolten Items aufgelistet.

Zusammenfassung einiger Itemstatistiken

(für verschiedene Transformationen: ITEM, DITEM, TRAI, TRADI)

Nr.	M ITEM	s <sup>2</sup>	M TRAI	M TRADI	M DITEM	Fit von ITEM (1)	Fit von DITEM	Fit von TRAI	Fit von TRADI	n
1	3.2	0.6	.0	.4	.9	.56	.49	.58	.53	467
2	2.7	0.8	-.0	.6	.6	.37	.35	.38	.30	461
3	3.4	1.1	-.1	.7	.8	.29	.27	.25	.23	472
4	3.6	0.6	-.1	.7	.9	.49	.42	.53	.55	469
5	3.6	0.5	-.1	.7	.9	.43	.32	.46	.50	467
6	3.4	0.6	-.0	.5	.9	.42	.38	.43	.39	474
7	3.4	0.7	-.1	.6	.9	.44	.33	.44	.44	471
8	3.0	1.0	-.1	.4	.7	.38	.34	.35	.31	468
9	3.8	0.4	-.1	.8	.9	.39	.31	.42	.41	474
10	3.3	0.6	-.1	.4	.9	.48	.45	.51	.49	467
11	3.6	0.5	.0	.7	.9	.54	.54	.54	.48	473
12	3.6	0.5	-.1	.8	.9	.56	.50	.56	.53	471
13	3.6	0.5	-.1	.7	.9	.62	.51	.63	.58	473
16	3.2	0.8	.0	.4	.8	.52	.48	.53	.46	474
17	3.2	0.8	-.1	.5	.8	.48	.41	.49	.46	472
18	3.1	1.2	.0	.5	.8	.23	.24	.19	.21	473
19	3.0	1.1	.0	.4	.7	.48	.45	.47	.42	446
21	3.8	0.4	-.1	.8	.9	.43	.35	.46	.43	471
22	3.4	0.7	.0	.6	.9	.43	.37	.44	.43	470
23	3.3	0.7	-.1	.5	.8	.54	.54	.54	.47	473
24	2.5	1.3	.0	.5	.5	.41	.43	.38	.35	464
25	3.6	0.5	.0	.7	.9	.39	.37	.41	.40	471
26	3.8	0.4	-.1	.8	.9	.32	.24	.36	.40	473
27	3.7	0.4	-.1	.8	.9	.59	.54	.58	.51	474
28	3.3	0.8	-.1	.5	.8	.57	.53	.56	.49	470
29	3.7	0.4	-.1	.8	.9	.43	.42	.39	.35	458
30	3.2	1.2	.0	.6	.8	.24	.25	.20	.19	469
31	3.1	0.9	-.1	.4	.8	.64	.54	.63	.50	441
32	3.5	0.5	-.1	.7	.9	.38	.35	.40	.38	452
33	3.5	0.8	-.1	.7	.9	.34	.23	.41	.48	449
35	3.0	1.0	.0	.4	.7	.40	.34	.40	.42	456
	3.4		-.06	.6	.8	(2)				

(1) alle Korrelationen ohne Part-Whole Korrektur

(2) in dieser Zeile steht jeweils der Mittelwert aller Mittelwerte einer bestimmten Transformation, d.h. die Spaltensumme dividiert durch die Anzahl der Items.

Faktorenanalyse TRAI, 4 Faktoren, Kommunalitäten und  
Ladungen

Nr.	L a d u n g e n				Kommunalität
	Faktor 1	Faktor 2	Faktor 3	Faktor 4	
1	.24	.25	.47	.15	.36
2	-.01	.38	.21	.01	.19
3	.10	.06	.04	.30	.11
4	.45	.22	.32	-.06	.35
5	.20	.41	.13	.06	.22
6	.14	.40	.15	.00	.20
7	.24	.40	.03	.12	.23
8	-.12	.13	.31	.47	.35
9	.41	-.03	.34	.07	.29
10	.26	.24	.38	.01	.27
11	.55	.33	.01	.15	.44
12	.62	.27	.09	.06	.48
13	.60	.18	.37	.11	.54
16	.21	.15	.60	-.02	.42
17	.18	.44	.19	.02	.26
18	.12	-.03	-.14	.55	.34
19	.06	.30	.35	.15	.24
21	.37	.31	.06	.11	.25
22	.23	.39	.07	.05	.21
23	.16	.19	.61	.01	.44
24	-.11	.18	.30	.49	.38
25	.45	.13	.11	-.02	.24
26	.07	.42	.06	-.01	.19
27	.51	.23	.31	.07	.41
28	.19	.50	.18	.19	.36
29	.26	.24	.10	.11	.15
30	.06	-.06	-.02	.57	.34
31	.31	.28	.47	.17	.42
32	.21	.29	.16	-.04	.16
33	.17	.29	.22	-.04	.17
35	.09	.29	.26	-.01	.16

Faktorenanalyse TEM, 4 Faktoren, Kommunalitäten und  
Ladungen

Nr.	L a d u n g e n				Kommunalität
	Faktor 1	Faktor 2	Faktor 3	Faktor 4	
1	.24	.27	.45	.15	.36
2	.01	.41	.16	.04	.19
3	.10	.07	.02	.30	.10
4	.43	.17	.31	-.07	.31
5	.19	.40	.11	.06	.21
6	.17	.38	.17	-.03	.20
7	.31	.32	.03	.11	.21
8	-.07	.12	.30	.45	.32
9	.35	-.04	.34	.08	.25
10	.26	.23	.35	.01	.25
11	.60	.27	.04	.12	.44
12	.67	.26	.06	.04	.53
13	.62	.13	.38	.10	.55
16	.17	.20	.61	-.02	.44
17	.20	.44	.17	.03	.26
18	.14	-.08	-.12	.54	.33
19	.07	.35	.30	.17	.25
21	.45	.19	.08	.11	.25
22	.24	.41	.02	.06	.23
23	.15	.23	.61	.04	.45
24	-.06	.21	.25	.50	.36
25	.43	.11	.14	-.03	.22
26	.10	.38	.04	-.01	.15
27	.57	.19	.30	.10	.46
28	.28	.45	.17	.19	.34
29	.36	.25	.07	.10	.21
30	.04	-.05	-.02	.59	.35
31	.33	.27	.46	.18	.43
32	.24	.23	.15	-.01	.13
33	.15	.24	.17	-.05	.11
35	.10	.29	.23	.01	.15

Diskriminanzanalyse TRAI, DIRECT: Varianzanteile und  
canonische Korrelationen

Function	(1) % der disk. Varianz	(2) kumulat. Prozent	(3) canonische Korrelation
1	20.3	20.3	.76
2	13.8	34.1	.70
3	11.4	45.4	.66
4	9.8	55.2	.64
5	8.0	63.2	.60
6	7.1	70.3	.57
7	5.4	75.7	.52
8	4.2	80.0	.47
9	3.4	83.2	.44
10	3.2	86.4	.43
11	3.0	89.4	.42
12	2.1	91.6	.36
13	1.7	93.3	.33
14	1.5	94.8	.31
15	1.2	96.0	.28
16	1.1	97.1	.26
17	1.0	98.1	.25
18	0.6	98.6	.19
19	0.5	99.1	.18
20	0.3	99.4	.15
21	0.3	99.7	.14
22	0.2	99.9	.12
23	0.1	100.0	.06
24	0.0	100.0	.05

Bei der Methode DIRECT werden alle 31 Fragen zugleich in der Diskriminanzanalyse betrachtet.

- (1) Jeder Funktion ist ein Eigenwert zugeordnet, die Summe aller Eigenwerte ist ein Maß für die totale diskriminierende Varianz; der Prozentanteil eines Eigenwertes ist somit ein Maß für die relative Bedeutung der korrespondierenden Funktion.
- (2) Diese Spalte gibt an, wieviel % der gesamten diskriminierenden Varianz durch eine bestimmte Anzahl von Diskriminanzfunktionen erklärt wird.
- (3) Hier ist die kanonische Korrelation einer Diskriminanzfunktion und dem Set von 24 dummy-Variablen -die die 25 Gruppen definieren- angegeben. Dies ist ein Maß dafür, wie eng die Funktionen und die Gruppenvariable zusammenhängen; das Quadrat kann interpretiert werden als der Varianzanteil in der Diskriminanzfunktion, der durch die Gruppen erklärt wird.



Diskriminanzanalyse TRAI, WILKS: Diskriminanzkoeffizienten und Korrelationen von 17 Fragen mit den 4 Dimensionen

(Bei einer schrittweisen Diskriminanzanalyse wird zuerst die Frage herausgesucht, die alleine am besten trennt; diese Frage wird mit allen übrigen kombiniert, und das beste Paar wird ausgewählt. Dieses Paar wird dann mit allen übrigen Fragen kombiniert, und es wird die beste Dreier-Kombination ausgewählt. Dieser Prozeß wird solange fortgesetzt, bis entweder alle Fragen aufgenommen wurden oder bei einem Schritt keine Hinzunahme einer Frage zu einer Verbesserung der Trennkraft führt. Als Kriterium wird der multivariate F-Test für Unterschiede zwischen Gruppennittelwerten genommen; jene Fragen, die dieses F maximieren, minimieren zugleich WILKS LAMBDA, so daß hier auf jedem Schritt die Frage mit dem geringsten LAMBDA ausgewählt wird (vgl. NIE u.a. 1975, S. 447). Eine Frage wird zudem nur dann aufgenommen bzw. beibehalten, wenn sie auf dem 0,1 % Niveau noch einen signifikanten Beitrag zur Trennung leistet (PIN/POUT=.001; ein weniger strenges Niveau von z.B. 5 % führt zu einer Auswahl von 24 Fragen, die aber ein prinzipiell vergleichbares Ergebnis darstellen). Es wurden in diesem Prozeß 17 Fragen ausgewählt. Eine Analyse mit 5 Diskriminanzfunktionen führt wieder zu Interpretationsschwierigkeiten, so daß wiederum nur 4 Funktionen interpretiert werden.)

Nr	FUNCTION 1			FUNCTION 2			FUNCTION 3			FUNCTION 4		
	Disk. Coef.	r mit F1	r <sup>2</sup>	Disk. Coef.	r mit F2	r <sup>2</sup>	Disk. Coef.	r mit F3	r <sup>2</sup>	Disk. Coef.	r mit F4	r <sup>2</sup>
2	-.16			.00			.20	.35	.12	.17		
3	-.58	-.59	.35	.08			.12			.47	-.50	.25
7	.16			.05	.24	.06	.03			.19		
8	-.24	-.34	.12	.04			.04			.36	.50	.25
11	.21			-.09	.32	.10	.16			.19		
13	-.04			.43	.62	.38	.40			.08		
16	.20			.32	.52	.27	.39			.07		
17	-.12			.15			.19	.20	.04	.28		
18	-.46	-.55	.30	.10			.27			.18		
22	.08			.01			.48	.44	.19	.13		
23	.13			.05	.43	.19	.07			.05		
24	-.27	-.36	.13	.05			.37			.55	.70	.50
27	.01			.50	.69	.48	.01			.11		
28	-.10			.29	.47	.22	.35	.34	.12	.12		
29	-.11			.12			.20	.20	.04	.05		
30	-.33	-.46	.21	.10			.30			.15		
32	.10			.26	.41	.17	.35	.34	.12	.08		

Bei den Korrelationen sind nur jene Korrelationen mit r größer/gleich .3 angegeben; ist keine Korrelation zu einer Dimension größer/gleich .3, so ist jeweils die höchste Korrelation zu einer Dimension angegeben.

Die 4 Funktionen klären zusammen 63,6 % der gesamten diskriminierenden Varianz auf; in der Reihenfolge der Funktionen: 24,4 %, 15,1 %, 12,8 % und 11,3%.

Faktorenanalyse TRAI, reduzierter Satz von 20 Items,  
4 Faktoren, Kommunalitäten und Ladungen

Nr.	L a d u n g e n				Kommunalität
	Faktor 1	Faktor 2	Faktor 3	Faktor 4	
11	.60	.05	.28	.16	.47
12	.71	.13	.16	.05	.55
13	.61	.41	.10	.07	.55
25	.50	.08	.10	-.01	.26
27	.47	.32	.25	.04	.38
1	.27	.51	.18	.12	.39
10	.25	.35	.31	-.01	.28
16	.17	.54	.27	-.07	.40
23	.12	.59	.22	-.02	.41
31	.32	.52	.20	.13	.42
5	.18	.14	.46	.06	.26
6	.17	.22	.32	-.01	.18
17	.23	.22	.36	.02	.23
26	.05	.06	.51	.03	.26
28	.19	.22	.46	.19	.33
3	.06	.05	.07	.30	.10
8	-.08	.41	-.01	.45	.38
18	.07	-.12	.04	.57	.34
24	-.02	.38	.02	.46	.36
30	.04	-.03	.01	.57	.33

Die Nummer der Faktoren entspricht der im Text beschriebenen Nummer der Skalen, d.h. Faktor 1 beinhaltet die Items, die Skala I bilden usw..

Skalen-Statistiken und Fragen-Skalen-Korrelationen

	$\bar{x}$	Median	s	n	Kurtosis	Skewness
SCALE 1	-.06	-.08	.87	478	-.33	-.68*
SCALE 2	-.01	-.08	.98	475	-.27	-.21
SCALE 3	-.02	.07	.94	477	-.43	-.24
SCALE 4	.00	.03	.98	476	-.48	-.08
GESAMT	-.01	.02	1.00	474	-.02	-.06
DIN 1	18.3	19.2	2.5	478	4.9	-2.2*
DIN 2	16.0	16.5	2.9	476	0.6	-0.8*
DIN 3	17.1	17.5	2.4	477	1.0	-1.0*
DIN 4	15.2	15.7	3.3	476	-0.5	-0.5*
GELTEN	66.5	67.8	7.8	474	1.7	-1.1*

Nach LIENERT 1969, S.172f. wird mit einem t-Test auf Normalität geprüft. Mit \* sind die Kurtosis- und Skewness-Koeffizienten bezeichnet, die auf dem 1 % Niveau von der Hypothese der Normalverteilung abweichen.

TRAI	SCALE 1	SCALE 2	SCALE 3	SCALE 4	Gesamt	gehört zu
11	.72	.30	.38	.16	.52	SCALE 1
12	.73	.34	.34	.10	.51	
13	.70	.50	.33	.16	.59	
25	.69	.26	.20	-	.39	
27	.63	.46	.34	.11	.52	
1	.40	.68	.33	.23	.59	SCALE 2
10	.38	.65	.34	-	.52	
16	.33	.73	.33	-	.54	
23	.27	.71	.32	.14	.52	
31	.42	.67	.35	.23	.59	
5	.31	.31	.60	-	.45	SCALE 3
6	.28	.28	.64	-	.44	
17	.32	.34	.62	-	.47	
26	.19	.22	.56	-	.36	
28	.33	.36	.63	.20	.52	
3	-	.12	-	.49	.28	SCALE 4
8	-	.23	.16	.64	.40	
18	-	-	-	.61	.26	
24	.13	.24	.17	.67	.45	
30	-	-	-	.63	.27	

(Aufgelistet sind nur solche Fragen-Skalenkorrelationen, die auf dem 1 % Niveau signifikant von Null verschieden sind.)

Die Skalen-Statistiken für die Skalen SCALE 1-4 fallen technisch gesehen besser aus; die Hypothese der Normalverteilung kann hier im Unterschied zu den Skalen DIN beibehalten werden.

Die Fragen-Skalen-Korrelationen (nicht Part-Whole korrigiert) fallen befriedigend aus; die höchste Korrelation hat eine Frage immer mit der zugehörigen Skala; daneben gibt es aber noch andere höhere Korrelationen, die auf eine hohe Skaleninterkorrelation hinweisen. Die Struktur der Korrelationen ist bei SCALE und DIN identisch.

Diskriminanzanalyse auf Skalenebene

Alle Skalen wurden gleichzeitig betrachtet; 3 Funktionen erklären bei SCALE 94,4 % und bei DIM 95 % der diskriminierenden Varianz. Die folgende Tabelle enthält die einfachen Korrelationen der Skalen zu den entsprechenden 3 Diskriminanzfunktionen

	FUNC 1	FUNC 2	FUNC 3
SCALE 1	-.03	.88	.34
SCALE 2	.08	.79	-.48
SCALE 3	.04	.49	.48
SCALE 4	.98	.19	.05
DIM 1	.06	-.77	.35
DIM 2	.11	-.89	-.35
DIM 3	.07	-.55	.61
DIM 4	.99	-.13	.02

In absteigender Folge erklären die 3 Diskriminanzfunktionen bei SCALE 54,1 %, 28,6 % und 11,7 % der diskriminierenden Varianz, bei DIM 56,4 %, 26,7 % und 11,8 %. Die Diskriminanzanalysen bei DIM und SCALE decken sich weitgehend (die unterschiedlichen Vorzeichen bei FUNCTION 2 kommen durch die Polung zustande). Bei DIM wird nur die 3. Diskriminanzfunktion noch stärker und deutlicher durch DIM 3 "Arbeitsklima" bestimmt; dies kommt der Analyse mit 31 bzw. 17 Fragen näher.

3 Diskriminanzanalysen erlauben bei SCALE 18,6 % und bei DIM 18,4 % richtige Klassifikationen (bei 4 Diskriminanzfunktionen 19,6 % bzw. 21,3 %). Gegenüber der Berücksichtigung von 31 Fragen (schlechtestes Klassifikationsergebnis bei 4 Funktionen = 34,7 %) oder 17 Fragen (schlechtestes Ergebnis bei 4 Funktionen = 31,2 %) sind also bei den Skalenwerten weniger richtige Klassifikationen möglich. Die Zusammenfassung der Fragen in Skalen beinhaltet also einen Informationsverlust: auf der Ebene der Fragen sind deutlich mehr diskriminierende Informationen enthalten, die teilweise auf der Skalenebene verloren gehen. Andererseits - betrachtet man die Wahrscheinlichkeit einer richtigen Klassifikation ohne Ausnutzung der Fragebogeninformationen (4,2 % richtige Klassifikationen) - so lassen sich mittels einer Diskriminanzanalyse mit den 4 Skalen immerhin noch bei DIM 4,4 mal mehr richtige Klassifikationen treffen. Auch auf Skalenebene ist der Fragebogen also noch für Unterschiede zwischen den Veranstaltungen sensibel.

Hans - Günther Roßbach

4400 Münster, den  
Hansaring 26 - 28  
Tel.: 0251 - 66 29 19

Betr.: Expertenbefragung zur Einschätzung der Gültigkeit eines Fragebogens zur Beurteilung von Fortbildungsveranstaltungen für Fachkräfte im Vorschulbereich

Sehr geehrte Damen und Herren!

Ich wende mich an Sie mit der Bitte um Unterstützung eines Forschungsvorhabens über "Instrumente zur Evaluation von Fortbildungsveranstaltungen für Fachkräfte im Vorschulbereich", welches ich im Rahmen meiner Dissertation am Institut für Erziehungswissenschaft der Universität Münster durchführe.

Im Bereich der Vorschuleraziehung, bzw. des Elementarbereiches, -wie auch in anderen sozialpädagogischen Bereichen- wird häufig auf die Notwendigkeit eines ständigen Weiterlernens der Fachkräfte (speziell der Erzieher) hingewiesen, da Mängel in der Ausbildung oder ein Verschleiß des alten Wissens festgestellt werden. In den letzten Jahren hat sich eine vielfältige Fortbildungspraxis entfaltet, die sich unterscheidet hinsichtlich Organisationsformen, Inhalte, Methoden, Medieneinsatz usw.. Gegenüber diesen erfreulichen Aktivitäten sind Versuche, die Güte, die Wirkung oder den Erfolg einer Fortbildungsveranstaltung zu erfassen, nur wenig entwickelt - sieht man einmal von den zuweilen praktizierten Möglichkeiten informeller Rückkopplungsgespräche oder ad hoc konstruierter Verfahren ab. Dies überrascht, da an vielen Stellen auf die Wichtigkeit und Notwendigkeit einer Erfolgskontrolle hingewiesen wird.

Auf diesem Hintergrund sollte versucht werden, ein Instrument zur Beurteilung von Fortbildungsveranstaltungen zu entwickeln, welches einmal für die Praxis und die Fortbildungspraktiker sinnvoll einsetzbar sein und zum anderen wissenschaftlichen Gütekriterien eines Beurteilungsinstrumentes genügen soll. Hierfür war eine zweifache Beschränkung nötig:

- auf mehrtägige Fortbildungsveranstaltungen, die an einem zentralen Lernort angeboten werden. Damit soll nicht die Relevanz von "Fortbildung vor Ort", "Praxisberatung" oder "selbstorganisierten Arbeitskreisen" verneint werden, aber in allen Konzepten wird zentralen Veranstaltungen eine besondere Bedeutung zugeschrieben.
- auf einen Fragebogen, der mittels Einschätzskalen konstruiert wird. Diese Auswahl wurde getroffen, da ähnliche Verfahren schon angewandt werden (aber eben nicht abgesichert sind) und da sich solche Verfahren in einem gewissen Rahmen in anderen Bereichen bewährt haben (z.B. Lehrereinschätzverfahren im anglo-amerikanischen Bereich).

Die bisherigen Entwicklungsschritte dieses Fragebogens waren zuerst eine möglichst umfassende Sammlung von Items. Dafür wurde die gängige Literatur (allgemeine Erwachsenenbildung, Unterrichtsforschung, Lehrerbewertungsforschung, Fortbildungsdiskussion) einer ausführlichen Analyse unterzogen. Daran schloß sich eine erste Erprobung an. Aufgrund der Ergebnisse einer Faktorenanalyse wurde die Anzahl der Items stark reduziert (weitere Ergebnisse der bisherigen Entwicklungsarbeit sind in Anlage I zusammengestellt; für die Expertenbefragung braucht

dieser Teil jedoch nicht gelesen zu werden). Der neue, verbesserte Fragebogen umfaßt nun 55 Items.

Um Ihnen einen Einblick in den Kontext des Fragebogensinsatzes zu geben, möchte ich hierzu noch einige Angaben machen. Der Fragebogen (siehe Anlage IV) wird am Schluß der mehrtägigen Veranstaltung von dem Leiter (falls kein Leiter vorhanden ist, von dem Referenten) ausgegeben und von den Teilnehmern ausgefüllt. Der Leiter gibt dabei einige Erläuterungen zum Fragebogen (siehe Anlage II); er selbst erhält vorher zusätzlich eine Projektbeschreibung (siehe Anlage III), um auf Rückfragen eingehen zu können.

In dem momentanen Entwicklungsstadium scheint sich damit eine praktikable Form eines Beurteilungsinstrumentes abzuzeichnen, so daß nun eine Einschätzung der Güte, Validität des Fragebogens unzugänglich wird. Für diese Einschätzung der Gültigkeit bin ich auf Sie als Experten angewiesen. In dem beigefügten Bogen "Vorlage zur Experteneinschätzung" habe ich für Ihre Experteneinschätzung einige Fragen zusammengestellt. Bitte beantworten Sie mir diese und schicken Sie mir den ausgefüllten Bogen an die oben angegebene Adresse bis zum                      zurück.

Ich möchte Ihnen vielmals für Ihre Mitarbeit danken.

Ihr

*Hans-Joachim Roßbach*

Vorlage zur Experteneinschätzung

Bitte geben Sie Ihre Einschätzung in folgenden Schritten ab:

1. Sehen Sie sich bitte den Fragebogen an -hier vor allem die Seiten 2 - 5 (siehe Anlage IV).

2. Halten Sie als Experten für Fortbildung diese Zusammenstellung der Items für eine gelungene Auswahl aus dem Gesamt der zur Zeit diskutierten Postulate zur Gestaltung einer Fortbildungsveranstaltung?

a. Geben Sie bitte eine Einschätzung zwischen 1 (sehr gute Auswahl) und 6 (sehr schlechte Auswahl) ab:

1	2	3	4	5	6
(sehr gute Auswahl)					(sehr schlechte Auswahl)

b. Welche Aspekte, Bereiche kommen Ihrer Meinung nach zu kurz?...

.....

.....

.....

3. Betrachten Sie nun bitte jedes Item einzeln und schätzen Sie den angesprochenen Aspekt nach seiner Bedeutung / Wichtigkeit für die Gestaltung einer Fortbildungsveranstaltung ein. Geben Sie bitte eine Einschätzung zwischen 1 (sehr wichtig) und 6 (völlig unwichtig) ab:

Item-Nr.	(sehr wichtig)					(völlig unwichtig)
1	1	2	3	4	5	6
2	1	2	3	4	5	6
3	1	2	3	4	5	6
4	1	2	3	4	5	6
5	1	2	3	4	5	6
6	1	2	3	4	5	6
7	1	2	3	4	5	6
8	1	2	3	4	5	6
9	1	2	3	4	5	6
10	1	2	3	4	5	6
11	1	2	3	4	5	6
12	1	2	3	4	5	6
13	1	2	3	4	5	6
14	1	2	3	4	5	6
15	1	2	3	4	5	6
16	1	2	3	4	5	6
17	1	2	3	4	5	6
18	1	2	3	4	5	6
19	1	2	3	4	5	6
20	1	2	3	4	5	6
21	1	2	3	4	5	6
22	1	2	3	4	5	6
23	1	2	3	4	5	6
24	1	2	3	4	5	6
25	1	2	3	4	5	6
26	1	2	3	4	5	6

Item-Nr.	(sehr wichtig)					(völlig unwichtig)
27	1	2	3	4	5	6
28	1	2	3	4	5	6
29	1	2	3	4	5	6
30	1	2	3	4	5	6
31	1	2	3	4	5	6
32	1	2	3	4	5	6
33	1	2	3	4	5	6
34	1	2	3	4	5	6
35	1	2	3	4	5	6

4. Die 35 Items ließen sich nach der Voruntersuchung in 5 Bereiche aufteilen (siehe auch Anlage I). Schätzen Sie bitte wieder die Bedeutung der einzelnen Bereiche für die Gestaltung einer Fortbildungsveranstaltung ein:

Bereich	(sehr wichtig)					(völlig unwichtig)
1. Methodischer Aufbau und Strukturierung einer Veranstaltung.....	1	2	3	4	5	6
2. Arbeitsklima in der Veranstaltung und Leiterverhalten.....	1	2	3	4	5	6
3. Lernertrag und globaler Eindruck von der Veranstaltung.....	1	2	3	4	5	6
4. Aktivität und Selbständigkeit der Teilnehmer.....	1	2	3	4	5	6
5. Äußerer / organisatorischer Rahmen der Fortbildungsveranstaltung.....	1	2	3	4	5	6

Der Fragebogen soll nicht nur ein Forschungsinstrument sein, sondern er soll auch für die konkrete Arbeit des Fortbildungspraktikers (z.B. Auswertungsgespräche mit Teilnehmern und Referenten; Protokollerstellung; Programmplanung; Absicherung der eigenen Arbeit) nützlich sein. Diese Frage nach der "Nützlichkeit" für die praktische Arbeit - um die es im Folgenden geht - ist zu trennen von der Frage nach der eher theoretisch - wissenschaftlichen Bedeutsamkeit, Wichtigkeit von Items.

5. Stellen Sie sich nun bitte die Situation eines Leiters einer Fortbildungsveranstaltung vor, in dessen Veranstaltung der Fragebogen ausgefüllt wurde und dessen Antworten er nun vor sich liegen hat. Glauben Sie, daß die möglichen Antworten auf die Fragen für die Arbeit des Leiters nützlich sein können?



- a. Geben Sie bitte eine Einschätzung zwischen 1 (sehr nützlich) und 6 (völlig unnütz) ab:

1	2	3	4	5	6
(sehr					(völlig
nützlich)					unnütz)

- b. Wo sehen Sie Schwierigkeiten in der Verwirklichung des angestrebten Zieles der Nützlichkeit?.....

.....

.....

.....

6. Betrachten Sie nun bitte einen potentiellen Teilnehmer einer Fortbildungsveranstaltung. Glauben Sie, daß das Ausfüllen des Fragebogens für den Teilnehmer "etwas bringen" kann? Welche Auswirkungen hat die Bearbeitung vermutlich auf die Teilnehmer?.....

.....

.....

.....

.....

- Falls Sie noch Anmerkungen haben, so wäre ich Ihnen dankbar, wenn Sie sie stichwortartig darlegen würden:.....

.....

.....

.....

.....

.....

Nochmals vielen Dank für Ihre Mitarbeit.

Erste Ergebnisse der Fragebogenentwicklung

Von September bis Dezember 1978 wurden 44 Items einer ersten Erprobung unterzogen. Aus 27 Veranstaltungen lagen insgesamt 467 ausgefüllte Fragebögen vor, von denen 445 auswertbar waren. Da sich zwischen Tagesveranstaltungen und mehrtägigen Fortbildungsveranstaltungen größere Unterschiede ergaben, werden im folgenden nur die mehrtägigen Veranstaltungen betrachtet. Diese Untergruppe bestand aus 368 Personen, die 20 Veranstaltungen besucht hatten.

6 Items wurden aus den weiteren Analysen ausgeschlossen, da in ihnen weniger als 5 % der Teilnehmer die Veranstaltung negativ einschätzten. Die restlichen Items wurden prozentrang-transformiert und anschließend in z-Äquivalente umgewandelt. Aus diesen Werten wurde ein Gesamtwert gebildet und jedes Item damit korreliert. Es ergaben sich Trennschärfen von .27-.75. In dem Gesamtwert unterscheiden sich die Veranstaltungen nach einer Einwegvarianzanalyse hochsignifikant (auf Itemebene unterscheiden sich die Veranstaltungen in 36 von 38 Items auf dem 1 % Niveau). Für den Gesamtwert wurde eine Reliabilitätschätzung durchgeführt; nach Kuder Richardson Formula 8 ergab sich z.B. ein  $r_{tt}$  = .92.

Verschiedene Faktorenanalysen sprachen für eine 5 Faktor-Lösung, die 39,7 % an der Gesamtvarianz aufklärte (für 5 Faktoren sprach der Eigenwertverlauf und vor allem das FÜRTRATT-Kriterium). Gemäß dieser 5 Dimensionen wurde der neue, hier vorliegende Fragebogen konstruiert (pro Dimension 7 Items). Im folgenden sind die Dimensionen sowie die dazu gehörenden Items des neuen Fragebogens aufgelistet:

<u>Dimension:</u>	<u>Item-Nr.:</u>
I. Methodischer Aufbau und Strukturierung der Veranstaltung	1,5,10,15,19,25,27
II. Arbeitsklima in der Veranstaltung und Leiterverhalten	6,7,26,32,33,34,35
III. Lernertrag und globale Einschätzung der Veranstaltung	4,9,13,16,20,23,31
IV. Aktivität und Selbständigkeit der Teilnehmer	2,11,12,17,21,28,29
V. Äußerer/organisatorischer Rahmen der Fortbildungsveranstaltung	3,8,14,18,22,24,30

Um aus den Ergebnissen der ersten Itemanalyse einen vorläufigen Hinweis auf die neue Zusammenstellung der Items und der 5 Dimensionen zu bekommen, wurden entsprechende Items der ersten Erprobung, die noch am ehesten die verschiedenen Gütekriterien erfüllten, ausgewählt (6 je Dimension) und zu Skalen durch einfaches Aufaddieren zusammengefaßt. Bei allen Skalen ergaben sich nach Einwegvarianzanalysen signifikante Gruppenunterschiede auf dem 1 % Niveau. 3 Skalen nähern sich einer Normalverteilung, die beiden anderen scheinen relativ rechtsschief verteilt. In jeder Dimension wurde eine Reliabilitätschätzung versucht; nach Kuder Richardson Formula 8 ergab sich ein  $r_{tt}$  von .72-.87.

Es scheint insgesamt die Hoffnung berechtigt, daß der neue, revidierte Fragebogen ähnlich hohe bzw. bessere Werte in den verschiedenen Gütekriterien erhält.

Experten- und Fortbildungsleiterbefragung: WichtigkeitTab. 1: Wichtigkeit der einzelnen Aspekte für die Gestaltung einer Fortbildungsveranstaltung (Punkt 3 der Vorlage zur Experteneinschätzung)

Frage Nr.	Experten		Leiter		Frage Nr.	Experten		Leiter	
	$\bar{x}$	s	$\bar{x}$	s		$\bar{x}$	s	$\bar{x}$	s
1	1.7	0.8	1.2	0.4	21	1.7	0.6	1.5	0.6
2	2.0	1.3	2.2	0.4	22	2.3	1.4	1.5	0.6
3	2.3	0.9	1.7	0.8	23	1.7	0.7	1.8	1.0
4	3.0	1.6	2.5	2.1	24	3.0	1.3	2.5	1.1
5	2.1	1.1	1.2	0.4	25	2.1	1.0	2.3	1.8
6	1.9	0.7	1.2	0.4	26	1.5	0.7	1.0	0.0
7	2.9	1.0	1.5	0.8	27	2.0	0.9	2.0	1.6
8	2.1	0.7	2.0	0.6	28	1.8	0.9	3.0	1.3
9	1.9	1.0	1.2	0.4	29	2.2	0.7	2.0	1.7
10	2.5	1.4	2.3	1.9	30	2.8	0.9	2.3	1.4
11	2.6	0.9	2.8	2.1	31	3.1	1.8	3.2	1.5
12	2.5	1.2	2.0	0.9	32	2.1	1.1	1.7	1.2
13	3.5	1.4	3.5	2.1	33	2.0	0.7	1.2	0.4
14	2.3	0.9	2.8	1.9	34	2.2	1.1	1.2	0.4
15	2.2	1.3	1.5	0.6	35	2.4	0.8	2.0	0.9
16	2.1	1.3	1.3	0.5					
17	2.1	1.3	2.5	1.9					
18	2.1	0.8	2.0	1.3					
19	2.5	1.3	2.6	1.5					
20	1.9	1.0	2.2	1.2					

Eine "1" betrachtet den entsprechenden Aspekt der Veranstaltung als "sehr wichtig", eine "6" als "völlig unwichtig". Bei den Experten liegen 15 Antworten vor, bei den Fortbildungs-

leitern nur 6; alle Mittelwerte und Streuungen müssen somit in Bezug auf das geringe n mit Vorsicht betrachtet werden. Wird anstatt d-s Mittelwertes der Median berechnet, so ändert sich an den Beurteilungen nichts.

Tab. 2: Wichtigkeit der 4 Skalen, gebildet nach den 4 Faktoren der Teilnehmereinschätzung

Skala	Experten		Leiter	
	$\bar{x}$	s	$\bar{x}$	s
Gesamteinsch./Meth.Aufbau	2.6	0.8	2.5	1.1
Lernertrag	2.2	0.9	1.8	0.8
Arbeitsklima	2.0	0.8	1.8	0.8
äußerer Rahmen	2.5	0.8	1.8	1.0
Gesamtwert	2.4	0.7	2.2	0.4

Die 5 Dimensionen der Voruntersuchung ließen sich nicht reproduzieren, so daß die entsprechende Frage für Experten (vgl. Punkt 4 der Vorlage zur Experteneinschätzung) und Leiter nicht interpretiert werden kann. Stattdessen wurden in Tabelle 2 aus den Einzeleinschätzungen der Wichtigkeit gemäß der Faktorenanalyse mit den 4 bekannten Dimensionen jeweils durch Zusammenfassung der Einzeleinschätzung ein Skalen- und Gesamtwert der Wichtigkeit gebildet.

Fortbildungsleiterbefragung: NützlichkeitTab. 1: Nützlichkeit der Informationen bezüglich jedes Einzelaspektes des Fragebogens (nur Fortbildungsleiter)

Frage Nr.	$\bar{Y}$	s	Median	Frage Nr.	$\bar{Y}$	s	Median
1	2.0	1.6	1.5	21	2.3	1.0	2.2
2	2.8	1.8	2.2	22	2.5	1.4	2.2
3	2.5	1.5	2.5	23	2.5	1.4	2.2
4	2.7	0.8	2.5	24	3.3	1.4	3.0
5	1.8	1.0	1.5	25	2.5	1.4	2.2
6	2.0	1.1	1.8	26	1.8	1.2	1.5
7	2.7	1.2	2.5	27	1.7	0.8	1.5
8	3.3	1.5	3.0	28	2.2	1.2	2.0
9	3.0	1.8	2.5	29	2.3	1.0	2.2
10	2.5	1.8	2.0	30	3.3	1.5	3.0
11	2.7	1.5	2.2	31	4.2	2.0	3.5
12	2.7	0.8	2.5	32	1.5	0.5	1.5
13	4.0	1.3	3.5	33	1.2	0.4	1.1
14	3.0	1.4	3.0	34	1.7	0.8	1.5
15	2.5	1.1	2.5	35	2.2	0.8	2.2
16	3.0	1.8	2.5				
17	3.3	2.1	3.5				
18	2.8	1.9	2.5				
19	3.2	1.6	2.5				
20	2.0	0.9	2.0				

Eine "1" betrachtet den entsprechenden Veranstaltungaspekt als "sehr nützlich", eine "6" als "völlig unnützlich". Die Kennzahlen in Tabelle 1 sind nur mit großer Vorsicht zu betrachten, da sie nur auf insgesamt 6 Fällen beruhen.

Tab. 2: Nützlichkeit der 4 Skalen, gebildet nach den 4 Faktoren der Teilnehmereinschätzung

Skala	$\bar{Y}$	s
Gesamteinsch./Meth. Aufbau	2.5	0.6
Lernertrag	2.8	1.0
Arbeitsklima	2.0	0.9
Äußerer Rahmen	3.2	1.2
Gesamtwert	2.7	0.8

In Tabelle 2 wurden wieder aus den Einzeleinschätzungen der Nützlichkeit gemäß der Faktorenanalyse mit den 4 bekannten Dimensionen jeweils durch Zusammenfassung der Einzeleinschätzungen ein Skalen- und ein Gesamtwert gebildet.

Vergleiche mit externen Einflußgrößen: Individualebene

Tab. 1: Statistiken

	$\bar{x}$	s	Minimum	Maximum
ALTER	27.3 J.	7.3 J.	17 J.	55 J.
BDAUER	5.8 J.	5.0 J.	0 J.	33 J.
FOZAHL	2.1	1.7	0	12
KENNEN	0.7	1.1	-	-

Die Inhalte der Veranstaltungen kannten vorher 17,8 % der Teilnehmer nicht, 66,9 % ein wenig und 15,3 % sehr gut. 11,9 % der Teilnehmer gaben an, Probleme beim Besuch der Veranstaltung gehabt zu haben.

Variablennamen:

ALTER - Alter der Teilnehmer in Jahren

BDAUER - Dauer der Berufstätigkeit im jetzigen Beruf

FOZAHL - Anzahl der in den letzten beiden Jahren besuchten Fortbildungsveranstaltungen

KENNEN - Durchschnittliche Bekanntheit des Referenten/Leiterteams

BEKANNT- Vorkenntnisse bezüglich des Themas der Veranstaltung

PROBLEM- Probleme beim Besuch der Veranstaltung

Tab. 2: Interkorrelationen

	ALTER	BDAUER	FOZAHL	BEKANNT	PROBLEM	KENNEN
ALTER	-	.73	-	-	.14	.21
BDAUER	-	-	-	.11	.11	.31
FOZAHL	-	-	-	.17	-	.31
BEKANNT	-	-	-	-	-	-
PROBLEM	-	-	-	-	-	-
KENNEN	-	-	-	-	-	-

Tab. 3: Partielle Korrelationen, Trägereffekte auspartialisiert

	ALTER	BDAUER	FOZAHL	BEKANNT	PROBLEM	KENNEN
ALTER	-	.73	-	-	.14	.24
BDAUER	-	-	-	.13	.12	.23
FOZAHL	-	-	-	.13	-	.25
BEKANNT	-	-	-	-	-	-
PROBLEM	-	-	-	-	-	-
KENNEN	-	-	-	-	-	-

In Tabelle 2 und 3 sind nur die auf dem 1 % Niveau signifikanten Korrelationskoeffizienten angegeben. In der folgenden Tabelle 4 werden die Daten auf Veranstaltungsebene aggregiert; es liegen dann 25 Fälle vor, hierfür werden nur die auf dem 5 % Niveau signifikanten Korrelationen angegeben. Tabelle 5 und 6 enthalten wieder die Werte auf Individualebene; hier beträgt die Anzahl n der gültigen Fälle ca. 475 Teilnehmer, was auch niedrige, hochsignifikante Koeffizienten erklärt.

Tab. 4: Partielle Korrelationen, Trägereffekte auspartialisiert, Daten auf Veranstaltungsebene aggregiert

	ALTER	BDAUER	FOZAHL	BEKANNT	PROBLEM	KENNEN
ALTER	-	.71	-	-	-	.42
BDAUER	-	-	.41	-	-	.47
FOZAHL	-	-	-	-	-	.37
BEKANNT	-	-	-	-	-	-
PROBLEM	-	-	-	-	-	-
KENNEN	-	-	-	-	-	-

Tab. 5: Korrelationen mit Skalen

	SCALE					DIM				
	1	2	3	4	GE-SAMT	1	2	3	4	GE-ITEM
ALTER	.12	.24	.19	.16	.26	.11	.23	.19	.16	.25
BDAUER	.15	.22	.17	.13	.25	.14	.22	.19	.15	.24
FOZAHL	.14	-	-	-	-	.15	-	-	-	-
BEKANNT	-	-	-	-	-	-	-	-	-	-
PROBLEM	-	-	-	-	-	-	-	-	-	-
KENNEN	.11	-	.13	.22	.20	.12	-	.14	.22	.21

Tab. 6: Partielle Korrelationen mit den Skalen, Trägereffekte auspartialisiert

	SCALE					DIM				
	1	2	3	4	GE-SAMT	1	2	3	4	GE-ITEM
ALTER	.13	.21	.19	.14	.26	.11	.20	.19	.14	.24
BDAUER	.16	.18	.17	-	.23	.13	.19	.18	-	.22
FOZAHL	-	-	-	-	-	-	-	-	-	-
BEKANNT	-	-	-	.12	-	-	-	-	.13	-
PROBLEM	-	-	-	-	-	-	-	-	-	-
KENNEN	-	-	-	-	-	-	-	-	-	-

Tabelle 5 und 6 enthalten nur die auf dem 1 % Niveau signifikanten Korrelationskoeffizienten.

Tab. 7: Partielle Korrelationen mit den Skalen, Trägereffekte auspartialisiert, auf Veranstaltungsebene aggregiert

	SCALE					DIM				
	1	2	3	4	GE-SAMT	1	2	3	4	GE-ITEM
ALTER	.31	.28	-	-	-	.31	.31	-	-	-
BDAUER	.34	.34	-	-	-	.30	.36	-	-	-
FOZAHL	.32	-	-	-	-	.30	-	-	-	-
BEKANNT	-	-	-	.37	-	-	-	-	.39	-
PROBLEM	.41	.48	-	-	.37	.36	.47	-	-	.33
KENNEN	-	-	.33	-	.33	-	-	.30	-	.33

Mit \* gekennzeichnete Koeffizienten sind auf dem 5 % Niveau signifikant; die anderen Koeffizienten haben ein Signifikanzniveau zwischen 5 % und 10 %.

Vergleiche mit externen Einflußgrößen: Veranstaltungsebene  
SCALE

Variable \	SCALE				GESAMT
	1	2	3	4	
TRÄGER	-	-	-	$F=10.9^{+++}$	$F=4.3^{+}$
INHALT	-	-	-	-	-
TZAHL	-	$r=.3^{\circ}$ $p=9\%$	-	$r=-.31^{\circ}$ $p=8\%$	-
VNRDAUER	-	-	-	-	-
SOZFAKT	-	-	-	$F=4.9^{+}$	-
SOZBED	-	-	-	-	-
ERFAHR	-	$r=-.3^{\circ}$ $p=9\%$	-	-	-
VERMITT	-	-	$r=.39^{+}$	-	-
METHODE	-	-	$r=.48^{+}$	-	-
PLANUNG	-	-	-	-	-
PART 1	-	-	$r=-.46^{+}$	-	-
PART 2	-	-	-	-	-
PART 3	-	-	-	-	-

Erläuterungen:

- TRÄGER = unterschiedliche Trägerzugehörigkeit/INHALT = Inhalte der Veranstaltung in 5 Kategorien/TZAHL = Teilnehmerzahl in der Veranstaltung laut Angabe der Veranstalter/VNRDAUER = Dauer der Veranstaltung in Tagen/SOZFAKT = Sozialform mit dem größten Zeitraum in der Veranstaltung/SOZBED = Sozialform, mit der größten Bedeutung hinsichtlich der Ziele der Veranstaltung/ERFAHR = durchschnittliche Anzahl der von den Referenten/Leitern in den letzten beiden Jahren durchgeführten Fortbildungsveranstaltungen/VERMITT = Kontinuum zwischen reiner Infor-

DIM

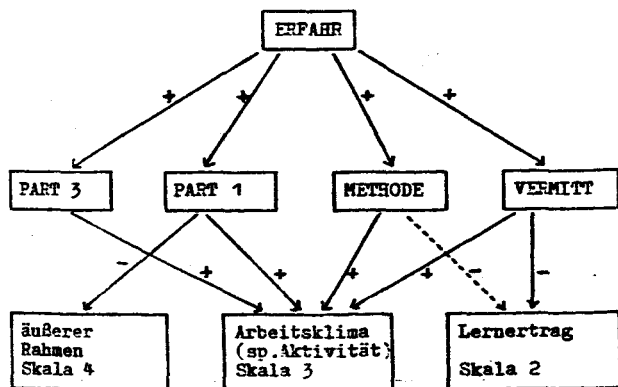
Variable	DIM				GEITEN
	1	2	3	4	
TRÄGER	-	$F=2.4^{\circ}$ $p=9\%$	-	$F=12.4^{+++}$	$F=6.8^{++}$
INHALT	-	-	-	-	-
TEILN.	-	$r=-.31^{\circ}$ $p=8\%$	-	$r=-.36^{\circ}$ $p=5.1\%$	-
VERDAUER	-	-	-	-	-
SOZFAKT	-	-	-	$F=4.7^{+}$	-
SOZBED	-	-	-	-	-
ERFAHR	-	$r=-.33^{\circ}$ $p=7\%$	-	-	-
VERMITT	-	-	$r=.42^{+}$	-	-
METHODE	-	-	$r=.53^{+++}$ $F=2.8^{\circ}$ $p=9\%$	-	-
PLANUNG	-	-	-	-	-
PART 1	-	-	$r=-.46^{+}$	-	-
PART 2	-	-	-	-	-
PART 3	-	-	-	-	-

(Forts.) mationsveranstaltung und reiner gruppendynamischer Veranstaltung/METHODE = darbietende vs. erarbeitende Vorgehensweise/PLANUNG = Ausmaß an exakter Vorplanung/PART 1 - PART 3 = Mitbestimmung der Teilnehmer bei Themenauswahl (1), Lernzielen (2) und Ablauf der Veranstaltung (3).

- Bei der Variable TRÄGER wurde eine Einwegvarianzanalyse berechnet; ansonsten meint r den Koeffizienten aus der partiellen Korrelation und F den entsprechenden F-Wert aus den Kovarianzanalysen, jeweils unter Kontrolle von Trägereffekten. Signifikanzniveau: + = 5 %; ++ = 1 %; +++ = 0.1 %; ° zeigt eine tendenzielle Beziehung an (p = 5 % - 10 %), das exakte Signifikanzniveau ist dann unter p angegeben.



# Beziehungsmuster auf Veranstaltungsebene (Ausschnitt)



## Erläuterung:

Mit zunehmender Erfahrung des Referenten/Leiterteams (ERFAHR) wird eher erarbeitend vorgegangen (METHODE) und gruppendynamische Übungen werden eingebaut (VERMITT); zugleich steigt das Ausmaß an Mitbestimmung der Teilnehmer an Entscheidungen über den Ablauf der Veranstaltung (PART 3) und der Themenauswahl (PART 1).

Das Arbeitsklima wird um so besser eingeschätzt, je mehr die Teilnehmer mitbestimmen können (PART 1 und PART 3), je eher erarbeitend vorgegangen wird (METHODE) und gruppendynamische Übungen eingebaut werden (VERMITT).

Werden aber gruppendynamische Übungen (VERMITT) eingebaut (schwächer wenn erarbeitend vorgegangen wird (METHODE)), so wird der Lernertrag von den Teilnehmern negativ eingeschätzt.

Bei Mitbestimmung an Entscheidungen über die Themenauswahl (PART 1) - in der vorliegende Stichprobe unklar zu interpretieren - sinkt zugleich die Einschätzung des Äußeren Rahmens (hier speziell des Freizeit- und Geselligkeitsaspektes).

Liebe Fortbildungsteilnehmerin!  
Lieber Fortbildungsteilnehmer!

Sie betrachten Fortbildung als eine wichtige Aufgabe, um sich auf dem "laufenden" zu halten und Ihren beruflichen Anforderungen gerecht zu werden; deswegen nehmen Sie an dieser Veranstaltung teil. Allerdings gibt es Unterschiede in der Qualität von Fortbildung, und nicht jede Veranstaltung erfüllt die Erwartungen der Teilnehmer in gleicher Weise.

Bitte beurteilen Sie mithilfe des beigefügten Fragebogens<sup>+</sup> den Verlauf und die Zweckmäßigkeit der jetzt laufenden (gerade abgelaufenen) Fortbildungsveranstaltung, die Sie als unmittelbar betroffene Teilnehmer mit am besten einschätzen können. Ihr Urteil und das der anderen Teilnehmer kann dann in einem gemeinsamen Auswertungsgespräch besprochen werden und evt. zu einer Verbesserung der Fortbildungsarbeit führen.

Der Fragebogen ist knapp gehalten und schnell zu beantworten. Gehen Sie bitte -für sich alleine- die einzelnen Punkte zügig durch und kreuzen Sie die Ihrer Meinung nach passende Antwort an.

Hier ein Beispiel:

	s t i m m t			
	nicht	ein wenig	ziemlich	völlig
0. In dieser Fortbildungsveranstaltung konnte ich aktiv mitarbeiten.	1	2	3	4
.....				
.....				

Kreuzen Sie bitte eine der Zahlen von 1 bis 4 an. Dabei bedeuten:

- 1 : Die Aussage stimmt nicht für diese Veranstaltung.  
2 : Die Aussage stimmt ein wenig für diese Veranstaltung.  
3 : Die Aussage stimmt ziemlich für diese Veranstaltung.  
4 : Die Aussage stimmt völlig für diese Veranstaltung.

Bitte bearbeiten Sie alle Aussagen. Bei jeder Aussage finden Sie auch noch zwei gepunktete Zeilen; hier können Sie kurz und stichwortartig Ihre Einschätzung zwischen "stimmt nicht (1)" und "stimmt völlig (4)" begründen.

+ Der Fragebogen wurde im Rahmen eines Forschungsvorhabens am Institut für Erziehungswissenschaften der Universität Münster entwickelt und erprobt.

© H.G. Roßbach, Münster 1980. Die Vervielfältigung zum Zwecke der Veranstaltungsbeurteilung ist gestattet. Der Verfasser bittet darum, ihm Ergebnisse und Erfahrungen mitzuteilen.

# Fragebogen zur Beurteilung von Fortbildungsveranstaltungen

ZUERST EINIGE AUSSAGEN ZUM AUFBAU DER VERANSTALTUNG UND  
ZU DEM, WAS SIE INHIN (BISHER) "GEBRACHT" HAT :

	s t i m m t			
	nicht	ein wenig	ziemlich	völlig
1. Der Aufbau der Fortbildungsveranstaltung war sehr Übersichtlich gegliedert. ..... .....	1	2	3	4
2. Die Veranstaltung wurde oft mit etwas Humor aufge- lockert. ..... .....	1	2	3	4
3. Der Besuch der Veranstaltung hat sich für mich sehr gelohnt. ..... .....	1	2	3	4
4. Aufbau und Stil der Veranstaltung entsprach sehr gut meinen Vorstellungen von einem "sinnvollen" Lernen. ..... .....	1	2	3	4
5. Die Anforderungen in der Veranstaltung waren für mich angemessen (d.h. weder zu "hoch" noch zu "nie- drig"). ..... .....	1	2	3	4
6. Die Auswahl des Themas der Veranstaltung traf die wichtigsten Probleme der Praxis. ..... .....	1	2	3	4
7. Ich habe durch den Besuch der Veranstaltung fach- lich sehr viel gelernt. ..... .....	1	2	3	4
8. In der Veranstaltung wurden die wichtigsten Aspekte des Themas gezielt herausgearbeitet. ..... .....	1	2	3	4

	nicht	ein wenig	ziemlich	völlig
9. Ich glaube, daß ich nach dem Besuch dieser Veranstaltung viele praktische Probleme besser angehen kann. ..... .....	1	2	3	4
10. Wenn ich diese Veranstaltung im Gesamt betrachte, so war sie (bisher) eine der besten Fortbildungsveranstaltungen, die ich besucht habe. ..... .....	1	2	3	4

ALS NÄCHSTES EINIGE AUSSAGEN ZUM "ARBEITSKLIMA" :

11. Während der Veranstaltung habe ich mich in der Gruppe sehr wohl gefühlt. ..... .....	1	2	3	4
12. Auf Verständnisschwierigkeiten der Teilnehmer wurde immer klärend eingegangen. ..... .....	1	2	3	4
13. Durch die Gestaltung der Veranstaltung wurde ich zu einer aktiven Mitarbeit angeregt. ..... .....	1	2	3	4
14. Abweichende Meinungen konnten in der Veranstaltung frei und ohne Angst geäußert werden. ..... .....	1	2	3	4
15. Bei Entscheidungen über den Ablauf der Veranstaltung wurden die Vorstellungen der Teilnehmer intensiv berücksichtigt. ..... .....	1	2	3	4

ZUM ABSCHLUSS EINIGE AUSSAGEN ZU DEM MEHR "ÄUSSEREN"  
 RAHMEN" DER VERANSTALTUNG :

	s n i c h t	e i n	z i e m l i c h	v ö l l i g
16. Für diese Veranstaltung war die Anzahl der Teilnehmer angemessen (d.h. weder zuviele noch zuwenig). ..... .....	1	2	3	4
17. Durch den äußeren Rahmen der Veranstaltung gab es sehr gute Möglichkeiten zu geselligem Beisammensein. ..... .....	1	2	3	4
18. Die Raumverhältnisse (z.B. Sitzgelegenheiten, Sichtverhältnisse, Lautstärke....) förderten die effektive Arbeit. ..... .....	1	2	3	4
19. Neben der Veranstaltung gab es sehr gute Möglichkeiten zur Gestaltung der Freizeit. ..... .....	1	2	3	4
20. Die äußeren Bedingungen der Veranstaltung (z.B. Anfahrtswege, Unterkunft, Veranstaltungsort....) waren sehr gut. ..... .....	1	2	3	4

VIELEN DANK !

Auswertungsschema

Frage	s t i m m t								Summe b a <sub>1</sub> -a <sub>n</sub>	x = b/n
	nicht(1)	a <sub>1</sub> = Anzahl(f) fx1	ein wenig(2)	a <sub>2</sub> = Anzahl(f) fx2	ziemlich(3)	a <sub>3</sub> = Anzahl(f) fx3	völlig(4)	a <sub>4</sub> = Anzahl(f) fx4		
1	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
2										
3										
4										
5										
Ska- la I	Summe 1-5 →	(a)	Summe 1-5 →	(b)	Summe 1-5 →	(c)	Summe 1-5 →	(d)	(e)	(f)
6										
7										
8										
9										
10										
Ska- la II	Summe 6-10 →		Summe 6-10 →		Summe 6-10 →		Summe 6-10 →			
11										
12										
13										
14										
15										
Ska- la III	Summe 11-15 →		Summe 11-15 →		Summe 11-15 →		Summe 11-15 →			
16										
17										
18										
19										
20										
Ska- la IV	Summe 16-20 →		Summe 16-20 →		Summe 16-20 →		Summe 16-20 →			

Erläuterung siehe folgende Seite

### Erläuterungen zum Auswertungsschema

Der Fragebogen zur Beurteilung von Fortbildungsveranstaltungen kann auf zwei Ebenen ausgewertet werden. Die inhaltsreichsten Informationen zur Vorbereitung von Auswertungsgesprächen bietet die Betrachtung jedes Aspekts, jedes Items einzeln. Werden aber die Bereiche einer Veranstaltung zusammenfassend betrachtet oder ist ein Vergleich zwischen verschiedenen Veranstaltungen oder verschiedenen Zeitpunkten einer Veranstaltung (z.B. in der Mitte und am Ende) vorgesehen, sollten -wegen der höheren Zuverlässigkeit- die Skalenwerte herangezogen werden. Beide Auswertungen sind mithilfe des Schemas schnell möglich.

### Auswertung auf Itemebene

Die Auswertung von z.B. Item 1 geschieht in folgenden Schritten (vgl. im Schema (1) bis (11)):

- In Zelle (1) wird aufgelistet, wieviele Teilnehmer das Item 1 mit "stimmt nicht" beantwortet haben; entsprechend wird in den Zellen (3), (5) und (7) vorgegangen. Auf diese Weise erhält der Auswerter einen Eindruck von dem in Item 1 angesprochenen Veranstaltungsaspekt, der die unterschiedlichen Teilnehmerbeurteilungen festhält. Für ein Auswertungsgespräch ist dies wichtig, z.B. wenn eine Hälfte der Teilnehmer mit "stimmt nicht" und die andere mit "stimmt völlig" antworten. Diese Unterschiede gehen in den folgenden Schritten bei der Berechnung des Mittelwerts verloren.
- Die Antworthäufigkeiten pro Kategorie werden mit dem Code der Kategorie multipliziert und in den Zellen (2), (4), (6) und (8) eingetragen; hätten z.B. 7 Teilnehmer Item 1 mit "stimmt völlig" eingeschätzt, so müßte in Zelle (8) eine 28 (=7x4) stehen.
- In Zelle (9) wird die Anzahl n der Teilnehmer aufgeschrieben, die das Item 1 beantwortet haben.
- Die Werte der Zellen (2), (4), (6) und (8) werden addiert und in Zelle (10) eingetragen. Dieser Wert (10) wird dann durch die Anzahl der Teilnehmer, die Item 1 beantwortet haben (9), dividiert und in Zelle (11) festgehalten. Hier steht nun der Mittelwert der Antworten für Item 1, d.h. die durchschnittliche Teilnehmereinschätzung des mit Item 1 angesprochenen Veranstaltungsaspekts.

Die entsprechende Prozedur wird dann für die Items 2 bis 20 durchgeführt.

### Auswertung auf Skalenebene

Im Auswertungsschema bedeuten:

- Skala I : Gesamteinschätzung/Methodischer Aufbau, Frage 1-5
- Skala II : Lernertrag, Frage 6-10
- Skala III : Arbeitsklima, Frage 11-15
- Skala IV : äußerer Rahmen, Frage 16-20

Bei der Auswertung auf Skalenebene -z.B. zum Vergleich der Antworten in der Mitte und zum Abschluß einer Veranstaltung- kann wie folgt vorgegangen werden (vgl. die dickumrandeten Zeilen,

sowie die Zellen (a) bis (g) ):

- Zuerst werden alle Werte, die in Spalte (1) stehen aufaddiert, das Ergebnis wird in Zelle (a) eingetragen; entsprechend gilt für die Spalten (4), (6), (8) und die Zellen (b), (c), (d).
- Die Summe aller Antworten auf die Fragen 1 bis 5 wird gebildet (Anzahl der Antworten auf Frage 1 plus Anzahl der Antworten auf Frage 2 plus....); dazu werden die Werte der Spalte (9) addiert, die Summe wird in Zelle (e) festgehalten.
- Anschließend werden die Werte der Zellen (a), (b), (c) und (d) addiert, das Ergebnis wird in Zelle (f) aufgeschrieben.
- Als letzter Schritt wird der Wert der Zelle (f) durch den Wert der Zelle (e) dividiert, das Ergebnis wird in Zelle (g) eingetragen; damit erhält der Auswerter den Mittelwert der Teilnehmereinschätzung in Skala I. Dieser Wert variiert zwischen 1 und 4; eine 1 bedeutet, daß im Durchschnitt die Teilnehmer in der Skala I die Veranstaltung negativ einschätzen (eine 1 kann nur dann erreicht werden, wenn alle Teilnehmer in allen 5 Fragen mit "stimmt nicht" antworten); eine 4 bedeutet, daß im Durchschnitt die Teilnehmer die Veranstaltung in Skala I positiv einschätzen (eine 4 kann nur erreicht werden, wenn alle Teilnehmer alle 5 Fragen mit "stimmt völlig" beantworten). Der berechnete Skalenmittelwert wird aber kaum diese Extremwerte annehmen.

Für Skala II, III und IV wird entsprechend vorgegangen.