

Konfliktfähigkeit der Evaluation und die „Standards für Evaluationen“

Wolfgang Beywl

*Gott aber ruhte am 7. Tag, betrachtete sein Werk
mit Wohlgefallen und sprach: „Es ist gut!“
Da trat sein Erzengel zu ihm und fragte:
Welche Kriterien legst Du Deinem Urteil zu Grunde?
Hast Du die Geschöpfe, die Du geschaffen hast,
an der Festlegung dieser Kriterien beteiligt?
Welche Daten hast Du herangezogen,
um Deine Schlussfolgerung zu begründen?
Wie sind die Daten zustande gekommen?
Sind Deine Erhebungsinstrumente gültig und zuverlässig?
Gott aber ging zur Ruhe und verbrachte grübelnd eine unruhige Nacht.
Früh am morgen rief er seinen Erzengel zu sich und sprach:
„Fahr zur Hölle, Luzifer!“*

1 Einführung: Konflikt als Regelfall der Evaluationspraxis

An diese Geschichte, die Michael Patton (1997, S. 1) gelegentlich zu Beginn seiner Evaluationsfortbildungen erzählt, werde ich durch Hendrik Faßmanns Beitrag „Forschungspraktische Probleme der Evaluation“ (in diesem Heft) erinnert. Widerstand und Konflikt sind in der Evaluation stets präsent. Das Machwort „Füge Dich oder weiche!“ wird zwar nur in Ausnahmefällen gesprochen – sei es vom Auftraggeber¹ der Evaluation oder anderen Mächtigen. Dabei hängt es beständig als Damoklesschwert über den spannungsrei-

¹ Der Gebrauch des Genus erfolgt in diesem Artikel wie im „Handbuch der Evaluationsstandards“: Zur Bezeichnung der evaluierenden Personen wird ausschließlich die weibliche Form genutzt; die anderen Akteure erscheinen in der Regel in der männlichen Form.

chen Fäden, welche sich im Geviert zwischen Auftraggebern, Durchführungsorganisation, Zielgruppen und Evaluationsteam spannen.

In der sozialwissenschaftlichen Grundlagenforschung, die häufig von akademischen, grundgesetzlich in ihrer forscherschen Autonomie geschützten, beamteten Hochschullehrern geleitet wird, sind Konflikte mit den Praktikern die zu vermeidende und oft auch vermeidbare Ausnahme. Sie stellen gewissermaßen einen Grenzfall der Forschungspraxis dar, während sie für die Evaluation allgegenwärtig sind.

Hendrik Faßmann nennt in seinem Beitrag vielfach negative Konfliktwirkungen, welche die Nutzung der Evaluation – also deren Effektivität – sowie deren Durchführbarkeit – also insbesondere deren Effizienz – mindern und rechtliche Sanktionen oder ethische Beeinträchtigungen bei den Beteiligten und Betroffenen nach sich ziehen. Diese offensive Thematisierung von Konflikten begrüße ich sehr. Selten liest man klare Worte darüber, in welcher Schärfe Interessengegensätze im Abhängigkeitsgeflecht von Auftraggeber, Evaluatoreninnen, Durchführungsorganisationen sowie Zielgruppen auftreten, welche Dilemmata sie besonders für Evaluationsteams bedeuten können. Es ist an der Zeit, dieses m. E. zentrale Thema der Evaluationspraxis zum Gegenstand der professionellen Selbstvergewisserung und schließlich auch der sozialwissenschaftlichen Forschung über Evaluation zu machen.² Die „Standards für Evaluationen“ bieten hierfür einen geeigneten Referenzrahmen.

2 Auf dem Weg zu deutschsprachigen Standards für Evaluationen

1975 begann in den USA ein interdisziplinäres Komitee mit der Entwicklung von Standards für die Evaluation in Bildung und Erziehung, das *Joint Committee on Standards for Educational Evaluation*. Es veröffentlichte 1981 die *Standards for Evaluations of Educational Programs, Projects, and Materials* (im Folgenden kurz JC-Standards genannt). Sie wurden 1989 vom American National Standards Institute (ANSI) anerkannt. 1994 ist nach umfassender fachöffentlicher Überprüfung eine von nun 16 Organisationen³ getragene Neuauflage erschienen. Diese Publikation führt den Titel *Program Evaluation Standards* und verdeutlicht den Anspruch, für einen breiten Bereich von Humandienstleistungen zu gelten. Im Titel der 1999 erstmalig erschienenen deutschen Übersetzung *Handbuch der Evaluationsstandards (Joint Committee 2000)* ist der Bezug zu einem speziellen Anwendungsbereich aufgelöst.

² Auf dem Hintergrund sozialpsychologischer Theorien und eigener empirischer Erhebungen liefert Taut (2001) hierfür einen wichtigen Grundlagenbeitrag.

³ Darunter befinden sich z.B. die American Educational Research Association, die American Evaluation Association oder die American Psychological Association.

Nach mehrjährigen Vorarbeiten, die wesentlich durch die Forschungen von Thomas Widmer geprägt sind,⁴ verabschiedete die *Schweizerische Evaluationsgesellschaft* (SEVAL) auf ihrer Mitgliederversammlung am 4. Mai 2001 eigene Evaluationsstandards, die eine Adaption der JC-Standards für die Schweizerische Evaluation darstellen.⁵ Die Besonderheit der SEVAL-Standards besteht darin, dass sie Geltung für das breite Feld staatlicher Politik und auch für Evaluationsgegenstände in privaten Organisationen und Unternehmen beanspruchen. Lediglich die Evaluation von Personal ist – wie auch in den JC-Standards – ausgenommen. Bemerkenswert ist, dass die SEVAL-Standards bereits parallel zu ihrem Erarbeitungsprozess Eingang fanden in rechtliche Bestimmungen und insbesondere die Ausschreibungspraxis der Schweizer Bundes- und kantonalen Behörden. Die Evaluation als professionell gerahmte, wissenschaftliche Dienstleistung hat damit in der Schweiz an Geltung gewonnen.⁶ Die Deutsche Gesellschaft für Evaluation (DeGEval) hat sich bereits 1997 in ihrer Gründungssatzung die Entwicklung von Evaluationsstandards zum Ziel gesetzt. Aufbauend auf eine im Jahr 2000 per Internet durchgeführte Mitgliederbefragung hat die DeGEval auf der Mitgliederversammlung 2000 eine „Kommission Standards für Evaluationen“ beauftragt, ein an die Standards der SEVAL/des Joint Committee angelehntes Papier zu erarbeiten, welches als vorläufiges allgemeines Standard-Set für Evaluationen in der Mitgliederversammlung 2001 verabschiedet werden soll.⁷ Dieses allgemeine Standard-Set soll für die Mitglieder der DeGEval klare Anhaltspunkte für die Planung, Beauftragung, und Durchführung von Evaluationen abstecken, so dass für die Mitglieder und auch für Auftraggeber und Evaluatorinnen außerhalb der DeGEval ein evaluationsfachlicher Orientierungsrahmen sowie erhöhte Handlungssicherheit entstehen. Auch für die Aus- und Weiterbildung in Evaluation und die Evaluation von Evaluationen sollen die Standards eine Orientierung bieten, und dabei theoretische und methodische Innovationen fördern. Die DeGEval strebt an, dass diese Standards Relevanz über breite Evaluationsarten und Evaluationsfelder und damit eine allgemeine Bedeutung erlangen.⁸ Im folgenden Kapitel werden Grundstruktur und formaler Aufbau der JC/SEVAL-Standards erläutert. Der vorgestellten Gliederung nach vier Standard-Gruppen wird voraussichtlich auch die DeGEval folgen. Die potentielle konfliktpräventive Funktion der Standards für Evaluationen soll anschließend beispielhaft verdeutlicht werden.

⁴ Bahn brechend für die Frage der intersektoralen Übertragbarkeit der JC-Standards waren seine umfassenden Meta-Evaluationen, z. B. Widmer (1996).

⁵ Eine aktuelle Fassung ist unter folgender Internet-Adresse einsehbar: <http://www.seval.ch/deutsch/stad/stad1.htm>

⁶ So hat das schweizerische Bundesamt für Justiz kürzlich das Verzeichnis von Evaluationsklauseln in Erlassen des Bundesrechts aktualisiert. Das Verzeichnis ist zugänglich über: <http://www.bj.admin.ch/d/index.html>

⁷ Eine ausführliche Dokumentation des Entwicklungsprozesses von Standards findet sich auf der Homepage <http://www.degeval.de>

⁸ Beschluss der Mitgliederversammlung vom 23.11.2000.

3 Kurze Einführung in die Evaluationsstandards⁹

Die JC-Standards bestehen aus 30 Einzelstandards;¹⁰ die SEVAL hat deren Zahl bei einer Neufassung sowie Streichungen/Zusammenfassungen auf 27 reduziert. Die Einzelstandards sind in folgende vier Gruppen gegliedert:

- Nützlichkeitsstandards, die sicher stellen sollen, dass sich eine Evaluation an den Informationsbedürfnissen der vorgesehenen Evaluationsnutzer ausrichtet, so dass diese die bereit gestellten Informationen tatsächlich nutzen.
- Durchführbarkeitsstandards, die gewährleisten sollen, dass eine Evaluation praxisnah, zielorientiert und realistisch geplant sowie diplomatisch und kostenbewusst ausgeführt wird.
- Korrektheitsstandards, die auslösen sollen, dass eine Evaluation rechtlich und ethisch korrekt durchgeführt wird und dem Wohlergehen der in die Evaluation einbezogenen und auch der durch die Ergebnisse betroffenen Personen gebührende Aufmerksamkeit widmet.
- Genauigkeitsstandards, die darauf zielen, dass eine Evaluation über die Güte und/oder die Verwendbarkeit des evaluierten Programms fachlich angemessene Informationen auf der Basis gültiger und zuverlässiger Daten hervorbringt und vermittelt.

Die Einzelstandards werden auf je ca. fünf Druckseiten in einheitlicher Form präsentiert (s. die Tabelle auf S. 155 f.):

⁹ Die übersetzte Kurzfassung der Joint-Committee-Standards ist einsehbar unter http://www.univation.org/eval_q1.html.

¹⁰ Das Buch umfasst ca. 230 Druckseiten.

Tabelle: Format der Standards mit Beispiel

Format der Standards	Beispiele
<p>„Kürzel/Standard-Name Standard-Formulierung“ in Form eines Sollens-Satzes</p>	<p>K2 Formale Vereinbarungen Die Pflichten der Vertragsparteien einer Evaluati- on (was, wie, von wem, wann getan werden soll) sollten schriftlich festgehalten werden, damit die Parteien verpflichtet sind, alle Bedingungen dieser Vereinbarung zu erfüllen oder aber diese erneut zum Gegenstand von formalen Verhandlungen zu machen.</p>
<p>„Überblick“: Eine Begriffsklärung/Einführung</p>	<p>„Unter einer schriftlichen Vereinbarung ... versteht man die gegenseitige Verständigung über die fest- gelegten Erwartungen und die Verantwortlich- keiten des Auftraggebers und der Evaluatorin. ...Die von ... Behörden erlassenen Richtlinien für externe Evaluationen verlangen häufig ... einen formellen Vertrag. Zu den Vertragsinhalten sollten zählen.“ „Aber auch wenn ein formaler Vertrag nicht ge- fordert wird – wie es in der Regel bei internen Evaluationen der Fall ist – sollten die Evaluations- parteien zumindest eine kurze Einigungsnote aus- arbeiten.“ ...</p>
<p>Richtlinien: Vorschläge, wie der Standard erfüllt werden kann</p>	<p>...„Im Verlauf der Arbeit sollten Änderungen der Vereinbarung ausgehandelt und dokumentiert wer- den, wenn durch sich ändernde Umstände Abwei- chungen beim Arbeitsumfang, bei den Kosten oder beim Zeitplan notwendig oder wünschenswert werden.“„Es sollte – in vertretbarem Rahmen – festgelegt werden, in welchen Fällen man von einer Nichteinhaltung der Vereinbarung durch eine der beiden Parteien sprechen kann und welche Konsequenzen damit verbunden sind.“...</p>

<p>Häufige Fehler: Liste typischerweise auftauchender Schwierigkeiten</p>	<p>... „In der schriftlichen Vereinbarung fehlen wichtige Vertragsbestandteile.“ ... „Es wird versäumt, vor Unterzeichnung des Vertrages mit denjenigen zu sprechen, die von der Evaluation unmittelbar betroffen sind, selber aber nicht Vertragspartner sind ...“ ...</p>
<p>Anschauungsbeispiel: ein oder zwei Beispiele dazu, wie der Standard angewendet werden kann</p>	<p>(1) Beschreibung: „Der Weiterbildungsleiter eines Unternehmens beauftragte eine Unternehmensberaterin damit, die Inhalte des firmeneigenen Führungskräfte Trainings zu evaluieren. Beide unterzeichneten eine „Einigungsnote“, die dem Evaluationsplan beigelegt war. Während die Umfrage im Gange war, wurde der Weiterbildungsleiter versetzt. Sein Nachfolger traf zwei Monate später ein. Der Weiterbildungsleiter akzeptierte die Rechnung, später jedoch weigerte er sich, die Kosten für die Datenverarbeitung anzuerkennen. Weder die vorhandene „Übereinkunftserklärung“ noch irgendein Mitarbeiter konnte die mündliche Vereinbarung zwischen dem früheren Leiter und der Unternehmensberaterin zur Durchführung der Umfrage bestätigen ...“</p> <p>(2) Analyse: „Dieser Fall macht deutlich, dass Verträge aktualisiert werden müssen, wenn vom Evaluationsplan abgewichen werden soll. Rückblickend kann man sagen, dass die geeignete Vorgehensweise gewesen wäre, die schriftliche Vereinbarung abzuändern.“</p>

4 Die Standards als Mittel der Konfliktprävention

Die vier Standard-Gruppen Nützlichkeit, Durchführbarkeit, Korrektheit und Genauigkeit haben in Bezug auf das Thema „Konflikte in der Evaluation“ eine unterschiedliche Bedeutung.

4.1 Genauigkeit als unverzichtbare Basis¹¹

Nicht selten entzünden sich Konflikte - besonders solche in der Berichts- und Vermittlungsphase von Evaluationen – an Aspekten, welche durch die *Genauigkeitsstandards* angesprochen sind:

- „Ist die realisierte Stichprobe wirklich repräsentativ? Was wären die Ergebnisse für diejenigen, die nicht geantwortet haben? Könnte dies die Tendenz der Schlussfolgerungen nicht in ihr Gegenteil verkehren? Hat es sich angesichts dieser unsicheren Ergebnisse überhaupt gelohnt, eine Evaluation zu beauftragen?“ Auf diese Einwände, die einen berechtigten Kern haben, bin ich nach einigen bitteren Erfahrungen gut vorbereitet: Ich vereinbare z. B. bei Befragungen im Vorhinein mit den relevanten Beteiligten, ab welcher Rücklaufquote ein Ergebnis ausreichende Sicherheit bietet, Schlussfolgerungen zu ziehen und Entscheidungen zum Evaluationsgegenstand zu treffen, und wie bei Nicht-Erreichen dieser Quote zu verfahren ist.
- „Die zentralen Prozesse heißen in unserem Qualitätsmanagement nicht ‚Schlüsselprozesse‘ sondern ‚Kernprozesse‘. Wir haben nicht 250 Mitarbeiter, sondern 238“. „Herr Guthorst wird in der Mitte mit zwei ‚t‘ geschrieben und ist nicht Qualitätsbeauftragter, sondern Qualitätsleiter, und dies bereits seit über 3 Jahren ...“Es ist peinlich für eine Evaluatorin, wenn sie sich in einer Abschlussbesprechung solche Kommentare anhören muss. Es ist nicht so einfach, dem vorzubeugen, zumal Bezeichnungen, Zahlen bis hin zu Schreibweisen in den wenigsten Organisationen durchgängig eindeutig genutzt werden. Dennoch: Das Streben nach Genauigkeit sollte die Evaluation auszeichnen, von der Nutzung angemessener statistischer Maßzahlen bis hin zur korrekten Orthographie. Beteiligte und Betroffene mit ihrer Feldkompetenz sollten als Gegenleser von Fragebogen-Entwürfen und von Berichten genutzt werden, um Fehler zu vermeiden. Die Grenzen der Genauigkeit, sei es z. B. wegen akuter Störungen in der Feldphase oder knapper Ressourcen für die Stichprobensicherung, sollen in angemessener Weise transparent gemacht werden.

Genauigkeit von Evaluationen – wie in der vierten Standardgruppe gefordert – ist eine notwendige, dabei nicht hinreichende Voraussetzung für die Konfliktprävention. Genauigkeitsmängel, auch vorgebliche, werden nicht selten als „Material“ für die Konfliktthematik genutzt, werden auch einmal vorgeschoben, um die Evaluatorin oder ihren Auftraggeber zu denunzieren, die Evaluationsergebnisse zu diskreditieren und Eva-

¹¹ Die Genauigkeitsstandards: G1 Dokumentation des Evaluationsgegenstandes, G2 Kontextanalyse, G3 Beschreibung von Zielen und Vorgehen, G4 Verlässliche Informationsquellen, G5 Valide Informationen, G6 Reliable Informationen, G7 Systematische Informationsüberprüfung, G8 Analyse quantitativer Informationen, G9 Analyse qualitativer Informationen, G10 Begründete Schlussfolgerungen, G11 Unparteiische Berichterstattung, G12 Meta-Evaluation

luation damit wertlos zu machen. Auch die Forderung nach übertriebener Genauigkeit kann zu Konflikten führen und eine Evaluation zum Stillstand bringen.

Die sorgsame Beachtung der Genauigkeitsstandards kann dem Evaluationsteam und seinen Auftraggebern den erforderlichen Rückhalt geben, mit Konfliktpotentialen angemessen umzugehen, die letztlich aus Macht- und Interessenkonstellationen – in der Organisation oder im Politikfeld – stammen, auf deren Regulierung die Evaluatorin kaum Einfluss hat.

4.2 Durchführbarkeit als Handlungselixier¹²

Wenn Evaluationen unsensibel sind für die Erfordernisse der täglichen Arbeit, wenn sie den Leistungsprozess der Organisation und ihrer Mitarbeiter z. B. durch umfangreichste Dokumentationspflichten oder Störungen häufig anwesender Evaluatorinnen über Gebühr belasten oder wenn sie sich über die einzuhaltenden „Dienstwege“ hinwegsetzen, mit dem Design „politisch“ hoch brisante Aspekte der Behörde, des Unternehmens oder des Programms berühren, dann wurde offensichtlich gegen die Durchführbarkeitsstandards verstoßen.

Der Widerstand gegen eine Evaluation kann sowohl top-down als auch bottom-up wachsen: letzteres, wenn z. B. die Mitarbeiter mit Instrumenten konfrontiert sind, die in keiner Weise aufnehmen, was zu ihrem beruflichen Grundwissen gehört. Andererseits kann Widerstand ganz oben, auf der Leitungsebene einer Organisation aufgebaut werden, wenn z. B. die „Kleiderordnung“ bei der Abstimmung von Vorgehensweisen der Evaluation nicht eingehalten wurde, (scheinbar) eine bestimmte opponierende Gruppe durch bestimmte Fragestellungen (einseitig) bedient wird.

Gute Kenntnis der Evaluatorin über die jeweilige Organisationskultur und freier Zugang zu „Eingeborenen“ dieser Kultur sind eine wichtige Voraussetzung, um frühzeitig Durchführungshindernisse zu erkennen und Evaluationsaktivitäten so einzuführen und umzusetzen, dass unproduktive Konflikte vermieden werden.

Ich trete in aller Regel als „Fremder“ in ein Evaluationsfeld ein. Dies hat Vorteile, da ich von Vorurteilen, Projektionen und blinden Flecken frei schauen kann. Besonders bei formativer Evaluation, die von den Mitgliedern der Durchführungsorganisation fordert, dass sie evaluative Informationen bald nutzen, muss ich mich auf die Codes und Muster des Klientensystems „einschwingen“ („align to the client!“). Hierfür ist eine Mindestzahl von kontaktreichen Besuchen und Gesprächen erforderlich. Auch können mir Mitglieder ähnlicher Systeme (eine andere Kommune, ein anderer Wohlfahrtsverband) helfen, die Durchführungsorganisation zu verstehen. Effektiv sein kann eine Evaluation dann, wenn Datenerhebungsinstrumente, Partizipationsformen, Rückmeldungen und Berichte genügend an die Adressatengruppen und ihre Organisationskulturen adap-

¹² Die Durchführbarkeitsstandards: D1 Praktische Verfahren, D2 Politische Tragfähigkeit, D3 Kostenwirksamkeit

tiert sind. Und wenn sie nicht überangepasst im Sinne einer Verschmelzung sind, so dass Evaluation keinen Unterschied mehr machen kann. Effizient wird die Evaluation dann sein können, wenn Kontakt schaffende Vorbereitungs- und Begleitmaßnahmen eng verknüpft sind mit Zentralaufgaben der Evaluation wie z. B. der Datenerhebung, und wenn sie in einem dem Evaluationsbudget angemessenen Maße erfolgen.

Jede Evaluation, besonders aber auch jede Folgeevaluation, ist darauf angewiesen, dass relevante Beteiligte und Betroffene zu Akteuren der Evaluation werden, eine Selbstverpflichtung eingehen, die Evaluationsaktivitäten aufmerksam zu verfolgen, die Evaluatortin dabei zu fordern, dass sie die Evaluation auf Kurs hält um schließlich die Evaluation als Prozess und dessen Ergebnisse zielgeführt zu nutzen.

4.3 Rechtliche und ethische Korrektheit als Richtschnur¹³

Evaluationen verlaufen – zumal in einer so hochgradig verrechtlichten Gesellschaft wie der deutschen – unter juristischen Rahmenbedingungen. Oft gehört es zu den ersten Aktivitäten einer Evaluatortin, sich die gesetzlichen Grundlagen des Evaluationsgegenstandes klar zu machen (z. B. Bundes-/Landes-/kommunale Zuständigkeit; subsidiäre Zuständigkeiten, Bestimmungen des Datenschutzes und Rechte der Zielgruppen, arbeitsrechtliche und mitbestimmungsrechtliche Situation der Mitarbeiterschaft, Haushalts- und Vertragsrecht usw.). Evaluationen müssen von Beginn an diese Grenzen berücksichtigen und einhalten, wollen sie nicht erhebliche Konflikte bis zu gerichtlichen Auseinandersetzungen riskieren oder in Kauf nehmen, dass ihre Schlussfolgerungen oder Empfehlungen an den rechtlichen Grenzen scheitern. Dies bedeutet nicht immer, das Maximum an Rechtssicherheit zu suchen. Nicht immer muss einer Mitarbeiterbefragung ein formelles Beteiligungsverfahren einer Mitarbeitervertretung voraus gehen. Wenn die zu erwartenden Ergebnisse jedoch z. B. Arbeitszeitregelungen oder Stellenbeschreibungen beeinflussen werden, ist diese Beteiligung unumgänglich. Bei anderen Themen mag es angeraten sein, die Mitwirkungsorgane der Beschäftigten einzuschalten, wenn es die Wahrscheinlichkeit erhöht, dass Evaluationsergebnisse genutzt werden. „Übertriebene“ Beteiligung einer unvollständig organisierten, wenig erfahrenen Mitarbeitervertretung kann aber auch zu mehr unnötigen Konflikten führen – Partizipation heißt nicht immer auch Konfliktreduktion!

Neben der rechtlichen gibt es eine Sphäre ethischer Fragen, die bei der Planung und Durchführung von Evaluationen zu beachten sind: In welchem Maße sind z. B. provokative, bedrängende oder auch intime Fragen, von denen man sich vielleicht starke Aufklärungskraft in Bezug auf die Fragestellungen der Evaluation erwartet, den zu Be-

¹³ Die Korrektheitsstandards: K1 Unterstützung der Dienstleistungsorientierung, K2 Formale Vereinbarungen, K3 Schutz individueller Menschenrechte, K4 Human gestaltete Interaktion, K5 Vollständige und faire Einschätzung, K6 Offenlegung der Ergebnisse, K7 Deklaration von Interessenkonflikten, K8 Finanzielle Verantwortlichkeit

fragenden zumutbar? Ist es zu rechtfertigen, ihnen durch bestimmte Fragen Hoffnungen zu machen, die mit hoher Wahrscheinlichkeit nicht einlösbar sind?

Ein Beispiel der harmloseren Sorte ist die Standardfrage nach Weiterbildungswünschen in Mitarbeiterbefragungen. In aller Regel wird nur ein Bruchteil der geäußerten Wünsche bei gegebenen Budgets auch befriedigbar sein. Wieder einmal nicht genutzte Evaluationsergebnisse werden sowohl die Glaubwürdigkeit der Leitung als auch die von Evaluation negativ beeinflussen. Realistischer ist es vielleicht, nach dem aus Sicht des Antwortenden wichtigsten Weiterbildungserfordernis für das Team zu fragen – es kommen wenige Vorschläge heraus; sie sind reflektiert auf die Arbeitsbedingungen; es wird sich ein größerer Anteil befriedigen lassen. Verantwortliches Leitungspersonal weiß nach meiner Erfahrung gut um die Untiefen solcher Fragen und ich nutze seine Kompetenz gerne, um zu einem „politisch tragfähigen“ Evaluationsdesigns zu kommen.

Ethische Aspekte tauchen dann auf, wenn einflusssschwache/-lose Zielgruppen beobachtet oder befragt werden, z. B. Schüler oder Sozialhilfeempfänger. Hier ist auch das Erhebungssetting ein wichtiger Schauplatz, an dem sich Evaluationsethik manifestieren muss. So haben wir es z. B. vermieden, Sozialhilfeempfänger im Sozialamt zu befragen und neutrale Orte, wie z. B. Verbraucherberatungen, gewählt. Wertschätzung der befragten Person gegenüber, die sich ausdrückt in freundlicher Begrüßung, Ansprechen mit dem Namen, Bewirtung und ggf. auch einer angemessenen Aufwandsentschädigung hat vielfältige positive Wirkungen. Wenn es gelingt, ethische Verletzbarkeiten frühzeitig zu erkennen und zu berücksichtigen, kann vielfach das Entstehen von Konflikten verhindert werden.

Eine ethische Grundspannung war uns immer dann sehr bewusst, wenn wir uns ganz offensichtlich in einem interkulturell geprägten Umfeld bewegten. Dann z. B., wenn wir ein Projekt evaluierten, in dem die überwiegend türkischstämmigen Mitarbeiter eines Projektes im Rahmen einer bundesweit tätigen, überwiegend von deutschstämmigen Mitgliedern getragenen Organisation ein Curriculum zur frühkindlichen Entwicklung für türkische Migranten entwickelten und umsetzten. Aber auch zwischen Behörden, z. B. den für Jugendhilfe und den für Sozialhilfe zuständigen, gibt es solche kulturellen Spannungen mit vielfältigen Vorurteilen, Projektionen und Zuschreibungen, welche für gemeinsam zu tragende Leistungsprozesse eine große Belastung sein können. Mir hat es oft geholfen, mir klar zu machen, dass ich als Evaluator in aller Regel in eine fremde Kultur eintrete, wenn ich mit meiner Arbeit beginne. Dies erlaubt mir auch, interkulturelle Konflikte leichter zu erkennen.

4.4 Nützlichkeit als Ziel

Evaluation beweist sich an ihrer Nützlichkeit, im Sinne genutzter Erkenntnisse, Informationen und Schlussfolgerungen und deren Einwirkung auf das Handeln der Adressaten in ihrer Praxis. Der – in der Abfolge der Nützlichkeitsstandards letzte – Standard N7 „Wirkung der Evaluation“ stellt die zentrale, empirisch zu prüfende Erfolgsdimension für Evaluationen dar. Genannte mögliche Nutzungen sind „Verbesserung der Program-

me, Umstellung auf kostengünstigere Produkte bzw. Ansätze oder Beendigung von unnützen, unproduktiven Tätigkeiten. Des Weiteren sollten die Evaluatorinnen den Beteiligten und Betroffenen helfen, durch die Evaluation neue Perspektiven zum Programm zu entwickeln“ (vgl. Joint Committee 2000, S. 83). Die vorgelagerten Nützlichkeitsstandards können so gelesen werden, dass sie Hinweise geben für Erfolg – im Sinne von Nutzung – versprechendes Konfliktmanagement:

- (N1) Die an einer Evaluation beteiligten oder von ihr betroffenen Personen sollten identifiziert werden, damit deren Interessen und Bedürfnisse berücksichtigt werden können.
- (N2) Wer Evaluationen durchführt, sollte sowohl vertrauenswürdig als auch kompetent sein, damit bei den Evaluationsergebnissen ein Höchstmaß an Glaubwürdigkeit und Akzeptanz erreicht wird.
- (N3) Die gewonnenen Informationen sollten von einem Umfang und einer Auswahl sein, welche die Behandlung sachdienlicher Fragen zum Programm ermöglichen und gleichzeitig auf die Interessen und Bedürfnisse des Auftraggebers und anderer Beteiligter und Betroffener eingehen.
- (N4) Die Perspektiven, Verfahren und Gedankengänge, auf denen die Interpretationen der Ergebnisse beruhen, sollten sorgfältig beschrieben werden, damit die Grundlagen der Werturteile klar ersichtlich sind.
- (N5) Evaluationsberichte sollten das evaluierte Programm einschließlich seines Kontextes ebenso beschreiben wie die Ziele, die Verfahren und Befunde der Evaluation, damit die wesentlichen Informationen zur Verfügung stehen und leicht verstanden werden können.
- (N6) Wichtige Zwischenergebnisse als auch Schlussberichte sollten den vorgesehenen Nutzern so zur Kenntnis gebracht werden, dass diese sie rechtzeitig verwenden können.

Ein „starker“ persönlicher Kontakt¹⁴ der über alle Evaluationsphasen glaubwürdig, unabhängig und kompetent handelnden Evaluatorinnen mit den Beteiligten und Betroffenen ist genau so zentral für das Gelingen der Evaluation wie für die Bewältigung auftauchender Konflikte.¹⁵ Interessen, Anforderungen und Bedürfnisse aus dem Evaluationsfeld sollen wahrgenommen und im Untersuchungsdesign angemessen berücksichtigt werden. Dies gilt für die Abgrenzung der Evaluationsfragestellungen genau so wie für die

¹⁴ Zum Konzept „Beraten mit Kontakt“: vgl. Schmidt/Berg (1995); dort auf S. 158 ff. „Schichtenmodell zur Bearbeitung von Konflikten“.

¹⁵ Die Studie von Patrizi/McMullan (1999, S. 33) zu den Evaluationsaktivitäten von 21 Stiftungen resümiert: „The most highly regarded evaluators to be appreciated by foundations invest heavily in face-to-face conversation with foundation staff and grantees. They bring well-honored skills in facilitation and reflection into the work of program evaluation.“

Wortwahl in Befragungsinstrumenten oder die Interpretation der Daten, bis hin zur Gestaltung der Berichte, die rechtzeitig, auf die konkreten Handlungsanforderungen der Adressaten hin erstellt werden müssen.¹⁶ Dieser Kontakt erfordert einerseits Persönlichkeit und Erfahrung auf Seiten der Evaluatorin, die sicher nur partiell „erlernbar“ sind. Zum anderen sind auch trainierbare Schlüsselqualifikationen hilfreich, angefangen bei Moderationstechniken über Techniken des kontrollierten Dialoges, der lösungsorientierten Beratung bis hin zur Konfliktklärung oder Mediation, die im Rahmen von Weiterbildungen erworben werden können.

5 Ausblick: Konflikte als Motor von Evaluation und organisationalem Lernen?

„By definition, bringing together people with different points of view will involve conflict” (Nevis et al. 1996, S. 119).

Konflikte sind in der Evaluation nicht bloße Störgrößen, sondern können organisatorische Entwicklungs- und Lernprozesse fördern, können auch notwendig sein für Wandel in Organisation und Programm. Zur Rolle von Konflikt oder Widerstand in Veränderungsprozessen bezieht die Organisationsberatung/-entwicklung ausführlich Stellung (z. B. Glasl 1999; Schwarz 2000). An den Berater, der vor allem Prozessspezialist ist, werden vielfältige Anforderungen gestellt, das System und die darin Verantwortlichen in Bewegung zu setzen, hin zu den Zielen, welche die Beratenen setzen. Ich glaube, dass Ed Nevis (1988, S. 187) Recht hat, wenn er dem nach Effektivität seines Handelns trachtenden Berater empfiehlt, „ein Bad im Widerstand zu nehmen, bevor Energien für die Veränderung freigesetzt werden können“.

Evaluatorinnen müssen je nach Auftrag und Adressaten unterschiedliche Rollen adaptieren. Diese reichen von der wissenschaftlichen Gutachterin/Forscherin über die Auditorin bis hin zur Moderatorin oder Beraterin (vgl. Patton 1997, S. 127 ff.). Besonders die „klärend“ oder „interaktiv“ tätige Evaluatorin, die Programmverantwortliche in der Startphase und während der Entwicklung eines neuen Programms durch empirisch gewonnene Informationen unterstützt (vgl. Beywl 2001; Owen/Rogers 1999), hat ähnliche Aufgaben wie der gestalttheoretisch orientierte Berater.

Es gibt deutliche Unterschiede im Ressourceneinsatz zwischen Organisationsberatung und Evaluation: Während der Berater hauptsächlich sich selbst als Instrument einbringt, handelt die Evaluatorin meist im Team – ist also für Kollegen verantwortlich –, hat in EDV-Hard- und Software, eine Anlage für Telefoninterviews u. v. m. investiert, ist also materiell nicht so konfliktfähig wie der „freelancende“ Berater. Gleichzeitig muss

¹⁶ Ausführlich beschreibt Shadish (1995) Anforderungen an Evaluatorinnen.

die Evaluatorin einen Auftrag auch kündigen können, wenn dessen Ausführung nur im Widerspruch zu den professionellen Standards gerät. Dieses Dilemma gilt es zu managen: Es ist eine wichtige Aufgabe der Weiterentwicklung der Evaluationsdisziplin, Regeln, Reflexionsverfahren und Sicherungen zu schaffen, so dass Evaluationen immer wieder hohe Konfliktfähigkeit erreichen – zum Nutzen der Systeme, in denen sie tätig sind. Die Standards für Evaluationen sind hierfür ein wichtiger Bezugspunkt.

Literatur:

- Beywl, Wolfgang, 2001: Mit zielgeführter Evaluation wirksames Stiftungshandeln gestalten. In: Stiftung & Sponsoring, Nr. 3, S. 13-15.
- Glasl, Friedrich 1999: Konfliktmanagement. Ein Handbuch für Führungskräfte, Beraterinnen und Berater. Stuttgart.
- Joint Committee on Standards for Educational Evaluation; Sanders, James R. (Hrsg.), 2000: Handbuch der Evaluationsstandards: Die Standards des „Joint Committee on Standards for Educational Evaluation“. Opladen.
- Owen, John M.; Rogers, Patricia J.; 1999: Program evaluation: Forms and approaches. Thousand Oaks.
- Patton, Michael Quinn, 1997: Utilization-focused evaluation. The new century text. Thousand Oaks.
- Nevis, Edwin C., 1988, Organisationsberatung. Ein gestalttherapeutischer Ansatz. Köln.
- Nevis, Edwin C.; Lancourt, Joan; Vassallo, Helen G., 1996: Intentional Revolutions. A seven-point strategy for transforming organizations. San Francisco.
- Patrizi, Patricia; McMullan, Bernhard J., 1999: Realizing the potential of program evaluation. A new study looks at how 21 foundations ‚do‘ evaluation. In: Foundation news & commentary, May/June, S. 30-35.
- Shadish, William. R.; Newman, Dianna. L.; Scheirer, Mary Ann; Wye, Christopher (Hrsg.), 1995: Guiding principles for evaluators. New Directions for Program Evaluation, 66. San Francisco.
- Schmidt, Eva Renate; Berg, Hans Georg, 1995: Beraten mit Kontakt. Handbuch für Gemeinde- und Organisationsberatung. Offenbach.
- Schwarz, Gerhard 2000: Konfliktmanagement. Konflikte erkennen, analysieren, lösen. Wiesbaden.
- Taut, Sandy, 2001: Psychologische Erklärungsansätze für Widerstand gegen Programmevaluation. Köln (Diplomarbeit, Institut für Wirtschafts- und Sozialpsychologie).
- Widmer, Thomas, 1996: Meta-Evaluation. Bern.

Dr. Wolfgang Beywl
Univation e.V.
Institut für Evaluation und wissenschaftliche Weiterbildung
Zülpicher Straße 58
50674 Köln
Tel: +49.221.4248071
<http://www.univation.org>
e-Mail: wolfgang.beywl@univation.org

Wolfgang Beywl, Dr., leitet Auftragevaluationen u.a. für Ministerien, Non-Profit-Organisationen und Unternehmen, Vorstandmitglied der Deutschen Gesellschaft für Evaluation und Vorsitzender der Kommission „Standards für Evaluationen“.