

Evaluation politischer Bildung

Katrin Uhl, Susanne Ulrich, Florian M. Wenzel (Hrsg.)

Evaluation politischer Bildung

Ist Wirkung messbar?

Bibliografische Information der Deutschen Bibliothek

Die Deutsche Bibliothek verzeichnet diese Publikation in der Deutschen Nationalbibliografie; detaillierte bibliografische Daten sind im Internet unter <http://dnb.ddb.de> abrufbar.

© 2004 Verlag Bertelsmann Stiftung, Gütersloh

Verantwortlich: Michael Seberich, Katrin Uhl

Lektorat: Michael Kühlen

Herstellung: Christiane Raffel

Umschlaggestaltung: HTG Werbeagentur, Bielefeld

Umschlagabbildung: Zefa

Satz: digitron GmbH, Bielefeld

Druck: Hans Kock Buch- und Offsetdruck GmbH, Bielefeld

ISBN 3-89204-720-0

www.bertelsmann-stiftung.de/verlag

Inhalt

Vorwort	7
Einleitung: Evaluation und politische Bildung – was kann man »messen«?	9
<i>Katrin Uhl, Susanne Ulrich, Florian M. Wenzel</i>	
Teil A	
(Wirkungs-)Evaluation politischer Bildung: Betrachtungen	13
Zivilgesellschaft und politische Bildung – Zur Evaluation von Programmen und Projekten	15
<i>Heinz Lynen von Berg, Andreas Hirsland</i>	
Partizipative Evaluation	27
<i>Susanne Ulrich, Florian M. Wenzel</i>	
Teil B	
(Wirkungs-)Evaluation politischer Bildung in der Praxis	49
Evaluation von Demokratie- und Toleranzerziehungsprogrammen – die Erfahrungen aus dem Projekt »Erziehung zu Demokratie und Toleranz«	51
<i>Katrin Uhl</i>	
Evaluation des Programms »Achtung (+) Toleranz« – ein Praxisbericht	57
<i>Andreas Hirsland, Oliver Dimbath, Julia von Hayek, Werner Schneider</i>	

Praxisbericht: Evaluation des Programms »Eine Welt der Vielfalt« <i>Jochen Kehr</i>	83
»Miteinander – Erfahrungen mit Betzavta«: Evaluationsprozess und Ergebnisse <i>Michael Bommes, Jochen Guter, Ulrike Wolff-Jontofsohn</i>	97
Schule und Ausbildung für Toleranz und Demokratie: Ein Projekt der Europäischen Jugendbildungs- und Jugendbegegnungsstätte Weimar und seine wissenschaftliche Begleitung <i>Stephan Eschler, Erich Schäfer, Sandra Uhl</i>	121
Teil C (Wirkungs-)Evaluation politischer Bildung: Methodenreflexion und Empfehlungen	145
Die Evaluation von »Achtung (+) Toleranz« – eine Methodenreflexion <i>Andreas Hirseland, Oliver Dimbath, Julia von Hayek, Werner Schneider</i>	147
Datenerhebungsmethoden der Evaluation »Eine Welt der Vielfalt« <i>Jochen Kehr</i>	161
»Miteinander – Erfahrungen mit Betzavta« – Methode der Evaluation <i>Michael Bommes, Jochen Guter, Ulrike Wolff-Jontofsohn</i>	167
Selbstevaluation wertschätzend gestalten – methodisches Vorgehen in sechs Schritten <i>Florian M. Wenzel</i>	177
Die Autoren	197

Vorwort

Ist Wirkung messbar? Kann man den Erfolg von politischer Bildung ermitteln, und was genau ist erfolgreiche politische Bildung? Diese Fragen stehen zunehmend im Zentrum des Interesses, wenn es um Projekte der politischen Bildung geht.

Die Bertelsmann Stiftung hat sich diesen Fragen im Rahmen ihres Projekts »Erziehung zu Demokratie und Toleranz« bewusst gestellt und sich über drei Jahre intensiv mit dem Thema Evaluation von politischer Bildung beschäftigt. Seit 1995 werden in diesem Kooperationsprojekt der Bertelsmann Stiftung und des Centrums für angewandte Politikforschung der Universität München Bildungsprogramme zur Demokratie- und Toleranz-erziehung entwickelt bzw. für den Einsatz in Deutschland adaptiert und Multiplikatoren der schulischen und außerschulischen politischen Bildung für die Arbeit mit den Programmen ausgebildet.

Die beiden Partner haben mit dem Projekt den Wunsch verbunden, durch neue, kreative Ideen zur Vermittlung von Demokratie- und Toleranzkompetenz einen nachhaltigen Beitrag zur Förderung einer Kultur der Toleranz, der Gewaltlosigkeit und des demokratischen Zusammenlebens zu leisten. Die drei Programme »Achtung (+) Toleranz«, »Eine Welt der Vielfalt« und »Miteinander – Erfahrungen mit Betzavta«, die den Kern des Projekts ausmachen, wurden in den letzten Jahren umfangreichen Untersuchungen im Hinblick auf ihre Wirkung auf Einstellungen und Verhalten des Individuums unterzogen.

Die Erfahrungen, die im Lauf der Evaluationen gemacht wurden sowie die Ergebnisse und Empfehlungen, die aus ihnen resultierten, sind der Ausgangspunkt für diesen Band, mit dem wir einen Beitrag zur Diskussion um Wirkungsevaluation politischer Bildung und ihrer Methoden leisten wollen.

Der Band bildet gleichzeitig den Abschluss des Projekts »Erziehung zu Demokratie und Toleranz«. Ich möchte an dieser Stelle allen danken, die dieses Projekt im Lauf der Jahre begleitet haben. Mein besonderer Dank gilt den Projektmanagern Michael Seberich und Katrin Uhl auf Seiten der Bertels-

mann Stiftung sowie Susanne Ulrich und Florian M. Wenzel als Projektkoordinatoren des Centrums für angewandte Politikforschung. Nicht zuletzt möchte ich aber den Autorinnen und Autoren dieses Bandes danken. Ich hoffe, dass die hier versammelten Beiträge Praktikern der politischen Bildung als Anregung für den Umgang mit dem Thema Evaluation dienen.

Prof. Dr. Marga Pröhl

Mitglied der Geschäftsleitung der Bertelsmann Stiftung

Einleitung:

Evaluation und politische Bildung – was kann man »messen«?

Katrin Uhl, Susanne Ulrich, Florian M. Wenzel

Der Begriff der Evaluation ist schwer greifbar und unscharf definiert, er bezeichnet unterschiedlichste Vorhaben, die von Teilnehmer¹-Feedback bis zum Total Quality Management (TQM) reichen, das eine komplette Organisation erfassen und bewerten möchte. Gegenwärtig wird er fast inflationär gebraucht und scheint sich somit beinahe selbst zu rechtfertigen.

Auch in der Bildungsarbeit werden zunehmend methodisch abgesicherte Evaluationen eingefordert. Ihre Ergebnisse bilden immer häufiger die Entscheidungsgrundlage für die Weiterführung von Projekten. Aber nicht nur die Notwendigkeit zur Legitimierung von Projekten gegenüber (potenziellen) Geldgebern ist Ursache für den wachsenden Ruf nach Evaluationen in der politischen Bildung.

Im Kontext von Qualitätssicherung und Qualitätsentwicklung wächst auch das Interesse von Programmentwicklern, Trägern, Projektleitern und Trainern, mehr über die eigenen Bildungsprogramme zu erfahren. Nicht nur Rahmenbedingungen und Prozesse von Bildungsarbeit rücken hierbei ins Zentrum des Interesses, sondern auch die Ergebnisse politischer Bildung.

Kritisch diskutiert wird jedoch, inwiefern Evaluation – ein Begriff, mit dem häufig zunächst »messen« assoziiert wird – den Ansprüchen und Wirklichkeiten von Bildungsarbeit überhaupt gerecht werden kann. Sind Evaluation, das »Messen« von Ergebnissen, und ein Verständnis von politischer Bildung als sozialem Lernen vereinbar oder Gegensätze, die nicht zusammenzuführen sind?

Diese Frage stellt sich insbesondere, wenn es um die Evaluation von Bildungsprogrammen geht, die bei ihren Teilnehmern Einstellungs- und Verhaltensänderungen bewirken wollen, etwa im Sinne toleranten Verhaltens oder friedlicher Konfliktlösung. Und in der Tat ist es keine unberechtigte Frage, ob man solche komplexen Veränderungsprozesse überhaupt erfassen,

¹ Der Band verwendet zur besseren Lesbarkeit in der Regel die männliche Schreibweise. Frauen sind dabei stets mitgemeint.

ja gar »messen« kann. Die Tatsache, dass die Wirkung solcher Bildungsmaßnahmen nicht standardisierbar ist und eine einfache lineare Kausalität von Ursache und Wirkung hier nicht greift, befreit jedoch nicht von der Notwendigkeit, sich der Frage nach ihren Ergebnissen zumindest zu stellen.

Evaluation im Bildungsbereich und eine entsprechende, praxisnahe Evaluationsforschung sind in Deutschland noch relatives Neuland. Im deutschsprachigen Raum befinden sich universitäre Ausbildungen zum Evaluator derzeit im Aufbau. Die Universität Bern bietet einen Zuzatzstudiengang Evaluation an, und die Universität Saarbrücken wird ab 2004 als erste Universität in Europa einen Regelstudiengang Evaluation anbieten. Standards für Evaluation werden bisher hauptsächlich aus der US-amerikanischen Diskussion entnommen und entwickeln sich in Deutschland erst langsam (zu Adaption, Entwicklung und Stand deutscher Evaluationsstandards vgl. die Internetseite der Deutschen Gesellschaft für Evaluation: www.degeval.de).

Auch im Praxisbereich findet eine kritische Auseinandersetzung mit der Theorie der Evaluation selten statt. Trotz der großen Zahl von Seminaren, Trainings und Programmen, die auf Einstellungs- und Verhaltensänderungen abzielen, kommt eine systematische Analyse und Bewertung der Wirkungen solcher Maßnahmen im Alltag der politischen Bildung oft zu kurz.

Häufig existieren Vorbehalte und Widerstände gegenüber Evaluationen. Bewertungsdruck, die mögliche Gefahr des Verlusts von Fördermitteln und Stellen oder Angst vor Veränderung können Ursachen hierfür sein. Aber auch ganz pragmatische Gründe können ein negatives Verständnis von Evaluation als »notwendigem Übel« bewirken: Oft sind die erforderlichen personellen und finanziellen Ressourcen für umfassende Untersuchungen nicht vorhanden. Evaluation im Bereich politischer Bildung ist jedoch keineswegs von vornherein zum Scheitern verurteilt. Sie muss sich allerdings, wenn sie aussagekräftig sein will, auf den Kontext einlassen, den politische Bildung ihr bietet, und in ihrer methodischen Anlage den Anspruch reflektieren, den politische Bildung an sich selbst stellt.

Politische Bildung möchte Inhalte vermitteln und findet sich dabei innerhalb eines bestimmten gesellschaftlichen Wertekontextes. Doch gerade in partizipatorischen und erfahrungsorientierten Lernarrangements wird deutlich, dass diese Inhalte nur Angebote und Rahmen sein können, innerhalb derer die Lernenden ihre eigenen Erfahrungen reflektieren und ihre Perspektiven und Bedürfnisse in gegenseitigen Austausch bringen.

Lernen und Bildung finden durch Reflexion von Inhalten in einem sozialen Kontext statt und sind nicht vollständig planbar. Ins Zentrum rückt der Prozess des Lernens, alltagsnah und erfahrungsorientiert. Fertige Lösungsangebote spielen eine untergeordnete Rolle.

Unterschiedliche Dimensionen des Bildungsprozesses kommen hierbei zum Tragen: Zum einen geht es um die Vermittlung von Wissen und Kompetenz, zum anderen um die Reflexion des Vermittelten – und in der Folge auch um Einstellungs- und Verhaltensänderungen. Diese unterschiedlichen Dimensionen politischer Bildung haben Konsequenzen für die Evaluation: Während die erste Dimension sich weitgehend mit den bekannten quantitativen und qualitativen Methoden der Evaluation erfassen lässt, sind für die zweite andere Ansätze nötig, um Evaluation gegenstandsadäquat zu betreiben.

Mehrere Fragen stellen sich in diesem Zusammenhang: Wie ist Erfolg konkretisierbar? Lassen sich bei Programmen, die auf individueller Ebene Einstellungs- und Verhaltensänderungen bewirken wollen, vorab Zielkriterien bestimmen? Wie können nichtintendierte Nebenwirkungen berücksichtigt werden? Wer ist das Evaluationsobjekt in einem interaktiven Bildungsprozess von Lehrenden, Lernenden und Programm? Wie kann bei der Wirkungsbestimmung mit dem Vorhandensein externer Einflüsse umgegangen werden?

Mit diesen und weiteren Fragen beschäftigen sich die Beiträge in dieser Publikation. Unterschiedliche Wirkungsevaluationen von Programmen der politischen Bildung, ihre Designs und Ergebnisse sind in diesem Band gesammelt. Den Kern bilden dabei Erfahrungen, die mit der Evaluation von Programmen der Demokratie- und Toleranzerziehung im Rahmen des Projekts »Erziehung zu Demokratie und Toleranz« der Bertelsmann Stiftung und des Centrums für angewandte Politikforschung der Universität München gemacht wurden.

Im ersten Teil des Bandes werden Chancen und Schwierigkeiten der gegenstandsadäquaten Evaluation politischer Bildung und die Möglichkeit eines partizipativen Ansatzes diskutiert. Der zweite Teil vermittelt Erkenntnisse und Erfahrungen aus der Bildungs- und Evaluationspraxis. Vor dem Hintergrund dieser Praxiserfahrungen schließt der Band im dritten Teil mit einer Methodenreflexion.

Der vorliegende Band versucht eine Brücke zu schlagen zwischen wissenschaftlich-theoretischer Reflexion des Themas und der praktischen Umsetzung im Arbeitsalltag der politischen Bildung. Im Zentrum steht hierbei ein gewandeltes Verständnis von Evaluation, das diese nicht als aufkotroyierte Maßnahme begreift, sondern als partizipativen Lernprozess. Evaluation, so verstanden, ist nicht nur in der Tat vereinbar mit den Grundsätzen und Prinzipien einer Bildungsarbeit, die dem Individuum Erfahrungs- und Lernräume eröffnen möchte – sie wird vielmehr elementarer Bestandteil der Bildungsarbeit selbst.

In diesem Sinne hoffen wir, dass die hier versammelten Beiträge nicht nur zur Auseinandersetzung mit dem Thema anregen, sondern auch dazu beitragen, Evaluation in der politischen Bildung von einer Ausnahmeerscheinung zur Normalität werden zu lassen.

Teil A

**(Wirkungs-)Evaluation politischer Bildung:
Betrachtungen**

Zivilgesellschaft und politische Bildung – Zur Evaluation von Programmen und Projekten

Heinz Lynen von Berg, Andreas Hirsland

Zivilgesellschaft und politische Bildung

Als Gegenreaktion auf den Anstieg fremdenfeindlicher Gewalttaten und den offen zu Tage tretenden Rechtsextremismus ist in Deutschland seit Ende der 90er Jahre eine breit gestreute Förderung von Maßnahmen für mehr Demokratie und Weltoffenheit festzustellen (vgl. u. a. das von der Bertelsmann Stiftung initiierte Programm »Erziehung zu Demokratie und Toleranz«; Lynen von Berg 2003; Rieker 2002).

Diese Maßnahmen und Aktivitäten werden weitläufig unter dem Etikett des »zivilgesellschaftlichen Ansatzes« zusammengefasst, der einen Paradigmenwechsel in der Auseinandersetzung mit dem Rechtsextremismus kennzeichnet. An die Stelle von und ergänzend zu staatlichen Maßnahmen und Vorgaben soll im Zuge einer zunehmenden Entstaatlichung des gesellschaftlich-öffentlichen Zusammenlebens das (lokale) Engagement der Bürger treten.

Bei dem Konzept der »Zivilgesellschaft« als politischer Leitsemantik handelt es sich um ein umkämpftes und erst noch mit Inhalt und Praxis zu füllendes Konzept. So ist es nicht verwunderlich, dass sich auch der wissenschaftliche Diskurs zur Zivilgesellschaft heterogen präsentiert: Was Zivilgesellschaft (und die in ihrem Umfeld diskutierten politischen, ethischen und moralischen Tugenden und Kompetenzen) ist oder sein soll, erscheint unscharf und wird mit unterschiedlichen Zielsetzungen aufgeladen, die ebenso unterschiedliche, häufig vermischte normative und empirische Bezugsebenen aufweisen (vgl. u. a. Klein 2001; Kocka 2003).

Dennoch: Ein auch im Wissenschaftsdiskurs verbreiteter Minimalkonsens umgrenzt das zivilgesellschaftliche Konzept als ein auf den basalen demokratischen Austragungsformen beruhendes diskursives Verfahren zur Regelung öffentlicher Angelegenheiten, bei der sich Bürger einmischen und versuchen,

den demokratischen Prozess »von unten« mitzugestalten (vgl. Keupp 2001). Dieses Postulat einer (Mit-)Gestaltung der Gesellschaft »von unten«, durch den Bürger als Citoyen, trifft hingegen auf eine Reihe von Problemen, die für die Theorie und Praxis einer angestrebten »Verzivilgesellschaftlichung« der politischen Kultur bedeutsam sind.

Zunächst steht der Umstand, dass viele zivilgesellschaftliche Projekte in ihren Grundlagen nicht autonom, sondern selbst auf staatliche Initiativen zurückzuführen oder auf öffentliche Fördergelder angewiesen sind, in einem Gegensatz zu der Forderung einer vom Bürger ausgehenden Zivilgesellschaft. So versucht das CIVITAS-Programm zu gesellschaftlichem Engagement in den neuen Ländern zu ermutigen und die dazu nötigen Strukturen mit aufzubauen. Die für die Bundesrepublik entwickelten Konzeptionen zum zivilgesellschaftlichen bzw. bürgerschaftlichen Engagement beziehen sich vorwiegend auf die Situation der alten Bundesländer und gehen implizit von funktionierenden, von der Bevölkerung mental internalisierten demokratischen Institutionen aus.

Die dabei insbesondere in den neuen Bundesländern festgestellten Desintegrationserscheinungen (vgl. Heitmeyer 2003) gehen vielfach mit hoher System- und Politikverdrossenheit einher, die sich in einer deutlichen Distanz zu den Organen der repräsentativen Demokratie niederschlagen (vgl. Stöss 2000). Vor diesem Hintergrund kann dort der stabilisierende Charakter von Institutionen nicht im gleichen Maße wie in den westlichen Bundesländern vorausgesetzt werden, so dass für die Entwicklung des »zivilgesellschaftlichen Sektors« insbesondere in den neuen Bundesländern eine andere Ausgangssituation vorliegt: Zu beobachten ist hier ein Mangel an zivilgesellschaftlichen Strukturen sowie eine geringe Dichte an Initiativen (vgl. Janer 2001) und qualifizierten freien Trägern, beispielsweise in der Jugendarbeit und in der politischen Bildung.

Um dem zu begegnen, soll hier eine Stärkung zivilgesellschaftlicher Strukturen mit staatlichen Programmen und Mitteln erreicht (vgl. z. B. Tolantes Brandenburg 1998, CIVITAS-Leitlinien), d. h. zivilgesellschaftliche Strukturen sollen »von oben« implementiert und gefördert werden.

Zudem lässt sich aus dem Blickwinkel eines durch die Bewegungsforschung geprägten Analyserasters beobachten, dass es unter dem Vorzeichen zivilgesellschaftlichen Engagements zur Ausdifferenzierung einer sozialen Bewegung gekommen ist, in der sich neben engagierten Initiativen und Bewegungskeimen auch so genannte »Bewegungsunternehmer«¹ gebildet ha-

1 Der Begriff »Bewegungsunternehmer« wurde zuerst von den amerikanischen Wissenschaftlern John D. McCarthy und Mayer N. Zald (1977) gebraucht. In

ben. Diese sind in der Lage, im öffentlichen Diskurs Begriffe oder Themen zu besetzen und somit Deutungsmacht und Einfluss innerhalb und außerhalb des »Bewegungslagers« zu erlangen.

Auch am Beispiel der Bewegung gegen Rechtsextremismus zeigt sich, dass soziale Bewegungen allgemeine und ein breites soziales Spektrum ansprechende Begriffe – etwa Stigmaworte (Herrmans 1994) wie »rechtsextrem« – als moralisch-affektive Mobilisierungsressource benötigen. Mit Dramatisierung und Emotionalisierung lässt sich temporär ein über das Bewegungslager hinausreichender Adressatenkreis erreichen und öffentliche Aufmerksamkeit erzeugen.

Dabei bildet die relative Unbestimmtheit und Offenheit der Zielsetzungen des zivilgesellschaftlichen Ansatzes eine Plattform für die praktische Zusammenarbeit unterschiedlicher Initiativen, Institutionen und Verbände mit unterschiedlichen inhaltlichen Ansätzen und politischen Positionen. Es besteht allerdings das Risiko einer Engführung des Konzepts der Zivilgesellschaft, wenn etwa in der aktuellen politischen Auseinandersetzung das zivilgesellschaftliche Engagement auf die »Bekämpfung« des Rechtsextremismus kanalisiert wird.

Angesichts der hier – ebenso unvollständig wie überzeichnet – skizzierten Problemlage (Besetzung zivilgesellschaftlicher Themenfelder durch deutungsmächtige Bewegungsunternehmer, Ressourcenkonkurrenz und institutionelles Vakuum) scheint eine Besinnung auf die Grundlagen einer – in Teilen erst noch zu schaffenden – zivilgesellschaftlichen Kultur angebracht. Diese besteht in ihrem Kern aus einer Identifikation mit demokratischen Spielregeln eines an Partizipation und Solidarität orientierten Gemeinwesens.

Daher ist für die Zivilgesellschaft ein Set an Sollwerten zur Regulierung von gesellschaftlichen Konflikten und Angelegenheiten konstitutiv. So sind die Akzeptanz und Förderung von unterschiedlichen Meinungen sowie Pluralität auch von Lebensstilen grundlegend für eine Zivilgesellschaft. Das heißt: Toleranz, die Akzeptanz und der (welt-)offene Umgang mit Differenzen jeder Art, Ambiguitätstoleranz und die Fähigkeit mit Ambivalenzen und Widersprüchen selbstreflexiv umzugehen, sind die – für gesellschaftlich-staatliche Institutionen ebenso wie für den einzelnen Bürger als Citoyen – maßgeblichen Eigenschaften und Ziele einer Zivilgesellschaft, die sich auf universelle und uneingeschränkte Menschenrechte gründet.

Deutschland ist der Begriff im Rahmen der Diskussion zu den neuen sozialen Bewegungen eingeführt worden und mittlerweile etabliert (vgl. u. a. Rucht 1994; Geiling 1995). Als klassisches Beispiel für die Entwicklung zu einem »Bewegungsunternehmer« wird immer »Greenpeace« angeführt (vgl. u. a. Geiling 1995). Zur Bewegungsforschung und zu sozialen Bewegungen allgemein vgl. u. a. Raschke 1985.

Insofern ist die Weiterentwicklung von Zivilgesellschaft nicht nur eine Frage des (immer auch notwendigen) Engagements gegen unerwünschte soziale und gesellschaftliche Tendenzen, sondern eine Frage der Ausbildung jener politisch-sozialen Handlungskompetenzen, die zivilgesellschaftlichen Umgang erst ermöglichen. Dies ist nun eine genuine Aufgabe der politischen Bildung, die sich neuen Anforderungen gegenüberstellt und sich nicht allein auf die Vermittlung historisch-politisch-sozialkundlichen Wissens beschränken kann.

In diesem Kontext ist auch das von der Bertelsmann Stiftung initiierte Projekt »Erziehung zu Demokratie und Toleranz« zu sehen, dessen Intention sich auf die Bildung zivilgesellschaftlicher Handlungs- und Urteilskompetenzen richtet und damit – etwa mit Bezug auf das nach wie vor virulente Problem eines erstarkten (häufig latenten) Rechtsextremismus – einen präventiven (Bildungs-)Ansatz verfolgt.

Ein derartiges Verständnis des zivilgesellschaftlichen Ansatzes verweist auf Folgendes: Es bedarf der Entwicklung demokratischer Konfliktlösungsmuster, der Verbesserung ziviler Konfliktaustragung und der diskursiven Auseinandersetzung vor der Folie der empirischen Tatsache, dass all dies in der Alltagspraxis – jedenfalls in Teilen der Bevölkerung und der gesellschaftlichen Institutionen – noch nicht durchgängig gelebte Realität ist. Insbesondere in den neuen Bundesländern müssen die Verfahren demokratischer Aushandlungsprozesse und die dazu notwendigen Räume und Strukturen oft erst noch zum Bestandteil der politischen Alltagskultur werden.

Zivilgesellschaftliches Engagement darf sich dabei nicht in einer Fixierung auf einen »Gegner« wie den »Rechtsextremismus« beschränken, sondern muss positive Zielsetzungen und Werteorientierungen entwickeln und »vorleben«. An diesem Punkt setzen die in diesem Band evaluierten Programme zur Demokratie und Toleranzbildung – »Achtung (+) Toleranz«, »Eine Welt der Vielfalt« und »Miteinander – Erfahrungen mit Betzavta« – an und müssen sich an diesem Anspruch messen lassen.

Evaluation zivilgesellschaftlicher Programme und Maßnahmen

Praxisbeobachtungen weisen darauf hin, dass die Konkurrenz innerhalb des zivilgesellschaftlichen Sektors um knappe Ressourcen wie Engagement, Aufmerksamkeit und Finanzmittel einerseits und die vielfältigen von Öffentlichkeit und Förderern an die Träger zivilgesellschaftlichen Engagements herangetragenen heterogenen Erwartungen andererseits das Feld einer Evaluation zivilgesellschaftlicher Programme und Maßnahmen zu einem um-

strittenen Terrain machen, in dem häufig unterschiedliche Interessenlagen, aber auch diffuse Besorgnisse zum Tragen kommen. Dies ist umso verständlicher, als von Evaluationen vielfach die Fortsetzung der eigenen Arbeit und der Erhalt von Arbeitsplätzen abhängig scheint oder auch ist, während auf Seiten etwa der Förderer ein angesichts knapper Mittel zunehmender Legitimationsdruck bezüglich einer »sinnvollen« Verwendung der vergebenen Mittel besteht.

Dabei wird »sinnvoll« meist mehr oder weniger ausdrücklich mit »wirksam« gleichgesetzt; mit Blick auf die vorherrschende Orientierung am Kriterium »Problembewältigung« – »des Rechtsextremismus« – werden eindeutige Erfolge gefordert. Das dort zu Grunde liegende Verständnis von Evaluation ist meist ein summatives und bilanzierendes. Der Umgang mit den Evaluationsergebnissen birgt – pointiert ausgedrückt – die Gefahr in sich, diese nach Art von Schulzensuren zu handhaben.

Unter derartigen Voraussetzungen wird von Evaluation oft ein eindeutiger Wirkungs- und Effizienznachweis der geförderten Maßnahmen verlangt, der wissenschaftlich haltbar meist nicht zu erbringen ist. Auch wird den Evaluatoren vielfach die Rolle letztinstanzlicher Wahrheitsfinder zugeschrieben – eine Erwartung, die sich weder mit den methodologischen Einsichten der Sozialwissenschaften deckt noch den neueren Debatten um die Rolle des Evaluators in der Steuerung von Prozessen der politischen Bildung angemessen ist (vgl. Lynen von Berg und Pallocks 2004).

In der gegenwärtigen wissenschaftlichen Diskussion zur Evaluation von im weiteren Sinn zivilgesellschaftlichen Programmen und Maßnahmen besteht weitgehende Einigkeit darüber, dass sich die Kausalwirkungen solcher »weichen« Programme nicht messen lassen (vgl. u.a. Hirsland und Dimbath 2003; Scherr 2003; Stöss 2003). Die Wirkungen von politischen Bildungsprogrammen, die wie die in diesem Band evaluierten meist über die Förderung oder Entfaltung individueller Handlungs- und Urteilskompetenzen auf eine Veränderung der politischen Kultur zielen, ließen sich nur dann eindeutig bestimmen, wenn es möglich wäre, eine Kausalität zwischen einer Intervention und möglichen Veränderungen als Folge dieser Intervention nachzuweisen (vgl. Stöss 2003). Das jedoch hätte zur Voraussetzung, alle neben der eigentlichen Intervention auf das Feld einwirkenden externen Faktoren zunächst isolieren zu können, sodann die im Feld auftretenden Effekte (Wirkungen) identifizieren und diese schließlich messen, d.h. entlang einer normierten Dimension quantitativ repräsentieren zu können.

Dies wäre – wenn überhaupt – nur experimentell möglich und ist daher für die Evaluierung der hier in Frage stehenden Programme aus ethischen, aber auch aus methodologischen und forschungspraktischen Gründen nicht

praktikabel, wie vergleichbare Erfahrungen aus experimentellen Evaluierungsdesigns beispielsweise zur Kriminalprävention zeigen (vgl. Schumann 2001). Bereits die Festlegung adäquater Messzeitpunkte stellt ein fast unüberwindliches Hindernis dar, da sich die Wirkung einer Intervention oder eines Interventionsbündels auch erst nach mehreren Jahren und durch das Zusammenwirken mit anderen Faktoren einstellen kann.

Ebenso kann nicht von Konzepten, Programm- und Projektzielen auf später erfolgende Handlungen geschlossen werden; umgekehrt lassen sich diese Handlungen nicht unmittelbar auf Konzepte, Programm- und Projektziele zurückführen. Handlungen sind immer das Resultat von Interaktionen und den diese Interaktionen beeinflussenden Kontextbedingungen (vgl. etwa Blumer 1973) und lassen sich somit immer nur am konkreten Fall – seinen Gelegenheitsstrukturen und deren situativen Bedingungsfaktoren – untersuchen. Die in diesem Band vorgestellten Praxisberichte und auf unterschiedlichen methodischen Pfaden gewonnenen Evaluationsergebnisse stützen diesen Befund.

Aus diesen Überlegungen heraus ergibt sich nach unserer Meinung, dass es für die Evaluation von Programmen und Maßnahmen der politischen Bildung keinen wissenschaftlichen »Königsweg« bzw. nicht die eine Evaluationsmethode geben kann. Das liegt zunächst daran, dass es ebenso wenig das Kriterium für den Erfolg oder Misserfolg von Programmen und Maßnahmen gibt. Vielmehr stellen Kriterien selbst begründungsfähige und daher begründungspflichtige Wahrnehmungsraster dar, die festlegen, worauf zu achten ist und worauf nicht. Somit sind Kriterien keine unabhängig vom Beobachtungs- bzw. Erkenntnisinteresse von der Wirklichkeit vorgegebenen Entitäten, sondern selbst immer schon Bestandteil einer Bewertung (vgl. Hirseland und Uhl 2004).

Vor diesem Hintergrund gewinnen Bestrebungen, einen »partizipativen Ansatz« der Evaluation (Ulrich und Wenzel 2003) zu entwickeln, an Bedeutung. Dieser zielt einerseits darauf, die unterschiedlichen Interessen und Perspektiven der von einem Programm oder einer Maßnahme (und ihrer Evaluation) betroffenen Stakeholder in den Evaluationsprozess zu integrieren. Andererseits sollen – etwa durch die Einbeziehung von Verfahren wie dem der »wertschätzenden Befragung« – nicht die (vermuteten) Schwächen und Probleme im Vordergrund stehen, sondern ebenso die Bewusstwerdung und Bewusstmachung vorhandener Stärken und Potenziale.

Dies wäre auch ein Beitrag zur Entwicklung realistischer, den (lokalen) Gegebenheiten angemessener Erwartungen und Beurteilungsmaßstäbe, die dann jedoch methodisch kontrolliert auf ihre Konsistenz und Gegenstandsangemessenheit hin reflektiert und empirisch geprüft werden müssten.

Erst wenn sich Evaluation der Feststellung, Dokumentation und kritischen Würdigung neuer Erfahrungen öffnet, kann sie auch zu neuen Einsichten führen und diese – formativ – für die prozessbegleitende Steuerung, Optimierung und Weiterentwicklung von Programmen und Maßnahmen fruchtbar machen. Dies wäre eine praktische Implikation der auch für den Bereich der Evaluation und Evaluationsforschung geführten »Konstruktivismusdebatte« (vgl. Guba und Lincoln 1989) und ihrer Argumentation für eine »naturalistische«, verstärkt auf qualitative Verfahren sowohl des Aushandelns von Zielen als auch der Datenerhebung und -interpretation setzende Evaluation.

Derartige Überlegungen verweisen – vor dem Hintergrund der eingangs skizzierten Kritik an eher »positivistisch« ausgerichteten Evaluationsverfahren – auf einen Zusammenhang zwischen der Ausgestaltung der institutionellen Rahmenbedingungen und den forschungsmethodischen Anforderungen an Evaluation. Angesichts der Heterogenität der Evaluationsgegenstände einerseits und des relativen Mangels an gesicherten Erfahrungs- und Vergleichswerten andererseits, ist es nicht möglich, bereits vorliegende Evaluationsmodelle und -designs einfach von einem Programm oder einer Maßnahme auf andere zu übertragen.

Aus den komplexen, hochgradig auf Interaktionsprozessen beruhenden Eigenheiten des zu evaluierenden Feldes ergibt sich, dass standardisierte, hypothesenprüfende Forschungsansätze die dort auftretenden Prozesse und Effekte (Wirkungen) allein schon deshalb nicht hinreichend zu erfassen vermögen, da diese vielfach unbekannt sind und erst »entdeckt« werden müssen. Unter methodischen Gesichtspunkten erscheinen daher für die Evaluation von Programmen, die auf die Entwicklung und Stärkung zivilgesellschaftlich relevanter Handlungs- und Urteilskompetenzen zielen, gegenstandsbezogene Forschungsansätze besonders geeignet. Hier bietet der von Glaser und Strauss (1979) entwickelte Ansatz der »Grounded Theory«² einen allgemeinen forschungspraktischen Rahmen, der darauf zielt, in methodisch kontrollierter Weise Prozesse der Daten- bzw. Materialanalyse und Hypothesengewinnung mit dem Ziel der Entwicklung einer »gegenstandsbezogenen Theorie« zu verknüpfen (Glaser und Strauss 1979).

Das ermöglicht ein »offenes« Evaluationsdesign, das durch eine hohe Anpassungsfähigkeit an die strukturellen und lokalen Gegebenheiten gekenn-

2 Die Grounded Theory geht als induktives und gegenstandsbezogenes Analyseverfahren nicht von einer Theorie aus, die bewiesen werden soll. Ausgangspunkt ist vielmehr ein Untersuchungsbereich – »was in diesem Bereich relevant ist, wird sich erst im Forschungsprozess herausstellen« (Strauss und Corbin zitiert nach Titischer u. a. 1998: 94; vgl. auch Mayring 1990: 78).

zeichnet ist und bei dessen Anwendung neben standardisierten oder teilstandardisierten Erhebungsverfahren vor allem auch auf erprobte Verfahren der qualitativen Sozialforschung (vgl. u.a. Hitzler und Honer 1994; Flick u.a. 1991) zurückgegriffen werden kann.

Anzustreben ist eine Form der Evaluation, die den »Schwerpunkt von tatsachenorientierten, vorgegebenen zu wertbezogenen, im Evaluationsprozess selbst zu entwickelnden Fragen« (Beywl 1988: 66) verlagert. Eine so ausgerichtete Evaluation ermöglicht eine nah an der Wirklichkeit des Projektalltags und den lokalen Bedingungen orientierte Einschätzung des Potenzials von Projekten und verhindert eine vorschnelle Beurteilung anhand praxisferner externer Kriterien, die oft z. B. von einem »pädagogischen Mythos« (Stake 1972: 96) oder durch die öffentliche Wahrnehmbarkeit der Projektarbeit (etwa aufgrund der Anzahl von Presseerwähnungen) beeinflusst sind.

Die Gegenstandsbezogenheit des Evaluationsansatzes zeigt sich u.a. gerade darin, bei der Beurteilung von Abläufen, Produkten und Entwicklungen die Normen und Werte der Beteiligengruppen, etwa von Projektmitarbeitern oder Kooperationspartnern der Projekte, als Daten mit einzubeziehen, da diese den immanenten Bezugspunkt der dem jeweiligen Arbeitshandeln zugrunde liegenden Absicht bilden. Diese sind durchaus einer kritischen Reflexion zugänglich, womit nun an die Stelle vorab deduzierter theoretischer Indikatorensysteme als »Messlatte« für vermeintlich »objektive« Beurteilungen ausweisbare und einer kritischen Diskussion zugängliche Werturteile als Entscheidungsgrundlage treten.

Ein derartiges Vorgehen steht dabei weniger in Widerspruch zu etablierten Evaluationsansätzen – sofern diese hinreichend komplex sind –, als dies auf den ersten Blick vermutet werden mag. So stellt schon das in den siebziger Jahren von Stufflebeam (1972; vgl. auch Grohmann 1997) zur Evaluierung von Schulprogrammen entwickelte CIPP-Modell das verallgemeinerbare Zusammenwirken verschiedener Rahmungen und Akteursebenen in Rechnung, denen verschiedene Erkenntnis- und Zugriffsebenen der Evaluation entsprechen.

Stufflebeam unterscheidet entsprechend zwischen Kontext-, Input-, Prozess- und Produktevaluation (CIPP):

- Dabei generiert die Kontextevaluation Aussagen zu den Bedarfslagen, Defiziten, unbefriedigten Bedürfnissen und Voraussetzungen im Vorfeld der Programmplanung und ermöglicht durch gezielten Einbezug der »Außenperspektive« von relevanten »stakeholdern« neben Aussagen zur Erreichung von Zielgruppen, zu Interaktionen und Resonanzen auch punktuelle Aussagen über die Einschätzung der Angemessenheit von Interventionsmaßnahmen.

- Die Inputevaluation sollte nach Möglichkeit vor Maßnahmebeginn durchgeführt werden. Sie betrachtet das Feld der zur Erreichung von Programmzielen relevanten Entscheidungen über den Einsatz und die Verteilung von Ressourcen, d. h. den dem jeweiligen Programmhandeln zugrunde liegenden Zusammenhang zwischen Zielen, Strategien und Plänen.
- Prozessevaluation bezieht sich unmittelbar auf die Projektpraxis und dort wiederum auf deren zeitlichen Ablauf. Das ermöglicht eine intensive Begleitung und formative Mitgestaltung des Projektverlaufes.
- Produktevaluation schließlich bezeichnet die summative Analyse und Bewertung der Ergebnisse bzw. Wirkung einer Maßnahme. Insbesondere bei fortlaufenden Maßnahmen muss »summativ« nicht gleichbedeutend mit »abschließend« sein. Vielmehr kann die Feststellung von Zwischenergebnissen, im Sinne einer Momentaufnahme zu einem gegebenen Zeitpunkt, durchaus wichtige Erkenntnisse und Einsichten für die Weiterentwicklung oder Umgestaltung einer Maßnahme liefern. Diese können eine Modifikation der Ziele (Was ist realistischerweise erwartbar?), der Art oder der Verteilung der eingesetzten Ressourcen und der institutionellen Abläufe beinhalten.

So verstanden öffnet das CIPP-Modell den Blick für systemische Zusammenhänge und dient so der Vermeidung voreiliger »pars pro toto«-Beurteilungen, die nur einen Einzelaspekt des Gesamtzusammenhangs zur Grundlage der Bewertung machen.

Der Schwerpunkt der in diesem Band vorgestellten Evaluationen liegt, legt man die eben ausgeführte Systematik zugrunde, eher auf dem Feld der Produktevaluation. Ihr Ziel war es, mit unterschiedlichen Vorgaben und auf unterschiedlichen methodischen Pfaden Programmwirkungen zu erfassen.

Ihre Gemeinsamkeit liegt darin, dabei eben nicht den Blick für das Ganze verloren zu haben, sondern die verschiedenen Zielsetzungen und empirischen Einsatzbedingungen der evaluierten Programme wenn nicht unmittelbar mit evaluiert, so doch reflexiv mit bedacht zu haben. Von daher ermöglichen sie durchaus auch einen Vergleich unterschiedlicher Ansätze zur Förderung von Zivilgesellschaft und ein wechselseitiges Lernen – auch wenn derartige Vergleiche, wie stets, ein wenig hinken mögen.

Was sich vor dem Hintergrund dieser Überlegungen anbietet, ist ein Verständnis von Evaluation und wissenschaftlicher Begleitforschung, das nicht von dem seriös kaum einlösbaren Desiderat der Herstellung einer nicht weiter hinterfragbaren Eindeutigkeit getragen ist. Wir plädieren damit nicht gegen Evaluation, sondern gerade für eine gegenstandsangemessene, nach den Standards der DeGEval »faire« Evaluation, die die Chancen einer Evaluation

politischer Bildungsmaßnahmen sowohl für die Gesellschaft als auch für die Träger von Programmen und die vor Ort tätigen Praktiker aufgegriffen.

Programme und Maßnahmen zur Förderung von Zivilgesellschaft stellen ein historisch neues und damit bislang einzigartiges politisches Bildungsprojekt bzw. -experiment dar, das es wert ist, flächendeckend dokumentiert, analysiert und hinsichtlich seiner möglichen Effekte auch wertgeschätzt zu werden. In diesem Sinne geht es darum, Wege zur weiteren Förderung zivilgesellschaftlicher Werte und Handlungsdispositionen zu finden, aus Erfahrungen und »Fehlern« zu lernen. Mit anderen Worten: Evaluation setzt zivilgesellschaftliche Umgangsformen auch im Umgang der von Evaluation Betroffenen untereinander voraus – also eine Evaluationskultur, zu der auch eine vielfach erst noch zu etablierende Kultur der Fehlertoleranz zählen muss.

Literatur

- Beywl, Wolfgang (1988): *Zur Weiterentwicklung der Evaluationsmethodologie*. Frankfurt am Main.
- Blumer, Herbert (1973): »Der methodologische Standort des symbolischen Interaktionismus«. In: Arbeitsgruppe Bielefelder Soziologien (Hrsg.): *Alltagswissen, Interaktion und gesellschaftliche Wirklichkeit*. Reinbek. 80–146.
- »CIVITAS – initiativ gegen Rechtsextremismus in den neuen Bundesländern. Maßnahmen zur Stärkung der demokratischen Kultur und zur Bekämpfung von Rechtsextremismus, Fremdenfeindlichkeit und Antisemitismus in den neuen Bundesländern« (= CIVITAS-Leitlinien 2002). www.civitas.st20.de/downloads/civitas2002.pdf.
- Flick, Uwe, und Ernst von Kardorff u. a. (Hrsg.) (1991): *Handbuch Qualitative Sozialforschung. Grundlagen, Konzepte, Methoden und Anwendungen*. München.
- Geiling, Heiko (1995): »Politische Kultur und neue soziale Bewegungen«. In: *agis-Info: Forschung zum gesellschaftlichen Strukturwandel*. Hannover.
- Glaser, Barney G., und Anselm L. Strauss (1979): Die Entdeckung gegenstandsbezogener Theorie. Eine Grundstrategie qualitativer Sozialforschung. In: Christel Hopf und Elmar Weingarten (Hrsg.): *Qualitative Sozialforschung*. Stuttgart. 91–111.
- Grohmann, Romano (1997): *Das Problem der Evaluation in der Sozialpädagogik*. Frankfurt am Main.
- Guba, Egon S., und Lincoln, Yvonna, S. (1989): *Fourth Generation Evaluation*. Newbury Park, Cal.
- Heitmeyer, Wilhelm (Hrsg.) (2003): *Deutsche Zustände*. Frankfurt am Main.

- Hermanns, Fritz (1994): »Schlüssel-, Schlag- und Fahnenwörter. Zur Begrifflichkeit und Theorie der lexikalischen Semantik«. In: *Arbeiten aus dem Sonderforschungsbereich 245: Sprache und Situation*. Bericht 81. Heidelberg und Mannheim.
- Hirseland, Andreas, und Oliver Dimbath (2003): »Lassen sich Wirkungen messen? – Zur Evaluation von Achtung (+) Toleranz.« Vortrag zur Tagung »Was können wir messen?« der Bertelsmann Stiftung und der Evangelischen Akademie zu Berlin. Unveröffentlichtes Manuskript. Berlin.
- Hirseland, Andreas, und Katrin Uhl (2004): »Evaluationsmethoden für Modelle zur Förderung von Zivilcourage in der politischen Bildung«. In: Gerd Meyer, Siegfried Frech, Günther Gugel und Ulrich Dovermann (Hrsg.): *Zivilcourage lernen – Analysen, Modelle, Arbeitshilfen*. Berlin. 384–399.
- Hitzler, Ronald und Anne Honer (1994): »Qualitative Methoden«. In: Jürgen Kriz, Dieter Nohlen und Rainer-Olaf Schulze (Hrsg.): *Lexikon der Politik. Bd. 2: Politikwissenschaftliche Methoden*. München. 389–395.
- Janer, Josefine (2001): »Freiheit klingt nach Ellenbogen. Nichts aus freien Stücken. Warum Amnesty International kaum Mitglieder in den östlichen Bundesländern hat«. In: *Frankfurter Allgemeine Zeitung*. 24.7.2001. 12.
- Keupp, Heiner (2001): »Zivilgesellschaftliches Engagement. Das Konzept gegen Extremismus«. In: *KJG* 1 2001. 3–12.
- Klein, Ansgar (2001): *Der Diskurs der Zivilgesellschaft. Politische Kontexte und demokratie-theoretische Bezüge der neueren Begriffsverwendung*. Opladen.
- Kocka, Jürgen (2003): »Zivilgesellschaft in historischer Perspektive«. In: *Forschungsjournal Neue soziale Bewegungen* 2 2003. 29–37.
- Lynen von Berg, Heinz (2003): »Paradigmenwechsel in der Auseinandersetzung mit Rechtsextremismus und Fremdenfeindlichkeit – der zivilgesellschaftliche Ansatz«. In: Hamburger Institut für Sozialforschung (Hrsg.): *Reader zum Symposium »Rechtsextremismusforschung. Aktuelle Tendenzen empirischer Untersuchungen (27.–29.08.03)«*. Hamburg. 137–148.
- Lynen von Berg, Heinz, und Kerstin Palloks (2004): »Die Evaluierung eines Programms zur Stärkung der Zivilgesellschaft. Teil I: Anlage, Konzeption und Vorgehensweisen der wissenschaftlichen Begleitforschung des CIVITAS-Programms. In: *Journal für Konflikt- und Gewaltforschung*« 1 2004 (i. E.).
- Lynen von Berg, Heinz, und Roland Roth (Hrsg.) (2003): *Maßnahmen und Programme gegen Rechtsextremismus wissenschaftlich begleitet. Aufgaben, Konzepte und Erfahrungen*. Opladen.
- Mayring, Philipp (1990): *Einführung in die qualitative Sozialforschung*. München.

- McCarthy, John D., und Mayer N. Zald (1973): *The Trend of Social Movements in America. Professionalization and Resource Mobilization*. Morristown, New Jersey.
- Ministerium für Bildung, Jugend und Sport: Koordinierungsstelle Tolerantes Brandenburg (Hrsg.) (1998): *Handlungskonzept der Landesregierung gegen Gewalt und Rechtsextremismus*. Potsdam.
- Raschke, Joachim (1985): *Soziale Bewegungen. Ein historisch-systematischer Grundriss*. Frankfurt am Main.
- Rieker, Peter (2002): »Aktionsprogramme gegen Rechtsextremismus und Fremdenfeindlichkeit. Überblick und Einschätzungen«. In: *Recht der Jugend und des Bildungswesens*.
- Rucht, Dieter (1994): *Modernisierung und neue soziale Bewegungen*. Frankfurt und New York.
- Scherr, Albert (2003): »Pädagogische Konzepte gegen Rechts. Was hat sich bewährt, was ist umstritten, was sollte vermieden werden?« In: Heinz Lynen von Berg und Roland Roth (Hrsg.): *Maßnahmen und Programme gegen Rechtsextremismus wissenschaftlich begleitet. Aufgaben, Konzepte und Erfahrungen*. Opladen. 249–264.
- Schumann, Karl F. (2001): »Experimente mit Kriminalprävention«. In: Günter Albrecht, Otto Backes und Wolfgang Kühnel (Hrsg.): *Gewaltkriminalität zwischen Mythos und Realität*. Frankfurt am Main. 435–456.
- Stake, Robert E. (1972): »Verschiedene Aspekte pädagogischer Evaluation«. In: Christian Wulf (Hrsg.): *Evaluation. Beschreibung und Bewertung von Unterricht, Curricula und Schulversuchen*. München. 92–112.
- Stöss, Richard (2000): *Rechtsextremismus im vereinten Deutschland*. 3. Auflage. Berlin.
- Stöss, Richard (2003): »Möglichkeiten und Grenzen der Evaluierung von persuasiven Programmen«. In: Heinz Lynen von Berg und Roland Roth (Hrsg.): *Maßnahmen und Programme gegen Rechtsextremismus wissenschaftlich begleitet. Aufgaben, Konzepte und Erfahrungen*. Opladen. 95–102.
- Stufflebeam, Daniel L. (1972): »Evaluation als Entscheidungshilfe«. In: Christoph Wulf (Hrsg.): *Evaluation. Beschreibung und Bewertung von Unterricht, Curricula und Schulversuchen*. München. 113–145.
- Titscher, Stefan, und Ruth Wodak u. a. (1998): *Methoden der Textanalyse. Leitfaden und Überblick*. Opladen. 320–363.
- Ulrich, Susanne, und Florian M. Wenzel (2003): *Partizipative Evaluation. Ein Konzept für die politische Bildung*. Gütersloh.

Partizipative Evaluation

Susanne Ulrich, Florian M. Wenzel

Von Evaluationen betroffene Menschen, Projekte und Organisationen stehen vor einer Herausforderung, die unbekannt ist und oftmals von außen in bedrohlicher Weise an sie herantritt. Evaluation wird als »Richter« wahrgenommen. Gerade im schwer messbaren Bereich des sozialen Lernens führen Evaluationen häufig zu Irritationen, wenn etwa Projektentwickler Zweifel haben, ob die Evaluationsfrage »Welche langfristigen Verhaltensänderungen zeigen Ihre Teilnehmer?« überhaupt so zu beantworten ist. Eine solche Frage konfrontiert Bildungsprozesse mit enormen Schwierigkeiten, wenn sie nicht in deren Selbstverständnis integriert ist. Widerstände und Konflikte wurden so in der Vergangenheit vielfach zum Merkmal von Evaluationen.

Das Verhältnis gerade von politischer Bildung zu Evaluationsfragen ist ungeklärt. Das Paradoxon einer Pädagogik, die sich an der Freiheit des Menschen und der Erziehung zu Autonomie und Verantwortung orientiert und diese durch geeignete Lernarrangements zielgerichtet erreichen möchte, verschärft sich, wenn die Frage von Evaluation auftaucht: Wie können Verantwortungsbewusstsein, Mündigkeit und Autonomie evaluiert werden? Konterkariert Evaluation politische Bildung und ihre pädagogischen Grundlagen durch stereotype Fragestellungen?

Häufig werden diese Fragen nicht theoretisch reflektiert. Pragmatische Arrangements herrschen vor. In der Folge werden die Ergebnisse von Evaluationen, ob positiv oder negativ, in ihren Konsequenzen nicht umfassend genutzt oder oftmals nur zur Befriedigung der Interessen der Auftraggeber verwendet. Wir möchten die »Herausforderung Evaluation« aufgreifen und erste Antworten für politische Bildner und ihr Umfeld anbieten. Aus der Erfahrung der dreijährigen Evaluation unseres Projektes zur Demokratie- und Toleranzerziehung¹ haben wir eine systematische Erfassung und Konzeptualisierung dieses Feldes vorgenommen.

¹ Informationen zu dem Projekt und seinen Programmen sind unter www.cap.uni-muenchen.de/bertelsmann/toleranz.htm zu finden. Als Downloads stehen die Eva-

Wir stellen hiermit die »Partizipative Evaluation« als ein Konzept vor, das für die Praxis politischer Bildung Orientierung und Einordnungsmöglichkeiten geben möchte. Es bietet weder eine fertige Lösung für konkrete Evaluationsvorhaben an, noch löst es die Dramatik von Evaluation in ihren Konsequenzen für die Arbeitswelt auf.

Es soll vielmehr eine neue Perspektive auf die Möglichkeiten und Chancen von Evaluation in der politischen Bildung eröffnen und insbesondere die Spannungen und Ängste, die mit Evaluation verbunden sind, in produktive Lern- und Reflexionsmöglichkeiten umwandeln. Das Konzept der partizipativen Evaluation ist aus einer systematischen Betrachtung der Praxis für die Praxis entstanden. Es soll Betroffene und Beteiligte in die Lage versetzen, informiert und selbstbewusst an Evaluationen teilnehmen zu können.

Was ist Evaluation?

Im gängigen Verständnis ist Evaluation eine ziel- und zweckgerichtete Handlung. Das übergeordnete Ziel ist die Erfassung von Daten zur Optimierung und Entscheidungsfindung in komplexen Handlungssituationen. Im Gegensatz zu allgemeiner Forschung will Evaluation keine generellen Aussagen treffen, sondern etwas über eine konkrete Praxis aussagen.

Es können drei Funktionen von Evaluation unterschieden werden, die besonders für den Bildungs- und Trainingsbereich relevant sind (vgl. Kinast 1998: 23 ff.):

- Erkenntnisfunktion: Es sollen wissenschaftlich fundierte Erkenntnisse über ein Evaluationsobjekt gewonnen werden.
- Legitimationsfunktion: Die Evaluation sollte im bestmöglichen Fall die Wirksamkeit von Interventionen bestätigen und deren Existenz rechtfertigen.
- Optimierungsfunktion: Es soll die Grundlage zur Steuerung der Verbesserung des Evaluationsobjektes geschaffen werden.

Häufig wird mit Hinblick auf die Verwertbarkeit der Ergebnisse der Blick auf die zweite und dritte Funktion gelenkt, auch um sich von wissenschaftlicher Forschung im Allgemeinen abzugrenzen. Im Vordergrund von Evaluation steht also eine konkrete Frage, die Konsequenzen für die Praxis haben wird.

An die Legitimations- und Optimierungsfunktionen knüpft die Unterscheidung von summativer und formativer Evaluation an:

luationsberichte der einzeln in ihrer Wirkung evaluierten Programme bereit.

- Summative Evaluation ist bilanzierend, zieht an einem bestimmten Punkt einen Schlusstrich, erhebt die Wirkungen des bisher Erreichten und kann so eine Legitimationsbasis bieten.
- Formative Evaluation ist projektbegleitend und gibt ihre Ergebnisse immer wieder in kleinen Schritten an ein Projekt zurück, um kontinuierlich Optimierung schon während des Projektverlaufs zu erreichen.

An das Evaluationsobjekt, d.h. an den Gegenstand, der untersucht werden soll, werden bestimmte Kriterien angelegt, die eine Bewertung ermöglichen. Diese Kriterien werden meistens aus den ursprünglichen Zielvorstellungen des jeweils untersuchten Projektes entwickelt. Durch geeignete Datenerhebungsmethoden soll dann die Ausgangsfrage der Evaluation beantwortet werden.

Das folgende Schema verdeutlicht dieses Grundverständnis von Evaluation und ist ein Beispiel einer summativen Evaluation, die abschließend bestimmte Wirkungen erheben möchte:

- Ausgangsfrage: Wie viel toleranter sind Teilnehmer eines Toleranzseminars geworden?
- Kontext: Bildungsprojekt mit Praxismaterialien und Multiplikatorenausbildungen
- Evaluationsobjekt: Seminarteilnehmer
- Evaluationskriterien:
 - Respekt gegenüber Minderheiten
 - Konfliktfähigkeit
 - Soziales Engagement
- Datenerhebung:
 - Fragebogen
 - Einzelinterviews
 - Alltagsbeobachtung

An dieser Darstellung zeigt sich die Problematik von Evaluation gerade im Bereich politischer Bildung. Die harmlos wirkende Ausgangsfrage impliziert eine Messbarkeit, die eine quantitative Aussage zum Ergebnis hat. Doch Evaluation kann hier kaum mit naturwissenschaftlichem Vorgehen verglichen werden.

Dem Wort nach ist Evaluation ein wertgebender Akt, d.h. der Wert oder die Qualität einer Sache sollen beurteilt werden. Schon hier zeigt sich, dass Evaluation nicht abstrakt oder gar neutral zu denken ist. Evaluation ist keine wissenschaftliche Technik oder neutrale Fragestellung, die unabhängig von bestimmten Norm- und Wertvorstellungen existiert. Evaluation benötigt zu-

mindest einen Vergleichsmaßstab, um eigene Bewertungen vornehmen zu können. Woher dieser Maßstab kommt und auf welchen Werten er gründet, ist aber in jedem praktischen Evaluationsvorhaben umstritten.

Die Benennung der Ziele eines Seminarprogramms und die daraus resultierenden Evaluationskriterien erweisen sich in der Bildungspraxis als höchst flexibel und sind abhängig von den unterschiedlichen Wertvorstellungen und Weltbildern derjenigen, die in der Praxis tätig sind. Die Konflikthaftigkeit von Evaluation ergibt sich somit aus dem Gegensatz der Forderung einer objektiven Bewertung und der Tatsache, dass jede Evaluation sich an häufig unausgesprochenen und unreflektierten Normen orientiert. Diese Tatsache ist im Laufe der Evaluationsgeschichte zunehmend in den Blick gerückt.

Historische Entwicklung – Vom Messen zum Aushandeln

Begriffe wie Effektivitätskontrolle, Qualitätssicherung und Qualitätsmanagement haben seit der Industrialisierung zunehmend an Bedeutung gewonnen. Sie standen zunächst vor allem in Zusammenhang mit betriebswirtschaftlichen und administrativen Entscheidungsgrundlagen. Egon Guba und Yvonna Lincoln (1989: 21–49) unterteilen die Geschichte der Evaluation in vier Generationen. Diese Entwicklung soll am Beispiel des Bildungsberreichs dargestellt werden.

Zunächst stand die Messbarkeit von Abläufen und Kennzahlen im Vordergrund. Das quantitative Erfassen von Fakten mit dem Ziel der Vergleichbarkeit zeigt sich in der jahrhundertealten Tradition der Methoden der Naturwissenschaften. Diese Methoden wurden aufgrund ihres enormen Erfolges im 19. Jahrhundert auch für die Humanwissenschaften übernommen und bildeten im Schulbereich die Grundlage für Leistungsprüfungen, in denen das gelernte Wissen reproduziert und somit der Lernerfolg dargestellt werden sollte.

Die zweite Generation der Evaluation ergab sich aus der Problematik des Messens. Als Evaluationsobjekte wurden nur die »Verwerter« von Bildungsprogrammen in den Blick genommen, während beispielsweise Lehrplankonzeptionen nicht hinterfragt wurden. Die Anwendung in der Praxis zeigte aber, dass die Umsetzung von Zielvorstellungen im Bildungsbereich nicht möglich war, weil der Zusammenhang von Schülern und Lehrplan nicht passte. Der Vergleichsmaßstab existierte wohl, war aber nicht an die Realität von Schülern angepasst.

Nun wurde der Blick auf die Beschreibung dieser Phänomene in der Praxis gerichtet, um Konsequenzen für die Programmentwicklung zu ziehen.

Evaluation als Praxisbeschreibung hatte hier Optimierungsfunktion, um die gewünschten und feststehenden Ziele mit anderen Mitteln zu erreichen. In den ersten zwei Evaluationsgenerationen hatten Messen und Beschreiben der Rolle des Evaluators noch die Funktion eines externen und neutralen Beobachters zugewiesen, der auf wissenschaftliche Art und Weise Daten erhob, aus denen sich quasi »automatisch« Konsequenzen ergaben, da ja die zu erreichenden Ziele feststanden.

In den 60er Jahren wurde schließlich von den Evaluatoren ein Urteil über ihre Beobachtungen erwartet. Dieses Urteil nahm nun auch kritisch die Kriterien und Zielvorstellungen von Programmen in den Blick: Nicht nur wie diese Ziele messbar sind und ob sie in der Praxis angemessen umgesetzt werden können, sondern ob diese Ziele selbst überhaupt angemessen sind, stand nun im Mittelpunkt der Überlegungen. Die Bewertung von Programmen und deren Anwendung wurde eingefordert.

Auch wenn Evaluatoren sich für diese Rolle allein oftmals fachlich nicht kompetent genug fühlten, wurden sie immer wieder als Richter über ein Projekt oder Programm angesehen, da sie vermeintlich am objektivsten urteilen würden. Sie sahen sich so als Wissenschaftler mit der Herausforderung konfrontiert, für ihnen fachfremde Disziplinen inhaltliche Empfehlungen zu geben. Evaluatoren wurden als Personen zu Legitimationsbeschaffern für Bildungsprojekte.

Gerade für den Bildungsbereich wurde jedoch deutlich, dass Fragen nach Programmwirkungen und Programmlegitimation immer noch nicht befriedigend genug beantwortet werden konnten. Es mussten vielmehr pädagogisch-psychologische Evaluationsmethoden entwickelt werden, die Aussagen ermöglichen, die nicht nur den angemessenen Einsatz von Ressourcen betreffen, sondern vor allem auch den Lernerfolg, die mögliche Transferleistung in den Alltag und die Akzeptanz der Methoden.

In den 80er Jahren kritisierten Evaluationsforscher aus den USA den damaligen Stand der Theoriebildung. Die zunehmende Problematisierung von Evaluationsobjekten, Zielen und Fragestellungen und die Ergebnisse interdisziplinärer Forschung (etwa Siebert 2002) zur Erkenntnistheorie führten die Forscher zum Konstruktivismus und seiner Grundannahme, dass Wirklichkeit nicht problemlos erfassbar ist, sondern sich je nach Kontext, Herkunft, Weltansicht, Glauben oder Profession anders darstellt. Dies hat auch Folgen für die Evaluation: »Der Konstruktivismus problematisiert die »ungebrochenen« Evaluationsmethoden. Die Wirksamkeit des Lehrens und Lernens lässt sich nur begrenzt feststellen.« (Siebert 2002: 192)

Evaluation kann aus dieser Sicht »nicht als Versuch gesehen werden, an dessen Ende eine Erfassung der Realität steht« und »Evaluationsergebnisse

stellen keine Fakten dar, die gesammelt zu einem erhöhten Wissen über Wirklichkeit führen.« (Grohmann 1997: 83) Auf der Basis der konstruktivistischen Erkenntnistheorie entstand die von Guba und Lincoln entwickelte vierte Generation, die sie schlicht als Fourth Generation Evaluation bezeichnen. Sie versucht ein »übergreifendes Verständnis und eine besondere Bezugnahme auf die menschlichen, politischen, sozialen, kulturellen und kontextuellen Elemente des Evaluationsprozesses« (Grohmann 1997: 83) herzustellen.

Dies geschieht, indem die Betroffenen und Beteiligten (Stakeholder) mit ihren unterschiedlichen Interessen von den Evaluatoren in einen demokratischen Aushandlungsprozess eingebunden werden, an dessen Ende die Einigung über die Interpretation der erhobenen Daten zu einer besonders hohen Relevanz der Resultate führt. Objektivistische Ansprüche der Auswertung und Erfolgskontrolle treten in den Hintergrund zugunsten anderer Faktoren, wie Glaubwürdigkeit, Anschlussfähigkeit und Handlungsrelevanz. Die notwendigen evaluatorischen Aktivitäten und Kompetenzen haben sich demnach hin zum Aushandeln unterschiedlicher Perspektiven verschoben. Das Urteilen über die Glaubwürdigkeit und die Relevanz der Daten steht in diesem Fall nicht nur den Evaluatoren, sondern allen Beteiligten zu.

Die Geschichte der Evaluation lässt sich also so zusammenfassen:

1. Messen: Evaluator als externer Datensammler
2. Beschreiben: Evaluator als neutraler Beobachter
3. Urteilen: Evaluator als Richter
4. Aushandeln: Evaluator als Moderator

Die Geschichte der Evaluation zeigt hier dreierlei:

Erstens wurden die Rahmenbedingungen, in denen Evaluation stattfindet, zunehmend hinterfragt und kritisch in den Blick genommen. Dies kann als eine Auflösung der Selbstverständlichkeit verstanden werden, mit der Evaluation eine kritisch-aufklärerische Position zugewiesen wird. Evaluation wird zum Symbol für die Schwierigkeit, die Welt linear zu verstehen, zu deuten und zu planen. Sie ist keine neutrale Messsonde mehr, sondern beeinflusst entscheidend das, was sie untersuchen möchte. Auf »ewige« wissenschaftliche Aussagen muss zunehmend verzichtet werden.

Zweitens änderte sich die Funktion des Evaluators in der Weise, dass ihm zunehmend mehr Verantwortung zugewiesen wurde: Zum Beschreiben von Praxis benötigt er ein umfassenderes methodisches Repertoire als beim rein quantitativen Messen. Um auch noch ein Urteil fällen zu können, benötigt er anerkannte Standards und Hintergrundkonzepte als Vergleichsfolie. Schließlich muss er in einer Evaluation, die sich als offener Prozess aller Beteiligten versteht, neben wissenschaftlichen Fähigkeiten auch soziale Kompe-

tenzen einbringen. Es wird anerkannt, dass Evaluation in einem grundsätzlich von unterschiedlichen Interpretationen und Konflikten geprägten Umfeld stattfindet.

Drittens fand – zumindest im Bildungsbereich – eine Verschiebung von quantitativer zu qualitativer Evaluation statt. Stand ursprünglich Messen und Zählen im Vordergrund, so rückten aufgrund des zunehmenden Komplexitätsbewusstseins qualitative Verfahren wie offene Interviews, Gruppendiskussionen und die Analyse von offenen Lernprozessen ins Zentrum der Betrachtung. Aufgrund der aktuell großen Forderungen nach Evaluation bei gleichzeitig begrenzten Mitteln findet allerdings häufig ein Rückgriff auf die vermeintlich einfacheren und besser darstellbaren quantitativen Evaluationsmethoden statt, die jedoch gerade im Bereich der Bildungsarbeit hinter das zurückfallen, was sich an Entwicklung in der Evaluationsforschung bereits gezeigt hat.

Die Entstehung von Evaluationsstandards

In der Betrachtung der Entwicklung von Evaluation stellte sich zunehmend die Frage, welchen Anspruch Evaluation selbst hat. Auf welchem Fundament steht sie, nach welchen Kriterien findet sie statt? In den USA werden seit einiger Zeit Evaluationsstandards für die Qualitätssicherung verwendet, die auch in Deutschland übernommen und adaptiert werden.

Die Qualität einer Evaluation bemisst sich im gängigen Verständnis vor allem in der Qualität der Datenerhebung und -auswertung, d.h. sie hängt davon ab, ob sie den Gütekriterien quantitativer bzw. qualitativer Forschung genügt. Im Falle einer quantitativen Evaluation sind dies die Kriterien

- Objektivität: neutral und durch keine subjektive Sicht beeinflusst;
- Reliabilität: wiederholbar und damit überprüfbar, vergleichbar;
- Validität: genaue Datenerfassung, Gültigkeit.

Im Falle qualitativer Forschung bzw. Evaluation gehen die Meinungen über die Gütekriterien stark auseinander und lassen sich vereinfacht drei Grundpositionen zuordnen (Steinke 2000):

- Quantitative Kriterien gelten auch für qualitative Forschung/Evaluation.
- Es werden eigene Kriterien für die Evaluation aufgestellt: Kriterien sind hier etwa kommunikative Validierung, d.h. die erhobenen und bearbeiteten Daten werden den Befragten vorgestellt und mit ihnen besprochen, oder Methodentriangulation, d.h. dasselbe Phänomen wird mit unterschiedlichen methodischen Herangehensweisen in den Blick genommen.

- Aus einer postmodernen oder konstruktivistischen Haltung heraus werden Gütekriterien abgelehnt: Die radikale Ablehnung von Gütekriterien hängt mit der Problematisierung der Reduktion von Evaluation auf Datenerhebung zusammen.

Wenn Evaluation zunehmend von der Auflösung der Selbstverständlichkeiten und Rahmenbedingungen betroffen ist, kann sie in ihrer Qualität nicht auf Datenerhebung reduziert werden. Deshalb verlieren Gütekriterien ihre Relevanz, zumal sie einem überholten naturwissenschaftlichen Verständnis von Evaluation angehören. In dieser Position bricht auch die Unterscheidung quantitativ/qualitativ zusammen.

Deshalb ergab sich die Notwendigkeit, Qualitätskriterien, Regeln oder Standards für »gute« Evaluationen insgesamt zu formulieren. Denn schließlich sollen Evaluationen »auf Akzeptanz und Glaubwürdigkeit bei den am evaluierten Programm Beteiligten und den Adressaten der Evaluationsberichte stoßen« (Beywl und Taut 2000: 358 ff.). Dies bemisst sich nicht nur am Umgang mit Daten, sondern an dem Vorgehen im Evaluationsprozess insgesamt.

So wurden bereits Anfang der 1980er Jahre in den USA vom Joint Committee on Standards for Educational Evaluation Standards zur Durchführung von Evaluationen entwickelt (deutsch Sanders 2000). In Deutschland hat die Deutsche Gesellschaft für Evaluation (DeGEval) davon adaptierte Standards veröffentlicht (2002, vgl. www.degeval.de). Diese Standards stellen einen ersten Schritt in Richtung von Professionalisierung von Evaluation dar. Sie enthalten vier Hauptforderungen:

- Nützlichkeit: Die Nützlichkeitsstandards sollen sicherstellen, dass die Evaluation sich an den geklärten Evaluationszwecken sowie am Informationsbedarf der vorgesehenen Nutzer orientiert.
- Durchführbarkeit: Die Durchführbarkeitsstandards sollen sicherstellen, dass eine Evaluation realistisch, gut durchdacht, diplomatisch und kostenbewusst geplant und durchgeführt wird.
- Fairness: Die Fairnessstandards sollen sicherstellen, dass in einer Evaluation respektvoll und fair mit den betroffenen Personen und Gruppen umgegangen wird.
- Genauigkeit: Die Genauigkeitsstandards sollen sicherstellen, dass eine Evaluation gültige Informationen und Ergebnisse zu dem jeweiligen Evaluationsgegenstand und den Evaluationsfragestellungen hervorbringt und vermittelt.

Hierauf aufbauend stellen wir in diesem Beitrag eigene Schwerpunkte im

Rahmen dieser allgemeinen Evaluationsstandards vor, die sich insbesondere als Orientierung für die politische Bildung eignen.

Warum partizipative Evaluation?

Evaluation sollte grundsätzlich, vor allem aber im Zusammenhang mit politischer Bildung demokratisch und damit partizipativ angelegt werden. So verstanden, wird Evaluation zu einem interaktiven Prozess, der durch das Einbeziehen unterschiedlicher Perspektiven blinde Flecken aufdecken kann und damit zu einer umfassenden Analyse des gesamten Interventionskontextes kommt. Es genügt nicht, aus einer einzigen externen Perspektive Daten zu erfassen und zu bewerten.

Geeignete Evaluationsverfahren bieten die Möglichkeit, »andere Perspektiven zu Worte kommen zu lassen, Urteile von anderen zu integrieren und Bewertungsmonopole zu öffnen. Evaluation ist kein Instrument zur technischen Überprüfung und Kontrolle (zumindest nicht ausschließlich) klar definierter Zielvorgaben am Ende einer Entwicklung ... Evaluation ist ein sozialer, kein messtechnischer Prozess.« (Herrmann und Höfer 1999: 102)

So verstanden hat Evaluation selbst Bildungswirkung, sie lässt sich nicht unabhängig von politischer Bildung denken. Genauso wie politische Bildung das Ursache-Wirkung-Verhältnis von Lernen problematisiert und in reflexive Zirkel überführt, so wird Evaluation auch zur Ursache von neuen Lernerfahrungen und hat eine zusätzliche Wirkung auf die erforschten Lernprojekte.

Statt es als vielleicht unvermeidbaren und im besten Falle herausdefinierbaren Faktor zu betrachten, stellen wir diesen Umstand in den Mittelpunkt eines neuen Verständnisses von Evaluation und betrachten Evaluation als Chance für erweiterte Partizipation und Pluralität. Damit fällt auch die klare Unterscheidung zwischen den verschiedenen Funktionen von Evaluation weg: Erkenntnis des eigenen Tuns, Anregungen für Verbesserung und Auftauchen gemeinsamer bilanzierender Erkenntnisse bedingen sich. Evaluation wird zum Anstoß für vertiefte Reflexion und damit dem Anspruch politischer Bildung selbst gerecht.

Der Verlauf eines partizipativen Evaluationsprozesses

Eine Evaluation beginnt immer »impulshaft« mit der Intervention eines Auftraggebers oder einer drängenden Evaluationsfrage, die sich stellt und zum Anlass für den Auftrag einer Evaluierung eines Projektes der politischen Bil-

dung wird. Evaluation konstruiert schon mit ihrer ersten Fragestellung eine Wirklichkeit eigener Art (vgl. Siebert 1999) und muss sich dessen bewusst sein: Die Fragestellung einer Evaluation konzentriert sich auf bestimmte Aspekte der Außenwelt und blendet damit andere aus. Sie erlaubt eine bestimmte Perspektive auf die Welt, die auch ganz anders ausfallen könnte.

Dieser Tatsache wird durch die Einbeziehung weiterer Perspektiven im Verlauf der Evaluation Rechnung getragen. Durch die demokratische, d.h. gleichberechtigte Beteiligung der von der Evaluation Betroffenen kommen verschiedene Perspektiven zur Sprache. Partizipative Evaluation bedeutet, konstruktiv mit diesen konflikthaften Perspektiven umzugehen.

Wichtig ist, dass eine konkrete Ausgangsfrage existiert und dass diese vertraglich festgehalten wird, um überhaupt einen handlungsfähigen Ausgangspunkt zu haben. Der Ausgangsvertrag für eine Evaluation sollte außerdem sicherstellen, dass die verschiedenen Stakeholder identifiziert werden und gleichberechtigt ihre Sichtweise nicht nur bei der Beantwortung von Evaluationsfragen, sondern auch zu diesen Fragen selbst einbringen können. Stakeholder sind alle diejenigen, die ein bestimmtes »Interesse« an der Evaluation haben, beteiligt oder auch von einer Bildungsmaßnahme ausgeschlossen, gleichsam »negativ« oder »invers« betroffen sind. Sie haben alle unterschiedliche Perspektiven und verschiedene Fragestellungen an die Evaluation, die berücksichtigt werden müssen.

Das wesentliche Augenmerk von Evaluation liegt deshalb auf der Einbeziehung möglichst vieler Stakeholder. Der Evaluationsprozess muss mit ihrer Identifizierung beginnen. Damit werden blinde Flecken einseitiger Beobachtung oder Bewertung durch neue Perspektiven erhellt. Politisch brisant wird dieses Vorgehen, weil es häufig auch diejenigen zu Wort kommen lässt, die bisher von der Programmentwicklung oder -teilnahme ausgeschlossen waren oder speziell zu diesem Bildungsprogramm eine konträre Meinung vertreten.

Ähnlich wie die zu evaluierende politische Bildung orientiert sich partizipative Evaluation hier eher an der Differenz als an übereinstimmenden Meinungen. Nachdem die Stakeholder identifiziert sind und sie ihre unterschiedlichen Perspektiven eingebracht haben, beginnen die ersten Aushandlungsprozesse. Hierbei ist es praktikabel, eine Begleitgruppe mit Vertretern der unterschiedlichen Stakeholder einzurichten und sie als intensiven Kern des Evaluationsprozesses umfassend mit einzubeziehen. So kann das Evaluationsvorhaben handlungsfähig bleiben und Teilschritte umfassend über die Begleitgruppe weitergeleitet werden.

Während des ersten Aushandlungsprozesses richtet sich der Blick zuerst auf das Finden akzeptabler Zielkriterien und geeigneter Evaluationsobjekte. Die ursprüngliche Ausgangsfrage der Evaluation steht im besten Falle hier

zur Disposition, da es keinen starren Bedingungsrahmen mehr geben kann, wenn ernsthaft partizipativ vorgegangen werden soll. Meist stößt der Auftraggeber mit einer bestimmten Intervention den Evaluationsprozess an, der im weiteren explorativ verläuft und gemeinsam mit den Beteiligten die relevanten Fragen erarbeitet. Das Augenmerk liegt so vorrangig auf dem Prozess der Evaluation und nicht wie bei einem naturwissenschaftlichen Experiment auf der Beantwortung vordefinierter Fragen.

Dies bedeutet auch, aus den unterschiedlichen Perspektiven evtl. vorgegebene oder schriftlich fixierte Zielvorstellungen von Programmen und Projekten zu diskutieren. Oft werden diese in der pädagogischen Praxis unreflektiert angewandt und sehr unterschiedlich in die eigenen Welt- und Menschenbilder von politischen Bildnern eingeordnet. Andere Beteiligte der Evaluation haben oft völlig andere Zielvorstellungen im Kopf, die sich unausgesprochen im Verlaufe eines Projektes gebildet haben. Für die politische Bildung macht sich dies z. B. an der Frage fest, welche Bedeutung Konflikte in Bildungsveranstaltungen haben oder haben sollen. Diese gilt es zu benennen und in die Aushandlung zu bringen, um eine gemeinsame Basis für die Datenerhebung der Evaluation zu haben.

Im weiteren Verlauf der Evaluation werden Daten erhoben. Es ist nicht entscheidend, ob es sich um quantitative oder qualitative Daten handelt. Die Erhebung der Daten ist nicht traditionell als wissenschaftlich unabhängiger Prozess zu verstehen, sondern als das Einholen unterschiedlicher Perspektiven in Bezug auf das Evaluationsobjekt. In diesem Sinne können jenseits der Evaluatoren auch andere an der Evaluation Beteiligte Daten erheben (vgl. auch den Methodenteil dieses Buches und Mohr und Jacobsgaard 2003).

Unerlässlich ist, dass die erhobenen Daten und die aus ihnen gezogenen Schlussfolgerungen immer wieder in den Aushandlungsprozess der unterschiedlichen Stakeholder gebracht werden. Wo Differenzen, Unterschiede oder Konflikte in der Bewertung auftauchen, sind Punkte gegeben, weiter zu forschen und mit neuen Teilfragestellungen Daten zu erheben.

Diese Art des emergenten Vorgehens schafft für alle Beteiligten die Möglichkeit, unterschiedliche Perspektiven und methodisches Vorgehen in einen Aushandlungsprozess zu bringen, der vom Evaluator begleitet und moderiert wird. Das Aushandeln der Interessen und Interpretationen der Daten ist als schleifenartiger Prozess zu verstehen, der kontinuierlich wiederholt wird.

Am Ende des Prozesses steht nicht die Veröffentlichung objektiver Daten, sondern eine sinnvolle gemeinsame Konstruktion, die aus der demokratischen Aushandlung der verschiedenen Sichtweisen erwachsen ist. Ziel von Evaluation ist also nicht, die »Wahrheit« zutage zu befördern. Evaluation hat hier einen konstruktivistischen Grundcharakter, der sich darin äußert, dass

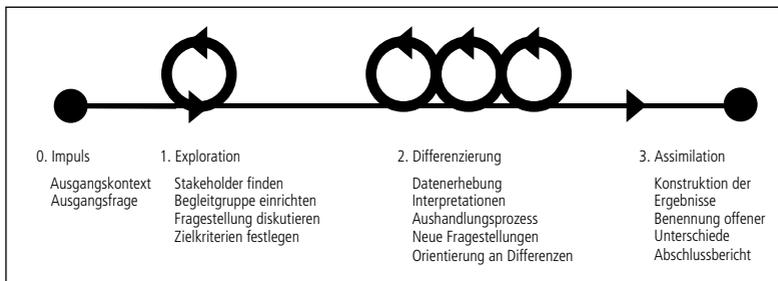
Sinnperspektiven nicht vorgegeben oder entdeckt, sondern in konkreten Prozessen je neu konstruiert werden. Gemeinsame Sinnzuschreibungen können in einem Abschlussbericht ebenso wie unterschiedliche und konflikthafte Perspektiven benannt werden. Wichtig ist dabei jedoch, einen demokratischen Grundrahmen als Orientierungsmaßstab anzulegen, um nicht zu relativistischem und beliebigem Vorgehen zu kommen.

Wann der Prozess beendet ist, hängt natürlich von finanziellen und personellen Ressourcen ab. Entscheidend ist jedoch die Authentizität des Vorgehens: Wenn Redundanzen in den gemeinsamen Perspektiven auftauchen und kein neuer Erkenntnisgewinn vom gemeinsamen Austausch zu erwarten ist, können diese Gemeinsamkeiten niedergelegt werden. Falls dies nicht möglich ist, besteht auch die Möglichkeit der gegenseitigen Anerkennung von Differenzen (»We agree to disagree«), die in einem Evaluationsbericht benannt werden können.

Der Abschlussbericht sollte auf jeden Fall alle Perspektiven integrieren und evtl. für unterschiedliche Stakeholder in Hinblick auf die jeweilige Brauchbarkeit unterschiedlich verfasst werden. Dies sollte keinesfalls geschehen, um unterschiedliche Interessen zu »bedienen«, sondern um sicherzustellen, dass die Handlungskonsequenzen, die aus den Diskussionen und Verhandlungen hervorgegangen sind, transparent und deutlich formuliert werden.

Zudem muss ein Abschlussbericht den Prozess der Evaluation verdeutlichen, auch seine Schwierigkeiten und Konflikte. Dies ist ein entscheidender Teil der Ergebnisse, denn die Art und Weise, wie der gemeinsame Weg beschritten wurde, ist wichtig für die Glaubwürdigkeit der Evaluation und ist auch Ausdruck der (existierenden oder fehlenden) Stimmigkeit zwischen Evaluation und dem Anspruch des untersuchten Projektes der politischen Bildung.

Abbildung 1: Die Phase des partizipativen Evaluationsansatzes



Der partizipative und damit zeitaufwändige Ansatz empfiehlt sich langfristig, weil er Ergebnisse in Form von gemeinsamen Handlungsempfehlungen verspricht, deren Umsetzung von allen Seiten mitgetragen wird. Der partizipative Ansatz muss deswegen die geringsten Reibungsverluste beim Transfer in die Praxis fürchten. Die Art des Vorgehens kann als partizipative Evaluation verstanden werden, da sie Beteiligung sichert und Praxisrelevanz in den Konsequenzen des Evaluationsberichtes schaffen will. Besonders im Hinblick auf den sozialen und interaktiven Kontext politischer Bildung als Demokratie-Lernen ist dieses Vorgehen wichtig, da es sich an der Freiheit des Menschen orientiert.

Das beschriebene Vorgehen möchte in seiner Gesamtheit dazu beitragen, dem weit verbreiteten Missbrauch von Evaluation als Legitimationsbeschaffer zur Fortführung oder Beendigung von Bildungsprojekten zu begegnen. Wenn eine authentische Einbeziehung aller Stakeholder in den Evaluationsprozess zu dem (vorläufigen) Ergebnis der Fragwürdigkeit eines Bildungsprojektes führt, können gemeinsam neue und kreative Lösungen gesucht werden.

In der Praxis ist in der Konsequenz einer Evaluation sicher von einer Ungleichheit der Verantwortungs- und Entscheidungsebenen (insbesondere auch im Hinblick auf Arbeitsplätze) auszugehen. Insbesondere die Einbeziehung von Auftrag- und Geldgebern kann zu einem neuen gegenseitigen Verständnis der unterschiedlichen Perspektiven und Handlungsnotwendigkeiten führen.

Die Rolle des Evaluators

Partizipative Evaluation verändert die Rolle des Evaluators in fundamentaler Art und Weise: Sie schreibt dem Evaluator nicht nur die Aufgabe zu, »wissenschaftliche« Daten zu erheben. Zusätzlich ist es seine Aufgabe, möglichst viele Beteiligte und Betroffene sowie ihre Perspektiven in einen kommunikativen Aushandlungsprozess zu bringen. In diesem Prozess, dessen Moderator er ist, werden die Ausgangsfragestellungen, das Vorgehen und schließlich die Interpretationen der erhobenen Daten abgestimmt.

Der Evaluator erkennt und thematisiert Konflikte. Seine Verhandlungskompetenz ist ausschlaggebend für das Gesamtergebnis. Damit alle ihre Perspektiven einbringen können, wird außerdem die Kenntnis geeigneter Methoden benötigt, die auch im visuellen oder gestalterischen Bereich liegen. Dies wird besonders deutlich, wenn die Beteiligten des Evaluationsprozesses aus unterschiedlichen Kulturen, Bildungsmilieus oder Sprachheimaten

kommen. Aufgabe des Evaluators ist es hier, aktivierend zu wirken und methodische Angebote zu machen, mit denen Daten gleichberechtigt erhoben werden können.

Der Evaluator wird so zum Advokaten unterschiedlicher gleichberechtigter Perspektiven. Dies ergänzt die klassische Datenerhebung und bedarf beim Evaluator einer erhöhten Sensibilität für die verschiedenen (Macht-)Positionen, Hierarchien und blinden Flecke. Der Evaluator trägt die Verantwortung für den erfolgreichen Prozess der Evaluation und muss immer wieder Offenheit und zielgerichtetes Vorgehen abwägen können. Dies kann als die Wahrung eines demokratischen Grundrahmens einer Evaluation verstanden werden.

Die persönliche Glaubwürdigkeit eines Evaluators zeigt sich darin, dass er jedes Anliegen ernst nimmt und die richtige Balance von Sicherheit und notwendiger Konfrontation findet. Glaubwürdigkeit, Autorität und Kompetenz eines Evaluators zeigen sich auch in der theoretischen Rekonstruktion (House und Howe 1999: 115) des zu untersuchenden Gegenstandes. Er versetzt sich dadurch in den historischen und sozialen Kontext eines Bildungsprojektes und gewinnt so eine eigene Perspektive.

Dies kann sicher nicht in Abstraktion, sondern muss im Verlauf der Einbeziehung der Stakeholder geschehen. Dennoch stellt es eine eigene wichtige Dimension dar, die den Evaluator vom vermeintlich neutralen Datensammler zum ebenfalls Beteiligten macht. Die theoretische Rekonstruktion erlaubt außerdem im Evaluationsprozess eine bessere Moderation der unterschiedlichen Perspektiven und Interessen jenseits relativistischer Gleichberechtigung.

Der Evaluator garantiert den Beteiligten ein Höchstmaß an Partizipation, ohne den zeitlichen und finanziellen Rahmen aus den Augen zu verlieren. Auf diese Weise sorgt er dafür, dass die gemeinsamen Konstruktionen der Evaluation für die unterschiedlichen Stakeholder brauchbar sind und damit quasi »automatisch« zu Handlungskonsequenzen führen.

Schließlich ist es die Aufgabe des Evaluators, den Abschlussbericht stimmig zu gestalten, d.h. die Gesamtatmosphäre des Evaluationsprozesses einzufangen. Hier sind weniger wissenschaftliche denn ästhetische Kompetenzen nötig.

Kompetenzen und Aufgaben des Evaluators (im Kreuzungsfeld konstruktivistischer Evaluation nach Guba und Lincoln und deliberativer demokratischer Evaluation nach House und Howe), spezifiziert für die politische Bildung, sind demnach:

- Anwaltschaft (»advocacy«): Suche nach und Einbindung möglichst vieler Stakeholder in den Evaluationsprozess

- Theoretische Rekonstruktion: Rekonstruktion der theoretischen Implikationen und Grundannahmen des zu evaluierenden Projektes oder Programms, Gewinnung einer eigenen Perspektive
- Verhandlungskompetenz: Klärung unterschiedlicher Perspektiven und Konstruktion gemeinsamer Bewertungsmuster
- Methodenkompetenz: Angebot geeigneter Methoden zur Erhebung von Evaluationsdaten für unterschiedliche Gruppen der Beteiligten
- Demokratiekompetenz: Wahrung eines demokratischen Grundrahmens durch Dialog und geteilte Werte
- ästhetische Kompetenz: Gestaltung eines Abschlussberichtes, der die Dynamik, Atmosphäre, Unterschiede und Gemeinsamkeiten des Evaluationsprozesses widerspiegelt

Qualität und Standards

Wir haben die Geschichte der Evaluation und unsere Konzeptualisierung bisher als Steigerung der Komplexität und Reflexivität verstanden. Nach diesem Verständnis stellt sich die Frage nach einer »guten« Evaluation auf eine völlig neue Art, wenn selbst überlieferte wissenschaftliche Standards wie Validität (Gültigkeit) nur als Konstrukte betrachtet werden, die zur Disposition stehen. Welchen »Rahmen« hat nun Evaluation und wie kann garantiert werden, dass Evaluatoren nicht im Sinne des »Anything goes« vorgehen und Evaluation scheinbar theoretisch legitimiert zu einem Verfahren der Beliebtheit wird?

Der Evaluator verstand sich bislang als Erhebender von Daten, deren Qualität zu sichern ist und die in einem Abschlussbericht festzuhalten sind. Evaluation von politischer Bildung, wie wir sie verstehen, verändert den Qualitätsbegriff. Im Folgenden werden Standards für Evaluation vorgestellt, die einen sinnvollen Rahmen bieten, um demokratisch und partizipativ vorzugehen. Sie sollen es ermöglichen, den Evaluationsprozess möglichst produktiv und sinnvoll zu gestalten, und die Aufgaben des Evaluators klären (orientiert an Guba und Lincoln – es werden im Folgenden diejenigen Kriterien aufgeführt, die sich von den veröffentlichten Standards für Evaluation unterscheiden):

- Fairness: Die verschiedenen Perspektiven werden im Evaluationsbericht, d.h. der gemeinsamen Sinnkonstruktion, berücksichtigt
- Weiterentwicklung: Die Verfeinerung und das komplexere Verständnis der eigenen Perspektive der Beteiligten im Rahmen des Evaluationsprozesses wird ermöglicht.

- Fortbildung: Das erweiterte Verständnis für die Bewertungen und Perspektiven der jeweils anderen am Evaluationsprozess Beteiligten wird gesichert.
- Relevanz: Die Verwertbarkeit der Ergebnisse und Relevanz der Evaluation für die Beteiligten führt zu Handlungskonsequenzen.
- Nachhaltigkeit: Die erkannten Handlungskonsequenzen können von den Beteiligten im Rahmen ihres institutionellen Kontextes auch durchgeführt werden.
- Glaubwürdigkeit: Die Evaluation geht authentisch und umfassend partizipativ vor, der Prozess als solcher wird von den Beteiligten akzeptiert.

An diesen Standards zeigt sich in konzentrierter Form die veränderte Stellung von partizipativer Evaluation. Statt externer Bewertung fördert sie Beteiligung und Handlungsorientierung und erhöht die Qualität des untersuchten Bildungskontextes im Hinblick auf Programm, Organisation und Interaktion.

Die Chancen und Konsequenzen partizipativer Evaluation

Partizipative Evaluation tritt in den Kontext von Aushandlungsprozessen unterschiedlicher Interessen und Meinungen und bekommt somit selbst politischen Charakter. In der Konsequenz bedeutet dies, dass eine Evaluation dieser Art, insbesondere wenn sie in hierarchische oder politisch einseitige Kontexte eintritt, auch system- und organisationsverändernden Charakter hat.

Evaluation zeigt sich in der Realität nicht nur als ein interessantes methodisches Vorgehen, sondern hat Implikationen für den Kontext selbst. Evaluation ist nicht neutral, sondern hat in der Verbindung mit politischer Bildung Bekenntnischarakter für ein demokratisches und partizipatives Vorgehen.

In diesem Sinne ist sie auch nicht abstrakte Wissenschaft, sondern hat einen politikberatenden Charakter: »Evaluation ist selbst Teil dieses politischen Prozesses, wird von ihm beeinflusst und beeinflusst ihn seinerseits – ob sie es bewusst will oder nicht. Auf diesem politischen Bezug beruht die disziplinäre Eigenart der Evaluation und daraus resultiert ihre Mittlerstellung zwischen Forschung und Politikberatung.« (Heiner 2001: o. S.) Partizipative Evaluation ist darüber hinaus eine neue Möglichkeit, zu einer weiteren Politisierung nicht nur der Zielgruppe eines Projektes, sondern des Organisationskontextes selbst beizutragen.

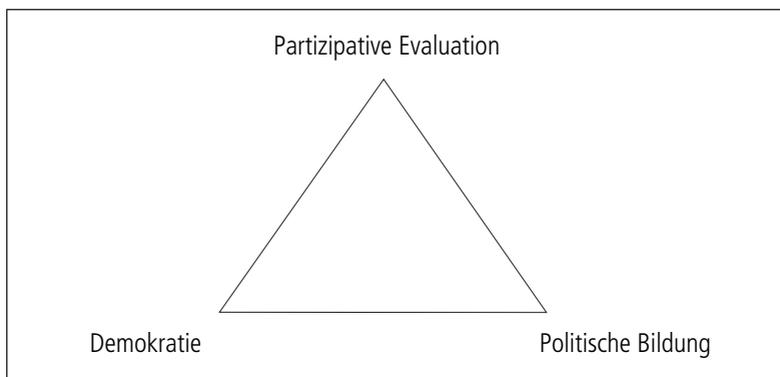
Evaluation in unserem Sinne bedeutet also eine Ausweitung demokratischen Vorgehens in Bereiche, die vormals vermeintlich neutraler Wissenschaft reserviert schienen. Dies muss am Anfang eines Evaluationsvorhabens

klar benannt und auch vertraglich festgehalten werden. Brisant kann dies im Kontext einer politischen Bildung werden, die sich mit ihren eigenen Wertvorstellungen auf sicherem Terrain zu befinden meint und andere von ihren Grundwerten überzeugen möchte. Diese Problematik ist in vielen Programmen »gegen rechts« erkennbar, die nicht grundsätzlich pluralistisch angelegt sind und sich durch den Ausschluss »falscher« Weltbilder definieren. Nach unserer Definition befinden sich solche Programme meist innerhalb der ersten Dimension politischer Bildung (vgl. Lynen van Berg und Roth 2003).

Evaluation kann in solchen Zusammenhängen als Einladung zu neuer Kommunikation dienen, muss sich aber ihrer Grenze bewusst sein: Wenn Bildungsprogramme auf der Grundunterscheidung von richtig und falsch operieren, so kann partizipative Evaluation sehr stark als (manipulative) Zerstörung dieser Grundunterscheidung verstanden werden. Aus Sicht von Programmverantwortlichen würde sich Evaluation so abermals zum Richter aufschwingen und über »besseres« Wissen verfügen. Dies wäre kontraproduktiv für das eigentliche Anliegen partizipativer Evaluation. Im Praxisfall muss ausgelotet werden, wo Ansätze für unterschiedliche Perspektiven bereits vorhanden sind und wie diese weitergetragen werden können.

So gesehen kann Evaluation nicht länger von außen an Bildungsprojekte herangetragen werden, sondern ist immer schon mit ihnen und ihren Kontexten verwoben. Evaluation tritt nicht objektiv auf, sondern verändert von Beginn an den sozialen Kontext und die Strukturen innerhalb eines Bildungsprojektes als eine Intervention, die grundsätzlich offen in Vorgehen und Konsequenzen ist und sich in gesamtgesellschaftliche Diskurse einfügt.

Abbildung 2: Wechselwirkung von partizipativer Evaluation, politischer Bildung und Demokratie



Das Schema verdeutlicht die Wechselwirkung von partizipativer Evaluation, politischer Bildung und Demokratie. Partizipative Evaluation orientiert sich an der Grundanlage politischer Bildung, die sich plural versteht. Sie wird aber ebenso von den Errungenschaften und Grundwerten von Demokratie bestimmt (vgl. House und Howe 1999: 97). Insbesondere der Aushandlungsprozess im Verlauf einer Evaluation ist nicht ohne Rückbezug auf den demokratischen Rahmen zu denken.

Dies bedeutet insbesondere, dass die Werte Gleichheit und Freiheit aller Beteiligten gewahrt werden und Ansichten oder Interpretationen, die diese Werte untergraben, auch nicht als gültige Perspektiven der Evaluation zu betrachten sind. Dies ist ein Unterscheidungsmerkmal zur rein konstruktivistischen Evaluation, die in der Gefahr steht, keinen Orientierungsrahmen für unterschiedliche Wertvorstellungen zu bieten: »Suppose that a participant advances a racist view. Are we to hold that this view is as good as other views? It is no use arguing that other stakeholders would disagree ... To say that we can have no way of resisting these views suggests why relativism is untenable.« (House and Howe 1999: 60)

Umgekehrt bedeutet die Orientierung partizipativer Evaluation an einem demokratischen Grundrahmen auch eine Rückwirkung auf politische Bildung bzw. konkrete Projekte und Bildungsmaßnahmen politischer Bildung, die evaluiert werden. Hier bietet partizipative Evaluation selbst die Chance, politische Bildung voranzubringen.

Evaluation kann zur Erweiterung und Vertiefung von bestehenden Projekten der politischen Bildung beitragen. Durch die systematische Einbeziehung aller Beteiligten wird durch Evaluation oft zum ersten Mal klar, welche unterschiedlichen expliziten und impliziten Annahmen über das eigene Tun und Erleben vorherrschen. Sie ermöglicht eine neue Reflexionsebene, die zu einem Qualitätssprung innerhalb eines Praxisprojektes werden kann. Theoretische Konzeptionen werden klarer gesehen; eine deutliche Einordnung und mögliche Abgrenzung von anderen Konzeptionen kann stattfinden.

Evaluation hat darüber hinaus auch Fortbildungsfunktion, indem sie nicht nur neutral Daten abrufen, sondern verschiedene Perspektiven in das gemeinsame Spiel immer wieder einbringt. Dies führt zur Neubewertung der eigenen Perspektive und der Entwicklung neuer Sichtweisen. Evaluation erweitert so in ihrem Stellenwert den Wirkungsbereich politischer Bildung, wie wir sie konzeptualisiert haben.

Partizipative Evaluation ermöglicht es den Beteiligten eines Bildungsprojektes, durch die gemeinsame Auseinandersetzung klarer nach außen zu beschreiben, wie ihr »Produkt« definiert ist und wie es sich von anderen Produkten politischer Bildung unterscheidet. Dies ist nicht zuletzt unter Wett-

bewerbsgesichtspunkten notwendig und produktiv. Demokratisches und partizipatives Vorgehen in einer Evaluation verbindet die Innen- und Außenwirkung von Bildung.

Die oft mangelhafte Umsetzung der externen Bildungsansprüche im Rahmen der eigenen Organisation kann mit partizipativer Evaluation verwirklicht werden. Dies erlaubt eine wesentlich höhere Transparenz und damit Glaubwürdigkeit bei der Darstellung der Binnenstruktur von Bildungsprojekten.

Partizipative Evaluation verabschiedet sich von der Vorstellung, neutral und objektiv vorgehen zu können. Sie beeinflusst nicht nur das, was sie untersucht, sondern bindet sich inhaltlich ein und tritt in eine Interaktion mit dem jeweiligen Evaluationsobjekt. Sie wird zum Bestandteil eines Projektes, ohne darin völlig aufzugehen, denn sie beginnt mit einer systematisierten Ausgangsfrage, die explorativ verändert werden kann, aber zu einer bestimmten gemeinsamen Aussage über ein Bildungsprojekt führen sollte.

Partizipative Evaluation hat Konsequenzen. Die Beteiligten und Betroffenen erhalten Antworten auf ihre Fragen und werden so befähigt, entsprechend zu handeln. Im Gegensatz zu vielen Evaluationen, die beeindruckende Statistiken und Zahlenwerke vermitteln, aber ohne Konsequenz bleiben, ist der partizipative Evaluationsansatz selbst darauf angelegt, Maßnahmen zu formulieren und gemeinsam umzusetzen.

Die Chancen partizipativer Evaluation sind zusammengefasst somit:

- Qualitätssprung: konzeptionelle Vertiefung und Einordnung eines Praxisprojektes
- Bildungscharakter: eigene Supervision und Weiterbildung der pädagogischen Praxis
- Konturierung: klarere »Produktbeschreibung« und Außendarstellung
- Demokratie: Transparenz und Glaubwürdigkeit des Vorgehensweise
- Interaktion: inhaltliche Einbindung in ein Bildungsprojekt
- Konsequenz: Handlungsrelevanz für alle Beteiligten

Wir haben in einem dreijährigen Evaluationsprozess unseres Projektes Erziehung zu Demokratie und Toleranz einiges von diesen Chancen erfahren dürfen. In der Auseinandersetzung ist partizipative Evaluation durchaus ein konflikthafter Prozess im Streiten um Deutungen und Sichtweisen. Irritationen, Verunsicherungen und die Suche nach eigenen neuen Wegen und Orientierungen sind Teil der Erfahrungen – und gerade dies schlägt den Bogen zu politischer Bildung, die sich der Freiheit und Verantwortung des Menschen verpflichtet weiß.

Literatur

- Altrichter, Herbert (1999): Evaluation als Alltäglichkeit, als Profession und als Interaktion. In: Josef Thonhauser und Jean-Luc Patry (Hrsg.): *Evaluation im Bildungsbereich. Wissenschaft im Dialog*. Studien zur Bildungsforschung und Bildungspolitik 22. Innsbruck.
- Lynen von Berg, Heinz, und Roland Roth (Hrsg.) (2003): *Maßnahmen und Programme gegen Rechtsextremismus wissenschaftlich begleitet. Aufgaben, Konzepte und Erfahrungen*. Opladen.
- Bertelsmann Forschungsgruppe Politik (Hrsg.) (2000): *Toleranz. Grundlage für ein demokratisches Miteinander*. Gütersloh.
- Beywl, Wolfgang (1999): *Zur Weiterentwicklung der Evaluationsmethodologie. Grundlegung, Konzeption und Anwendung eines Modells der responsiven Evaluation*. Frankfurt am Main.
- Beywl, Wolfgang, und Sandy Taut (2000): »Standards. Aktuelle Strategie zur Qualitätsentwicklung in der Evaluation«. In: *Vierteljahreshefte zur Wirtschaftsforschung* (69) 3 2000. 358–370.
- Chelimsky, Eleanor, und William R. Shadish (Hrsg.) (1999): *Evaluation for the 21st Century. A Handbook*. Thousand Oaks.
- Denzin, Norman K., und Yvonna S. Lincoln (2000): *Handbook of Qualitative Research*. 2. Auflage. Thousand Oaks, Cal.
- Deutsche Gesellschaft für Evaluation (Hrsg.) (2001): *Evaluation. Reformmotor oder Reformbremse*. Köln.
- Deutsche Gesellschaft für Evaluation (Hrsg.) (2002): *Standards für Evaluation*. Köln.
- Dunn, Seamus, K.-Peter Fritzsche und Valerie Morgan (Hrsg.) (2003): *Tolerance Matters. International Educational Approaches*. Gütersloh.
- Fengler, Jörg (1997): »Lernziel, Design und Evaluation«. In: Oliver König (Hrsg.): *Gruppendynamik. Geschichte, Theorien, Methoden, Anwendungen, Ausbildung*. München. 151–194.
- Fetterman, David M. (2001): *Foundations of Empowerment Evaluation*. London.
- Flick, Uwe, Ernst von Kardorff und Ines Steinke (2000): *Qualitative Forschung. Ein Handbuch*. Reinbek.
- Georgi, Viola, Susanne Ulrich und Florian M. Wenzel (2003): »Education for Democracy and Tolerance in Germany.« In: Seamus Dunn, K.-Peter Fritzsche und Valerie Morgan (Hrsg.): *Tolerance Matters. International Educational Approaches*. Gütersloh. 85–115.
- Grohmann, Romano (1997): *Das Problem der Evaluation in der Sozialpädagogik. Bezugspunkte zur Weiterentwicklung der evaluationstheoretischen Reflexion*. Frankfurt am Main.

- Guba, Egon G., und Yvonna S. Lincoln (1989): *Fourth Generation Evaluation*. Newbury Park, Cal.
- Guba, Egon G., und Yvonna S. Lincoln (o.J.): »Guidelines and Checklist for Constructivist (a.k.a. Fourth Generation) Evaluation.« www.wmich.edu/evalctr/checklists/constructivisteval.htm (Download 16.6.2003).
- Heiner, Maja (Hrsg.) (1996): *Qualitätsentwicklung durch Evaluation*. Freiburg im Breisgau.
- Heiner, Maja (2001): »Evaluation«. In: Hans-Uwe Otto und Hans Thiersch (Hrsg.): *Handbuch Sozialarbeit/Sozialpädagogik*. Neuwied. 481–495.
- Herman, Joan L., Lynn Lyons Morris und Carol Taylor Fitz-Gibbon (1987a): *Evaluator's Handbook*. Newbury Park, Cal.
- Herman, Joan L., Lynn Lyons Morris und Carol Taylor Fitz-Gibbon (1987b): *How to Measure Attitudes*. 2. Auflage. Newbury Park, Cal.
- Herrmann, Joachim, und Christoph Höfer (1999): *Evaluation in der Schule. Unterrichtsevaluation. Berichte und Materialien aus der Praxis*. Gütersloh.
- House, Ernest R., und Kenneth R. Howe (1999): *Values in Evaluation and Social Research*. Thousand Oaks, Cal.
- House, Ernest R., und Kenneth R. Howe (o.J.): »Deliberative Democratic Evaluation Checklist.« Veröffentlicht auf den Internet-Seiten des Evaluation Center. www.wmich.edu/evalctr/checklists/dd_checklist.htm (16.6.2003).
- Kinast, Eva-Ulrike (1998): *Evaluation Interkultureller Trainings*. Lengerich.
- Kirkpatrick, Donald L. (1998): *Evaluating Training Programs. The Four Levels*. 2. Auflage. Berkeley, Cal.
- Kromrey, Helmut (2000): »Evaluation. Ein vielschichtiges Konzept. Begriff und Methodik von Evaluierung und Evaluationsforschung. Empfehlungen für die Praxis«. In: *Sozialwissenschaften und Berufspraxis* (24) 2 2000. 105–129.
- Mohr, Bernhard, und Mette Jacobsgaard (2003): *Appreciative Evaluation*. Taos, N.M.
- Morscher, Edgar (2003): »Evaluation aus Sicht der Ethik.« In: Josef Thonhauser und Jean-Luc Patry (Hrsg.): *Evaluation im Bildungsbereich. Wissenschaft im Dialog*. Studien zur Bildungsforschung und Bildungspolitik 22. Innsbruck. 35–65.
- Patton, Michael Quinn (2000): *Qualitative Research and Evaluation Methods*. 3. Auflage. Thousand Oaks, Cal.
- Pawson, Ray, und Nick Tilley (1997): *Realistic Evaluation*. London.
- Reichertz, Jo (o.J.): »Zur Gültigkeit von Qualitativer Sozialforschung«. In: *Forum Qualitative Sozialforschung/Forum Qualitative Social Research* (Online Journal) (1) 2 (o.J.). <http://qualitative-research.net/fqs-texte/2-00/2-00reichertz-d.htm> (16.6.2003).

- Reischmann, Jost (2002): *Weiterbildungs-Evaluation. Ein Lehrbuch für Praktiker mit zahlreichen Übungsaufgaben.* Neuwied.
- Rogge, Klaus (2001): »Qualitätsmanagement: Zwischen (Selbst-)Evaluation und Qualitätssicherung«. In: *Außerschulische Bildung. Materialien zur politischen Jugend- und Erwachsenenbildung* 2 2001. 393–402.
- Sanders, James R. (1999): *Handbuch der Evaluationsstandards. Die Standards des »Joint Committee on Standards for Educational Evaluation«.* Opladen.
- Schiele, Siegfried und Herbert Schneider (Hrsg.) (1996): *Reicht der Beutelsbacher Konsens?* Stuttgart.
- Shaw, Ian F. (1999): *Qualitative Evaluation: Introducing Qualitative Methods.* London.
- Siebert, Horst (1999): *Pädagogischer Konstruktivismus.* Berlin.
- Siebert, Horst (2002): *Der Konstruktivismus als pädagogische Weltanschauung.* Frankfurt am Main.
- Steinke, Ines (2000): »Gütekriterien qualitativer Forschung«. In: Uwe Flick, Ernst von Kardorff und Ines Steinke (Hrsg.): *Qualitative Forschung. Ein Handbuch.* Reinbek bei Hamburg. 319–331.
- Thonhauser, Josef, und Jean-Luc Patry (Hrsg.) (1999): *Evaluation im Bildungsbereich. Wissenschaft im Dialog.* Studien zur Bildungsforschung und Bildungspolitik 22. Innsbruck.
- Weiss, Carol H. (1998): *Evaluation. Methods for Studying Programs and Policies.* 2. Auflage. Upper Saddle River, N.J.
- Wottawa, Heinrich, und Heike Thierau (1998): *Lehrbuch Evaluation.* 2. vollständig überarbeitete Auflage. Bern.

Teil B

(Wirkungs-)Evaluation politischer Bildung in der Praxis

Evaluation von Demokratie- und Toleranzerziehungsprogrammen – die Erfahrungen aus dem Projekt »Erziehung zu Demokratie und Toleranz«

Katrin Uhl

Das Projekt »Erziehung zu Demokratie und Toleranz«

Mit dem Ziel, einen nachhaltigen Beitrag zur Förderung einer Kultur der Toleranz, der Gewaltlosigkeit und des demokratischen Zusammenlebens zu leisten, hat die Bertelsmann Stiftung 1995 das Projekt »Erziehung zu Demokratie und Toleranz« ins Leben gerufen. Sie arbeitet in diesem Projekt eng zusammen mit der »Bertelsmann Forschungsgruppe Politik« am Centrum für angewandte Politikforschung der Universität München.

Der Schwerpunkt des Projekts liegt darauf, internationale Erfahrungen und erfolgreiche Modelle zur Toleranzerziehung zusammenzutragen, handlungsorientierte Trainingskonzepte für Deutschland zu entwickeln bzw. adaptieren und Multiplikatoren in diesen Programmen auszubilden. Die Bildungsprogramme verfolgen einen erfahrungsorientierten Ansatz und versuchen mit Hilfe neuer pädagogischer Methoden dem Einzelnen einen aktiven Zugang zu den Themen Demokratie und Toleranz zu ermöglichen.

Die Zielgruppe der Programme ist im Laufe der Jahre ständig gewachsen und umfasst mittlerweile Schulen, Unternehmen, Universitäten, die außerschulische politische Bildung, die öffentliche Verwaltung und die Polizei. Im Zentrum der Arbeit stehen die folgenden drei Trainingsprogramme:

– *Achtung (+) Toleranz*

Das von der Bertelsmann Forschungsgruppe Politik entwickelte Programm »Achtung (+) Toleranz« setzt seinen Schwerpunkt auf die Förderung einer partnerschaftlichen Kommunikation in Konfliktsituationen und die Entwicklung von Zivilcourage. In einer Zusammenstellung von bewährten Methoden und speziell kreierten Aktivitäten schafft es Übungsräume, mit interkulturellen und gesellschaftlichen Konflikten deeskalierend und friedlich umzugehen. Die Dialog- und Vermittlungskompetenz der Teilnehmer wird in Rollen- und Planspielen, gruppendynamischen Gesprächsrunden und in Kleingruppen gefördert und trainiert. Das Programm ist mittlerwei-

le neben der klassischen außerschulischen politischen Bildung auch in öffentlichen Verwaltungen und bei der Polizei eingeführt worden.

– *Eine Welt der Vielfalt*

Das Programm »Eine Welt der Vielfalt« wurde von der 1913 in Chicago gegründeten Anti-Defamation League (ADL) konzipiert. 1985 entwickelte die ADL das Bildungs- und Seminarprogramm »A WORLD OF DIVERSITY®«, das bald in Schulen, Universitäten, Behörden und Unternehmen eine große Anzahl von Teilnehmern erreichte. Im Projekt »Erziehung zu Demokratie und Toleranz« wurde das erfolgreiche Programm für die Schule in Deutschland adaptiert, um wirksam Vorurteilen und Diskriminierung präventiv zu begegnen. Über Trainings wird das Programm an Lehrer weitervermittelt. Es kommt in der Grundschule und in der Sekundarstufe I zum Einsatz und thematisiert altersgerecht und interaktiv die eigene Identität, den Umgang mit (kultureller) Vielfalt, die Entstehung von Vorurteilen und Diskriminierung und ihre Konsequenzen.

– *Miteinander – Erfahrungen mit Betzavta*

»Miteinander« ist ein Programm zur Demokratieerziehung, das von Uki Maroshek-Klarman (Adam Institute for Democracy and Peace, Jerusalem) entwickelt und von der Bertelsmann Forschungsgruppe Politik adaptiert wurde. Das Programm basiert auf dem Grundsatz der Sicherung des Rechts aller Menschen auf Freiheit. In Übungen und spielerischen Aktivitäten soll ein »qualitatives Demokratieverständnis« geweckt werden. Hierzu werden auf vielfältige Weise externe Konflikte in interne Dilemmata umgewandelt. Diese Form der Bewusstseinsbildung steht im Mittelpunkt des Programms. Das Programm wird vorwiegend in den Bereichen außerschulische politische Bildung, Polizei und Schule angewendet.

Alltagsrelevantes Toleranzverständnis

Grundlage einer Erziehung zu Demokratie und Toleranz, wie sie in diesen pädagogischen Programmen verfolgt wird, ist ein praxisnahes und alltagsrelevantes Konzept von Toleranz, das Orientierung für verantwortungsbewusstes Handeln im Konfliktfall bietet. Toleranz definiert sich dort als die Maxime, in gegenseitiger Anerkennung einen Konflikt auszuhalten oder ihn friedlich zu regeln. So verstanden ist Toleranz keine Haltung, die – einmal erworben – als vorhanden betrachtet werden kann. Vielmehr besteht »Toleranzkompetenz« in der Bereitschaft und der Fähigkeit, in Konfliktsituationen mit den Bedürfnissen aller Beteiligten verantwortungsbewusst umzugehen (Bertelsmann Forschungsgruppe Politik 2001).

Toleranz macht es einem dabei nicht einfach – sie fordert das Wertesystem und die eigene Identität heraus. Toleranzerziehung muss deshalb beim Individuum ansetzen, den Einzelnen in seiner Identität ernst nehmen und stärken, bevor sie ihn mit Normen, Werten und Lebensentwürfen konfrontiert, die ihm fremd sind. Toleranzerziehung will vermitteln, dass in einem Konflikt stets mehrere Positionen und Handlungsoptionen zur Verfügung stehen.

Bei der Umsetzung dieses pädagogischen Anspruches gilt es, gezielt diejenigen sozialen Kompetenzen zu fördern, die tolerantes Handeln ermöglichen, etwa Empathie, die Fähigkeit zu Perspektivwechsel, partnerschaftlichem Dialog und friedlicher Konfliktlösung. Didaktische Konzepte, die auf mehreren Ebenen ansetzen und kognitive, emotionale und handlungsorientierte Elemente miteinander verbinden, haben sich hier als besonders erfolgreich erwiesen. Einen solchen Ansatz des erfahrungsorientierten Lernens verfolgen alle der hier vorgestellten Bildungsprogramme.

Was bleibt übrig? – Die Frage nach der Wirkung

Bereits die Entwicklung der Seminarkonzepte wurde durch eine kontinuierliche Evaluation der Testworkshops begleitet, um ein ausgereiftes Programm veröffentlichen zu können. Auch im Verlauf des Projekts wurden wiederholt Maßnahmen der begleitenden formativen Evaluation durchgeführt mit dem Ziel, Methoden, Didaktik und Zielbestimmungen der Programme zu verbessern.

Die stete Nachfrage nach den Programmen, ihre Verbreitung und Rezeption in der Fachwelt ließen vermuten, dass die Bertelsmann Stiftung und das Centrum für angewandte Politikforschung mit ihrer Arbeit hier eine richtige Richtung eingeschlagen hatten. Dennoch: Unbeantwortet blieb die Frage nach der tatsächlichen Wirkung der Programme.

Um diese Frage zu klären – oder eher, um sich einer Klärung dieser Frage anzunähern –, wurden für die Programme »Achtung (+) Toleranz«, »Eine Welt der Vielfalt« und »Miteinander – Erfahrungen mit Betzavta« jeweils unterschiedliche externe Evaluatorenteam¹ mit der Entwicklung eines Evalua-

¹ Die Evaluationen wurden durchgeführt von: Münchner Projektgruppe Evaluation (Prof. Dr. Werner Schneider; Dr. Andreas Hirseland; Dr. Oliver Dimbath; Julia von Hayek, Institut für Soziologie der Universität München) (»Achtung (+) Toleranz«); Univation – Institut für Evaluation und wissenschaftliche Weiterbildung e.V, Köln (»Eine Welt der Vielfalt«); Prof. Dr. Michael Bommes, Dr. Ulrike Wolff-Jontofsohn, Jochen Guter, Pädagogische Hochschule Freiburg (»Miteinander – Erfahrungen

tionsdesigns beauftragt. Die Evaluatoren erhielten den Auftrag zu untersuchen, inwiefern die Programme gemäß ihrer Intentionen wirken, sollten dabei jedoch auch nichtintendierte Nebenwirkungen in den Blick nehmen.

Im Mittelpunkt des Interesses der einzelnen Evaluationen standen die Aspekte Programmakzeptanz, Lernerfolg und Transfer – insbesondere die Wirkung der Programme auf Einstellungen und Verhalten –, Praxisbezug sowie die Nachhaltigkeit von Wirkungen. Neben dem prüfenden Blick auf die pädagogischen Programme sollten die Evaluationen auch einen Beitrag zur Diskussion um die Rollenbestimmung von Evaluation in der politischen Bildung leisten. Am Ende stehen sollten eine Programmbewertung, Empfehlungen zur Weiterentwicklung und die Erprobung und Reflexion neuer Wege der Evaluation politischer Bildung.

In Deutschland fehlt es bisher an einem umfassenden Konzept der Wirkungsevaluation von Demokratie- und Toleranzerziehung. Sowohl die Auftraggeber wie auch die Evaluatoren betraten mit ihrem Vorhaben somit relatives Neuland. Unstrittig war, dass eine gegenstandsadäquate Evaluation von Demokratie- und Toleranzerziehung nicht durch das reine Messen von Daten und eine neutrale, externe Beschreibung von Effekten erreicht werden kann. Ein solches Vorgehen würde weder dem Grundverständnis der Programme entsprechen noch relevante Ergebnisse produzieren.

Die Wirkungen von Programmen – oder gar deren Erfolg – ist nicht anhand vorgegebener Kriterien messbar. Sie sind stets an die (unterschiedliche) Perspektivität der Beteiligten, die darin enthaltenen Erwartungen, Interessen, Legitimations- und Handlungszwänge geknüpft. Die Evaluation von Bildungsmaßnahmen, deren Ziel Einstellungs- und Verhaltensänderungen beim Individuum sind, werfen somit sowohl methodische Fragen der Messbarkeit, wie auch normative Fragen der Bewertung bzw. Bewertbarkeit auf (Hirsland und Uhl 2004).

Vor diesem Hintergrund sollten bei der Entwicklung der Evaluationsdesigns die folgenden Punkte Berücksichtigung finden:

- partizipativer Charakter der Evaluation: Vertreter aller für die zu evaluierenden Programme relevanten Beteiligtegruppen (Stakeholder) sollten frühzeitig aktiv in den Prozess einbezogen werden, etwa bereits bei der Klärung der Zielkriterien (Ulrich und Wenzel 2003).
- Transparenz des Evaluationsprozesses: Offenlegung des Vorgehens und kontinuierliche kommunikative Validierung der Zwischenergebnisse mit den Beteiligtegruppen.

mit Betzavta«). Die Gutachten sind als Download erhältlich unter www.cap.uni-muenchen.de/bertelsmann/toleranz.htm.

- Wertpluralismus: Die Anerkennung eines Wertpluralismus, der den Bildungsprogrammen zugrunde liegt, und die damit verbundene Offenheit unterschiedlicher Wirkungen auf Individuen sollten reflektiert werden.

Die im Folgenden vorgestellten Evaluationen setzen diese Punkte auf unterschiedliche Weise um. Gemeinsam ist ihnen allen der Versuch eines umfassenden, gegenstandsadäquaten und partizipativen Vorgehens und die Bereitschaft, unterschiedliche Evaluationsmethoden zu erproben. Sie zeigen sich auf diese Weise sensibel für den Kontext des sozialen Lernens, in dem es keine einfachen Antworten auf die Frage nach dem Erfolg gibt.

Literatur

- Bertelsmann Forschungsgruppe Politik (Hrsg.) (2001): *Toleranz. Grundlage für ein demokratisches Miteinander*. 2. Auflage. Gütersloh.
- Bertelsmann Stiftung und Forschungsgruppe Jugend und Europa (Hrsg.) (2001): *Eine Welt der Vielfalt. Ein Trainingsprogramm des A WORLD OF DIFFERENCE®-Institute der Anti-Defamation-League, New York, in der Adaption für den Schulunterricht. Praxishandbuch für Lehrerinnen und Lehrer*. 2., überarbeitete Auflage. Gütersloh.
- Hirseland, Andreas, und Katrin Uhl (2004): »Evaluationsmethoden für Modelle zur Förderung von Zivilcourage in der politischen Bildung«. In: Gerd Meyer, Siegfried Frech, Günther Gugel und Ulrich Dovermann (Hrsg.): *Zivilcourage lernen. Analysen, Modelle, Arbeitshilfen*. Berlin. 384–399.
- Miteinander – Erfahrungen mit Betzavta. Praxishandbuch für die politische Bildung*. Ein Praxishandbuch auf der Grundlage des Werks »Miteinander« von Uki Maroshek-Klarman, Adam Institute, Jerusalem, in der Adaption von Susanne Ulrich, Thomas R. Henschel und Eva Oswald. 3. Auflage. Gütersloh 2001.
- Ulrich, Susanne (2001): *Achtung (+) Toleranz. Wege demokratischer Konfliktregelung*. Praxishandbuch für die politische Bildung. Unter Mitarbeit von Jürgen Heckel, Eva Oswald, Stefan Rappenglück und Florian M. Wenzel. 2. Auflage. Gütersloh.
- Ulrich, Susanne, und Florian M. Wenzel (2003): *Partizipative Evaluation. Ein Konzept für die politische Bildung*. Gütersloh.

Evaluation des Programms

»Achtung (+) Toleranz«

– ein Praxisbericht

Andreas Hirsland, Oliver Dimbath, Julia von Hayek, Werner Schneider

Einleitung

»What is wrong with Tolerance?« – so fragte jüngst Slavoj Žižek in einem Beitrag zu den Antinomien der postmodernen Gesellschaft. Žižek wendet sich pointiert gegen jene Form von postmoderner Toleranz, die, auf ihrer Vorderseite als Multiperspektivität mit oberflächlicher Kompromissbereitschaft daherkommend, auf ihrer Rückseite – z. B. im Gewand der »political correctness« – mit konsequenter Intoleranz gegenüber jeglichem Beharren auf verschiedenen Formen charakteristischer Andersartigkeit vorgeht (Žižek 2001: 339 ff.).

Dieser Gedanke ist insofern von Interesse, als er deutlich zeigt, dass der Begriff Toleranz zu einem flottierenden Signifikanten geworden ist, auffüllbar mit vielfältigen Bedeutungsgehalten und instrumentalisierbar für die verschiedensten politisch-praktischen Ziele: z. B. für den Schutz der Nichtraucher ebenso wie für eine Rettung des Genussrauchens usw. – bis hin zum zirkulären Diktum der »zero tolerance« gegenüber Intoleranz. Vor diesem Hintergrund wird die besondere Problemlage der Evaluation eines Programms wie dem im Rahmen eines Kooperationsprojekts der Bertelsmann Stiftung und des Centrums für angewandte Politikforschung entwickelten Trainingsprogramm »Achtung (+) Toleranz« deutlich, das sich die Förderung von Toleranz und Konfliktfähigkeit zum Ziel gesetzt hat.

Im Falle der Toleranz und ihrer Förderung durch Maßnahmen der politischen Bildung ist gerade wegen der relativen Unbestimmtheit des zu untersuchenden und in seinen Auswirkungen zu bewertenden Konstrukts also keineswegs davon auszugehen, dass das, was die Programmentwickler mit ihrem Konzept verbinden, von vornherein deckungsgleich ist mit dem, was sich eine interessierte Öffentlichkeit und auch die Evaluatoren darunter vorstellen mögen.

Im Gegensatz etwa zur Feststellung von Lernerfolgen einer Softwareschulung, deren Erfolgskriterium in Form der Beherrschung des zu vermitteln-

den Computerprogramms relativ eindeutig und messbar erscheint, ist die Feststellung einer Fähigkeit zu Toleranz und Konfliktfähigkeit als soziale Kompetenzen in hohem Maße von historischen, kulturellen und situativen Faktoren sowie vor allem von (manifesten und latenten) subjektiven Situationsdefinitionen abhängig, die sich einer einfachen Operationalisierung und Messbarkeit zunächst prinzipiell verweigern.

Diese Ausgangslage war auch der Bertelsmann Stiftung als Auftraggeberin der hier vorgestellten Evaluation bewusst, als sie der Münchner Projektgruppe Evaluation im Spätsommer 2001 den Auftrag erteilte, die Wirkungen des Bildungsprogramms »Achtung (+) Toleranz« zu evaluieren.

Gegenstand der Evaluation

»Achtung (+) Toleranz« ist ein Angebot zur individuellen Demokratie- und Toleranzbildung, das sich an ältere Jugendliche und Erwachsene richtet. Es zielt darauf ab, über eine kurzfristige Symptombekämpfung hinaus präventiv zu wirken, indem einerseits soziale Interaktionskompetenz gefördert und ausgebaut, andererseits zu einer selbstreflektierten Orientierung in Konfliktsituationen befähigt werden soll.

Damit wird dem Umstand Rechnung getragen, dass die vielfältigen alltäglichen Toleranzprobleme sich in Form von offenen und verdeckten Konflikten zeigen. Konflikte bilden somit zugleich die Voraussetzung als auch den konkreten und handlungsrelevanten Anwendungs- und Bewährungsfall für das Auftreten toleranten Handelns. Dementsprechend setzt das die Programmkonzeption leitende Toleranzverständnis von »Achtung (+) Toleranz« am Vorliegen eines Konflikts an:

»1. Voraussetzung: Nur in einem Konfliktfall stellt sich die Frage nach Toleranz, denn nur dann werden eigene Deutungsmuster, Werte oder Normen durch die Konfrontation mit Andersartigkeit angezweifelt, verletzt oder es treffen konkurrierende Interessen aufeinander.

2. Vorgehensweise: Das zweite Kriterium zur Bestimmung von Toleranz ist die Abwesenheit von Gewalt im Konfliktfall. Das gewaltlose Vorgehen kann entweder von einer Seite allein – im Sinne des Aushaltens des Konflikts – oder von den Konfliktpartnern miteinander erfolgen.

3. Motivation: Die Motivation, die dem Denken und Handeln in einem Konflikt zugrunde liegt, stellt das dritte und wesentlichste Unterscheidungskriterium für die Identifizierung von Toleranz dar. Nur auf der Grundlage der prinzipiellen Anerkennung des gleichen Rechts auf freie Entfaltung wird es dem Einzelnen möglich, aus Einsicht Andersartigkeit auszuhalten oder die

gemeinsame Regelung eines Konflikts zu suchen.« (Feldmann, Henschel und Ulrich 2000: 15)

Damit verweist das dem Programm zugrunde liegende Toleranzverständnis über das bloße Vorhandensein toleranter Einstellungen hinaus auf die Notwendigkeit der Fähigkeit zu tolerantem Handeln. Davon ausgehend räumt das methodisch-didaktische Konzept von »Achtung (+) Toleranz« dem Erfahrungslernen einen zentralen Ort im Seminaaraufbau ein.

Nicht die Präsentation theoretischer Einsichten und begrifflich gefasster Wissensbestände steht im Vordergrund, sondern die in und durch die entsprechenden Übungen vermittelten Erfahrungen, die sich in widersprüchlichen und widerstreitenden, d.h. auch konflikthafter Wahrnehmungen eigener und fremder Standpunkte äußern. Ergänzt wird diese Erfahrungsebene durch zwischengeschaltete Informations- und Reflexionsphasen, in denen das in den Übungen und Trainingseinheiten Erlebte und Erfahrene zu (Selbst-)Erkenntnissen verarbeitet werden und zugleich auch Anwendung finden soll. Das Seminar, begriffen als eine Situation, in der gehandelt werden muss, ist somit zugleich durch »learning by doing« ein Anwendungsfeld des zu Erlernenden.

Dabei setzt die Fähigkeit zu tolerantem Verhalten voraus, dass Individuen bei der Beurteilung einer Handlungs- oder Konfliktsituation die Perspektive des anderen in einer nicht-instrumentellen Weise in Rechnung stellen können – das ist der Aspekt der Achtung. Darüber hinaus müssen sie zugleich in der Lage sein, die dem eigenen und fremden Urteilen und Handeln zugrunde liegenden, vielfach nichtbewussten Vorannahmen und Vorurteile zu erkennen und wertreflexiv zu befragen. Durch eine entsprechende Sensibilisierung sollen die Teilnehmer in die Lage versetzt werden, offen und teils auch verdeckt diskriminierende oder anderweitig benachteiligende konflikthafte Handlungsstrukturen zu erkennen und sich unter Anwendung der im Seminar vermittelten Handlungs- bzw. Kommunikationskompetenzen tolerant-zivilcouragiert zu verhalten.

Vor diesem Hintergrund besteht das Ziel von »Achtung (+) Toleranz« darin, Seminarteilnehmern in Gruppenübungen eine erfahrungsbezogene Auseinandersetzung mit dem skizzierten Toleranzkonzept zu ermöglichen, ihnen Wege der gewaltfreien Konfliktbewältigung aufzuzeigen und damit toleranzfördernde Kompetenzen zu vermitteln. Im Einzelnen sind dies:

- Dialog und Kommunikationskompetenz
- die Fähigkeit, sich in den Standpunkt eines anderen zu versetzen
- die Fähigkeit, Modelle konstruktiver und demokratischer Konfliktregelung anzuwenden (Feldmann, Henschel und Ulrich 2000: 25)

Evaluationsauftrag

Vor dem Hintergrund der genannten Voraussetzungen umfasste der Evaluationsauftrag eine doppelte Zielvorgabe: Es sollte sowohl die Wirkung des Toleranz- und Demokratiebildungsprogramms in Bezug auf die teilnehmenden Individuen ermittelt, als auch nach neuen Wegen und Verfahren einer gegenstandsangemessenen Wirkungsforschung gesucht werden. Damit bewegt sich die Evaluation von »Achtung (+) Toleranz« im Spannungsfeld zwischen Evaluation und Evaluationsforschung.

Am Ende des Projekts sollten – neben einem Erfahrungsbericht zum methodischen Vorgehen – nicht nur eine Programmbewertung erfolgen, sondern darüber hinaus Empfehlungen zur Weiterentwicklung des Programms gegeben werden.

Damit verfolgte die Evaluation sowohl summative als auch formative Interessen, die durch ein in enger Zusammenarbeit mit dem Auftraggeber und den Programmentwicklern erfolgtes dialogisch-diskursives Vorgehen eingelöst wurden. Darunter ist ein Vorgehen zu verstehen, das über den partizipativen Austausch und Abgleich der Interessen der am Evaluationsprozess Beteiligten hinaus eine Komponente wissenschaftlicher Beratung und aufklärerischer Analyse beinhaltet.

Als Ergebnis dieses gemeinsamen, bereits zu Beginn der Evaluation eingeleiteten Aushandlungsprozesses richtete sich das Evaluationsinteresse auf folgende Bereiche:

- auf die Annahme des Programms durch die Teilnehmer (Akzeptanz)
- auf die im Seminar angeeigneten Inhalte (Wissens- und Kompetenzerwerb)
- auf die Relevanz der Lerninhalte für die alltäglichen Lebenswelten der Teilnehmer (Transfer)
- auf mögliche Effekte bei den Teilnehmern, die von den Programmanbietern nicht erwartet wurden (nicht intendierte Nebenfolgen)
- und – gewissermaßen als Querschnittsfrage – auf nachhaltige, also die unmittelbare Seminarteilnahme zeitlich überdauernden Wirkungen des Programms (Nachhaltigkeit)

Eine diesen Ansprüchen angemessene Wirkungsevaluation des Programms steht somit nicht nur vor dem Problem, wie die von den Programmentwicklern und -autoren intendierten Wirkungen zu erfassen sein könnten. Sie muss vielmehr gleichermaßen auch die mit der Programmteilnahme verbundenen subjektiven Erfahrungen und Bedeutungszusammenhänge durch die Teilnehmer erfassen.

Evaluationsdesign und -ablauf

Bei Toleranz handelt es sich um ein vieldeutiges Phänomen, das von den einzelnen Teilnehmern unterschiedlich gedeutet und so mit eigensinniger alltagspraktischer Relevanz versehen werden kann. Eine angemessene Beschreibung und Dokumentation möglicher Wirkungen von »Achtung (+) Toleranz« lässt sich daher durch ein offenes und mehrstufiges Forschungsdesign erreichen (für eine ausführlichere kritische Diskussion des Evaluationsdesigns unter methodischen und anwendungsbezogenen Gesichtspunkten siehe den Methodenbeitrag in diesem Band).

Die als erster Schritt durchgeführte Analyse vorliegender Programmmaterialien ermöglichte im Zusammenhang mit Expertengesprächen (Auftraggeber, Programmentwickler, Trainer) die Rekonstruktion der Entstehungsgeschichte und, damit verbunden, der theoretischen und didaktischen Grundlagen und Postulate (etwa zum Toleranz- und Kommunikationsverständnis) des Programms.

Zudem gewährten die Experteninterviews einen ersten Einblick in die Seminarpraxis und die aus Expertensicht wahrgenommenen Besonderheiten von »Achtung (+) Toleranz«. Dies eröffnete einen Zugang zu wichtigen, für das Handeln von Trainern typischen Intentionen und Relevanzsetzungen, d.h. dazu, wie Seminarleiter aus ihrer Sicht ihr Erfahrungswissen einsetzen, um das Seminargeschehen zu dem von ihnen verfolgten Konzept in Beziehung zu setzen.

Auf Grundlage der Kenntnis der durch einen Experten-Filter wahrgenommenen Seminarwirklichkeit und der Analyse der Seminarmaterialien wurde dann in einem zweiten Schritt ein Mehrebenendesign entwickelt. Es wurde erstens ein teilstandardisierter Fragebogen an alle erreichbaren Teilnehmer bisheriger Seminare verschickt. Zudem wurden zweitens in einer qualitativen Fallstudie mit allen Teilnehmern einer Seminargruppe (elf Teilnehmer) jeweils kurz vor sowie etwa drei Monate nach der Seminarteilnahme problemzentriert-diskursive Interviews und drittens ein Video-»Experiment« mit anschließenden Gruppengesprächen durchgeführt.

Die retrospektive Fragebogenerhebung zielte mit einer Mischung aus geschlossenen und offenen Fragen insbesondere auf die Überprüfung der Nachhaltigkeit des erworbenen Seminarwissens sowie auf Akzeptanzurteile gegenüber dem Seminarablauf und -konzept. Es wurden 169 ehemalige Seminarteilnehmer angeschrieben. Die um die unzustellbaren Sendungen bereinigte Rücklaufquote betrug 46 Prozent.

Die Auswertung des Fragebogens erfolgte für die geschlossenen Fragen mittels deskriptiv statistischer sowie faktoranalytischer Verfahren, für die of-

fenen Fragen durch ein an den Prinzipien der Grounded Theory (Glaser und Strauss 1998) angelehntes Codierverfahren.

Die mit den insgesamt elf Teilnehmern eines eigens für Evaluationszwecke veranstalteten Seminars durchgeführte qualitative Fallstudie gliedert sich in zwei Schritte: eine Vorab-Erhebung und eine im Abstand von etwa drei Monaten nach Seminarbesuch durchgeführte Nacherhebung.

Im Rahmen der Vorab-Erhebung wurden alle Teilnehmer jeweils kurz vor dem Seminar in problemzentriert-diskursiven Interviews zu ihrem Toleranz- und Konfliktverständnis sowie zu ihren Vorerfahrungen mit ähnlichen Seminaren und nach ihren Erwartungen an das kommende Seminar befragt. Entsprechende Belegerzählungen wurden dokumentiert und in einem diskursiven Teil durch konfrontative Meinungsfragen explizite Bewertungen und (latente) Deutungsmuster erhoben.

Zudem nahmen die Teilnehmer unmittelbar vor Seminarbeginn an einem Videoexperiment teil, dessen Inhalt darin bestand, ein gemeinsames Handlungsproblem zu lösen. Der Verlauf wurde auf Video dokumentiert und in einer anschließenden (wörtlich protokollierten) Gruppendiskussion reflektiert.

Die Nacherhebung wies einen umgekehrten Ablauf auf: Ungefähr drei Monate nach dem Seminarbesuch wurde die Gruppe erneut gebeten, in einer Gruppendiskussion Seminarerfahrungen zu reflektieren. Anschließend erfolgte eine Vorführung der Dokumentation des ersten Videoexperiments mit einer anschließenden weiteren Gruppendiskussion und der Bitte, das Videoexperiment mit einer formal ähnlichen, aber inhaltlich anderen Aufgabenstellung zu wiederholen und die gemachten Erfahrungen in einer erneuten Diskussionsrunde auszutauschen.

Den Abschluss der Nacherhebungsphase bildeten wiederum Einzelinterviews mit allen Seminarteilnehmern in einem zeitlichen Abstand von vier bis fünf Monaten nach Seminarbesuch. Gegenstand dieser Interviews waren in und nach dem Seminar gemachte Erfahrungen, die in Zusammenhang mit den Seminarinhalten standen, Re-Interpretationen von im Erstinterview berichteten Überlegungen und Begebenheiten sowie ein weiterer diskursiver Interviewteil.

Insgesamt betrug der zeitliche Rahmen der Evaluation von der Entwicklung des Designs bis zur Erstellung und Präsentation des Abschlussberichtes zwölf Monate. Einen Überblick über Aufbau und Ablauf der Evaluation gibt die folgende Abbildung:

Ausgewählte Ergebnisse der teilstandardisierten postalischen Umfrage

Im Folgenden werden die Evaluationsergebnisse anhand ausgewählter Aspekte diskutiert. Aufgrund der Komplexität der Evaluation kann an dieser Stelle nicht im Einzelnen auf alle Ergebnisse eingegangen werden. Eine detaillierte statistische Beschreibung und Interpretation der erhobenen Daten findet sich in unserem zweiten Zwischenbericht sowie dem Abschlussbericht, der unter www.cap.uni-muenchen.de/bertelsmann/toleranz.htm eingesehen werden kann. An dieser Stelle werden wichtige Befunde exemplarisch vorgestellt.

Allgemeine Akzeptanz

Die Fragebogenbefragung der ehemaligen Teilnehmer von »Achtung (+) Toleranz«-Seminaren machte deutlich, dass sowohl das Seminarkonzept, dessen konkrete Durchführung und die spätere alltagspraktische Anwendbarkeit der Seminarinhalte von einer großen Mehrzahl der Respondenten positiv eingeschätzt werden – das Programm mithin bei den Teilnehmern hohe Akzeptanz erfahren hat. So liegen etwa die Zustimmungswerte zu Statements wie »Ich würde an einem Seminar dieser Art nochmals teilnehmen« oder »Die Teilnahme an diesem Seminar kann ich weiterempfehlen« bei jeweils weit über 70 Prozent.

Eine detaillierte Analyse der etwas geringeren, aber immer noch sehr hohen Zustimmungswerte zu Fragen nach der Zufriedenheit mit dem eigenen Seminarerleben legt jedoch eine etwas differenziertere Sicht der Akzeptanz von Seminaraufbau und -inhalten nahe. Zwar bewertet eine Mehrzahl der Teilnehmer den Seminarbesuch als Beitrag zur persönlichen Weiterentwicklung, jedoch liegt die Zahl derer, die ihre Erwartungen an das Seminar insgesamt und die Lern- und Wissensinhalte im Besonderen »voll und ganz« erfüllt sahen, etwas niedriger.

Eine plausible Interpretation dieser Diskrepanz ist, dass sich hier eine retrospektive Bewertungsambivalenz ausdrückt: Zwar war die Seminarteilnahme eine sehr positive persönliche Erfahrung, jedoch hätte aus Sicht der Teilnehmer die konkrete Seminargestaltung inhaltlich in Bezug auf die Vermittlung manifester Wissensbestände noch konturierter sein können.

Das sich hier artikulierende Teilnehmerbedürfnis scheint jedoch nur geringen Einfluss auf die wahrgenommenen Auswirkungen der Seminarteilnahme auf das eigene Alltagsleben zu haben. Denn hier wiederum erzielen

Fragen nach der Anwendbarkeit der Seminarinhalte auf die Felder Beruf, Privates (Familie/Verwandtschaft/Partnerschaft) und öffentliche Konfliktsituationen hohe Zustimmungswerte, allerdings mit einer starken Ausrichtung auf das berufliche Feld. Dies ist insofern wenig überraschend, als ein Großteil der Befragten Teilnehmer von Trainerausbildungen waren, also selbst potenzielle Anbieter von »Achtung (+) Toleranz«-Trainings.

Da ablehnende Äußerungen (»Kann ich überhaupt nicht anwenden«) bei allen Items nur geringe Werte erzielten, lässt sich durchaus eine von den Teilnehmern wahrgenommene Praxistauglichkeit des erworbenen Seminarwissens feststellen, die sich durch gezielte Modifikationen der Übungseinheiten in Bezug auf verschiedene Praxisfelder sicher noch steigern ließe.

Jedoch gilt es zu bedenken, dass derartige itembasierte Selbstauskünfte noch keinen hinreichenden Aufschluss über die tatsächlichen Transfers von Seminarinhalten in den Alltag geben. Sie können nach unserer Auffassung vielmehr als Indikatoren für die retrospektive Akzeptanz der eigenen Teilnahme und des inhaltlichen Seminarangebots als sinnvolle Bildungsinvestition, die »etwas gebracht hat«, gelten.

Wissens- und Kompetenzerwerb/Nachhaltigkeit

Lassen sich die zuletzt vorgestellten Daten durchaus als einen noch wenig spezifischen, nur an der den Befragten reflexiv zugänglichen Oberfläche ansetzenden Hinweis auf einen andauernden (nachhaltigen) Wissens- und Kompetenzerwerb deuten, so ergibt die Auswertung der offenen Fragen des Fragebogens genaueren Aufschluss über Art und Richtung der Bildungserfahrung. Von besonderer Bedeutung hierfür war die im Fragebogen enthaltene offen gestellte Frage, ob die Teilnehmer im Verlauf des Seminars »Aha-Erlebnisse« oder »Schlüsselerfahrungen« hatten.

Diese Frage zielt auf die Feststellung nachhaltiger Seminarerfahrungen, denn im Alltagsverständnis werden »Aha-Erlebnisse« mit scheinbar unvermittelt auftretenden Erkenntnissen und Einblicken gleichgesetzt, nach deren Eintreten man sich selbst und die Welt (oder zumindest Teile von beidem) in einem »anderen Licht« sieht. Daher bildet die Frage einen guten Indikator für die subjektiv wahrgenommene Nachhaltigkeit der Seminarteilnahme bzw. von Seminarerfahrungen und dies unabhängig von den jeweiligen berichteten Erfahrungsinhalten. Deren Bandbreite wiederum gestattet Rückschlüsse auf die Vielfalt subjektiver Wirkungen.

Die oben genannte Frage wurde von über 80 Prozent bejaht. Die Gruppe derjenigen, die angaben, im Seminar keine derartigen Erlebnisse und Erfah-

rungen gemacht zu haben, bestand größtenteils aus Teilnehmern, die an diesem oder einem vergleichbaren anderen Programm bereits mehrfach teilgenommen hatten. Die hohe Zustimmung zur Frage weist darauf hin, dass das Seminar nachhaltig etwas bewirkt hat. Eine detaillierte, an das Verfahren der Grounded Theory angelehnte Auswertung der einzelnen Antworten gestattet Aussagen über die inhaltliche Richtung der Seminarwirkungen.

Ganz im Sinne der Programmintention von »Achtung (+) Toleranz« ist der Befund zu verstehen, dass mehr als doppelt so viele Antworten vom Standpunkt des Eigenerlebens (Selbstbezug) aus formuliert wurden als aus einer neutralen Beobachterperspektive (Fremdbezug). Dies kann als Indiz dafür gedeutet werden, dass ein Großteil der Befragten im Sinne des didaktischen Seminarkonzepts, den Einzelnen mit seiner jeweiligen Erfahrungswirklichkeit einzubinden und Wirkungen mit Relevanz für das jeweilige Selbsterleben hervorzurufen, erreicht wurden.

Bei den Antworten stehen die durch bestimmte Seminarteile hervorgerufenen Gruppenprozesse im Vordergrund, die den Ausgangspunkt allgemeinerer Reflexionen über Ablauf und Dynamik (konflikthafter) Interaktionsprozesse, über Vorurteilsstrukturen und Identitäten sowie über die Logik des Seminarablaufs bilden. Zudem dokumentieren die Antworten Bewusstwerdungsprozesse, die unmittelbar an eines der Seminarziele – die Vermittlung eines Unterschiedes zwischen Toleranz und Scheintoleranz – anknüpfen und sich in Form eines veränderten, differenzierteren und subjektiv als tiefer empfundenen Toleranzverständnisses zeigen. Ebenso werden Verhaltensänderungen berichtet, die zu reflexiv veränderten Handlungsweisen in Konfliktsituationen oder zu einer Neuausrichtung von Handlungsmotiven entlang einer Unterscheidung von ziel- und bedürfnisorientiertem Handeln führen.

Im Zusammenhang mit derartigen Bewusstwerdungsprozessen kam es auch zu Irritationen bislang interpretations- und handlungsleitender Grundüberzeugungen, die sich reflexiv in Form einer Einsicht in moralische Dilemmata artikulieren. So wird davon berichtet, das Seminar habe einerseits dazu geführt, eine relativistische moralische Perspektive einzunehmen, dies sei aber andererseits Grundlage für ein erweitertes, liberale(re)s Toleranzverständnis: »Es gibt kein Richtig & kein Falsch war eine Erkenntnis, die mir die Toleranz gab, auch völlig andere Lebensweisen zu akzeptieren« (Fall 57/C3.1).

Ein weiterer Schwerpunkt der berichteten selbstbezogenen Erlebnisse und Erfahrungen liegt jedoch im Bereich der Selbsterfahrung bzw. Selbsterkenntnis. Hier spielen emotionale Prozesse eine gewichtige Rolle, die durch intensive, mittels gemeinsamer Übungen in den verschiedenen Seminarteilen vermittelte Erfahrungen von Vergemeinschaftung hervorgerufen werden.

Ihren Ausdruck finden sie in berichteten emotionalen Erlebnissen wie Geborgenheit und der Möglichkeit, Anerkennung zu erfahren: »Es gab eine unheimliche Nähe innerhalb der Gruppe, Intensität, ich habe mich aufgehoben gefühlt und inhaltlich viel mitgenommen und gelernt« (Fall 47/C3.1). Mit dem angesprochenen Themenfeld werden nicht nur Erfahrungen als Seminarwirkungen angesprochen, sondern zugleich indirekt auch die Wirkungsweise des Seminars.

Wirkungsweise des Seminarkonzepts

Im Kontext einer Evaluation von Programmen der politischen Bildung ist es methodisch erforderlich, die beobachtbaren oder dokumentierbaren Wirkungen in einen Zusammenhang mit dem zu untersuchenden Programm zu bringen. Aus einer Vielzahl methodischer, forschungspraktischer und forschungsethischer Gründe ist es in diesem Feld nicht möglich, strikt kausale, im logischen Sinne eindeutige Beziehungen zwischen der Teilnahme an einem Bildungsprogramm und den beobachteten oder berichteten Wirkungen festzustellen. Jedoch ist es durchaus möglich, auf der Grundlage rekonstruierbarer, von den Teilnehmern explizit oder implizit vermittelten Deutungszusammenhänge nach theoretisch plausibilisierbaren Beziehungen zwischen Programm und Wirkung zu suchen.

Erfahrungslernen bedeutet im Zusammenhang der Konzeption von »Achtung (+) Toleranz«, dass es der Seminarleitung gelingen muss, über die Vermittlung explizierten Wissens hinaus durch eine vertrauensvolle Gruppenatmosphäre einen Handlungsrahmen zu schaffen, der es den Teilnehmern im Gegensatz zu den meisten alltäglichen Handlungsfeldern ermöglicht, einander etwa durch gezielte Feedback-Übungen ihre wechselseitigen Eindrücke zu vermitteln, ohne dass dies als Angriff auf die Integrität der eigenen Persönlichkeit erfahren würde.

Gelingt dies, entsteht ein Freiraum, der es den Teilnehmern ermöglichen kann, sich selbst eigene Handlungsgrundsätze explizit zu machen oder ohne Gesichtsverlust eigene blinde Flecken sowohl im Umgang mit anderen als auch in der eigenen Weltsicht sowie vorhandene Stereotype und Vorurteile zu erkennen und zum Gegenstand einer Auseinandersetzung mit sich selbst zu machen. Dabei spielen die im Seminarprozess auftretenden Konflikte eine wichtige Rolle, da sie Differenz erfahrbar machen und zugleich Gegenstand der Einübung eines toleranten Umgangs mit Differenz darstellen.

Aus diesem Grund umfasste der Fragebogen eine Reihe von Frageblättern, die sich auf das Seminargeschehen selbst bezogen und die Rückschlüsse

auf die Wirksamkeit des dem didaktischen Konzept unterlegten Erfahrungslernens im Seminar zulassen. Wir haben versucht, das Seminarsgeschehen z.B. in den wechselseitig aufeinander verweisenden Dimensionen »Teamwahrnehmung« (wie haben die Teilnehmer das Trainerteam wahrgenommen) und »Konfliktwahrnehmung« (wie haben die Teilnehmer die im Seminarablauf auftretenden Konflikte wahrgenommen) zu erfassen.

Teamwahrnehmung

Mit Hilfe der explorativen Faktoranalyse konnten aus den erhaltenen Antworten zu Fragen nach der erinnerten Wahrnehmung des Leitungsteams (»Wie hat das Leitungsteam in Ihren Augen gearbeitet?«) zwei Dimensionen extrahiert werden: eine inhaltlich-didaktische (Inhaltsdimension) und eine »atmosphärische«, auf den Umgang zwischen Team und Teilnehmern sowie an der Ausgestaltung des Seminarablaufs hin orientierte (Beziehungsdimension).

Für eine »good practice« der Seminarleitung sind beide Dimensionen gleichermaßen von Bedeutung. Die inhaltlich-didaktische Dimension steht für die Qualität des sachlichen »Inputs«, die Vermittlung von Inhalten, während die Beziehungsdimension mit dazu beiträgt, eben jene Atmosphäre zu schaffen, die es mit ermöglicht, dass sich die Teilnehmer auf die Inhalte und das angesprochene Erfahrungslernen einlassen können. Im Idealfall also müssten folglich beide Dimensionen hoch bewertet werden.

Da die beiden genannten Dimensionen nur analytisch voneinander zu trennen sind, in Wirklichkeit aber stets in unterschiedlichen Kombinationen miteinander auftreten, lassen sich vier Typen der Teamwahrnehmung unterscheiden:

Tabelle 1: Typen der Teamwahrnehmung

		<i>Inhaltsdimension</i>	
		niedrig	hoch
<i>Beziehungsdimension</i>	niedrig	Typ 1: Ablehnung 13 Prozent	Typ 2: inhaltlich- didaktische Akzeptanz 22 Prozent
	hoch	Typ 3: soziale Akzeptanz 24 Prozent	Typ 4: volle Akzeptanz 41 Prozent

Typ 1 (Ablehnung):

Sowohl auf der Inhalts- als auch auf der Beziehungsebene – dem persönlichen Umgang mit der Gruppe – wird das Team als wenig kompetent wahrgenommen. Aus Sicht der Teilnehmer wirkt das Team auf beiden Ebenen wenig überzeugend.

Typ 2 (didaktisch-inhaltliche Akzeptanz):

Auf der inhaltlich-didaktischen Ebene wird das Team als kompetent wahrgenommen, jedoch wird die Beziehungsebene eher negativ bewertet. Dies entspricht dem Typ eines eher kühl-professionellen Teams, bei dem der empathische Umgang mit Stimmungen und Bedürfnissen der Teilnehmer vermisst wird.

Typ 3 (soziale Akzeptanz):

Hier wird das Team auf der inhaltlich-didaktischen Ebene als wenig kompetent empfunden, findet jedoch auf der für die Seminaratmosphäre wichtigen Beziehungsebene Anerkennung. Aus Sicht der Teilnehmer wirkt das Team »nett« und »freundlich«, aber Darstellung und Vermittlung von Inhalten werden als weniger überzeugend empfunden.

Typ 4 (volle Akzeptanz):

Sowohl auf der inhaltlich-didaktischen als auch auf der Beziehungsebene liegen hohe Bewertungen vor. Das Team wird auf beiden Dimensionen als (hoch) kompetent wahrgenommen. Hier liegt der Typ eines aus Sicht der Teilnehmer sehr gut arbeitenden Teams vor.

Aus der Tabelle geht hervor, dass die Mehrzahl (41 Prozent) der Respondenten das Team auf der Inhalts- wie der Beziehungsebene hoch bewertet (Typ 4). Die beiden Typen 2 und 3, die durch eine diskrepante Kompetenzgewichtung charakterisiert sind, sind mit 24 Prozent bzw. 22 Prozent im Sample ungefähr gleich häufig vertreten, während lediglich 13 Prozent der Befragten dem Team auf beiden Ebenen eine geringe Kompetenz zuschreiben (Typ 1).

Die Auswertung hat ergeben, dass die Teilnehmerbewertungen zwischen beiden Dimensionen zum Teil stark unterschiedlich ausfallen. Interessant ist, dass diese Unterschiede keine durchschnittliche (etwa für Frauen oder Männer typische) Tendenz der Teilnehmergruppen widerspiegeln. Die Detailanalyse hat gezeigt, dass Unterschiede in der Wahrnehmung hauptsächlich mit der Teilnahme an bestimmten Seminarveranstaltungen zusammenhängen, mithin Unterschiede der Interaktion zwischen Trainerteams und der jeweiligen Teilnehmergruppe repräsentieren.

Konfliktwahrnehmung

Das didaktische Konzept von »Achtung (+) Toleranz« stellt darauf ab, dass im Seminarverlauf auftretende Konflikte und ihre jeweilige Bearbeitung im Gruppenprozess einen zentralen Aspekt des angestrebten Erfahrungslernens darstellen. So sind einzelne Programmteile darauf ausgelegt, Meinungsverschiedenheiten unter den Teilnehmern hervorzurufen, die einen konflikthafte Verlauf nehmen können und im Gruppenprozess bewältigt werden sollen. Zudem können im Seminarverlauf zusätzliche nichtintendierte Konfliktsituationen zwischen einzelnen Teilnehmern sowie zwischen Teilnehmern und Seminarleitung entstehen. Diese Konflikte bilden bereits einen »Anwendungsfall« für das im Seminar vermittelte reflexive und praktische Wissen.

Aus diesem Grund haben wir in den Fragebogen eine Reihe von Fragen aufgenommen, die sich auf die Erinnerung an im Seminarverlauf aufgetretene und durchlebte Konflikte beziehen. Aus den Antworten hierzu ließen sich Rückschlüsse auf die Wirkung des methodisch-didaktischen Programmkonzepts ziehen.

Die Mehrzahl der Untersuchungsteilnehmer (78 Prozent) gab an, sich an Konflikte erinnern zu können, die in den von ihnen besuchten Seminaren auch thematisiert wurden. Dies kann als erster Hinweis dafür gelten, dass das Seminarziel einer konfliktbezogenen Wahrnehmungssensibilisierung gegriffen hat.

Darüber hinaus zeigte die faktoranalytische Untersuchung der erhaltenen Antworten, dass diese Konflikte nicht einfach eindimensional – etwa als »gut« oder »schlecht« – wahrgenommen wurden, sondern entlang zweier sich gegenüberstehender Dimensionen. Deren inhaltliche Interpretation ergab, dass die im Seminarverlauf aufgetretenen Konflikte einmal als persönlich unangenehme »Störung« wahrgenommen wurden – was für sich allein genommen bezüglich einer subjektiven Umsetzung der Seminarziele als eher dysfunktional zu werten wäre –, zum anderen jedoch als »konstruktive Hilfe« für die Vermittlung der Seminarinhalte.

Anhand der beiden so identifizierten Dimensionen der Konfliktwahrnehmung ließen sich folgende vier Typen des erinnerten Konflikterlebens bzw. der zum Befragungszeitpunkt berichteten subjektiven Konflikteinschätzungen bilden:

Tabelle 2: Typen der Konfliktwahrnehmung

		Konflikt als »konstruktive Hilfe«	
		niedrig	hoch
Konflikt als »Störung«	niedrig	Typ 1: Indifferenz 20 Prozent	Typ 2: Empathie 36 Prozent
	hoch	Typ 3: Aversion 20 Prozent	Typ 4: Ambivalenz 24 Prozent

Typ 1 (Indifferenz):

Die Angaben der Befragten liegen sowohl bei der Konflikteinschätzung als »konstruktiv« als auch bei der Bewertung als »störend« unter dem Gruppenmittel. Sie geben damit vor, dem Konflikt eher keinen Nutzen im Sinne eines konstruktiven Hilfsmodus abgewinnen zu können. Gleichzeitig nehmen sie ihn aber auch nicht als besondere Störung wahr. Zusammengenommen deutet dies auf eine gewisse Indifferenz gegenüber funktionalen und dysfunktionalen Konfliktdimensionen hin; Konflikte sind in der Erinnerung des Seminars eher eine Randerscheinung geblieben.

Typ 2 (Empathie):

Die Angaben bei diesem Typus liegen beim Bewertungsaspekt »konstruktiv« über und bei der Wahrnehmung als »störend« unter dem Gruppenmittelwert. Dies weist darauf hin, dass hier Konflikte kaum gestört haben und als dem Seminar zuträglich interpretiert werden. Dieser Typ entspricht, verglichen mit den drei anderen, am ehesten dem in der Konzeption angedachten »idealen« Wahrnehmungs- und Erlebenstypus – das empathische Einlassen auf die funktional-konstruktiven Aspekte von Konflikten überwiegt die wahrgenommenen Störungen.

Typ 3 (Ambivalenz):

Die Angaben dieses Typs liegen sowohl bei der Konflikteinschätzung als »konstruktiv« wie auch bei der Bewertung als »störend« über dem Gruppenmittelwert. Hier handelt es sich um Fälle, die Konflikte insofern als »ambivalent« wahrgenommen haben, als trotz der als intensiv erinnerten »Störung« die funktional-konstruktiven Aspekte von Konflikten dennoch positiv bewertet werden.

Typ 4 (Aversion):

Die Angaben der Befragten liegen bei der Konfliktbewertung als »konstruktiv« unter, bei der Einschätzung als »störend« über dem Gruppenmittelwert. Dies weist darauf hin, dass die betreffenden Personen Konflikte erlebt haben, die in ihren Augen nicht nur als eindeutig »störend« empfunden wurden, sondern zudem die konzeptionell intendierte konstruktive Wirkung von Konflikten nicht oder als gering wahrgenommen wurde. Hier überwiegt die Abneigung (Aversion) gegenüber dem erinnerten Konfliktgeschehen.

Betrachtet man nun die Verteilung dieser vier Typen unter den befragten Teilnehmern, so zeigt sich, dass sich eine deutliche Mehrheit von 60 Prozent der Befragten an die im Seminar erlebten Konflikte zum Teil auch gegen eigene innere Widerstände (Dimension »Störung«) als konstruktive Hilfe erinnert. Der Anteil derer, die diese Konflikte in erster Linie als Störung interpretieren, ist mit 44 Prozent wesentlich geringer.

Geht man davon aus, dass die alltägliche Einstellung gegenüber Konflikten überwiegend konfliktaversiv ist, so lässt sich die mehrheitliche Betonung der auch konstruktiven Wirkungen von (gelösten) Konflikten als deutlicher Hinweis darauf werten, dass das Seminarkonzept seiner programmatischen Vorgabe weitgehend gerecht wird, Konflikte als wertvollen Erfahrungsraum begreifen zu lernen. Dies wird auch durch den Befund gestützt, dass 74 Prozent der Befragten angegeben haben, in den nach jeder Seminareinheit angebotenen Reflexionsphasen wichtige Denkanstöße dazu erhalten zu haben.

Die Varianz zwischen den Typen ist nicht von Alter, Geschlecht oder Bildung der Teilnehmer abhängig. Dies ist ein Hinweis darauf, dass das didaktische Konzept von »Achtung (+) Toleranz« innerhalb des befragten Sample alters-, geschlechts- und bildungsunabhängig¹ wirkt. Die Verteilung der Typen ist auch hier von der jeweils besuchten Seminarveranstaltung und damit von der Art der jeweiligen dort gemeinsam mit Team und Teilnehmergruppe erfahrenen Konflikte abhängig.

Ausgewählte Ergebnisse der qualitativen Fallstudie

Einen weitaus differenzierteren Einblick in typische Seminarwirkungen aus Teilnehmerperspektive gestattete die oben beschriebene qualitative Fallstudie, die aus vor und nach der Teilnahme an einem »Achtung (+) Toleranz«-Seminar geführten Intensivinterviews und einem ebenfalls vor und nach der Seminarteilnahme durchgeführten Videoexperiment bzw. -beobachtung bestand.

Der Vergleich der vor und nach der Seminarteilnahme geführten Interviews gestattet Rückschlüsse auf Transformationen etwa der Toleranzkonzepte von Teilnehmern und auf Reflexions- und Handlungsmuster im Umgang mit privaten und/oder öffentlichen, politischen Konfliktsituationen. In-

1 Insbesondere bezüglich der Vermutung der Bildungsunabhängigkeit von Achtung (+) Toleranz ist Vorsicht angebracht, da hier die vermutete Selektivität des Umfragesamples besonders zum Tragen kommen dürfte. Zudem ist davon auszugehen, dass die Grundgesamtheit der Seminarteilnehmer als eher bildungs- und reflexionsbereit zu gelten hat.

soweit derartige Veränderungen in den nach dem Seminar geführten Zweitinterviews in einem Sinnzusammenhang mit den Erfahrungen und Erlebnissen der Seminarteilnahme stehen, lassen sich typische mögliche Wirkungen von »Achtung (+) Toleranz« rekonstruieren und – aufgrund des zeitlichen Abstands der Zweitinterviews von der Seminarteilnahme – deren Nachhaltigkeit feststellen.

Das Feld der so rekonstruierten Wirkungen erstreckt sich nicht nur auf Effekte, die vom Seminarprogramm intendiert waren, sondern aufgrund der Offenheit der Erhebungsmethode ebenso auf nichtintendierte Wirkungen, die im Sinne der Programmatik sowohl erwünscht als auch weniger wünschenswert sein können.

Problemzentriert-diskursive Einzelbefragung prä-post

Lernerfolg (Nachhaltigkeit)

Aus der Diskussion mit den Auftraggebern und Programmentwicklern wurde als Kriterium der Wirksamkeit einer Seminarteilnahme festgelegt, dass bereits die nachhaltige Erinnerung an einzelne Seminarbausteine und Übungen als Erfolg gewertet werden könne. Die Auswertung der qualitativen Interviews unserer Fallstudie kam zu dem Ergebnis, dass sich die Nachhaltigkeit einzelner Seminarbausteine über das bloße Erinnern hinaus durch folgende Typologie aufeinander aufbauender Aneignungsstufen von Seminarinhalten und Bausteinen beschreiben lässt:

- Ebene 1: Sinnzumessung bzw. Bedeutungsverleihung
- Ebene 2: Perspektivenzuwachs
- Ebene 3: aktive Weitergabe an Dritte
- Ebene 4 praktische Umsetzung

Wichtig für die Validität der Typologie ist, dass die ihr zugrundeliegenden Aussagen ohne explizite oder suggestive Nachfrage des Interviewers getroffen werden, die Interviewten also von sich aus auf die entsprechenden Seminarinhalte zu sprechen kamen. Daher ist bereits die bloße Bezugnahme auf eine einzelne Übung oder Trainingseinheit als Indikator dafür zu werten, dass diese (nachhaltigen) Eindruck hinterlassen hat und für die Befragten bedeutsam war, wobei zunächst offen bleibt, in welcher Hinsicht: Auch eine explizite Erwähnung von etwas als »irrelevant« ist eine Bedeutungszuweisung (Ebene 1).

Komplexere Formen der Relevanzverleihung finden sich dort, wo über die

bloße Sinnzumessung hinaus berichtet wurde, dass sich entweder als Folge von oder unter Bezugnahme auf eine Übung die eigene Sicht der Dinge geändert bzw. erweitert hätte (Perspektivenzuwachs, Ebene 2). Im Sinne der Programmausrichtung »stärkere« Indikatoren für den Lernerfolg und dessen Nachhaltigkeit bilden explizite Hinweise auf das praktische Wirksamwerden von Seminarinhalten: etwa die Weitergabe an Dritte (Ebene 3) oder Berichte über das Einbeziehen von Seminarinhalten in eigene Handlungspläne bzw. -strategien (Ebene 4).

Aus der Gesamtauswertung der Interviews ergab sich, dass insbesondere fünf von insgesamt zwölf Übungen jeweils von mehr als der Hälfte der Befragten genannt wurden. Die nur selten thematisierten Übungen wiesen zudem, wenn sie Gegenstand des Interviews waren, in der Wirkungsanalyse nur geringe Effekte aus.

Eine im Konzept von »Achtung (+) Toleranz« zentrale Seminareinheit ist die so genannte »Toleranzampel«, die zur Vermittlung eines dem gesamten Programm unterlegten praxisbezogenen Toleranzschemas dienen soll. Entlang eines Ampelschemas sollen die Teilnehmer lernen, selbst zu prüfen, ob sie sich in gegebenen Konfliktsituationen im Bereich der »Intoleranz« (gewaltförmige Durchsetzung eigener Interessen – rot), der so genannten »Scheintoleranz« (Erduldung einer Störung ohne Lösung der Konfliktsituation – gelb) oder der »Toleranz« (gewaltfreies Konfliktlösen auf Grundlage der Anerkennung gleicher Rechte zur freien Entfaltung – grün) bewegen.

Von daher stellten die kognitive Präsenz dieses Schemas und seine alltagspraktische Relevanz bzw. Umsetzung einen wichtigen Indikator für die Wirksamkeit des Programms insgesamt dar. Die Analyse der qualitativen Interviews zeigt, dass im Vergleich mit anderen im Seminarkonzept enthaltenen Modulen am stärksten auf die »Toleranzampel« Bezug genommen wird (100 Prozent) und dass die Befragten diese auch noch drei Monate nach Seminarende in vielfacher Weise zur Bewältigung von Alltagssituationen bzw. -konflikten heranziehen, hier also auch die komplexeren Formen der inhaltlichen Aneignung überwiegen.

Vor diesem Hintergrund kann festgehalten werden, dass mit der »Toleranzampel« ein praxisorientiertes Toleranzschema angeboten wird und die diesbezüglichen Selbstwahrnehmung (z. B. Reflexivität in Konfliktsituationen, Prüfen der eigenen Toleranzmöglichkeiten und -grenzen) bei den Teilnehmern gefördert wurden. Die »Toleranzampel« erscheint damit gerade aus Sicht der Teilnehmer in deren Alltagserfahrung als ein durchaus hilfreiches Werkzeug für die Entwicklung individueller Konfliktbewältigungsstrategien.

Solchen intendierten Wirkungen stehen jedoch auch nichtintendierte Effekte gegenüber, die mit Blick auf die Programmkonzeption mitunter sogar

als nicht erwünscht eingeschätzt werden können. So kann das Ampel-Schema leicht zu einem unflexiblen und reduktionistischen »Kochrezept« werden, das gerade nicht dazu beiträgt, die Komplexität von Situationen und die Möglichkeit unterschiedlicher Perspektiven zu erkennen. Vielmehr besteht die Gefahr, dass Chancen einer transparenteren Konflikteinschätzung ungenutzt bleiben und die Vielschichtigkeit von Konfliktsituationen auf eine rigide »Ampel-Polarität« von rot oder grün reduziert wird.

Es lässt sich festhalten, dass die »Toleranzampel« nachhaltig in Erinnerung bleibt, dass aber manche Teilnehmer deren Botschaft im Alltag selektiv interpretieren und im Sinne der Programmentwickler womöglich »verkürzt« anwenden. Im Weiteren lässt sich nicht ausschließen, dass solche »instrumentalisierenden« (Miss-)Verständnisse – womöglich unter dem Label »Achtung (+) Toleranz« – weiter verbreitet werden und schließlich genau das befördern, was das Programm nicht transportieren soll: Schwarz-Weiß-Denken.

Transfer

Das Beispiel »Toleranzampel« zeigt, wie unterschiedlich ein Seminarbaustein von den Teilnehmern in ihren Alltag eingebaut werden kann – und verweist damit auf die Variabilität des Transfers von Seminarinhalten und -erkenntnissen. Auf der Basis unserer qualitativen Interviews konnten vier typische Umsetzungsmuster ausgemacht werden, die sich – idealtypisch zugespitzt – auf die Ebenen der Selbstreflexion und der Gestaltung sozialer Beziehungen richten.

- »Selbst-zentriert« (nach innen gerichtet): individualistische Selbstreflexion bis hin zur permanenten Selbstproblematisierung. Selbstreflexion anhand der im Seminar vermittelten Konzepte und Inhalte kann dazu dienen, eigene Verhaltensmuster in Frage zu stellen, was eine Voraussetzung für einen veränderten Umgang mit konflikthaft erfahrenen Situationen darstellt. Zentral ist hierfür der Umgang mit der Frage, wo man in einer Situation steht und wie man sich selbst verhalten muss.
- »Selbst-extrovertiert« (nach außen gerichtet): Vergewisserung der eigenen Person in der Außenwirkung – damit wird das andere Subjekt zum Ziel-Objekt in der Adressierung der eigenen Selbstdarstellung und Beeinflussung. Im Seminar vermittelte Konzepte und Inhalte werden dementsprechend darauf bezogen, das Verhalten anderer einzuschätzen und zu bewerten. Zentral ist die Frage, wo man in einer Situation steht und was man falsch/richtig gemacht hat.
- »Relational-dialogisch«: Ich- und Du-Bezug, gekennzeichnet durch Pers-

pektivenübernahme mit Antizipation wechselseitiger Handlungsgrenzen und -potentiale. Im Seminar vermittelte Konzepte und Inhalte führen zu einem Sich-Hineinversetzen in das Gegenüber und einer entsprechenden Beurteilung der Interaktionssituation aus den unterschiedlichen Perspektiven. Zentral ist die Frage, was der andere macht, wenn man selbst so und nicht anders handelt (Perspektivenübernahme).

- »Kontext-/Strukturorientiert«: sozial-strukturelle Kontextuierung von Ich und Du im Zusammenhang mit Norm- und Regelreflexion. Dieses Muster stellt eine Erweiterung des »relational-dialogischen« Umsetzungsmusters dar, da die jeweils (situativ) geltenden Interaktionsregeln und normativen Strukturen in Hinblick auf ihre Toleranzförderung hinterfragt und berücksichtigt werden. Zentral ist die Frage, ob dieser Kontext einer Situation in irgendeiner Form verändert werden muss.

Im Ergebnis ist festzuhalten, dass in den Interviews die beiden erstgenannten Umsetzungsmuster deutlicher vertreten waren als die beiden letztgenannten. Insbesondere das kontext-/strukturorientierte Muster trat in seiner »Reinform« nur in einem Falle auf, in dem aus dem im Seminar erworbenen Toleranzwissen eine Transformation von Kooperations- und Kommunikationsregeln in einer Gruppe von Verwaltungsangestellten angestrebt und wohl auch in Ansätzen erreicht wurde.

Derartig weit reichende Wirkungen sind im Sinne von »Achtung (+) Toleranz« als positiv zu werten und weisen auf das Potential des Konzepts hin. Die demgegenüber »schwächeren« Effekte, die von anderen Teilnehmern berichtet wurden, lassen sich vor dem Hintergrund der eingangs erwähnten entwicklungslogischen Annahmen interpretieren. Der Schwerpunkt der Übungen liegt darauf, die Reflexion eigenen Verhaltens und Deutens bei den Teilnehmern anzuregen. In diesem Sinne stellen die Befunde also durchaus die Wirkung des Seminars heraus.

Ergänzendes Videoexperiment

Das Evaluationskonzept umfasste neben den beiden oben berichteten »Pflichtteilen« auch den experimentellen »Kürteil« eines Videoexperiments, der aus Ressourcengründen nicht völlig abgeschlossen werden konnte (womit bereits bei Auftragserteilung zu rechnen war). Dennoch erbrachte auch dieser Zugang zur Maßnahmenevaluation interessante erste Ergebnisse und gab darüber hinaus Anlass zu grundlegenden inhaltlichen und methodologischen Fragestellungen (vgl. den Methodenbeitrag in diesem Band).

Das – hier stark verkürzt dargestellte – Videoexperiment beruhte auf einem Prä-/Post-Untersuchungsdesign. Unmittelbar vor und einige Monate nach Seminarteilnahme sollte die Seminargruppe gemeinsam ein formal vergleichbares Handlungsproblem (Erstellen eines Videospots über Toleranz, eines Werbespots für »Achtung (+) Toleranz«) innerhalb von 30 Minuten lösen. Dieser Prozess wurde u.a. mittels einer von den Evaluatoren fest installierten Aufzeichnungskamera dokumentiert.

Inhaltlich bedeutsam ist nun nicht das eigentliche Ergebnis der Aufgabe (Erstellen eines Videospots), sondern der Lösungsprozess selbst. In diesem treffen die unterschiedlichen Intentionen, Ansichten, Normen- und Rollenverständnisse der zu Beginn des vor dem Seminar stattfindenden Experiments einander unbekanntes Teilnehmer aufeinander und müssen, sofern die Gruppe die gestellte Aufgabe zu lösen beabsichtigt, aufeinander abgestimmt und in einen Kooperationsprozess überführt werden.

Aus Sicht der Evaluatoren auffällig war nun, dass sich die Gruppe vor dem Seminar relativ schnell und demokratisch auf formale Verfahren der Entscheidungsfindung einigte, die es jedem Gruppenmitglied ermöglichten, seine Ansichten zu äußern und zur Geltung zu bringen. Diese Verfahrensweise wurde durch solche Gruppenmitglieder in den Prozess eingebracht, die – berufsbedingt – bereits über moderative Vorerfahrungen verfügen. Zudem achteten die Mitglieder darauf, dass jeder zum Zuge kam und sich niemand aus dem Gruppenprozess zurückzog.

Nach dem Seminar verlief der Gruppenprozess in einer vollkommen anderen Richtung. Der von einigen Teilnehmern unternommene Vorstoß, das Handlungsproblem nach dem erfolgreichen Muster des ersten Zusammentreffens zu lösen, schlug fehl, weil die Gruppe dies ignorierte. Stattdessen bildeten sich mehr oder weniger stabile Untergruppen, die sich – ohne dass es zu einer formellen Verabredung dieses Vorgehens gekommen wäre – nach Neigung oder Sympathie zusammensetzten. So plante etwa eine Gruppe Inhalt und Durchführung, ohne dass dies einerseits bei anderen Teilnehmern auf erkennbaren Widerstand stieß oder andererseits Partizipation der Unbeteiligten eingeklagt wurde.

Oberflächlich besehen, verwundert dieses – hier nur skizzierte – Ergebnis, könnte es doch darauf verweisen, dass sich die Gruppe nach dem Seminar scheinbar weniger »tolerant« verhält als davor. Erst bei näherem Hinsehen zeigt sich, dass die Gruppe im Prozess der Differenzierung ihrer internen Rollen und Beziehungen in Folge der verallgemeinschaftenden Seminarteilnahme sehr wohl auf ihre Weise tolerant bleibt, womit die Vermutung aufkommt, dass »Toleranz« zu beiden Zeitpunkten auf zwei Typen der Toleranzherstellung/-sicherung verweist. Das vor dem Seminarbesuch beobacht-

bare Toleranzverständnis bezieht sich auf die Garantie formaler Partizipationschancen, während nach dem Seminarbesuch Toleranz mehr auf der Beziehungsebene anzusiedeln ist (jeder kann informell den Kreis verlassen und gerät nicht in den Ruch, ein Spalter zu sein) – alltagssprachlich wäre hier vielleicht der Begriff gegenseitiger Achtung treffender.

Dieses Ergebnis mag erstaunlich und auf den ersten Blick paradox anmuten, aber die Erwartung einer »linearen« Wirkung des Seminars auf der Handlungsebene wäre – abgesehen davon, dass eine solche aus methodischen Gründen kaum feststellbar sein dürfte – ein Anspruch, den keine Veranstaltung, die nicht nach behavioristischer Konzeption vorgeht, einlösen kann und wohl auch nicht soll.

Im Vordergrund steht mit Blick auf »Wirkungen« der interaktionale Kontext, in dem sich die Trias aus Gruppe, Leitung und Individuum befindet und durch den die Inhalte übermittelt werden – erst dahinter kann das wirken, was im Seminar vermittelt werden soll. Die Konzeption erfahrungsorientierter Erziehungsprogramme ist darauf ausgerichtet, Spuren zu hinterlassen – welche, das ist hochgradig von ihren jeweiligen Rezipienten und (gerade das fördert die Analyse sozialer Prozesse z. B. per Video-Beobachtung zutage) den konkreten Gruppenprozessen abhängig. Toleranz ist demzufolge Effekt latenter sozialer Prozesse und nicht einzelner individueller Selbstdefinitionen (»Ich bin tolerant«).

Abschließende Bewertung und Empfehlungen

Die Evaluation von »Achtung (+) Toleranz« zeigt generell, dass das Programm in der Gesamtschau durchaus gemäß seinen konzeptionellen Intentionen wirkt. Es erzielt nicht nur insgesamt hohe Zustimmungswerte bei den Respondenten der schriftlichen Befragung, sondern darüber hinaus erinnern sämtliche Seminarteilnehmer des Evaluationsseminars programmbezogenes Wissen nachhaltig.

Auf der manifesten Ebene konkreter Äußerungen in den zweiten Interviews benennen die Teilnehmer verschiedene Bedeutungen der Teilnahmeerfahrung für ihren Alltag (z. B. Veränderungen in der Kommunikation mit Arbeitskollegen); auf der latenten Ebene lassen sich – zwischen den Zeilen – eine Vielzahl weiterer toleranzförderlicher Wirkungen erfassen (z. B. Reflexion über Toleranzkonzepte und eigene Deutungsroutinen, Perspektivenübernahme, kommunikative Kompetenz).

Neben solchen erwünschten Wirkungen lassen sich jedoch auch Wirkungen und Effekte der Seminarteilnahme feststellen, die nicht eindeutig mit

den Absichten und Zielen von »Achtung (+) Toleranz« in Einklang stehen (z. B. Verunsicherung in Konfliktsituation, Schwarz-Weiß-Denken im Zusammenhang mit der »Toleranzampel«).

Vor dem Hintergrund dieser und weiterer, hier aus Platzgründen nicht referierter Evaluationsergebnisse können als Diskussionsvorschlag zur Weiterentwicklung des Programms nachstehende Empfehlungen gegeben werden.

Empfehlungen zur Programmstruktur

Das Programm »Achtung (+) Toleranz« ist so aufgebaut, dass es – ähnlich wie ein Baukasten – den jeweiligen Erfordernissen und spezifischen Problembezügen der einzelnen Seminargruppe angepasst werden kann. Der Seminarleitung wird die Möglichkeit gegeben, neben festen, mehr oder weniger »gesetzten« Bestandteilen wie der »Toleranzampel« zusätzliche modulare Bausteine anzubieten. Gerade mit Blick auf den modularen Ansatz erscheint es sinnvoll, die Programmstruktur dort zu ergänzen, wo Evaluationsergebnisse gezeigt haben, dass

- Nachhaltigkeit aus der Sicht der Teilnehmer noch weiter gefördert werden kann,
- zusätzliche Konfrontationen hilfreich sind und
- nichtintendierte Umsetzungsmuster mit Blick auf einzelne Seminarbausteine womöglich vermieden werden können.

Daraus leiten wir folgende Empfehlungen ab:

- *flexible Seminarkonzeption*: Das Antwortenspektrum zur Motivation der Teilnehmer verdeutlicht insbesondere ein auf konkrete Anwendungsbezüge (berufliche Motivation) gerichtetes Bildungsinteresse, darüber hinaus jedoch auch stark auf Persönlichkeitsentwicklung gerichtete Erwartungen. Die Gestaltung des konkreten Seminarkonzepts sollte dementsprechend zum einen unterschiedliche Motivationslagen der Teilnehmer, zum anderen die (teilnehmer-)spezifischen Anwendungsfelder berücksichtigen – d.h. es sollte eine stärkere Ausrichtung der im Programm enthaltenen Übungen und methodisch-didaktischen Bausteine an spezifischen Anwendungsfeldern erfolgen.
- *mehr Gruppenerfahrung*: Aus unserer Sicht sollte es das Ziel des Seminars sein, kontext-strukturorientierte Umsetzungsmuster im Alltag zu fördern. Somit erscheint es sinnvoll, über die überwiegend auf der Ebene der Selbsterfahrung und -erkenntnisse beobachtbaren Wirkungen hinauszugehen.

gehen. Daher empfehlen wir einen konfrontativ-gruppenbezogenen Einstieg, der die Selbsterfahrung zugunsten der Gruppenerfahrung zurücktreten lässt und gleich zu Beginn des Seminars die interaktive Herstellbarkeit, Gestaltbarkeit von Handlungs- und Kommunikationszusammenhängen erfahrbar macht.

- *Nachtreffen*: Um die Nachhaltigkeit der Seminarerfahrungen weiter zu unterstützen, soll ein mit den Teilnehmern abzusprechendes, thematisch fokussiertes Nachtreffen angeboten werden. Unbearbeitete (Seminar-) Konflikte und Alltagserfahrungen, die von den Teilnehmern zu Seminarinhalten und -erfahrungen in Bezug gesetzt werden, können hier zur Sprache kommen. Darüber hinaus kann ein nochmaliges Erinnern der Seminarveranstaltung durch den Austausch in der Gruppe und mit den Seminarleitern die Nachhaltigkeit der Wirkungen unterstützen.

Empfehlung zur seminar- und konzeptionsorientierten Qualitätssicherung

Weil sich der individuelle Lernprozess im Seminar bei einzelnen Teilnehmern sehr unterschiedlich gestalten kann, erscheint eine Erweiterung von Feedbackmöglichkeiten sinnvoll, die es den Teilnehmern ermöglichen, sich ihres persönlichen Fortschritts zu vergewissern und diesen ggf. mit Dritten bzw. in der Gruppe zu erörtern (z. B. als formloses Abendgespräch). Da »Achtung (+) Toleranz« ein Programm mit vielfältigen Wirkungen ist, könnte diese Praxis – sofern dokumentiert und mit Erlaubnis der Teilnehmer – zusätzlich zur Weiterentwicklung des Programms herangezogen werden.

Konkret lautet die Empfehlung: Einführung eines auf Videoaufzeichnungen dieser Abendgespräche beruhenden Instruments formativ-summativ-formativer Evaluation, das der konsequenten Qualitätskontrolle, -sicherung und Weiterentwicklung einer formativen Evaluation des gesamten Seminar-konzepts dienen kann. Formativ meint in diesem Zusammenhang zunächst eine fortlaufende, den Teilnehmern die eigene Entwicklung spiegelnde Bewertung – hier wird die Anpassung der Konzeption im Veranstaltungsablauf ermöglicht. Summativ meint, dass auf der Basis der Gesamtschau einzelner Abschnitte abschließend ein möglicher Fortschritt gemeinsam mit den einzelnen Teilnehmern rekonstruiert und bewertet werden kann. Formativ im zweiten Sinn meint eine sich aus dem Vergleich mehrerer Veranstaltungen ergebende sensible Datengrundlage zur Modifizierung des Gesamtkonzepts unter direktem Rückgriff auf die Äußerungen einzelner Teilnehmer.

Empfehlungen zur Verbreitung des Programms

Das Programm »Achtung (+) Toleranz« ist auf Breitenwirkung angelegt – der Programmordner kann im Buchhandel erworben werden. Somit wird jedem prinzipiell die Möglichkeit gegeben, selbst Seminare anzubieten. Weil jedoch mit dem Programm ein hoher Anspruch an die Seminarleitung verbunden ist, die Fehlinterpretationen von Programmbausteinen einerseits und Schulmeisterei andererseits vermeiden muss, und weil den Programmentwicklern klar ist, dass der Vermittlung der Inhalte von »Achtung (+) Toleranz« ein gewisser demokratischer Wertekonsens zugrunde liegen muss, bietet das Zentrum für angewandte Politikforschung ergänzende Kurse für Multiplikatoren an, die als Trainer, Lehrer etc. mit dem Programm arbeiten.

Damit ist eine gewisse Kontrolle über die Weitervermittlung der Idee des Programms gegeben. Über die tatsächliche Verbreitung können allerdings ebenso wenig Informationen eingeholt werden, wie kontrolliert werden kann, welche Botschaften die Multiplikatoren, die nicht weiter mit den Programmentwicklern in Verbindung stehen, in welcher Weise tatsächlich weitergeben. Wir empfehlen daher:

- *Seminarwerbung*: Um unterschiedlichen Motivationslagen und Anwendungsinteressen der Teilnehmer deutlicher Rechnung tragen zu können, sollten Seminarwerbung und Verbreitungswege stärker differenziert werden, um so die Zusammensetzung der Teilnehmergruppen besser vorstrukturieren zu können.
- *Adressenverwaltung*: Um zukünftig eine regelmäßige Seminarkritik der Teilnehmer auf der Basis eines eigens für »Achtung (+) Toleranz« entwickelten Fragebogens zu erhalten, muss eine zuverlässige Adressenverwaltung sichergestellt werden, die dann auch einen sicheren Überblick über den institutionalisierten Verbreitungsweg des Programms ermöglicht.
- *Erfahrungsaustausch*: Um den Erfahrungsaustausch mit dem Programm zu unterstützen sowie die sozialen Erwartungen der Teilnehmer zu berücksichtigen, kann ein dokumentiertes (Online-)Forum bzw. ein Newsletter eingerichtet werden.
- *Intensivierung der Trainerausbildung*: Infolge der komplexen Gruppendynamik in solchen Seminaren, die mitunter auch zu persönlichen Krisenerfahrungen bei den Teilnehmern führt, erscheint es unabdingbar, eine entsprechende Intensivierung der Ausbildung der Trainer mit Ergebniskontrolle und Verständnissicherung, Auffrischungsseminaren bzw. -tagungen und Supervisionsangeboten zu implementieren.
- *Zertifizierung*: Durch die Einführung eines Zertifikats wird nicht nur den Absolventen der Ausbildung ihre neu erworbene Kompetenz nachgewie-

sen, sondern darüber hinaus auch die »Marke« »Achtung (+) Toleranz« öffentlichkeitswirksamer eingeführt. Personen, die zwar mit dem Ordner arbeiten, aber nicht über das von den Programmentwicklern ausgestellte Trainerzertifikat verfügen, können damit dem Ruf des Programms (im Falle etwaiger Fehlinterpretationen), des Konzepts und seiner didaktischen Ausrichtung weniger Schaden zufügen.

Literatur

- Dimbath, Oliver, Julia von Hayek, Andreas Hirsland und Werner Schneider (2002): »Evaluationsprojekt »Achtung (+) Toleranz«. Gesamtbericht«. www.cap.uni-muenchen.de/download/2003/2003_toleranz_endbericht_toleranz.pdf.
- Bertelsmann Forschungsgruppe Politik (Hrsg.) (2000): *Toleranz. Grundlage für ein demokratisches Miteinander*. Gütersloh.
- Glaser, Barney G., und Anselm L. Strauss (1998): *Grounded Theory. Strategien qualitativer Forschung*. Bern und Göttingen.
- Žižek, Slavoj (2001): »What is wrong with Tolerance?« In: Jörn Rüsen (Hrsg.): *Kulturwissenschaftliches Institut. Jahrbuch 2000/2001*. Essen. 339–348.

Praxisbericht: Evaluation des Programms »Eine Welt der Vielfalt«

Jochen Kehr

Einführung

Auftragsklärung

Im Sommer 2001 wurde Univation e. V. von der Bertelsmann Stiftung beauftragt, das Trainingsprogramm »Eine Welt der Vielfalt« zu evaluieren. Zentrale Ziele von »Eine Welt der Vielfalt« sind, bei Schülern veränderte Einstellungen und neues oder verändertes Handeln im Verhalten untereinander und insbesondere gegenüber Mitgliedern anderer Kulturen hervorzubringen, Stabilisierungen erwünschter Verhaltensweisen zu erreichen und einen Wissenszuwachs zu Themen interkulturellen Lernens zu fördern oder zu initiieren. Die Evaluation sollte den Fokus auf die Resultate des Programms legen, d. h. beschreiben und bewerten, inwieweit die Ziele des Programms erreicht werden.

Während der Auftragsklärung wurde das Evaluationsvorgehen auf die Besonderheiten des Gegenstandes abgestimmt. Das zu beschreibende und zu bewertende Programm »Eine Welt der Vielfalt« ist ein seit mehreren Jahren eingesetztes, etabliertes Programm, das den Toleranzgedanken als wesentliches Leitziel in all seinen Bemühungen mit sich trägt.

Damit auch die Evaluation in ihrem Vorgehen diesen tragenden Gedanken unterstützt, orientierte sich das Evaluationsteam auf den besonderen Wunsch der Auftraggeber und im Einklang mit den Standards der Deutschen Gesellschaft für Evaluation (2001) an einem beteiligtenorientierten (partizipativen) Ansatz. Partizipativ zu arbeiten bedeutet, die Anliegen der wichtigen Beteiligtengruppen in allen Phasen der Evaluation so weit wie möglich zu berücksichtigen und für Transparenz der Arbeitsschritte und Aktivitäten der Evaluation zu sorgen.

Der Gegenstand der Evaluation:

Das Trainingsprogramm »Eine Welt der Vielfalt«

Die Anti-Defamation-League und das von ihr gegründete A WORLD OF DIFFERENCE®-Institute in New York haben ein Trainingsprogramm entwickelt mit der Zielsetzung, Lehrer dabei zu unterstützen, den Toleranzgedanken an Schüler zu vermitteln. Dieses Programm wurde im Rahmen des Projekts »Erziehung zu Demokratie und Toleranz« der Bertelsmann Stiftung für Deutschland adaptiert. Von der Bertelsmann Stiftung und dem Centrum für angewandte Politikforschung wurden daraus eine auf die deutsche Situation angepasste Programmkonzeption und eine Materialsammlung mit Praxishandbuch entwickelt. Diese erscheint als Band 2 in der Reihe »Internationale Programme zur Demokratie und Toleranzerziehung«.

Im Rahmen des Gesamtprogramms führen ausgebildete Trainer, die insbesondere an einigen Landesinstituten für Lehrerfortbildung tätig sind, schwerpunktmäßig Lehrer in das Trainingsprogramm ein. Diese sollen dabei unterstützt werden, den Demokratie- und Toleranzgedanken mit Hilfe der Materialsammlung an ihre Schüler zu vermitteln. Die angebotenen Praxismaterialien richten sich an Schüler der Grundschule und der Sekundarstufe I.

Das Unterrichtsmaterial besteht aus folgenden fünf Lektionen:

- Lektion 1: Schüler sollen sich selbst als Individuen und als Mitglieder unterschiedlicher Gruppen schätzen lernen.
- Lektion 2: Schüler sollen Ähnlichkeiten und Unterschiede bei Menschen erkennen lernen.
- Lektion 3: Schüler sollen Beispiele kultureller Vielfalt untersuchen und verschiedene Völker, Sprachen und volkstümliche Traditionen kennen lernen.
- Lektion 4: Schüler sollen das Wesen von Klischees, Vorurteilen und Diskriminierungen und deren Auswirkungen auf Individuen und Gruppen verstehen lernen.
- Lektion 5: Schüler sollen Strategien entwickeln, um Vorurteile und Diskriminierung zu bekämpfen.

Das Programm arbeitet somit auf unterschiedlichen Ebenen:

- Die vermittelnde Zielgruppe sind Lehrer, die in Seminaren/ Trainings für die Arbeit mit dem Toleranzprogramm sensibilisiert und vorbereitet werden sollen.
- Die Endzielgruppe sind Schüler, die im Unterricht mit diesem Programm arbeiten. Die Schüler sind somit in einer Kette von Maßnahmen die vorgesehenen Endnutzer des Programms.

Ziele der Evaluation

Gemäß dem Auftrag, die Resultate des Programms zu evaluieren, hat die Evaluation von »Eine Welt der Vielfalt« das Ziel, gesicherte Aussagen über die Umsetzungswirkungen des Programms zu machen. Damit bewegt sich die Evaluation im Bereich eines »outcome measurement«. »Outcomes« oder Ergebnisziele eines Programms beschreiben Erträge oder Veränderungen für Teilnehmende während oder nach deren Einbezug in das Programm; diese Veränderungen können Wissen, Einstellungen oder Verhalten betreffen.

Mit dem Ziel, die Informationsgrundlage hinsichtlich der Umsetzungswirkungen des Programms zu verbessern, nimmt die Evaluation von »Eine Welt der Vielfalt« eine bilanzierende, summative Funktion wahr. Wichtig war dabei der Grundsatz der Nutzungsorientierung, der eine Evaluation, die keinen Nutzen bringt, als Verschwendung ansieht. Ziel der Evaluation war es insofern, nützliche, weiterführende Informationen zu gewinnen, die die Programmverantwortlichen unterstützen, eine Entscheidung über Fortführung oder Verbreitung des Programms zu treffen.

Folgende zentrale Fragestellung war dabei leitend: In welchem Umfang kommt es in den exemplarisch ausgewählten Klassen/Gruppen, in denen das Trainingsprogramm eingesetzt wird, zu Verbesserungen bzw. Stabilisierungen gewünschter Einstellungen, Absichten und Handlungen bei den Schülern?

Die Verwirklichung des Ansatzes der partizipativen Evaluation

Die Wahl des pragmatisch-partizipativen Ansatzes

Bei der Wahl des partizipativen Ansatzes stellte sich die entscheidende Frage, in welchem Umfang die Anliegen der wichtigen Stakeholder berücksichtigt werden können.

Es standen zwei unterschiedliche Ansätze der partizipativen Evaluation zur Auswahl: Der aufklärerisch-partizipative Evaluationsansatz, auch Empowerment-Evaluation genannt, wird eingesetzt, um Betroffenen dabei zu helfen, sich selbst zu helfen und ihre Programme zu verbessern. Insofern werden die Anliegen insbesondere der benachteiligten Gruppen aufgenommen, die durch die Evaluation »gestärkt« werden sollen.

Bei der Evaluation von »Eine Welt der Vielfalt« wurde dieser Ansatz nicht verfolgt, weil Grenzen des aufklärerisch-partizipativen Evaluationsansatzes deutlich wurden:

- Das Programm ist bundesweit verbreitet, daher können nicht alle Betroffene beteiligt werden.
- Der Auftraggeber hat eine klare Evaluationsfragestellung als Ausgangspunkt der Evaluation festgelegt, die die Mitbestimmung der Beteiligten über das Evaluationsdesign einschränkt.
- Außerdem ist die summativ-bilanzierende Hauptfragestellung für die Betroffenen nicht zentral, da unter dieser Fragestellung Informationen dazu, wie das Programm verbessert werden kann, weniger relevant sind.

Daher wurde der pragmatisch-partizipative Ansatz für die Evaluation von »Eine Welt der Vielfalt« gewählt, der einerseits die verschiedenen Anliegen der wichtigen Stakeholder integriert (und damit auch dem Toleranzgedanken Rechnung trägt), andererseits die vorhandenen Rahmenbedingungen beachtet und die Nützlichkeit der Evaluation betont.

Beim pragmatisch-partizipativen Ansatz werden die anzulegenden Wertgrundlagen der relevanten Nutzer der Evaluationsergebnisse von den Evaluatoren geklärt. Die Interessen- und Wertposition der Auftraggeber findet dabei – im Ausmaß variierend je nach Partizipationsgrad der Evaluation – besondere Berücksichtigung. Mit diesen geklärten Grundlagen kann die Informationsgewinnung effizient durchgeführt werden. Außerdem können die gewonnenen Informationen auf Basis dieser Wertgrundlagen gemeinsam mit den wichtigen Beteiligten interpretiert werden.

Begleitung der Evaluation durch eine Steuerungsgruppe aus Vertretern aller Beteiligten

Für die Umsetzung des partizipativ-pragmatischen Ansatzes wurde eine Steuerungsgruppe aus Vertretern aller Beteiligten gebildet. Diese Gruppe setzte sich aus einer Vertreterin des Auftraggebers, einer Evaluatorsin, zwei Koordinatoren, zwei Lehrern, zwei Schülern und einer Multiplikatorin zusammen.

Im ersten Treffen der Steuerungsgruppe wurden zentrale Informationsbedürfnisse und Anliegen der Beteiligungsgruppen ermittelt. Während des gemeinsam vorgenommenen Arbeitsschrittes, aus der Ausgangsfragestellung des Auftraggebers die konkreten Teilfragestellungen herauszuarbeiten, wurde deutlich, dass gemeinsame und auch unterschiedliche Werthaltungen in Bezug auf das Programm bei den unterschiedlichen Beteiligungsgruppen vorliegen.

Die eine These besagt: Primär entscheidend für den Erfolg des Pro-

gramms sind Intensität und Umfang der Umsetzung der Übungen. Unterstützend wirkt außerdem die Koordination und Stimmigkeit des Gesamtprogramms, d.h. die Schulung der Lehrer sowie die Unterstützung und Koordination bei der Umsetzung an den Schulen.

Die andere These lautet: Primär entscheidend für den Erfolg des Programms ist, mit welcher Haltung Lehrer das Programm an die Schüler vermitteln. Dabei unterstützend können eine aufgeschlossene Schulkultur, pluralistische Werthaltungen und weitere günstige Rahmenbedingungen an den Schulen wirken.

Den beiden Thesen liegen unterschiedliche Perspektiven zugrunde, die mit jeweils anders pointierten Werthaltungen verbunden sind:

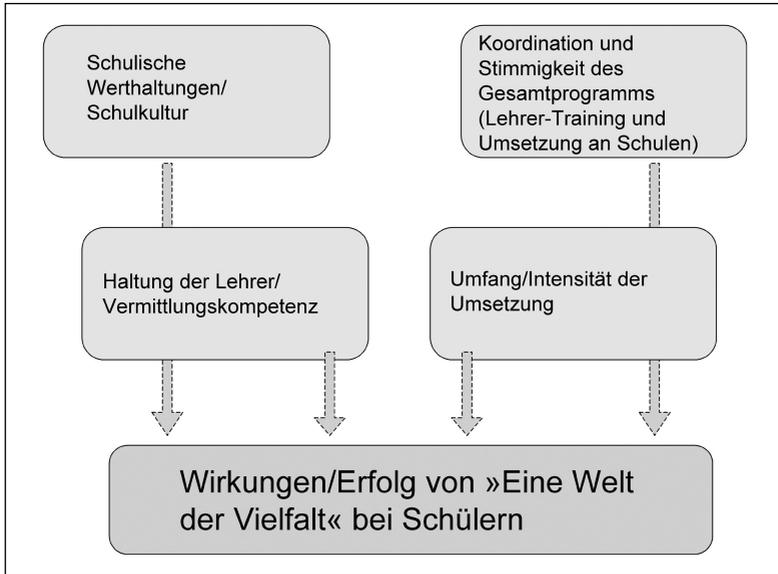
- Die eher erziehungswissenschaftlich geprägte Position sieht in gut entwickelten, wissenschaftlich und didaktisch abgesicherten Curricula einen hohen Wert, da diese ein generelles Potenzial für Toleranzförderung in den Schulen in Deutschland/im deutschsprachigen Raum bieten. Der Wert liegt in der Möglichkeit zur Verallgemeinerbarkeit und in der guten wissenschaftlichen Absicherung des Programms.
- Auf der anderen Seite steht die Position, die vorrangig auf ganz konkrete Schul- und Unterrichtspraxis und die Rahmung der Schulpraxis durch konkrete soziale, ökonomische und politische Bedingungen an jedem Einsatzort des Curriculums ausgerichtet ist. Die lokale Anpassungsfähigkeit, Flexibilität und der unmittelbare Nutzen der Materialien und Handreichungen des Curriculums stehen im Vordergrund. Sein Wert liegt in der Nutzbarkeit unter sehr unterschiedlichen Rahmenbedingungen.

Die im Folgenden aufgelisteten gemeinsam herausgearbeiteten Teilfragestellungen spiegeln die unterschiedlichen Informationsinteressen, Grundannahmen und Werthaltungen der Beteiligten wider.

- In welchem Umfang wird das Programm an den exemplarisch ausgewählten Schulen umgesetzt?
- Welche Rahmenbedingungen sind an den ausgewählten Schulen besonders fördernd bzw. hemmend für die Umsetzung des Programms?
- Wie wird das Programm durch die Schüler in den ausgewählten Klassen im Programmverlauf akzeptiert?
- Inwiefern zeigt sich bei den Personen in den ausgewählten Klassen, die mit dem Programm gearbeitet haben, eine erhöhte Sensibilität (Diskriminierung hören/fühlen) für das Thema Toleranz?
- Inwiefern trägt das Programm dazu bei, den Prozess der (Selbst-)Reflexion in Bezug auf das Thema Toleranz bei den ausgewählten Schülern und Lehrern zu initiieren bzw. zu fördern?

Abbildung 1: Wertbasierte Perspektiven auf das Programm

»Eine Welt der Vielfalt«



Quelle: www.univation.org

- Inwiefern trägt das Programm zu einem veränderten Umgang zwischen den ausgewählten Beteiligten vor Ort bei?
- Inwiefern trägt das Programm dazu bei, Wissen und Informationen bei den ausgewählten Schülern zu den Themen Toleranz und Demokratie zu vermehren?
- Inwiefern trägt das Programm dazu bei, dass die Beteiligten der ausgewählten Klassen Anregungen zu einem Vorgehen gegen Diskriminierung bekommen?
- Inwiefern treten nichtintendierte Wirkungen auf?

Beim zweiten Treffen der Steuerungsgruppe befand sich die Evaluation in der Endphase der ersten Untersuchungsphase. Die Steuerungsgruppe unterstützte das Evaluatorteam dabei, aus den untersuchungsleitenden Fragestellungen erste Fragen für die Fallstudien zu formulieren. Neben ganz konkreten Vorschlägen, welche Sachverhalte mit Blick auf die zuvor formulierten Teilfragestellungen bei den Schülern abgefragt werden könnten, äußerten die Beteiligten ihre Einschätzungen und Erfahrungen mit den zur Diskussion stehenden Untersuchungsmethoden. Dabei waren die Erfahrun-

gen der Lehrer und der anwesenden Schüler entscheidend für die methodische Anlage der Fallstudien. Insgesamt wirkte die Zusammenarbeit mit der Steuerungsgruppe sehr unterstützend.

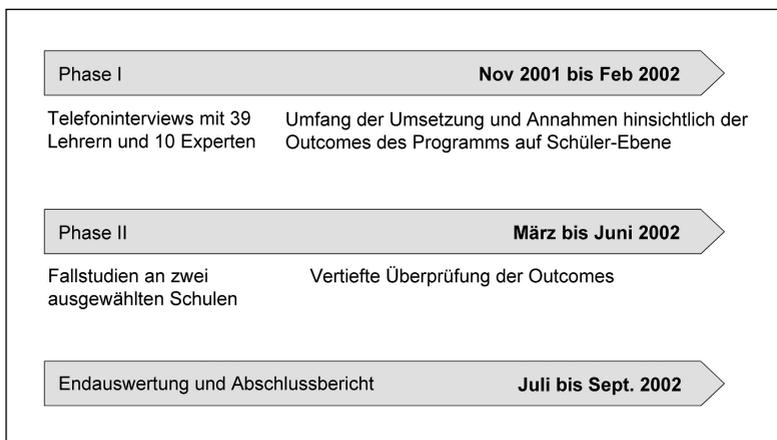
Das Gesamtdesign der Evaluation: Anlage und zeitlicher Verlauf

Die Evaluation wurde in zwei große Untersuchungsphasen aufgeteilt:

In Phase I wurde durch eine Befragung von 39 Lehrern, die an einem Sensibilisierungsseminar zu »Eine Welt der Vielfalt« teilgenommen hatten, zunächst der Umfang der Umsetzung des Programms an Schulen erhoben. Diese Lehrer, die mit dem Programm die Zielgruppe Schüler ansprechen, wurden danach befragt, welche »outcomes« sie für wünschenswert halten und welche »outcomes« sie konkret bei ihren Schülern beobachten können. Ergänzt wurde dies durch die Befragung von zehn Experten in ihrer Funktion als Länderkoordinatoren.

In Phase II der Untersuchung wurden an zwei ausgewählten Schulen (Rouanet-Gymnasium Beeskow in Brandenburg, Fachschule für Sozialpädagogik in Niedersachsen) die Schüler in Bezug auf die Akzeptanz des Programms und die vorliegenden »outcomes« befragt. Dabei wurden die Erkenntnisse aus Phase I für den Zuschnitt der Erhebungen an den Schulen genutzt.

Abbildung 2: Zeitlicher Verlauf der Evaluation



Quelle: www.univation.org

Zentrale Ergebnisse der Evaluation

Als Ergebnis der durchgeführten Erhebungen zeigte sich, dass mit dem Programm eine Basis für den Prozess der Toleranzentwicklung bei Kindern und Jugendlichen gelegt wird. Eine intensive Auseinandersetzung mit sich selbst als Person wird durch das Programm sehr stark angeregt. Die Jugendlichen lernen, sich selbst als Individuum zu verstehen und verbessern somit ihre Voraussetzungen, auch die Anliegen ihres Gegenübers wahrzunehmen und zu respektieren.

Dass der Respekt vor dem Gegenüber und der Gemeinschaft von den Schülern tatsächlich auch praktiziert wird, wird durch die sehr positiven Aussagen zum Umgang der Schüler miteinander und zum Umgang zwischen Schülern und Lehrern deutlich. Das heißt, sowohl Verhaltensbereitschaft wie das Überdenken eigener Wertvorstellungen, Einstellungen und/oder eigener Wahrnehmungsmuster und Interpretationen der wahrgenommenen Welt als auch Verhalten werden durch das Programm positiv beeinflusst.

Damit wird durch das Programm ein wesentlicher Beitrag zur Förderung der sozialen Kompetenz geleistet. Diese ist ein wesentlicher Faktor für ein friedliches tolerantes Miteinander und stellt darüber hinaus (neben der fachlichen Kompetenz) auch für die berufliche Zukunft der Schüler eine wichtige Ressource dar.

Die folgende Übersicht zeigt die kategorisierten Äußerungen der Lehrer in den Telefoninterviews zu der offenen Frage, welche Wirkungen die Lehrer tatsächlich festgestellt haben.

- Reflexion/Sensibilisierung: 28 Nennungen
- anderer Umgang miteinander: 21 Nennungen
- Wissenszuwachs: 9 Nennungen
- Prävention: 5 Nennungen
- anderes Klassenklima: 5 Nennungen

Akzeptanz und Praxisbezug

Die Evaluation zeigt, dass das Programm schwerpunktmäßig in den Klassen 7 bis 10 eingesetzt und von den Schülern sehr gut angenommen wird. Die Akzeptanz des Lehr- und Arbeitsmaterials des Programms liegt bei 75 Prozent der Befragten. Der überwiegende Teil findet das Thema sehr wichtig und wünscht sich dessen Behandlung im Rahmen des Unterrichts: »Das Programm lohnt sich, man lernt sich selbst und andere besser kennen. Große und alltägliche Vorurteile werden abgebaut. Die eigenen Grenzen werden

überprüft und eventuell erweitert.« 22 der 39 befragten Lehrer setzen das Material regelmäßig im Unterricht ein, 18 nutzen das Programm insbesondere in Konfliktsituationen (Mehrfachnennungen).

Wissens- und Kompetenzvermittlung: Lernerfolg und Nachhaltigkeit

Als wichtigstes vermitteltes Element nennen die Befragten die Aufklärung über Vorurteile und deren Entstehung. Über zwei Drittel der Schüler benennen ein besseres Wissen über Vorurteile und deren Auswirkungen. Insbesondere der Aspekt Entstehung von Vorurteilen wird von den Schülern als wichtige Lernerfahrung betrachtet. Hierzu berichten die Schüler in den Fallstudien: »Ich habe lange über meine Vorurteile und Einstellungen nachgedacht und mir die Frage gestellt ›Warum?‹. Warum möchte ich so etwas, warum habe ich Vorurteile, woher kommen die? Ich bin doch nicht mit Vorurteilen auf die Welt gekommen.« Die Mehrheit der Lehrer, die mit dem Programm arbeiten, sind davon überzeugt, dass das Programm eher langfristig als bereits kurzfristig bei den Schülern etwas auslöst.

Transfer: Einstellungs- und Verhaltensänderungen

Wichtiger Wirkungsbereich des Programms ist die Selbstreflexion und Sensibilisierung in Bezug auf das Thema Toleranz. Alle befragten Schüler berichten von einer positiven Weiterentwicklung ihrer Persönlichkeit durch das Programm und sind dazu angeregt worden, über ihre eigenen Einstellungen, Urteile und Werte nachzudenken.

Die meisten Schüler sind sich einig, dass ihre Wahrnehmung bewusster und differenzierter geworden ist: »Während des Programms habe ich viel über mich und mein Verhalten nachgedacht. Dabei ist mir vieles bewusst geworden, das früher für mich völlig anders war und das ich als selbstverständlich angenommen habe. Für meine differenzierte Sichtweise bin ich jetzt dankbar.«

Diese gestärkte Auseinandersetzung mit der eigenen Person dient als Ausgangsbasis für den toleranten und fairen Umgang mit anderen. Lehrer nennen Aspekte wie: »Lehrer-/Schüler-Verhältnis positiv, man nimmt sich gegenseitig ernst«, »einfühlsam miteinander umgehen«, »Respekt, friedfertiger Umgang«, »besserer Umgangston«, »es war eine andere Atmosphäre zu spüren: freundliches Begrüßen, mit dem Vornamen anreden«, »einfühlsamer Umgangston in der Klasse, weniger Schimpfwörter«.

Schüler geben für den Bereich Umgang miteinander vor allem eine veränderte Gesprächskultur an, die sich in Diskussionsfreude und effektiveren Auseinandersetzungen zeige. Außerdem sehen sie »größeren Zusammenhalt«, »bessere Gemeinschaft«, »weniger Gruppenbildung« und »Verständnis« in ihrer Klasse.

Das Programm leistet außerdem einen wichtigen Beitrag im Bereich des Konfliktverhaltens. Die große Mehrheit der Schüler (83 Prozent) fühlt sich durch die Arbeit mit dem Programm unterstützt, Konflikte zu thematisieren und Dinge anzusprechen, wozu ihnen vorher der Mut fehlte.

Eine große Stärke des Programms ist nach Auffassung der Schüler seine direkte Wirkung auf das Klassenklima: »Die Gemeinschaft wird gestärkt und ein besserer Zusammenhalt entsteht. In der Klasse entsteht eine verbesserte Gesprächskultur mit einem intensiven Meinungs austausch. ... Der Mikrokosmos Schulklasse als Lernraum für die Gemeinschaft/Gesellschaft wird mit dem Programm intensiv genützt.« Zwei Drittel der Schüler empfinden das Verhalten der Mitschüler nach Einsatz des Programms als rücksichtsvoller und verantwortungsvoller. Es verbessert sich nicht nur das Verhältnis der Schüler untereinander, sondern gerade auch das von Lehrern und Schülern.

Die beiden Fallstudien zeigen klar, dass es durch »Eine Welt der Vielfalt« in erhöhtem Maße zu Prozessen der Reflexion, der Einstellungsveränderung oder Stabilisierung von Einstellungen gekommen ist.

Nichtintendierte Nebenwirkungen und Gründe für das Nichterreichen bestimmter Wirkungen

Fast alle befragten Lehrer gehen von nichtintendierten Nebenwirkungen aus. Vor allem die betonte Emotionalität und die Mittelpunktstellung der Schüler in den Übungen werden als spannungsreich angesehen. Schüler betonen, dass sie es zu Anfang als ungewohnt empfinden, so offen zu arbeiten, jedoch inzwischen vielfach deutlich die Chancen sehen, die darin liegen. »Den Ansatz solcher Diskussionen bzw. solcher Besprechungen, mehr über sich zu erzählen, finde ich sehr gut. Allerdings wäre so etwas in kleineren Kreisen von etwa 10 bis 15 Personen eher angebracht als in einer Klasse von 30.«

Aus Sicht der Experten werden Widerstände bei den Schülern dann hervorgerufen, wenn Lehrkräfte das Programm unangemessen nutzen. Unangemessenheit sei dann gegeben, wenn die Übungen mit einer nichtwert schätzenden Haltung von Seiten der Lehrer praktiziert werden, oder wenn sie seltene Ausnahmen im Sinne von »highlights« darstellen, im Regelunterricht dann aber wieder zu einer autoritären Haltung zurückgekehrt wird.

Da nach Einschätzung der Lehrer konkrete Strategien gegen Diskriminierung von dem Programm weniger vermittelt werden, wünschen sie sich hier einen stärkeren Fokus der Arbeitsmaterialien. Ansonsten werden Gründe für das Nicht-Erreichen bestimmter Wirkungen weniger im Bereich des Materialordners von »Eine Welt der Vielfalt« als in gesellschaftlichen und sozialen Umfeldbedingungen der Schüler und der Schule gesehen. Die Koordinatoren sehen die strukturellen Bedingungen von Schule und Unterricht vielfach als Hindernis an. Auch die Schüler empfinden die Gruppen in der Klasse als zu groß und empfinden eine Benotung der mit »Eine Welt der Vielfalt« behandelten Themen unangemessen.

Als weitere Problemfelder werden genannt:

- die Diskrepanz zwischen dem geschützten Lernraum und den sonstigen Alltagserfahrungen
- strukturelle Schwierigkeiten: Lehrer werden kaum für Trainings freigestellt
- die fehlende Einbeziehung der Eltern

Kontinuität wird von den Lehrern als entscheidender Faktor betrachtet, um bei Themen wie Toleranz und Demokratie Wirkungen zu erzielen. Fördernd wie hemmend können sich die Haltung des Lehrers sowie Unterrichtsstil und Methodik auf die Effekte des Programms auswirken. Auch Schüler halten diese Aspekte für wichtig.

Schlussfolgerungen und Empfehlungen

Das Programm »Eine Welt der Vielfalt« wurde in einer Zeit evaluiert, in der die Aktualität des Themas immer wieder durch dramatische Ereignisse aufgezeigt wurde. Die Terroranschläge in New York und der Amoklauf eines Schülers in Erfurt sind zwei erschreckende Ereignisse, die in diese Zeit fielen.

Alle Beteiligte, Programmverantwortliche, Koordinatoren, Multiplikatoren, Lehrer und vor allem die Schüler selbst sind sich einig, dass das Thema außerordentlich wichtig ist. Die Akzeptanz bei den Schülern, die ja eine zentrale Voraussetzung ist, um überhaupt Erfolge erzielen zu können, ist deutlich ausgeprägt. Viele Schüler betonen, wie wichtig sie es finden, dass diese Themen Platz im Schulalltag haben.

Sowohl Schüler als auch Lehrer sind sich einig, dass durch das Programm im Bereich der Sensibilisierung einiges bewegt werden kann. Die beiden Fallstudien zeigen, dass es durch »Eine Welt der Vielfalt« in erhöhtem Maße

zu Prozessen der Reflexion, der Einstellungsänderung oder Stabilisierung von Einstellungen gekommen ist. Dies ist ein Bereich, der ansonsten nur schwer beobachtbar ist.

Da sich das Programm jedoch in einem multikausalen Wirkungsbereich bewegt, d.h. eine ganze Reihe anderer, interner persönlicher wie externer Faktoren berücksichtigt werden müssen, bleiben einige Lehrer vorsichtig in ihrer Deutung. Daraus leiten sie die Anforderung ab, so viele externe Wirkungsfaktoren (z. B. Eltern, Freunde, Gruppen in der Gemeinde) in die Umsetzung von »Eine Welt der Vielfalt« mit einzubeziehen wie möglich.

Eine Empfehlung an dieser Stelle ist, einen Austausch (landesweit oder länderübergreifend) im Sinne einer Verbreitung von »good practice« und einem Erfahrungsaustausch speziell zu dieser Perspektive zu initiieren. Die Aussagen einiger Lehrer in den Interviews zeigen sehr eindrückliche Beispiele, wie mit »Eine Welt der Vielfalt« auch außerhalb der Schule gearbeitet wird.

Schwierig werden von Seiten der Schüler und der Lehrer die Rahmenbedingungen angesehen, die die Schule vorgibt. Die Schüler sehen es als problematisch an, für persönliche, offene Inhalte Noten zu bekommen. Außerdem steht der prozessorientierten Methodik der Übungen einerseits der 45-minütige Stundentakt, andererseits die Größe der Klassen bzw. Gruppen entgegen. Um diesen ungünstigen Bedingungen zu entgehen, wenden viele Lehrer das Programm eher außerhalb des normalen Unterrichts an, an Projekttagen, in Projektwochen oder auf Klassenfahrten.

In Schulen, in denen in Lernfeldern gearbeitet wird, erleichtert dies, eine solche Thematik auf mehrere Fächer zu verteilen und daher mit dem fachspezifisch geforderten Stoff schnell genug voranzukommen.

Die Evaluation zeigt, dass Nischen außerhalb des Unterrichts bereits vielfach genutzt werden und eine andere Umgebung als die Schule auch als sehr förderlich für das Erreichen von Wirkungen eingeschätzt wird.

Das Material wird von fast allen befragten Lehrern nur teilweise (Elemente, einzelne Übungen) genutzt und selbst modifiziert, da es für eine durchgängige Bearbeitung zu umfangreich sei. Eine Empfehlung an dieser Stelle ist, auch kompaktere Lektionen anzubieten, die gerade für einen projektartigen Einsatz geeignet sind, und die Übungen alters- und zielgruppenspezifisch zu modifizieren.

Sehr deutlich geworden ist, dass die Lehrer das Material und das Training, an dem sie teilgenommen haben, sehr schätzen. Anschließend fühlen sie sich jedoch an ihren Schulen häufig als Einzelkämpfer allein gelassen. Kaum eine der befragten Schulen hat mehr als einen ausgebildeten Lehrer. Sehr deutlich wird der Wunsch nach Vernetzung, Peergruppen und Unterstüt-

zung auch nach der Zeit des Trainings. Hier ist es empfehlenswert, Möglichkeiten des organisierten Austauschs derjenigen, die mit dem Programm an Schulen arbeiten, zu entwickeln. Dabei würden zentral geführte aktualisierte Adressdateien eine gute Hilfe darstellen.

Insgesamt wird deutlich, dass mit den Unterlagen für »Eine Welt der Vielfalt« akzeptiertes, positiv bewertetes Material vorliegt, um an einem wichtigen Thema zu arbeiten. Wenn es den Lehrern gelingt, das Material häufiger und regelmäßiger einzusetzen, sind »outcomes« bei den Schülern klar feststellbar. Ein veränderter Umgang miteinander, ein Aufmerksam-Werden, eine Sensibilisierung für Thematiken der Toleranzerziehung – zu allen Punkten zeigt die Evaluation, dass Wirkungen bei den Schülern auftreten.

»Miteinander – Erfahrungen mit Betzavta«: Evaluationsprozess und Ergebnisse

Michael Bommes, Jochen Guter, Ulrike Wolff-Jontofsohn

»Erziehung zur Demokratie sollte auf das engste mit aktuellen politischen Ereignissen verbunden sein und Instrumente anbieten, um mit diesen umzugehen, denn ansonsten verliert die Erziehung zur Demokratie ihre Bedeutung.« (Maroshek-Klarman 1996)

Einleitung

Im Folgenden werden die Ergebnisse der Evaluation eines von der Bertelsmann Stiftung und dem Centrum für angewandte Politikforschung (Universität München) geförderten politischen Bildungsprogramms, des Programms »Miteinander – Erfahrungen mit Betzavta« vorgestellt. Diese Evaluation wurde im Zeitraum von August 2001 bis Ende 2002 durchgeführt.¹

Gegenstand der Evaluation war das auf der Basis des israelischen Programms Betzavta (Adam Institute, Jerusalem) entwickelte Programm »Miteinander – Erfahrungen mit Betzavta« (im Folgenden »Miteinander«), das seit 1997 in der Bundesrepublik in der außerschulischen und schulischen Bildungsarbeit als präventives Programm zur Demokratie- und Toleranzerziehung eingesetzt wird.

Die Initiatoren und Träger des Programms, die Bertelsmann Stiftung und das Centrum für angewandte Politikforschung, adaptieren seit Mitte der 90er Jahre im Rahmen des Kooperationsprojekts »Erziehung zu Demokratie und

1 Den Evaluatoren Michael Bommes und Ulrike Wolff-Jontofsohn (zu diesem Zeitpunkt beide Pädagogische Hochschule Freiburg) wurde von der Bertelsmann Stiftung und dem Centrum für angewandte Politikforschung auf der Grundlage eines von ihnen eingereichten Evaluationskonzeptes der Auftrag zur Evaluation des Programms erteilt. Der ausführliche Bericht ist online erhältlich unter: www.cap.uni-muenchen.de/download/2003/2003_toleranz_endbericht_betzavta.pdf.

Toleranz« internationale Kursprogramme und Materialien für die Verwendung in der politischen Bildungsarbeit in Deutschland und bilden Multiplikatoren für die Arbeit mit diesen sowie selbst entwickelten Programmen aus. Das Ende 1996 fertig gestellte Praxishandbuch »Miteinander« hatte die Reihe »Internationale Programme zur Demokratie- und Toleranzerziehung« der Bertelsmann Stiftung eröffnet.

Mit dem Praxishandbuch legten die Initiatoren und Träger der Forschungsgruppe ein vom Jerusalemer Adam Institute for Democracy and Peace entwickeltes und von den Mitarbeitern des Centrums für angewandte Politikforschung für die Anwendung in deutschen Bildungseinrichtungen angepasstes Konzept zur politischen Bildung vor.

Aufgrund des Interesses, das Lehrer aller Schularten an den Materialien äußerten, wurde das Handbuch mit der zweiten Auflage im Jahre 1999 um ein eigenes Schulkapitel erweitert, das einen Leitfaden zur Integration des Programms in den Fachunterricht und in den fächerübergreifenden bzw. projektorientierten Unterricht enthält. Dieser Leitfaden wurde in Zusammenarbeit mit dem Adam Institute for Democracy and Peace entwickelt, das zahlreiche Erfahrungen in Israel mit dem Einsatz des Programms in Schulen gesammelt hat.

Nach fünf Jahren des Einsatzes von »Miteinander« in der inner- und außerschulischen politischen Bildungsarbeit und der Ausbildung von Multiplikatoren für diese Arbeit sollte die Evaluation Einblicke vermitteln, was mit dem Einsatz von »Miteinander« bislang erreicht worden ist, welche Art von Problemen bestehen und wie auf diese reagiert werden kann.

Vorüberlegungen zur Evaluation

Vorgaben der Auftraggeber

Nach den Vorstellungen der Auftraggeber sollte die Evaluation des Programms mittels einer Befragung der im Laufe der letzten Jahre in diesem Programm rekrutierten Gruppe der Multiplikatoren durchgeführt werden. Zu diesem Zweck stellte das Centrum für angewandte Politikforschung eine Adressdatei von 256 Multiplikatoren zur Verfügung.

Die Einschätzung des Programms sollte sich auf die Aussagen der Multiplikatoren zu ihren eigenen Erfahrungen mit dem Programm und der pädagogischen Arbeit damit sowie auf deren Deutung und Abschätzung der Programmwirkungen auf die Teilnehmer stützen. Methodisch sollten in der Evaluation quantitative und qualitative Verfahren Verwendung finden. Die Eva-

luation sollte – dem »Geist des Programms« entsprechend – »gegenstandsadäquat, interaktiv und partizipativ« angelegt sein.

Das von wissenschaftlichen Mitarbeitern des Centrums für angewandte Politikforschung erstellte Grundlagenpapier (Ulrich und Wenzel 2001) zur Evaluation von Demokratie- und Toleranzerziehungsprogrammen sollte als Orientierungsrahmen bezüglich der Lernzieldefinitionen und Wirkungserwartungen dienen. Als zentrale Wirkungsdimensionen des Programms werden dort Kompetenz- und Wissensvermittlung sowie Reflexions- und Perspektivwechselvermittlung aufgeführt. Aufgabe der Evaluation war also, zu untersuchen, ob der je individuelle Kompetenzaufbau zu individuellen Handlungsveränderungen und in der Summierung der individuellen Veränderungsleistungen zu Effekten auf die Gesellschaft führt.

Evaluationen – ein Modus der Bearbeitung von Unsicherheit

Keine Unternehmung, sei sie ökonomischer, politischer oder erzieherischer Art, wird heute noch auf den Weg gebracht, ohne bereits mit ihrem Anfang den Zeitpunkt vorzusehen, zu dem ihre Evaluation anzusetzen ist. Evaluation scheint eine Bedrohung und ein Versprechen zugleich zu bezeichnen: eine Bedrohung, sofern sich die Evaluierten immer auch unter verschiedenen Gesichtspunkten kontrolliert sehen, betreffend insbesondere Mittelverwendung und Zielerreichung im Rahmen ihrer Tätigkeit; ein Versprechen, sofern ihnen zugleich in Aussicht gestellt wird, dass diese Kontrolle, durchgeführt im methodischen Soft-Pack von Interaktion und Partizipation, zu ihrem eigenen Besten sei, da sie ihnen Potenziale der permanenten Selbstverbesserung zur Verfügung stelle.

Zugleich lassen sich mit Evaluationen, so scheint es, die Dauerprobleme moderner Organisationen in den Griff bekommen: Organisationen sollen und müssen in einer turbulenten und unvorhersehbaren Umwelt lernen und sich permanent entwickeln. Evaluation scheint ihnen den reflexiven Mechanismus zur Verfügung zu stellen, der genau dies erlaubt.

Für das Übergreifen dieser Formen der Reflexion, die zunächst in marktorientierten Unternehmen ausprobiert worden sind, auch auf den Bereich der politischen Bildung gibt es viele Gründe, auf die hier nicht einzugehen ist. »Der Prozess des Organisierens« (Weick 1985) scheint mit Problemstellungen verbunden, die gleichermaßen verschiedene Organisationen betreffen, wenn sie auch mit je anderer Spezifizierung zur Geltung kommen, abhängig von der Funktionsausrichtung jeweiliger ökonomischer, politischer, rechtlicher, gesundheitlicher oder auch erzieherischer Organisationen. Vor diesem

Hintergrund soll hier nur auf einen Zusammenhang hingewiesen werden, der es möglicherweise erlaubt, etwas mehr Distanz zu den oftmals sicherlich zu hohen Erwartungen an Evaluationen zu gewinnen.

Evaluationen haben an Bedeutung gewonnen im Kontext der Umstellung von einer so genannten Input- auf eine so genannte Output-Orientierung mit Blick auf die Leistungserwartungen an Organisationen. Input-Orientierung meint dabei die Erwartung, dass die Steigerung der Ressourcenausstattung von Organisationen auch zur Steigerung von deren Leistungsfähigkeit führt.

Das glaubt so heute niemand mehr, und entsprechend erwartet man zum Beispiel (spätestens seit PISA) auch nicht mehr, dass der Geldbetrag, den ein Bildungssystem verbraucht, unmittelbar etwas über seine Leistungsfähigkeit aussagt. Die Leistungsfähigkeit und der Ertrag von Organisationen sind unsicher geworden, ihre Resultate scheinen erheblich von ihrer inneren, keineswegs transparenten Verfassung abzuhängen.

Man kann das auch so formulieren: Organisationen, die zentrale Erfindung der modernen Gesellschaft zur Absorption von Unsicherheit, erzeugen mit ihrer gesellschaftsweiten Expansion – kein Bereich, in dem Organisationen nicht eine zentrale Bedeutung zukommt – und dem damit verbundenen Mittelverbrauch Unsicherheit in Bezug auf ihre intransparente Struktur, ihre Kosten und ihren unsicheren (keineswegs nur ökonomischen) Ertrag. Es handelt sich offensichtlich nicht um intern rational strukturierte Einrichtungen, orientiert an der Erreichung deklarer Ziele, als die sich Organisationen unvermeidlich präsentieren müssen.

Output-Orientierung bedeutet vor diesem Hintergrund vereinfacht gesprochen die Orientierung am faktischen Output von Organisationen und den Vergleich mit extern oder intern formulierten Output-Erwartungen. Evaluationen untersuchen den erreichten Output (z. B. Gewinne, behandelte Patienten, Bildungsabschlüsse, Wissen etc.) in einem definierten Zeitraum, fragen nach den internen und externen Ursachen für den je erzielten Output (die bereitgestellten Leistungen) und versuchen, Möglichkeiten der Verbesserung zu benennen.

Ausgangspunkt für diese Entwicklung war, wie gesagt, die Erfahrung von Unsicherheit. Vor diesem Hintergrund verblüfft es, wenn mit dem Versuch, diese Unsicherheit dadurch wieder abzubauen, dass Organisationen mit Erwartungsenttäuschung und dem Zwang zur Reflexion durch Evaluation konfrontiert werden, mittlerweile oftmals die Annahme verknüpft zu sein scheint, dadurch ließe sich die Transparenz und Rationalität von Organisationen, ausgerichtet an Zielerreichung, sowie Organisationslernen und Reflexivität auf Dauer gewährleisten.

Gegen diesen gewissermaßen von hinten durch die Ausgangstür wieder-

eingeführten Organisationsoptimismus sprechen vielfältige Einsichten der Organisationssoziologie (für viele Brunsson 1989, Luhmann 2000). Dies gilt insbesondere für solche Organisationen (wie im Erziehungs- und Bildungsbereich), die durch so genannte Zielprogramme (im Unterschied zu Konditionalprogrammen; zu dieser Unterscheidung vgl. Luhmann 1981) und »weiche Technologien« in dem Sinne gekennzeichnet sind, dass die Zusammenhänge zwischen Mittel und Zielerreichung unsicher sind.

Für unseren Zusammenhang ist diesbezüglich bedeutsam: Evaluationen sind Formen der Fremd- oder der Selbstbeobachtung von Organisationen, mit denen sie intern oder extern kommunizierte Irritationen abarbeiten und sich gegebenenfalls Möglichkeiten verschaffen können, wie sie – bis zu den nächsten Irritationen – in anderer (oder auch gleicher) Weise weitermachen können. Damit sollten aber keine überzogenen Hoffnungen auf Rationalität oder Transparenz verbunden werden, sondern eher die Chance zur Eröffnung je anderer Möglichkeiten, den Stein den Berg hochzurollen – er wird nicht oben liegen bleiben.

Mit Blick auf die durchgeführte Evaluation, deren Ergebnisse hier zusammenfassend vorgestellt werden, bedeutet dies, dass es sich im Wesentlichen um eine Form der Fremdbeobachtung der Art und Weise handelt, in der sich die Resultate der Implementation eines politischen Bildungsprogramms denen darstellt, die daran wesentlich beteiligt waren – also eine Fremdbeobachtung der Selbstbeobachtung. Auf dieser Grundlage ist der Versuch unternommen worden, Strukturkontexte zu identifizieren, aus denen heraus der Verlauf und die Resultate dieser Implementation sowie damit verbundene Probleme erklärbar werden.

Zum Schluss geben wir einige Hinweise darauf, wie manche dieser Probleme möglicherweise behoben werden können. Sofern damit mehr Sicherheit gewonnen werden kann für das, was zukünftig getan und was nicht mehr getan werden soll, ist damit immer zugleich der Weg zu künftigen, sich dann neu stellenden Problemen und damit verbundenen Unsicherheiten gebahnt. Evaluation wird also hier verstanden als ein Beitrag zum reflexiven Umgang damit, dass Unsicherheit nicht vermieden, sondern nur bearbeitet werden kann.

Merkmale der deutschen Adaptierung – Programmprofil

»Miteinander – Erfahrungen mit Betzavta« ist ein »Praxishandbuch für die politische Bildung« auf der Grundlage eines in Israel entwickelten Bildungsprogramms. Für die Anlage der Evaluation und die Auswertung der Daten

sind zunächst einige Hinweise auf das im Folgenden zugrunde gelegte Verständnis von politischer Bildung und auf das Programmprofil der deutschen Adaptierung des Programms aus seinem israelischen Entstehungszusammenhang erforderlich.

Politische Bildung in demokratischen Staaten – eine Begriffsklärung

Zunächst eine knappe Festlegung unserer Verwendung des Begriffs politischer Bildung. Dabei sind wir uns angesichts des kaum überschaubaren Bestands an Literatur zur politischen Bildung des Risikos einer solchen Festlegung bewusst – der Versuch, unsere Redeweise transparent zu machen, impliziert also nicht den Anspruch, das einzig mögliche Verständnis politischer Bildung zu formulieren.

Der Versuch politischer Bildung verweist darauf, dass für den Bereich der Politik nicht davon ausgegangen wird, dass Individuen selbstverständlich die Voraussetzungen zur Teilnahme an Politik, insbesondere in demokratischen Staaten mitbringen. In Deutschland geht die Institutionalisierung politischer Bildung auf die Nachkriegsgeschichte und den Kontext der »re-education« als Versuch der politischen Erziehung eines »verführten Volkes« zurück. Politische Bildung zielt daher formal gesehen auf die Vermittlung der sozialen Strukturvoraussetzungen zur Herstellung politischer Entscheidungen im demokratischen Staat.

Diese Voraussetzungen umfassen in der Perspektive der politischen Bildung sowohl die Vermittlung von Wissen über die organisatorischen Voraussetzungen der Politik (Gewaltenteilung, föderaler Staatsaufbau, Bedingungen der Teilnahme: aktives und passives Wahlrecht, Parlament, Aufbau von Verwaltungen) als auch die Vermittlung der Handlungs- und Kommunikationsfähigkeiten auf Seiten der Individuen zur Erfüllung der Teilnahmevoraussetzungen sowie der dafür als erforderlich erachteten Einstellungen. Dies impliziert für zahlreiche politische Bildner zugleich die Vermittlung einer so genannten Handlungs- und Demokratisierungsorientierung. Wir lassen hier offen, ob dies notwendig mit einem Begriff der politischen Bildung verbunden sein muss.

Die Tradition der politischen Bildung ist durch unterschiedliche Schwerpunktsetzungen gekennzeichnet (Massing 1997; Himmelmann 2001). Konjunkturen der Wissensvermittlung über politische Strukturen im engeren Sinne und Konjunkturen, die stärker auf individuelle Teilnahmevoraussetzungen abheben und deshalb stärker auf so genannte Erfahrungs- und Lebensweltzusammenhänge und die Veränderung von Einstellungs- und Ver-

haltensmustern zielen (Edelstein et al. 2001), lösen sich wechselseitig ab. Vor diesem Hintergrund bewegen sich die Varianten politischer Bildung zwischen Konzepten, die enger auf Strukturen des politischen Systems bezogen sind, und solchen, die den Erwerb allgemeiner sozialer Kompetenzen und von Reflexionsvermögen unter Bedingungen sozialer Unsicherheit durch soziales Lernen als Voraussetzung zur adäquaten Ausfüllung der Bürgerrolle zu fördern suchen.

Original und Adaptierung – unterschiedliche soziopolitische Rahmenbedingungen

Das dem Programm »Miteinander« zugrunde liegende Handbuch ist im Hinblick auf beide Varianten offen und eröffnet damit einen Spielraum der Schwerpunktsetzung im Rahmen seiner Verwendung. Dies zeigt sich beim Durchsehen seiner einzelnen Bausteine, die zu einem Teil politisch-organisatorische Strukturen und die dilemmatischen Bedingungen des Handelns darin fokussieren und zu einem anderen Teil eher allgemeine soziale Bedingungen der Konfliktenstehung und die komplexen Voraussetzungen ihrer Handhabung und Bewältigung behandeln.

Diese Offenheit des Programms »Miteinander« findet keine Entsprechung in den ursprünglichen Programmen, die der Adaptierung für Deutschland zugrunde lagen. Weder die ursprüngliche Version »Betzavta« von 1988, aus der die Mehrzahl aller adaptierten Module stammt, noch die Module aus den zur Unterstützung des Friedensprozesses entwickelten Programmen »Building Blocks of Democracy« (1995) und »Peace among Equals« (1994) lassen eine derartige anlassfreie entpolitisierte Interpretation zu. Diese drei genannten Programme wurden – wie übrigens alle Programme des Adam Institute – im Kontext spezifischer sozialer und politischer Konflikte zur expliziten Auseinandersetzung mit den jeweiligen Konfliktsituationen und deren Veränderungsmöglichkeiten entwickelt.

Am deutlichsten wird diese Kontextorientierung vielleicht im Titel eines Programms von 1989 »Coping with the Events in the Territories«, das sich mit den Menschenrechtsverletzungen in den besetzten Gebieten und mit den Widersprüchen zwischen Sicherheitserwägungen und Vergeltungsmaßnahmen gegen palästinensische Zivilisten befasste. Wie politisch die Bildungsarbeit des Instituts ist, zeigte sich daran, dass dieses Programm nicht die Zensur des Unterrichtsministeriums passierte und daher nicht in den Schulen eingesetzt werden konnte.

Demgegenüber stellt das deutsche Programm eine Zusammenstellung

von Modulen dar, die das zugrunde liegende Bildungs- und Interventionskonzept in sehr allgemeiner Form vermitteln, was zum einen mit der Auswahl der Module und zum anderen mit der geringen Kontextualisierung der Bausteine zu tun hat. Während die ursprünglichen Programme eine Art didaktischen Dreischritt vorschlagen von der zunächst allgemeinen Auseinandersetzung mit moralischen Prinzipien über deren Umsetzungsmöglichkeiten im sozialen Nahraum (Klasse, Kollegium, Schule, Kibbuz etc.) bis hin zur Auseinandersetzung mit real-politischen Konflikten, tritt dieser Transferaspekt in der deutschen Version eher in den Hintergrund.²

Zu den keineswegs folgenlosen Veränderungen gehört auch, dass der Programmdidaktik im Sinne einer zielgruppen- und teilnehmerspezifischen Aufbereitung der Programme offensichtlich weniger Bedeutung zugemessen wird. Zumindest enthält der Programmordner, abgesehen von den Hinweisen für den Einsatz in der Schule, keine speziellen Seminarkonzepte und thematischen Vorschläge. Nach unserer Kenntnis finden auch während der Multiplikatorenschulung keine Einführungen zum Umgang mit dem Programm in unterschiedlichen Kontexten statt.

Das Adam Institute ist eine nichtstaatliche pädagogische Organisation, die sich durch internationale Sponsoren erhält und die ihre in einem bestimmten ideologischen Kontext angesiedelten Bildungsangebote an staatliche und nichtstaatliche Bildungsträger »verkauft«. Um die in eigener Regie entwickelten Bildungsprogramme den eigenen Intentionen gemäß zu implementieren, bildet das Institut einen Mitarbeiterstab aus, der in Kooperation mit diversen externen Bildungsträgern die unterschiedlichen Programme zielgruppen- und kontextsensitiv durchführt. In diesem organisatorischen Kontext wird der Qualifizierung des Personals, das diese Programme in der schulischen und außerschulischen Bildungsarbeit einsetzen soll, eine hohe Bedeutung beigemessen.

Hier verfolgt das Adam Institute offensichtlich eine andere Strategie als das Centrum für angewandte Politikforschung. Das Multiplikatorenkonzept wird vom Adam Institute eher kritisch hinterfragt, weil für die Durchführung der Programme nach Auffassung des Instituts nicht nur ein »commitment«

2 Die israelischen Programme gehen von der Annahme aus, dass Einstellungs- und Verhaltensänderungen nur sehr langsam vonstatten gehen. Deshalb nimmt der Anteil an Transferübungen einen großen Anteil ein; insbesondere das ursprüngliche Betzavta – wie auch das Peace-Programm – bieten zahlreiche Modul-Varianten zur Festigung und Anwendung von Erlerntem. Dabei wird ein und dasselbe Grundprinzip in unterschiedlichen Kontexten identifiziert und auf seine Tauglichkeit überprüft, in der Hoffnung, dass bestimmte normative Orientierungen verinnerlicht werden.

zum Institut und dessen politischer Philosophie, sondern auch ein hohes Maß an Professionalität erforderlich ist.

Mit sehr wenigen Ausnahmen werden sämtliche Programme des Instituts durch institutsinternes Personal durchgeführt. Die Mitarbeiter des Instituts werden nach einem klar definierten Auswahlverfahren rekrutiert und nehmen an einer einjährigen Ausbildung teil, in der die kontinuierliche Supervision der Anfänger sowie der Aufbau von Teams eine wichtige Rolle spielen. Die Personalschulung des Adam Institute wird nach der Ausbildung kontinuierlich in den regionalen Zentren fortgesetzt und umfasst sowohl die gemeinsame Weiterbildung in neuen Programmen als auch die Seminarplanung für bestimmte Zielgruppen, aber auch die Nachbereitung von Kursen und Einzelfallbesprechungen sowie zielgruppenspezifische Adaptierungen der Programminhalte.

Die Unterschiede zwischen dem israelischen Ausgangsprogramm und dem adaptierten Programm sind sicher vor dem Hintergrund zu sehen, dass in Israel der israelisch-palästinensische Dauerkonflikt die Relevanz des Programms unmittelbar einsichtig macht – es bedarf keiner Erläuterung, denn der Alltag stellt permanent Anlässe zur Verfügung, die nach Formen der Konfliktbearbeitung verlangen. In Deutschland ist der Ausgangspunkt für die Implementierung von Programmen wie »Miteinander« eher einer allgemeinen Sorge über wachsende Konfliktpotentiale und gegebenenfalls mangelnde Bewältigungskompetenzen geschuldet. Die Verwender solcher Programme finden aber nicht in der gleichen Weise wie in Israel einen unmittelbar einsichtigen Kontext für ihren Einsatz, ihnen fehlen oftmals spezifische Konflikthanlässe. Einzig die Auseinandersetzungen um die Castor-Transporte scheinen einen solchen konkreten Anlass geboten zu haben.

Mit anderen Worten: Die Entwicklung und Verwendung der Programme des Adam Institute waren und sind Reaktionen auf den landesimmanenten Dauerkonflikt, die Programme des Centrums für angewandte Politikforschung haben demgegenüber vorgehenden, präventiven Charakter, veranlasst durch von Intellektuellen und Pädagogen öffentlich artikulierten Sorgen über mangelnde Freiheits-, Demokratie- und Toleranzkompetenz der Staatsbürger.³

3 Dieses Verständnis wird auch aus den einleitenden Worten zum »Praxishandbuch Miteinander – Erfahrungen mit Betzavta« deutlich: »Demokratische Gesellschaften leben von ihrer Wachsamkeit gegenüber Intoleranz und undemokratischen Tendenzen. Fremdenangst, Ausgrenzung von Minderheiten und Gewalt haben auch in Deutschland eine Dimension erreicht, die die demokratischen Errungenschaften unserer Gesellschaft grundlegend in Frage stellen. Die Konsequenzen der derzeitigen Entwicklungen (Individualisierung der Lebensstile, Pluralität der

Vergleicht man vor diesem Hintergrund die Ausgangslage in Israel und die im Rahmen der Adaptierung in Deutschland zu beobachtenden Veränderungen, so wird deutlich, dass die deutsche Version offener und unbestimmter ist und damit einen Kontingenzbereich eröffnet, der den Anbietern, Verwendern und Klienten einen weiten Spielraum zur eigenständigen Interpretation des Programms bietet. Bei der Auswertung geht es daher darum zu untersuchen: Was machen die Befragten auf dem Hintergrund ihrer Qualifikation und unter den organisatorischen Bedingungen, unter denen sie agieren, aus dem Programm »Miteinander« und welche Formen politischer Bildung und davon ausgehender Wirkungen resultieren daraus?

Forschungsdesign der Evaluation

Unter Berücksichtigung der von den Auftraggebern formulierten Vorgaben für die Evaluation wurde folgendes Evaluationsdesign vereinbart:

- Entwicklung eines Fragebogens zur schriftlichen Befragung aller Multiplikatoren aus der Adressdatei
- Durchführung der Erhebung
- Durchführung einer Telefonerhebung zur Validierung der zur Verfügung gestellten Adressdatei der Multiplikatoren
- Auswertung der eingegangenen Fragebögen
- Entwicklung eines Leitfadens für elf Intensivinterviews mit Multiplikatoren
- Durchführung und Auswertung der Intensivinterviews
- Teilnahme am Jahrestreffen der Multiplikatoren und Durchführung einer Gruppendiskussion
- Gespräche mit Experten zur Situation der politischen Bildung
- Auswertung von Sekundärquellen und Literatursichtung zu Problemstellungen der politischen Bildung
- Aufbau der Forschungsbegleitgruppe und Zusammenarbeit
- Kooperation mit Evaluatoren israelischer Bildungsprogramme und israelischen Bildungsinstituten
- Zusammenfassung der Ergebnisse und Gesamteinschätzung

Dieser Anlage der Untersuchung unterliegen folgende Annahmen: Evaluationen auf der Grundlage von Teilnehmerbefragungen, die eine alternative Möglichkeit der Untersuchung gewesen wären, sind bekanntlich mit hohen

Lebensformen) lassen sich auf den Nenner bringen, dass wir heute mehr und andersartige Freiheiten verkraften müssen, als dies bislang der Fall war.« (S. 9).

Unsicherheiten hinsichtlich ihrer Aussagekraft verbunden. Die Befragung von Multiplikatoren bietet gegenüber dem möglichen Nachteil, dass so nur ein mittelbarer Zugang⁴ zu den Wirkungen des Programms erreicht werden kann, einige Vorteile: Im Unterschied zu Teilnehmern besitzen Multiplikatoren praktische Erfahrungen im Hinblick auf die organisatorischen, personellen und programmatischen Bedingungen der Implementierung, Verwendung und Wirkung des Programms.

Dies ist insofern bedeutsam, als damit auch die im Grundlagenpapier des Centrums für angewandte Politikforschung (Ulrich und Wenzel 2001) angelegte Beschränkung der Evaluationsperspektive auf einen rein individualpsychologischen Ansatz erweitert werden konnte und auf allgemeine Bedingungen jeder Bildungsarbeit, damit auch politischer und somit auch der Durchführung des Programms »Miteinander« verwiesen wird. Solche Bedingungen stellen den Rahmen zur Verfügung, in dem jeweils spezifische Zielsetzungen politischer Bildungsarbeit erreicht oder auch verfehlt werden können:

- Bildungsarbeit findet in einem spezifischen organisatorischen Rahmen statt (sei es Schulen, Volkshochschulen, Akademien, staatliche, kirchliche, gewerkschaftliche, betriebliche o.ä. Bildungsträger), in dem formale Bedingungen bezüglich finanzieller, personeller, zeitlicher und räumlicher Ressourcen sowie der Teilnahme mehr oder weniger stark reguliert sind. Im Beispiel: Schulen sind hochgradig regulierte Organisationen hinsichtlich der zugelassenen Teilnehmer, des Personals, der finanziellen, zeitlichen und räumlichen Ressourcen. In vielen Bereichen der politischen Bildung sind demgegenüber die Bedingungen der Bereitstellung der erforderlichen organisatorischen Randbedingungen und Ressourcen weniger strikt reguliert und werden oftmals stärker anlassbezogen ausgerichtet.
- Bildungsarbeit stützt sich auf mehr oder weniger qualifiziertes Personal. Die Bedingungen des Zugangs zu jeweiligen Qualifikationen sowie der lizenzierten Verwendung und des Einsatzes von Bildungsprogrammen sind mehr oder weniger reguliert und formalisiert (wie z. B. im Falle von Leh-

4 Auch deren rückblickende Bewertung der eigenen Erfahrungen als Teilnehmer an einem »Miteinander«-Seminar ist nur bedingt aussagekräftig. Diese Wirkungszuschreibungen erfolgen in zum Teil erheblicher zeitlicher Distanz zur Seminarerfahrung und lassen sich nur schwer von anderen Einflussfaktoren trennen. Besonders deutlich wird dies bei der Betrachtung unserer Interviewensample: Hier finden sich sowohl Multiplikatoren, die unter »Dauereinfluss« des Programms »Miteinander« und dessen Thematik stehen, als auch Befragte, die das Programm lediglich aus der Teilnehmerperspektive kennen. Darüber hinaus gilt auch hier: Jede Rede über die Vergangenheit ist ein Modus (und kein Tempus), sich zur Gegenwart in ein Verhältnis zu setzen.

ren, Erwachsenen- und Jugendbildnern, Sozialpädagogen, Jugendreferenten, Pfarrern, so genannten freien Trainern usw.).

Organisationen schneiden ihre Stellen mit Bezug auf Qualifikationsprofile zu, es besteht aber kein Eins-zu-eins-Verhältnis zwischen Stellen in Organisationen und Qualifikationen. Diese werden im Umfeld des Erziehungs- und Ausbildungssystems erzeugt. Es ist, abhängig von der jeweiligen Organisation, von einer mehr oder weniger engen Kopplung von Stellen und Qualifikationen auszugehen: Im Falle von Schulen hat man es zum Beispiel mit einer engen (Lehrer), im Falle der Erwachsenenbildung mit einer vergleichsweise losen Kopplung zwischen Stelle und Qualifikation zu tun, denn hier eröffnet offensichtlich eine Mehrzahl von Qualifikationsprofilen den Zugang zu mehr oder weniger formalisierten Stellen (z. B. Planstellen oder freie Mitarbeiter).

- Bildungsarbeit verwendet mehr oder weniger ausformulierte Bildungsprogramme. Solche Programme legen die Inhalte, Zielsetzungen und Wege der didaktisch methodischen Realisierung mehr oder weniger explizit fest, ohne damit vollständig die Bedingungen ihrer Interpretation und Verwendung mit festlegen zu können. Dabei kann man solche Programme wieder danach unterscheiden, inwieweit sie spezifische organisatorische Rahmenbedingungen ihrer Umsetzung voraussetzen. Im Beispiel: Bildungspläne und Curricula, wie sie in Kultusministerien formuliert werden, setzen die Schule als organisatorischen Rahmen voraus. Politische Bildungsprogramme treffen oftmals diesbezüglich keine Festlegungen.

Geht man von der erläuterten Unterscheidung von Organisation, Personal und Programm aus, dann sind bei der Implementierung des Programms »Miteinander« in den drei Dimensionen verschiedene Akzentsetzungen zu beobachten.

Die finanzielle Unterstützung von »Miteinander«-Kursen zielt, soweit wir dies überblicken können, vor allem darauf, die Multiplikatoren bei ihrem Versuch der Implementierung des Programms in verschiedene Bildungsorganisationen zu unterstützen. Von zentraler Bedeutung für die Implementierung ist aber wohl der programmatische Aspekt. Es wird vor allem darauf gesetzt, dass das Programm seine Wirksamkeit durch Teilnahme an Kursen und die Distribution des Praxishandbuchs »Miteinander« in den drei genannten Dimensionen entfaltet:

- Organisation: Es wird erwartet, dass relevante Organisationen das Programm in ihr Repertoire übernehmen.
- Personen: Es wird erwartet, dass ein dauerhaft relevantes Bildungsprogramm und eine damit verbundene Qualifikation vermittelt werden.

- Programm: Es wird erwartet, dass das Programm als Bildungsprogramm die Teilnehmer verändert.⁵

Vor dem Hintergrund dieser Überlegungen erschließen sich die Chancen des gewählten Evaluationsdesigns, denn die Befragung der Multiplikatoren eröffnet den Zugang zu dem empirischen Personal des Programms und dessen Qualifikation ebenso wie zu den organisatorischen Bedingungen der Implementierung und Durchführung. Letzteres in dem doppelten Sinne, dass sie nicht nur Erfahrungen mit unterschiedlichen Organisationen machen, in denen sie das Programm zur Geltung bringen, sondern darüber hinaus oftmals die organisatorischen Bedingungen und Ressourcen erst selbst beschaffen müssen und nicht – wie zum Beispiel Lehrer – voraussetzen können.

Schließlich sind die Befragten durch Teilnahme an dem Programm »Miteinander« zu Multiplikatoren geworden und repräsentieren insofern eine mögliche Wirkung des Programms – die Hervorbringung von Multiplikatoren. Zugleich formulieren sie ausgehend davon Vorstellungen über die Bedeutung des Programms für die politische Bildung und gegebenenfalls die allgemeine Lebensführung.

Aber die befragten Multiplikatoren eröffnen darüber hinaus einen weiteren, vermittelten Zugang zur Frage der Wirkung des Programms: Indem wir sie danach gefragt haben, in welcher Weise sie das Programm verwenden, worin sie die Bedingungen des Erfolgs pädagogischer Arbeit im politischen Bildungsbereich sehen und woran sie die Beobachtung von Erfolg oder Misserfolg des Einsatzes des Programms »Miteinander« festmachen, artikulieren sie zugleich Erwartungen im Hinblick auf die Zielsetzungen des Programms »Miteinander« und die Bedingungen ihrer Realisierung.

Geht man davon aus, dass kein Bildungsprogramm in seinem Zuschnitt vollständig die Bedingungen seiner Interpretation und Verwendung zu bestimmen vermag, dann wird über die Darstellung der Ziele, des Einsatzes der Programmbestandteile und die Bedingungen ihrer Realisierung auch deutlich, welche spezifischen Interpretationen und Zielsetzungen durch die Multiplikatoren vorrangig zur Geltung zu bringen versucht werden. Dies erlaubt dann abhängig von jeweiligen Präferenzen zu fragen, ob dies den Zielsetzungen der Adaptierung des Programms durch das Adaptionsteam entspricht.

5 Hier sieht man, dass es zu einer Überschneidung mit der Dimension Personal kommt, denn die Teilnahme an dem Programm »Miteinander« scheint nicht klar unterschieden zu sein von der Qualifikation des Personals zur Durchführung des Programms.

Gegenstand der Untersuchung sind daher hier nicht unmittelbar Wirkungen, sondern Wirkungserwartungen. Man kann von solchen Wirkungserwartungen nicht unmittelbar auf Wirkungen schließen, aber kein Bildungsprogramm und keine politische Bildungsabsicht wird für seine Chancen der Implementierung und Umsetzung die Erwartungen und Einschätzungen der Programmverwender über erzielte, erzielbare oder zu vermeidende Wirkungen für gleichgültig erachten.

Ergebnisse der Evaluation im Überblick

Die Evaluation des Programms »Miteinander« legt einen Bezugsrahmen zugrunde, in dem zur Beurteilung abgelaufener Prozesse und ihrer Resultate nach ihren formalen Strukturmerkmalen und den Auswirkungen ihrer spezifischen Ausprägung gefragt wird. Jede Form von Bildung, so haben wir argumentiert, stützt sich auf programmatische, organisatorische und personale Ressourcen. Dabei lassen sich unterschiedliche Akzentsetzungen beobachten: Es kann stärker auf Organisation, Personal oder Programm gesetzt werden – dies hat dann Auswirkungen auf die durch Bildung erzielbaren Ergebnisse.

Wenn wir im Folgenden eine Gesamteinschätzung der durch die Befragung der Multiplikatoren des Programms »Miteinander« erzielten Ergebnisse vorlegen, dann möchten wir folgende Punkte vorab klarstellen: Es geht uns nicht um eine Beurteilung des Programms »Miteinander« und seine Verwendung durch die Multiplikatoren. Ob man mit diesem Programm politische Bildung in engerem Sinne oder aber Prozesse sozialen Lernens veranlassen sollte, ist von uns nicht zu entscheiden. Dies hängt ab von den Präferenzen der Auftraggeber und Programmentwickler.

Wir haben mit unserer Analyse vor allem versucht deutlich zu machen, welche Ergebnisse mit der Art und Weise der Adaptierung und Implementierung des Programms »Miteinander«, der spezifischen Art und Weise seiner Verwendung und durch den Modus der Rekrutierung des »Personals«, der Multiplikatoren, erzielt worden sind, soweit dies mit Mitteln der Befragung, wie wir sie durchgeführt haben, durchsichtig gemacht werden kann.

Im Folgenden fassen wir die erzielten Ergebnisse entlang der dargelegten Systematik zusammen. Wir gehen zunächst darauf ein, in welcher Weise das Programm »Miteinander« umgesetzt worden ist und welche Wirkungen erzielt worden sind. Wir fragen dann nach den personellen und organisatorischen Rahmenbedingungen, unter denen dies geschehen ist, und welche Bedeutung diese für die Erreichung und/oder Verfehlung der Zielsetzungen

des Programms besaßen. Vor dem Hintergrund dieser Gesamteinschätzung formulieren wir zum Schluss einige Empfehlungen.

Diese Zusammenfassung der Ergebnisse stützt sich auf eine schriftliche Befragung von 69 Multiplikatoren und elf Intensivinterviews, die im Anschluss an die schriftliche Befragung und ihre vorläufige Auswertung geführt wurden. Die Ergebnisse sind im Detail im Abschlussbericht zugänglich gemacht; auf sie kann hier nicht im Einzelnen eingegangen werden.

Programm

Das Programm »Miteinander« wird von seinen Verwendern in der überwiegenden Mehrzahl als Programm zum sozialen Lernen verstanden. Dies lässt sich insbesondere an ihrer Präferenz für bestimmte Übungen des Programms, an ihren Lernzielbeschreibungen und ihren Kriterien für die Beobachtung pädagogischen Erfolgs erkennen. Fragt man nach den organisatorischen und personellen Randbedingungen, die für diese überwiegende Lesart des Programms ausschlaggebend sind, dann sind folgende Punkte von Bedeutung:

- Das Centrum für angewandte Politikforschung hat bei der Implementierung wesentlich auf die Form der Programmdistribution durch Schulung von Individuen und die Produktion des Praxishandbuchs gesetzt und kaum auf die Gewinnung von Bildungsträgern, die das Programm fest in ihr Repertoire übernehmen. Daher ist die Interpretation des Programms – sei es als politisches Bildungsprogramm im engeren Sinne, sei es als Programm zum sozialen Lernen – nicht durch Organisationen reguliert, die das Personal, die Multiplikatoren, auf eine spezifische Umsetzung verpflichten würden. Die Interpretation des Programms ist den Multiplikatoren also freigestellt.
- Die Interpretation von »Miteinander« als Programm zum sozialen Lernen schränkt seine Verwendbarkeit weniger ein. Mit dieser Interpretation des Programms sind weniger Wissensvoraussetzungen verbunden als mit einer Anwendung zur Auseinandersetzung mit einer konkreten Konfliktsituation: Zum einen ist weniger spezifisches Wissen über politische Strukturen erforderlich, zum anderen muss ein geringerer Aufwand der Wissensvermittlung betrieben werden. Das Programm verstanden als Programm des sozialen Lernens schränkt somit den Kreis der potentiellen Teilnehmer weniger ein.

Die Frage nach der Wirkung des Programms haben wir in zweifacher Weise untersucht: Zum einen sind die Befragten, die Multiplikatoren, gewissermaßen selbst eine Wirkung des Programms. Befragt nach der Bedeutung des Programms für sie, wird deutlich, dass für die Mehrheit nach ihrer Ansicht die Teilnahme an dem Programm Auswirkungen auf ihre allgemeine Lebenspraxis hatte und ihre Lebensführung verändert hat – und dies unabhängig von der Frage, welchen Umfang das Seminar hatte, an dem sie teilgenommen haben, und wie erfahren sie selbst in der Verwendung des Programms als Multiplikatoren sind. Dabei machen die Befragten deutlich, dass in ihrer Interpretation eine sinnvolle Verwendung des Programms kaum möglich ist, wenn es nicht auf die gesamte Lebensführung der Multiplikatoren selbst ausstrahlt.

Dieses auf den ersten Blick beeindruckende Ergebnis, das die intensive Wirkung der Teilnahme an dem Programm vor Augen führt, ist aber noch in einem anderen Sinne aufschlussreich für die Frage nach der Wirkung des Programms: Aus den Antworten der Befragten wird nämlich deutlich, dass für sie mit der Verwendung des Programms »Miteinander« ein tendenzielles Einreißen der Differenz zwischen Handeln in beruflichen Rollen und der sonstigen Lebensführung als Person verbunden ist. Dies steht aber quer zu zentralen Strukturmerkmalen der modernen, differenzierten Gesellschaft, für die die Trennung zwischen Rolle und Person konstitutiv ist – insbesondere für die Ausdifferenzierung moderner Organisationen.

Man muss weder für noch gegen diese Differenzierung sein. Für die Frage der Wirksamkeit des Programms ist dies dennoch bedeutsam, da man davon ausgehen muss, dass Individuen ihren sozialen Habitus entsprechend den Sozialstrukturen, in denen sie ihr Leben führen, ausbilden. Sofern das Programm »Miteinander« und seine kompetente Handhabung in der Interpretation der Befragten das Einreißen der benannten Differenz erforderlich macht, kann man fragen, ob es nicht Grenzen seiner Wirkung da erreicht, wo jeweils adressierte Individuen nicht bereit sind, dieses mitzuvollziehen.

Diese Überlegung kann man für spekulativ halten – sie findet aber eine Stütze darin, dass die Multiplikatoren ja selbst eine Wirkung des Programms sind: Es scheint, als wenn vor allem solche Individuen sich durch das Programm in einer Weise angesprochen fühlen, dass sie zu Multiplikatoren werden, die ihre Verwendung des Programms nicht vorrangig als berufliches Handeln und damit als eine formale Qualifikation betrachten, sondern auch als Teil ihrer persönlichen sozialen und politischen Überzeugungen.

Zu vermuten ist dann, dass vor allem solche Individuen als Multiplikatoren angezogen werden, die bereit sind, die benannte Differenz einzureißen, bzw. diese Differenz für problematisch halten. Dies verweist also auf eine

mögliche Grenze der Wirksamkeit des Programms in der spezifischen Verwendungsvariante, wie wir sie bei der Mehrzahl der von uns Befragten gefunden haben.

Eine direkte Übertragung der Wirkung des Programms von den Multiplikatoren auf die Teilnehmer ist nicht möglich – in besonderem Maße, wenn es sich dabei um unfreiwillige Teilnehmer handelt wie zum Beispiel Schüler oder bestimmte Berufsgruppen im Rahmen allgemeiner Fortbildungsmaßnahmen.

Die zweite untersuchte Wirkungsdimension des Programms umfasst die Wirkungserwartungen und -zuschreibungen, die die Multiplikatoren mit der praktischen Durchführung des Programms verbinden. Hierbei wird deutlich, dass pädagogischer Erfolg von den Befragten vorwiegend an Kriterien der Ablaufstrukturen (wie z. B. Mobilisierung der Teilnehmer) und der Teilnehmerveränderung hinsichtlich ihrer Einstellungen und Verhaltenstendenzen und weniger am Erwerb politischen Wissens gemessen wird – ein weiteres Indiz für die Interpretation und Verwendung des Konzepts als Programm zum sozialen Lernen. In einigen Antworten der Multiplikatoren, die das Programm in diesem Sinn verwenden, wird zudem eine quasi-therapeutische Interpretation des Programms »Miteinander« sichtbar.

Die in »Miteinander« formulierten Lernziele werden nach der Einschätzung der Multiplikatoren bei den Teilnehmern ihrer Veranstaltungen in einem sehr hohen Maße erreicht. Diese Bewertung treffen die Multiplikatoren jedoch unabhängig davon, wie häufig und wie lange sie selbst das Programm verwenden (dazu weiter unten).

Zugleich gibt es einen auffälligen Unterschied zwischen den Antworten auf die schriftliche Befragung und den Antworten in den Intensivinterviews. In der schriftlichen Befragung äußern die meisten Befragten die Überzeugung, dass die Teilnahme am Programm langfristige Wirkungen zur Folge hat. In den Einzelinterviews sind die intensiv befragten Multiplikatoren hinsichtlich Wirkungszuschreibungen deutlich zurückhaltender. Letztere bringen ein höheres Maß an Ungewissheit hinsichtlich der Beobachtung von Wirkungen und den daraus möglichen Schlussfolgerungen zum Ausdruck.

Die Untersuchung ergab jedoch insgesamt, dass für die meisten der Befragten sowohl die Übertragung eigener positiver Erfahrungen auf die Teilnehmer als auch ein grundsätzliches Vertrauen in die Wirksamkeit des Programms bei diesen Einschätzungen zum Tragen kommen.

Personal

Das Programm »Miteinander« wird vorwiegend von Multiplikatoren mit pädagogischer Ausbildung verwendet. Dies spiegelt sich auch in den Berufen der Befragten wider: Insgesamt mehr als drei Viertel gehen einer anderen pädagogischen Arbeit nach, davon ein Viertel der Befragten als Lehrer.

Hinsichtlich der Erfahrung im Umgang mit dem Programm zeigten sich erhebliche Unterschiede innerhalb der befragten Multiplikatorengruppe. Die Verwendungshäufigkeit macht deutlich, dass ein Teil der Befragten intensiv mit dem Programm arbeitet, anderen ist es bis zum Zeitpunkt der Befragung noch nicht gelungen, »Miteinander« im Rahmen ihrer jeweiligen Tätigkeit zum Einsatz zu bringen. Die Hälfte der Multiplikatoren hat das Programm bzw. einzelne Bausteine daraus nicht mehr als fünfmal verwendet.

Ein Erfahrungsindex, der von uns auf Grundlage der Variablen Verwendungsdauer und -häufigkeit erstellt wurde, verdeutlicht, dass rund 60 Prozent über wenig Erfahrung im Umgang mit dem Programm verfügen und demgegenüber etwa ein Viertel das Programm intensiv verwendet. Unter Berücksichtigung der Gesamtzahl aktiver Multiplikatoren scheint also ein vergleichsweise kleiner Kreis an Multiplikatoren zu bestehen, der das Programm »Miteinander« intensiv nutzt. Dabei stellen sich berufsspezifische Unterschiede heraus, da es vorwiegend Pädagogen im außerschulischen Bereich sind, die als erfahren gelten können, während insbesondere Lehrer offensichtlich weniger auf »Miteinander« zurückgreifen und dabei – das zeigen die Angaben zum Repertoire verwendeter Übungen – im Vergleich weniger Bausteine einsetzen.

Die Mehrheit der Befragten äußert ihre Zufriedenheit mit der Schulung. Überraschenderweise ist diese Bewertung unabhängig von der Dauer der absolvierten Schulung. Diese schwankte von wenigen Tagen bis zu mehreren Wochen, wobei die Hälfte der Befragten nicht länger als eine Woche teilgenommen hat. Im Zusammenhang mit anderen Ergebnissen sind wir zu dem Schluss gekommen, dass einige der Befragten die Schulung weniger als formale Qualifikation denn als Vermittlung einer Überzeugung und einer veränderten Sicht der Welt verstehen.

Zu den Zielgruppen des Programms zählen sowohl Jugendliche als auch Erwachsene. Auffallend ist bei Letzteren, dass auch hier wiederum pädagogische Berufsgruppen und vor allem Lehrer benannt werden. Während man bei einem Einsatz von »Miteinander« an Schulen und in der Jugendbildung noch von einer sozialen Heterogenität der vom Programm erreichten Individuen ausgehen kann, stellt sich bei der Zielgruppe der Erwachsenen die Frage, ob das Programm hinsichtlich der sozialen Reichweite seine Wirkung vor

allem in pädagogische Milieus hinein entfaltet. Hier scheint eine Art Zirkularität zu bestehen, in der Pädagogen andere Pädagogen als Zielgruppe gewinnen.

Die Frage nach der regionalen Reichweite ist allein aufgrund der vorhandenen Daten nicht eindeutig zu entscheiden, da davon auszugehen ist, dass einige Multiplikatoren – und hier ist vor allem an die freien Trainer zu denken – das Programm an verschiedenen Orten einsetzen. Allerdings fällt bei der Betrachtung der regionalen Verteilung der Befragten die häufige Nennung der Bundesländer Baden-Württemberg, Bayern und Niedersachsen auf. Wir haben darauf hingewiesen, dass diese regionale Konzentration abhängig von den Zielsetzungen der Implementierung bewertet werden muss und somit nicht zwangsweise problematisch ist. Aus den Ergebnissen wird jedoch deutlich, dass die spezifische Rekrutierungsform des Personals und dessen Qualifizierung Folgen für die soziale Reichweite und die Wirksamkeit des Programms haben.

Organisation

Aus den dargestellten Befunden deutet sich bereits an, dass bei der Implementierung von »Miteinander« die Akzentsetzung auf dem programmatischen Aspekt lag. Unter anderem wurde dies durch die unterschiedliche Kennzeichnung des Begriffs »Schulung« deutlich, der offensichtlich nicht eindeutig definiert ist.

Die untergeordnete Bedeutung organisatorischer Regulation zeigt sich darüber hinaus auch daran, dass die Mehrheit der Multiplikatoren nicht in der Lage ist, unterstützende oder erschwerende Rahmenbedingungen zu identifizieren bzw. angibt, dass solche Rahmenbedingungen in ihrem Arbeitskontext nicht vorhanden sind. Allerdings werden solche Aussagen vorwiegend von Multiplikatoren mit wenig Erfahrung im Umgang mit dem Programm gemacht, während die erfahrenen Befragten vergleichsweise oft auf positive und negative Rahmenbedingungen hinweisen. Dies legt die Vermutung nahe, dass eine stabile organisatorische Verankerung mit entscheidend für eine erfolgreiche und kontinuierliche Verwendung des Programms ist.

Vor diesem Hintergrund kann man die Frage stellen, ob das Fehlen von Organisation mitverantwortlich ist für die Polarisierung in solche Multiplikatoren mit Erfahrung im Umgang mit dem Programm gegenüber solchen ohne bzw. mit nur wenig Erfahrung, denn das Fehlen von Ressourcen schränkt offensichtlich die Aussichten einer erfolgreichen Anwendung des Programms erheblich ein.

Empfehlungen für die zukünftige Arbeit

Die nachfolgenden allgemeinen Empfehlungen sind im Bezugsrahmen der zuvor erläuterten Dimensionen Programm, Personal und Organisation formuliert. Diese Empfehlungen wurden von uns nach einer Phase der kommunikativen Validierung mit Experten der politischen Bildungsarbeit und Multiplikatoren des Programms »Miteinander« erstellt. Dabei wurden auch konkrete Vorschläge einbezogen, die die befragten Multiplikatoren selbst in der schriftlichen und mündlichen Befragung machten.

- Vor dem Hintergrund der genannten verschiedenen Interpretationsmöglichkeiten des Programms kann die Frage gestellt werden, ob das Programm in der vorliegenden offenen Form beibehalten werden soll oder ob bestimmte Verwendungsmodi explizit präferiert oder auch ausgeschlossen werden sollen. Kontingenz ist auf der Programmebene allein sicher nicht einschränkbar. Weitere Einschränkungen können durch Vorkehrungen auf den Ebenen von Organisation und Personal vorgenommen werden (dazu auch nachfolgend).

Gleichwohl kann man mit der vorliegenden Evaluation nunmehr um empirisch verschiedene Interpretations- und Verwendungsweisen des Programms wissen und darüber nachdenken, ob diese mit den angestrebten Zielsetzungen übereinstimmen oder davon abweichen.

- Die Qualifikationsvoraussetzungen für die Teilnahme an einer Schulung und die Arbeit als Multiplikator sollten klarer definiert werden. Um eine Verwendung des Programms zu gewährleisten, die sich stärker an Standards professionellen Handelns orientiert, wäre zu überlegen, ob die Angebote des Programms »Miteinander« nicht in dem Sinne differenziert werden müssten, dass Kurse für »Teilnehmer« stärker von solchen Kursen unterschieden werden, die die Schulung künftiger »Multiplikatoren« einerseits und das Training zukünftiger »Ausbilder« andererseits zum Ziel haben. Dies beinhaltet die Entwicklung unterschiedlicher Seminarkonzepte für »Teilnehmer«, zukünftige »Multiplikatoren« und zukünftige »Ausbilder«. Über entsprechende unterschiedliche Zertifizierungen wäre gegebenenfalls nachzudenken.
- Für die Multiplikatorenschulung ist zu bedenken, dass zukünftige Multiplikatoren des Programms – wie aus den Ergebnissen deutlich wurde – unterschiedliche Zielsetzungen verfolgen. Aus diesem Grund sollten schon im Vorfeld der Schulung der berufsspezifische Fortbildungsbedarf ermittelt und dementsprechend teilnehmerorientierte Fortbildungsangebote konzipiert werden.

Beispielsweise sollten zukünftige Multiplikatoren, die das Programm im

schulischen Bereich einsetzen, eine entsprechende Fortbildung zur Integration des Programms in den schulischen Kontext erhalten. Multiplikatoren, die im Bereich der politischen Bildung tätig sind, sollten darin geschult werden, das Programm mit spezifischen Themen der politischen Bildung zu verknüpfen.

- Es wäre darüber hinaus zu überlegen, ob Multiplikatoren in der ersten Phase des Einsatzes des Programms nicht durch Supervision und Praxisbegleitung unterstützt werden sollten.
- Einige der erfahrenen Multiplikatoren betonen, dass die Ausbilder eines derartig voraussetzungsreichen Bildungskonzepts über fundierte theoretische Kenntnisse in den Bereichen Politologie, Rechtsphilosophie und Sozialpsychologie sowie über weit reichende pädagogisch-psychologische Qualifikationen verfügen sollten (und über diese oftmals nicht verfügen). Daher wird auch empfohlen, dass Ausbilder sich auch in diesem Bereich regelmäßig weiterbilden.
- Aus der Untersuchung geht hervor, dass allgemein die pädagogische Arbeit nur in geringem Maße evaluiert wird. Im Rahmen der Schulungen sollten die zukünftigen Multiplikatoren über Voraussetzungen, Möglichkeiten und Methoden der Selbstevaluation ihrer Arbeit informiert und in der Durchführung solcher Selbstevaluationen angeleitet werden.
- Die Ergebnisse zeigen deutlich, dass die organisatorische Verankerung bzw. ihr Fehlen einen bedeutenden Faktor für den regelmäßigen Einsatz des Programms, seine Verwendungsweise und seine Reichweite darstellt. Um hier Verbesserungen zu erreichen, kann versucht werden, in regelmäßiger Kooperation mit Organisationen der politischen Bildung, der Jugendarbeit, Schulen, Verwaltungen und Unternehmen kontextsensitive Implementierungsstrategien zu entwickeln. Damit würde zugleich stärker auf Organisation als ein tragendes Element zur Verankerung des Programms in der bildungspolitischen Landschaft gesetzt.
- Neben verstärkten Kooperations- und Vernetzungsbemühungen wäre darüber hinaus eventuell auch über eine Koppelung mit anderen Projekten (nicht nur) der Bertelsmann Stiftung nachzudenken – wie zum Beispiel den Einbezug von »Miteinander« in die von der Stiftung initiierten und unterstützten Schulentwicklungsprogramme.
- Für zukünftige Einschätzungen von politischen Bildungsprogrammen sollte bei ihrer Implementierung und Durchführung darauf geachtet werden, dass zuverlässige Formen der Dokumentation der geleisteten Arbeit betreffend Personal, Teilnehmer und Organisationsformen entwickelt werden.

Literatur

- Altrichter, Herbert et al. (1997): »PraktikerInnen als ForscherInnen. Forschung und Entwicklung durch Aktionsforschung«. In: Barbara Friebertshäuser und Annedore Prengel (Hrsg.): *Handbuch qualitative Forschungsmethoden in der Erziehungswissenschaft*. Weinheim und München. 640–660.
- Bargal, David, Martin Gold und Miriam Lewin (1992): »The Heritage of Kurt Lewin. Theory, Research, and Practice«. In: *Journal of Social Issues* (48) 2.
- Bommes, Michael (1997): *Ausländerbeauftragte in den Kommunen. Eine Untersuchung zu ihrer Funktion, ihren Aufgaben und ihrer Stellung in den kommunalen Verwaltungen, Abschlussbericht über ein Forschungsprojekt, gefördert von der Beauftragten der Bundesregierung für die Belange der Ausländer*. Osnabrück und Florenz.
- Bommes, Michael, Frank-Olaf Radtke und Hans-Erich Webers (1995): *Schulpraktische Studien an der Johann-Wolfgang-Goethe-Universität*. 2. Auflage. Bielefeld.
- Bommes, Michael, Bernd Dewe und Frank-Olaf Radtke (1996): *Sozialwissenschaften und Lehramt. Der Umgang mit sozialwissenschaftlichen Theorieangeboten in der Lehrerbildung*. Opladen.
- Brunsson, Nils (1989): *The Organization of Hypocrisy*. New York.
- Creswell, John W. (1998): *Qualitative Inquiry and Research Design. Choosing Among Five Traditions*. London.
- Edelstein, Wolfgang, Fritz Oser und Peter Schuster (Hrsg.) (2001): *Moralische Erziehung in der Schule*. Weinheim.
- Deutsche Gesellschaft für Evaluation e.V. (2002): *Standards für Evaluation*. Köln.
- Eser Davolio, Myriam (1999): *Prävention und Bekämpfung von fremdenfeindlichen, rassistischen und gewaltaffinen Einstellungen bei BerufsschülerInnen*. Bern.
- Friebertshäuser, Barbara, und Annedore Prengel (Hrsg.) (1997): *Handbuch qualitative Forschungsmethoden in der Erziehungswissenschaft*. Weinheim und München.
- Himmelmann, Gerhard (2001): *Demokratie Lernen als Lebens-, Gesellschafts- und Herrschaftsform*. Schwalbach/Ts.
- Kohlberg, Lawrence (1987): »Moralische Entwicklung und demokratische Erziehung«. In: Georg Lind und Jürgen Raschert (Hrsg.): *Moralische Urteilsfähigkeit*. Weinheim und Basel. 25–43.
- Kvale, Steinar (1996): *InterViews. An Introduction to Qualitative Research Interviewing*. London.

- Luhmann, Niklas (1964): *Funktion und Folgen formaler Organisation*. Berlin.
- Luhmann, Niklas (1981): *Politische Theorie im Wohlfahrtsstaat*. München und Wien.
- Luhmann, Niklas (2000): *Organisation und Entscheidung*. Wiesbaden.
- Luhmann, Niklas, und Karl Eberhard Schorr (1992): *Zwischen Absicht und Person. Fragen an die Pädagogik*. Frankfurt am Main.
- Lynen von Berg, Heinz und Roland Roth (Hrsg.) (2003): *Maßnahmen und Programme gegen Rechtsextremismus wissenschaftlich begleitet*. Opladen.
- Maoz, Ifat (2000): *Evaluation of the Joint Israeli-Palestinian Peace Curriculum*. Unveröffentlichtes Manuskript. Jerusalem.
- Maroshek-Klarman, Uki (1995): *Education for Peace among Equals Without Compromises and Without Concession*. Jerusalem.
- Maroshek-Klarman, Uki (1996): *Miteinander. Erziehung zur Demokratie*. Jerusalem.
- Massing, Peter (1996): »Plädoyer für einen politischen Politikunterricht«. In: Dorothea Weidinger (Hrsg.): *Politische Bildung in der Bundesrepublik*. Opladen. 124-127.
- Massing, Peter, und Georg Weißeno (Hrsg.) (1997): *Politische Urteilsbildung*. Schwalbach/Ts.
- Monreal, Guido (2002): *Der Bildungsansatz »Betzavta« zwischen Moralerziehung und politischer Bildung. Diplomarbeit im Fachbereich Politik- und Sozialwissenschaften*. Freie Universität Berlin.
- Nevo, David (1995): *School-based Evaluation. A Dialogue for School Improvement*. Tel Aviv.
- Piaget, Jean (1979): *Das moralische Urteil beim Kinde*. Stuttgart.
- Seberich, Michael, und Florian Wenzel (2002): »Handeln reflektieren – Demokratie erleben. Das Projekt Erziehung zu Demokratie und Toleranz«. In: Gotthard Breit und Siegfried Schiele (Hrsg.): *Demokratie-Lernen als Aufgabe der politischen Bildung*. Schwalbach/Ts. 241–256.
- Ulrich, Susanne, Thomas R. Henschel und Eva Oswald (2001). *Miteinander – Erfahrungen mit Betzavta. Praxishandbuch für die politische Bildung*. Ein Praxishandbuch auf der Grundlage des Werks »Miteinander« von Uki Maroshek-Klarman, Adam Institute, Jerusalem. 3. Auflage. Gütersloh.
- Ulrich, Susanne, und Florian Wenzel (2001): *Evaluation von Demokratie- und Toleranzerziehung*. München. Internes Manuskript.
- Ulrich, Susanne und Florian Wenzel (2003): *Partizipative Evaluation. Ein Konzept für die politische Bildung*. Gütersloh.
- Weick, Karl E. (1985): *Der Prozess des Organisierens*. Frankfurt am Main.
- Wolff-Jontofsohn, Ulrike (1999): »Das Adam-Institut Jerusalem. Friedenserziehung durch Sicherung und Verbreitung demokratischer Werte«. In:

Ulrike Wolff-Jontofsohn: *Friedenspädagogik in Israel*. Schwalbach/Ts. 2. Auflage. 342–396.

Wolff-Jontofsohn, Ulrike (2000): »Friedenspädagogik in Israel. Chancen und Grenzen«. In: Deutsch-Israelischer Arbeitskreis für Frieden im nahen Osten e.V. (DIAK) (Hrsg.): *Israel & Palästina. Themenheft Friedensforschung und Friedenspädagogik in Israel* 58. Schwalbach/Ts. 4–23.

Wolff-Jontofsohn, Ulrike (2001): »Empfehlungen für den Schulunterricht«. In: Susanne Ulrich, Thomas R. Henschel und Eva Oswald (Hrsg.): *Miteinander – Erfahrungen mit Betzavta. Praxishandbuch für die politische Bildung*. Ein Praxishandbuch auf der Grundlage des Werks »Miteinander« von Uki Maroshek-Klarman, Adam Institut, Jerusalem. 3. Auflage. Gütersloh. 151–173.

Schule und Ausbildung für Toleranz und Demokratie: Ein Projekt der Europäischen Jugendbildungs- und Jugendbegegnungsstätte Weimar und seine wissenschaftliche Begleitung

Stephan Eschler, Erich Schäfer, Sandra Uhl

Das Projekt

Der Alltag in einer Bildungsstätte ist oftmals nicht ganz einfach, besonders wenn es um Erwartungshaltungen aus anderen Bereichen des Bildungssystems geht. Ein durchaus fiktives, im Kern aber immer wieder erlebtes Beispiel: Der Anruf einer Lehrerin: »Meine Schüler sind so fremdenfeindlich – können Sie die nicht mal toleranter machen? Ich hatte so an einen halben Tag bei uns in der Schule im Rahmen der nächsten Projektwoche gedacht.« Manchmal gelingt es, mit einer entsprechenden Förderung, die Klasse in die Bildungsstätte zu holen und mit den Jugendlichen ein bis zwei Tage zu arbeiten. Manches kann angesprochen, bedacht, entwickelt werden. Am Ende steht ein freundliches Feedback und die Frage für uns als Pädagogen: Hätte nicht noch mehr erreicht werden können?

Aus diesen Erlebnissen und den positiven Erfahrungen in der Arbeit mit den Trainingsprogrammen zur Demokratie-, Toleranz- und Menschenrechts-erziehung, die beim Centrum für angewandte Politikforschung an der Universität München und der Bertelsmann Stiftung adaptiert bzw. entwickelt worden sind (»Achtung (+) Toleranz«, »Eine Welt der Vielfalt« und »Miteinander – Erfahrungen mit Betzavta«), entstand zunächst eine Vision: mehrjährige Projekte mit Schulen zu initiieren, in denen ab dem 8. Jahrgang die Klassen jedes Jahr Gelegenheit haben, an einem Demokratie- und Toleranz-training teilzunehmen. Vielleicht könnte es gelingen, so der Gedanke, über diese mittelfristige pädagogische Intervention den Jugendlichen tatsächlich neue Erkenntnisse im Umgang miteinander, aber auch gegenüber Fremden oder Fremdem zu vermitteln.

Mitten in diese Zeit des Phantasierens über ideale Modelle der außerschulischen politischen Bildung kam das Förderprogramm XENOS der Bundesregierung (»XENOS – Leben und Arbeiten in Vielfalt«) mit der Chance, diese Vision umzusetzen. Die Förderrichtlinien eröffneten die Möglichkeit, über

drei Jahre lang in einem Projekt zu erproben, ob dieser Ansatz von Bildungsarbeit etwas mehr bringt als die üblichen kurzzeitpädagogischen Angebote im Alltag einer Jugendbildungsstätte.

Am 1. Januar 2002 konnte in der Europäischen Jugendbildungs- und Jugendbegegnungsstätte Weimar (EJBW) das Projekt »Schule und Ausbildung für Toleranz und Demokratie (SAFT)« beginnen. Ein großer Vorteil ist die dreijährige Absicherung durch Mittel aus dem Europäischen Sozialfonds (neben dieser Förderung wird das Projekt durch das Thüringer Kultusministerium, die Bundesanstalt für Arbeit und die Bertelsmann Stiftung und aus Eigenmitteln der Bildungsstätte finanziert), die den zuverlässigen Rahmen für dieses Modellprojekt darstellt. Zum ersten Mal wird in der Bundesrepublik der Versuch unternommen, mit den Trainingsprogrammen der Bertelsmann Stiftung in einem längerfristigen Projekt kontinuierlich mit Schul- und Ausbildungsklassen zu arbeiten.

Die Inhalte des Projektes

Das Projekt setzt sich aus mehreren Bestandteilen zusammen: Schwerpunkt sind drei fünftägige Trainingskurse mit verschiedenen Schul- und Ausbildungsklassen unterschiedlicher Thüringer Schulen. Im Laufe von etwa zwei Kalenderjahren (aber in drei Schul- bzw. Ausbildungsjahren) kommen dieselben Jugendlichen nach Weimar und nehmen an einem Demokratie- und Toleranztraining teil.

Neben den Seminaren für Jugendliche werden Fortbildungen für Lehrer durchgeführt und Seminare für die Eltern der teilnehmenden Jugendlichen angeboten. Im Sommer 2004 ist zudem ein Seminar auf Grundlage des Konzeptes »Ohne Angst verschieden sein« (Khanide und Giebeler 2003) als Ergänzungsangebote geplant.

Die vorrangigen Ziele des Projektes sind neben dem Erwerb bzw. der Vertiefung von sozialen Schlüsselqualifikationen wie Kommunikations-, Kooperations-, Toleranz- und Demokratiefähigkeit auch die Entwicklung eines demokratischen Klimas an den Schulen bzw. im Ausbildungsbetrieb. Im Rahmen des Projektes werden zudem Studentinnen sehr praxisnah auf eine künftige Tätigkeit als Trainerinnen für Demokratie- und Toleranzernziehung vorbereitet. Neben der eigenen Qualifizierung in Ausbildungskursen für die Trainingsprogramme haben sie parallel die Möglichkeit, sich selbst in der Praxis auszuprobieren.

Die Seminare für Jugendliche

Im ersten Projektjahr fand für alle Klassen ein Einführungsseminar statt. Grundlage waren die erwähnten Trainingsprogramme der Bertelsmann Stiftung und des Centrums für angewandte Politikforschung. Mit dem ersten Training sollten die Jugendlichen mit den Methoden und Lernformen im Rahmen der Demokratie- und Toleranzerziehung vertraut gemacht werden.

Thematisch stand in diesem Seminar das Miteinander im Klassenverband im Mittelpunkt. Ausgehend von einigen biographisch orientierten Übungen haben wir mit den Teilnehmern u. a. folgende Fragen bearbeitet: Wie gehen wir miteinander um, auch (und besonders) in Konflikt- oder Entscheidungssituationen? Was wollen wir an unserem Miteinander verbessern?

In diesem Seminarkonzept gab es einen großen Anteil an erlebnispädagogischen Interaktionsübungen. Diese Übungen stellten eine Ergänzung der Methoden aus den Trainingsprogrammen dar. Damit wird erstmals der Versuch unternommen, Demokratie- und Toleranzerziehung und Methoden der Erlebnispädagogik miteinander zu verknüpfen. So stehen neben den eher kognitiven Übungen der Trainingsprogramme erfahrungs- und erlebnisorientiertere Übungen im Seminarkonzept nebeneinander. Letztere sollten den Jugendlichen auf einer anderen Ebene als in Diskussionsrunden im Seminarraum Reflexionen über das eigene Verhalten ermöglichen und die Stärken und Schwächen ihrer Klasse verdeutlichen.

Das zweite Seminar bestand aus Übungen des Trainingsprogramms »Achtung (+) Toleranz«. Ein Schwerpunkt dieses Seminars war die Vermittlung von Definitionen der Begriffe »Konflikt«, »Gewalt« und »Toleranz«. Diese sind nötig, um das von der Bertelsmann Forschungsgruppe Politik am Centrum für angewandte Politikforschung entwickelte operationalisierbare Toleranzmodell für die politische Bildung zu verstehen.

Zusammen mit der so genannten »Toleranzampel«, einem zentralen didaktischen Element des Programms »Achtung (+) Toleranz«, kann es ein Werkzeug zur Selbstüberprüfung in schwierigen zwischenmenschlichen Beziehungen sein und helfen, den Begriff »Toleranz« für den Alltag handhabbarer zu machen (Bertelsmann Forschungsgruppe Politik 2001).

Zum anderen geht es im Programm »Achtung (+) Toleranz« um das Kennenlernen von Grundlagen der Kommunikation und das Einüben von Elementen der partnerschaftlichen Kommunikation als einer Möglichkeit, eine »Anerkennung des gleichen Rechts auf freie Entfaltung« zu entwickeln und damit tolerantes Verhalten im Alltag zu ermöglichen.

Die im zweiten Seminar zusätzlich verwendeten Interaktionsübungen hatten einen stärkeren thematischen Bezug zu den inhaltlichen Themen als im

Einführungsseminar. Mit ihrer Hilfe wurden Brücken zu den theoretischen Inputs gebaut, um den Teilnehmern damit den Zugang zu für sie ungewohnten Themen zu erleichtern.

Der dritte und letzte Trainingskurs für die Jugendlichen wird mit Übungen aus dem Trainingsprogramm »Miteinander – Erfahrungen mit Betzavta« durchgeführt. Zu den im veröffentlichten Handbuch enthaltenen Übungen sind im Laufe der letzten Jahre neue Übungen hinzugekommen, die das Spektrum der thematischen Zugänge zu einem qualitativen Demokratieverständnis deutlich erweitern und das Trainingsprogramm gerade für Jugendliche attraktiver machen, da diese neuen Übungen handlungs- und erlebnisorientierter sind.

Mit dem Seminarkonzept soll ein Bogen geschlagen werden zwischen dem Miteinander innerhalb des Klassenverbandes zu politischen Fragestellungen der Gegenwart. Im Rahmen des Seminars sollen die Teilnehmer Demokratie als Lebensform entdecken und gemeinsam gestalten, von diesen Erfahrungen ausgehend den Blick auf die Gesellschaft richten und so die Chancen der Demokratie als Staatsform erkennen lernen.

Design und Methode der Evaluation

Dem vielfältigen und breiten Erfahrungswissen praktischer Bildungsarbeit im Bereich der Erziehung zu Toleranz, Menschenrechten und Demokratie steht kein entsprechend breit fundiertes Wirkungsverständnis gegenüber. Eine Auseinandersetzung mit den Fragen der Wirkungen der Bildungsangebote und der dabei eingesetzten Instrumente sowie vermehrte Anstrengungen zur Evaluierung und wissenschaftlichen Begleitung dieser Bildungsarbeit sind somit dringend geboten und Voraussetzung zur Entwicklung von umsetzbaren Konzepten der Qualitätssicherung und -entwicklung in diesem Bereich.

Noch sind Versuche, die politische Bildungsarbeit empirisch zu untersuchen, sehr vereinzelt (vgl. Rudolf und Zeller 2000: 39). Praxisnahe Vorarbeiten und Bestandsaufnahmen hat u. a. das von der Bertelsmann Stiftung initiierte International Network Education for Democracy, Human Rights, and Tolerance (www.tolerance-net.org) vorgelegt.

Die Trainings- und Seminarkonzepte zur Toleranzerziehung, die im Rahmen des Projektschwerpunktes »Erziehung zu Demokratie und Toleranz« der Bertelsmann Stiftung entwickelt bzw. adaptiert und erprobt wurden, fußen auf der Überzeugung, dass eine demokratische Gesellschaft auf Toleranz angewiesen ist.

Intoleranz – in welcher Form auch immer – »schafft zunehmend Konflikte, die im Vorfeld der Rechtsordnung bzw. der Anwendung des staatlichen Gewaltmonopols nur durch Toleranz zu regeln sind.« Toleranz wird zur existentiellen Bedingung von Demokratie und Erziehung und Toleranz damit zur weltweiten Zukunftsaufgabe.

»Demokratielernen« bedeutet mehr, als das Lernen von demokratischen Strukturen und geht über das übliche Verständnis der Demokratie als Staatsform hinaus. Die Frage nach Demokratie und demokratischen Verhaltensweisen stellt sich nicht nur in der Politik, sondern auch in der Schule, in der Familie und in jeder alltäglichen Interaktion. Demokratie als Lebensform zu erlernen bedeutet somit, zu lernen, wie alltägliche Konflikte demokratisch und gleichberechtigt gelöst werden können. Dieser Prozess ist kaum durch Wissen zu vermitteln, sondern vielmehr durch Übung zu erfahren.

Die Evaluation von Methoden der Erziehung zu Toleranz und Demokratie als Projekt der politischen Bildung stellt erhebliche Anforderungen an die wissenschaftliche Begleitung. Zum einen erweist sich die »Messung« von Toleranz- und Demokratievariablen als komplexe und zum Teil umstrittene Herausforderung. Zum anderen begleitet das Projekt der Europäischen Jugendbildungs- und Jugendbegegnungsstätte Weimar sowohl Auszubildende wie Schüler in einem mehrjährigen Lernprozess, der den spezifischen politisch-sozialisatorischen Kontext der neuen Bundesländer berücksichtigen muss und zugleich den Auftrag der Projektinitiative der Bundesregierung »XENOS – Leben und Arbeiten in Vielfalt« erfüllen will: Toleranz und demokratischen Respekt gegenüber Angehörigen anderer Kulturen und Nationen zu fördern. Eine wissenschaftliche Begleitung, die diese Mehrdimensionalität der Problemlage adäquat abbilden möchte, erfordert einen erheblichen Ressourceneinsatz, der an dieser Stelle aber nur bedingt zur Verfügung steht.

Konzeption der wissenschaftlichen Begleitung

Die Evaluation von Angeboten der politischen Bildung sollte dem Bildungsziel angemessen sein: Eine auf demokratischen und partizipativen Prinzipien beruhende Bildungsarbeit, die Menschen zu selbstverantwortlichem Handeln als politischen Subjekten befähigen will, darf diese in der Evaluation nicht zu Objekten machen, sondern muss sie beteiligen.

Wenngleich der Anspruch an eine partizipative Evaluation (Ulrich und Wenzel 2003) prinzipiell aufrechtzuerhalten ist, so stehen ihm Probleme der praktischen Realisation entgegen, die mit der Ressourcenausstattung verbunden sind. Insbesondere eine formative, d.h. prozessbegleitende Evalua-

tion, die den zeitgemäßen Ideen einer Qualitätssicherung und -entwicklung nahe steht, ist unter den gegebenen Bedingungen nur bedingt möglich.

Um dennoch den demokratiepolitischen Impulsen des Projektes auch in der wissenschaftlichen Begleitung eine komplementäre Strategie anzubieten, soll die summative Evaluation durch Methoden der Selbstevaluation ergänzt werden, die die beteiligten Akteure handlungsorientiert bei der Selbstreflexion unterstützt.

Strategien der Selbstevaluation erscheinen darüber hinaus auch insoweit methodisch angemessen, dass sie eine langfristige Selbstbeobachtung der mit dem Projekt eingesetzten und weiterentwickelten Instrumentarien erlauben. Darüber hinaus wird Wert darauf gelegt, die im Vordergrund stehenden quantitativen Methoden durch qualitative zu ergänzen, also eine Methoden-triangulation durchzuführen.

Wenn im Rahmen des Projektes auch keine reine partizipative Evaluation realisiert werden kann, so findet aber eine kooperative Evaluation zwischen dem Projektträger der Europäischen Jugendbildungs- und Jugendbegegnungsstätte Weimar und dem Fachbereich Sozialwesen der Fachhochschule Jena statt.

Kooperation bedeutet im Rahmen dieser Evaluation eine Kombination aus Fremd- und Selbstevaluation des Projektes. Dabei werden die grundsätzlichen Fragestellungen der Erhebungsinstrumente zwischen den Partnern abgesprachen. Es wird das Ziel verfolgt, das Evaluierungsinstrumentarium als Hilfsmittel projektbezogener Planungs-, Monitorings-, Evaluierungs- und Selbstlernprozesse in Kooperation mit den Projektverantwortlichen weiterzuentwickeln und ihnen zur weiteren eigenverantwortlichen Nutzung zur Verfügung zu stellen. Damit ist neben den summativen Ergebnissen des Evaluationsprozesses gleichzeitig auf der formativen Ebene die Möglichkeit gegeben, regulierend in die Planung der einzelnen Trainingsseminare einzugreifen.

Diesem Zweck dient auch die kommunikative Validierung. Ziel ist nicht nur, die Projektverantwortlichen und -beteiligten über Prozesse und Resultate der Evaluation zu informieren, sondern diese darüber hinaus auch mit ihnen kritisch zu diskutieren: Einerseits soll diesem Personenkreis die Möglichkeit geboten werden, anhand der vorliegenden Ergebnisse ihre Arbeit zu überprüfen und gegebenenfalls zu modifizieren. Andererseits soll die Validität der Daten und ihre Interpretation durch diesen Personenkreis einer kritischen Betrachtung unterzogen werden – so kann möglichen Fehlinterpretationen entgegengewirkt und eine ergänzende Sichtweise auf die Untersuchungsbefunde gewonnen werden.

Das Projekt und seine Evaluation sind in mehrfacher Hinsicht ein Pilot-

projekt: Zum einen gibt es bisher keine empirischen Erfahrungen für den längerfristigen Einsatz der Trainingsprogramme bei Jugendlichen, zum anderen ist die Kombination von Demokratie- und Toleranzerziehung mit der Erlebnispädagogik ein noch wenig bearbeitetes Feld (vgl. Boock 2002). Für beide Herausforderungen sind die Ergebnisse des Evaluationsprozesses von großer Bedeutung.

Die Ziele der Evaluation

Die Ziele der Evaluation werden durch die Zielstellung des Projektes bestimmt. Es hat sich zur Aufgabe gemacht, soziale Schlüsselkompetenzen der Schüler zu fördern, das demokratische Miteinander in den Seminaren erlebbar zu machen und das demokratische Miteinander an der Schule der Projektteilnehmer zu stärken. Es geht also um den sowohl kognitiven wie emotionalen Erwerb von Kenntnissen und Fähigkeiten sowie ihren Transfer in den Schulalltag. Damit steht die Frage der Wirkung der Seminare im Zentrum der Evaluation.

Wenn hier die Rede von Wirkung ist, so bedarf dieser Begriff einer Erklärung. Mit Wirkung sind nicht objektive Entitäten gemeint, die in einem kausalen Ursache-Wirkung-Verhältnis eindeutig und zweifelsfrei auf bestimmte Bildungsinhalte zurückzuführen sind. Wirkungen haben u.E. vielmehr einen konstruktivistischen Charakter; sie entstehen in kommunikativen Prozessen als eine gemeinsame Konstruktion sozialer Wirklichkeit, die sich zwar wesentlich aus inneren Deutungsmustern der Individuen speist, sich aber gleichzeitig in ihren Handlungen manifestiert. In diesem Sinne geht es um die Untersuchung der realen und potenziellen Wirkungen der Seminare auf unterschiedlichen Ebenen. Wir nehmen dabei folgende Wirkungen in den Blick:

- Die individuelle Ebene der Persönlichkeitsentwicklung erschließt sich uns über die quantitativen und qualitativen Befragungsergebnisse.
- Die institutionelle Ebene, gemeint ist damit der Kontext der Schule, bildet sich ebenfalls in den Befragungsdaten ab.
- Die gesellschaftliche Ebene, also die Auswirkungen der Bildungsarbeit auf das Bürgerverhalten, bleibt zunächst außerhalb der Reichweite der uns zugänglichen Betrachtung.

Auf der dritten Ebene tut sich ein Hiatus auf zwischen dem Anspruch nach verlässlichen Aussagen über die (Miss-)Erfolge politischer Bildungsprogramme in der Gesellschaft und dem, was eine zeitlichen, materiellen und

personellen Friktionen unterliegende Evaluationsforschung zu leisten vermag. Allerdings bestehen berechnete Hoffnungen, dass identifizierte Wirkungen auf individueller und institutioneller Ebene nicht ohne Folgen auf der gesellschaftlichen bleiben werden. Ansatzpunkte dafür zu identifizieren bleibt das Bestreben dieser Evaluation.

Um verstehen zu können, was sich auf den drei Ebenen vollzieht, gilt es, die Impulse der Lernprozesse zu identifizieren. Deshalb soll der Lernprozess der Projektteilnehmer nachgezeichnet werden. Auf dieser Basis können methodische und didaktische Einflüsse der Seminare verstanden werden.

In der Zusammenschau der qualitativen und quantitativen Daten sollen Aussagen über die Möglichkeiten und Grenzen der Verwendung der drei eingesetzten Programme für spezielle Zielgruppen gemacht werden; außerdem geht es um die Bewertung der Kombination der ausgewählten Programmteile mit den erlebnispädagogischen Übungen. Der Evaluationsprozess ist so angelegt, dass der Blick für nichtintendierte Effekte offen bleibt.

Konflikte als zentraler Ansatzpunkt der Evaluation

Wirkungen der Seminare können sich an ganz unterschiedlichen Orten manifestieren. In unserem Fall geht es darum, sie an einem bestimmten Ereignis der alltäglichen Lebenswelt zu identifizieren, und zwar einem solchen, das in enger Beziehung zu den Bildungszielen der Seminare steht.

Eine spezifische Toleranzkompetenz basiert auf der persönlichen Bereitschaft, individuellen Fähigkeiten und entsprechender Sachkenntnis, sich in Konfliktsituationen angemessen, d.h. insbesondere gewaltlos zu verhalten. Aus diesem Grund konzentrieren wir uns auf die von den Schülern erlebten Konflikte und deren Umgang damit. Die Konflikte werden auf drei Ebenen beleuchtet:

- Auf der emotionalen Ebene geht es um die während des Konflikts wahrgenommenen Gefühle.
- Auf der kognitiven Ebene steht die Frage im Mittelpunkt, welche rationalen Schlüsse aus dem Konflikt gezogen und welche Einstellungen gewonnen werden.
- Auf der handlungsorientierten Ebene sollen die praktischen Verhaltensweisen in Konflikten analysiert werden.

Einerseits sollen die genannten drei Ebenen in ihrem wechselseitigen Bezug untersucht und andererseits Entwicklungen während des Projektverlaufes aufgezeigt werden.

Das Forschungsinteresse richtet sich gleichermaßen auf Konflikte in der Schule, im familiären Bereich sowie in der »peer group« als den zentralen Bezugspunkten der Lebenswelt von Auszubildenden und Schülern.

Um der Tatsache Rechnung zu tragen, dass die Sicht auf Konflikte bei den davon Betroffenen recht unterschiedlich sein kann, werden in einigen Fallstudien neben den Schülern auch Eltern, Lehrer und Freunde zu den von den Schülern geschilderten Konflikten befragt. So entsteht ein multiperspektivisches Bild auf das Geschehen.

Sowohl der Fragebogen als auch der Interview-Leitfaden schenken den aufgezeigten Dimensionen breiten Raum.

Die Instrumente im Evaluationsprozess

Im Verlauf des Projektes werden unterschiedliche quantitative und qualitative Erhebungen durchgeführt. Der Schwerpunkt liegt dabei aus forschungsökonomischen Gründen im Bereich der standardisierten Befragungen. Vertiefende qualitative Fallstudien in Form von studentischen Diplomarbeiten vertiefen einzelne Untersuchungsfragen.

Alle Teilnehmer werden mit einem standardisierten Fragebogen zu vier Zeitpunkten (Eingangsbefragung vor Einstieg in das Projekt; vor Beginn des zweiten und dritten Seminars; Ausgangsbefragung ca. drei Monate nach dem letzten Seminar) befragt. Der Schwerpunkt dieser Befragung liegt auf der Konfliktbeurteilung und Konfliktregelungskompetenz der Teilnehmer. Der insgesamt 94 Fragen umfassende Fragebogen nimmt u. a. folgende Untersuchungsdimensionen in den Blick:

- Konfliktverständnis und Beurteilung von Konflikten
- Konfliktursachen, -anlässe und -konstellationen
- Gefühle und Empfindungen bei Konflikten
- Verhalten im Konfliktfall
- Potenziale der Konfliktlösung
- eigene Konfliktregelungskompetenzen
- Klassen- und Schulklima
- Sozialdaten

Zum Einsatz kommt ein Mix unterschiedlicher Fragetypen (geschlossene, halbstandardisierte, offene). Verwendet werden Skalen, Statements, Polaritätsprofile, Einfach- und Mehrfachantwortvorgaben. Der verwendete Mix soll mit dazu beitragen, auf der Ebene der Auswertung die notwendige Differenzierung zu erreichen.

Die Auswertung der Daten soll es möglich machen zu untersuchen, ob sich im Verlauf des Projektes die Wahrnehmungen, Einschätzungen, Kompetenzen, Kenntnisse sowie die Handlungsweisen der Schüler und Auszubildenden verändern. Unser Augenmerk ist dabei – wie bereits erwähnt – auf drei Ebenen gerichtet: die kognitive, die emotionale und die handlungsbezogene.

Die ersten drei Befragungen werden mit dem gleichen Fragebogen durchgeführt, die letzte Erhebung wird auch rückblickende Fragen zum Projektverlauf, der Einschätzung der Seminare (im Blick auf Inhalte, Übungen, Methoden) und Fragen zu potenziellen und faktischen Wirkungen der Seminare enthalten.

Parallel zur Gruppe der Teilnehmer wird eine Kontrollgruppe zu zwei Zeitpunkten (Frühjahr 2003 und zeitgleich zur Ausgangsbefragung der Projektteilnehmer) mit dem gleichen Fragebogen befragt. Diese Kontrollgruppe ist von der Zusammensetzung (Anzahl, Alters-, Schul-/Ausbildungs-, Herkunftsstruktur) der Gruppe der Projektteilnehmenden sehr ähnlich. Mit der Kontrollgruppenuntersuchung sollen die Ergebnisse der Teilnehmerbefragung mit Blick auf die Interpretation des Zusammenhangs zwischen der Teilnahme an den Seminaren und deren möglichen Wirkungen abgesichert werden.

Die gemessene Veränderung eines Indikators in der Gruppe der Projektteilnehmer vor und nach den Seminaren ist dann besonders aussagekräftig, wenn sie als Effekt der Intervention identifiziert werden kann. Hierfür ist es notwendig, das Messergebnis der Kontrollgruppe heranzuziehen. Nur so kann annähernd sichergestellt werden, dass außer der Seminarteilnahme keine anderen Faktoren als Erklärung in Frage kommen. Vorausgesetzt wird dabei, dass beide Gruppen (Projektteilnehmer und Kontrollgruppe) ansonsten identischen gesellschaftlichen Einflussfaktoren unterliegen.

Als methodisches Problem bleibt allerdings der Umstand, dass wiederholte Messungen die Messwerte selbst verändern können. Für die Kontrollgruppe ist diese Tatsache aufgrund des zeitlichen Abstandes der beiden Befragungen (ca. 18 Monate) zu vernachlässigen.

Für die Teilnehmer sind die intensiven Befragungen jeweils am Anfang und Ende der drei Seminare, die Abschlussbefragung und die gegebenenfalls noch zusätzlichen qualitativen Erhebungen sicherlich nicht zu unterschätzen, veranlassen sie doch die Teilnehmer, sich abermals mit den Gegenständen der Seminare gedanklich intensiv auseinander zu setzen. Wir sehen hierin aber keine Verfälschung der Ergebnisse, sondern betrachten die Befragungen selbst als Teil der Bildungsarbeit, von der aktivierende Potenziale der Motivation auszugehen vermögen.

Nach jedem der drei Trainingsseminare werden die Teilnehmer mit einem Reflexionsbogen (unmittelbar am Ende des Seminars) nach ihrer Einschätzung des jeweiligen Seminars befragt. Die Schwerpunkte sind dabei:

- das Verständnis und seine möglichen Veränderungen zu den vier im Rahmen des Projektes zentralen Begriffen: Demokratie, Toleranz, Gewalt, Konflikt
- Fragen zur Einschätzung einzelner Themen, Elemente und Übungen aus dem Seminarverlauf unter den Gesichtspunkten Nutzen der Übung und Spaß am Lernen
- Fragen zur Gesamteinschätzung des Seminars bezogen auf das Individuum, die Gruppe, die Inhalte, die Rahmenbedingungen und das Leitungsteam

Die Daten dieser Erhebung werden bei der weiteren Planung der konkreten Projektinhalte, vor allem der konkreten Übungen berücksichtigt.

In Ergänzung zu den beiden Befragungen mit Frage- und Reflexionsbögen bei der Teilnehmergruppe werden ab Mitte des zweiten Projektjahres (Sommer 2003, nachdem die Teilnehmer das zweite Trainingsseminar absolviert haben) qualitative Methoden der Evaluation eingesetzt. Ein erster Schritt ist dabei die Durchführung von leitfadengestützten Interviews mit etwa je zwei Teilnehmern aus allen Klassen, die am Projekt teilnehmen. In diesen Interviews wird auf drei Ebenen gefragt nach:

- Wirkungsarten (Sinnzusammenhang der Themen der Seminare, aktive Weitergabe des Erlebten und Gelernten, Fähigkeit zum Perspektivenwechsel, aktive Umsetzung des Erlebten und Gelernten)
- Wirkungsräumen (Klasse; Schule, persönliches Umfeld, gesellschaftliches und politisches Umfeld)
- Wirkungsweisen (Konflikte als Chance oder Bedrohung, Empathie oder Abwehr im Konflikt)

Der Gegenstand, auf den sich die Fragen zu den Wirkungen beziehen, sind jene Konflikte, an denen die Befragten selbst beteiligt sind bzw. waren. Dabei wird auch auf die Seminare selbst geschaut: Welche Konflikte gab es hier? Wie wurde damit umgegangen? Welche Übungen waren besonders konfliktreich?

Mit demselben Leitfaden wie die Teilnehmer sollen anschließend auch Eltern und Lehrer interviewt werden. Neben den schon angesprochenen Untersuchungsdimensionen geht es dabei um die Frage, ob Lehrer und Eltern Veränderungen bei den Jugendlichen wahrnehmen, die sie auf die Teilnahme an den Projektseminaren zurückführen.

Die Anzahl dieser Interviews ist allerdings aus Ressourcengründen sehr begrenzt. Nach Abschluss des dritten Durchgangs der Trainings ist eine Gruppendiskussion mit den Trainern geplant, die sich schwerpunktmäßig auf deren Einschätzungen zu den beiden Ebenen Wirkungsarten und Wirkungsweisen im Blick auf die Inhalte der Seminare konzentriert.

Neben der Evaluation wird das Projekt dokumentarisch begleitet. Dazu werden aus jeder Gruppe Teilnehmende im Verlauf des Projekts kontinuierlich vor der Kamera zu ihrer Einschätzung der Seminare sowie zum Alltag in ihren Klassen während der Trainings und zwischen den Seminaren befragt. Ob und inwieweit diese Aufnahmen und die Aufzeichnungen, die die Teilnehmer in einem Projekttagebuch v. a. während der Seminare in Weimar für sich machen, noch mit in die Evaluation einbezogen werden, ist derzeit noch nicht geklärt.

Die »Produkte« der Evaluation

Eingangs haben wir bereits beschrieben, dass es uns wichtig ist, ein dem Bildungsprozess angemessenes Evaluationskonzept umzusetzen. Dieses Postulat der Angemessenheit gilt ebenso in Bezug auf die zentralen Inhalte der Seminare, nämlich Demokratie und Toleranz. Bei dem zu evaluierenden Kernbestand des Toleranzkonzeptes der Bertelsmann Forschungsgruppe Politik geht es im Wesentlichen um die Erfassung von Motivationen (vgl. im Folgenden Bertelsmann Forschungsgruppe Politik 2000): Toleranz fußt demgemäß auf einer »prinzipiellen Anerkennung des gleichen Rechts auf Freiheit«. Die so genannte »Scheinbare Toleranz« ist die Folge von »Nutzen- oder Risikoabwägung«, während Intoleranz durch die »vorrangige Durchsetzung eigener Interessen« motiviert wird.

Der Fragebogen bezieht sich als überwiegend summatives Instrument auf die Kernelemente des Toleranzkonzeptes. Der Reflexionsbogen hingegen dient in erster Linie als Instrument der Selbstevaluierung und der formativen Evaluation und damit zum Monitoring des Projektablaufes und der Qualitätsentwicklung des Projektes (Gestaltung, Steuerung). Beide Instrumente, »Fragebogen« und »Reflexionsbogen«, sind voneinander weitgehend unabhängig und können als jeweils eigenständige Informationsquelle genutzt werden.

Die Anlage der standardisierten Fragebogenuntersuchung macht es möglich, echte Panel-Daten (Längsschnittstudie) über den gesamten Verlauf des fast dreijährigen Bildungszyklus zu gewinnen. Dies erbringt eine viel höhere Datenqualität als reine »Zeitreihen«-Analysen, die auf der jeweils gleichen

Befragung, aber je unterschiedlichen Befragten beruhen. Da sich Motivationen durch standardisierte Befragungen nur unzureichend erfassen lassen, werden die quantitativen Erhebungen durch qualitative ergänzt.

Zusammen mit den Ergebnissen der Leitfadeninterviews und der Gruppendiskussion sollte ein Informationsraum entstehen, der es ermöglicht, gehaltvolle Thesen über die Wirkung des Projektes zu formulieren. Durch das Design und die zum Einsatz kommenden Methoden und Instrumentarien wird es möglich,

- Aussagen zu kurzfristigen, mittelfristigen und – unter Angabe der Unsicherheits-/Unschärfebereiche – längerfristigen Wirkungen der Seminare bei den Teilnehmern zu gewinnen;
- Erkenntnisse über die Wirkung der eher impliziten und der eher expliziten Intentionen und Zielsetzungen der Seminare zu gewinnen, auch über deren Umsetzung im Umgang der Teilnehmer miteinander (Selbstbezug, Entsprechung der Mittel und Ziele); sowie
- empirisch gehaltvolle Hinweise zur Bedeutung (Gewichtung, Relevanz, Selbstbezug, Kontexte) der Themen »Konflikt, Gewalt, Toleranz und Demokratie« für die Schüler in der Vergangenheit, für die Gegenwart und die Zukunft zu geben.

Die Ergebnisse der Evaluation werden auf verschiedene Weise der fachlichen und fachpolitischen Diskussion zugänglich gemacht. Es wird ein Evaluationsbericht erarbeitet, der sich weitgehend auf die Daten der quantitativen und qualitativen Erhebungen stützt und zum Ende des Projektes im Dezember 2004 fertig gestellt sein wird.

Neben den hier dargestellten Elementen der Evaluation des Projektes »Schule und Ausbildung für Toleranz und Demokratie«, die alle zum »offiziellen« Evaluationsauftrag an die Fachhochschule Jena gehören, werden weitere Teiluntersuchungen, z.B. im Rahmen von Diplomarbeiten an der FH Jena durchgeführt, die in zeitlicher Nähe zum Ende des Projektes vorgelegt werden. Dafür werden z.T. noch einmal eigene und spezifische Erhebungen gemacht.

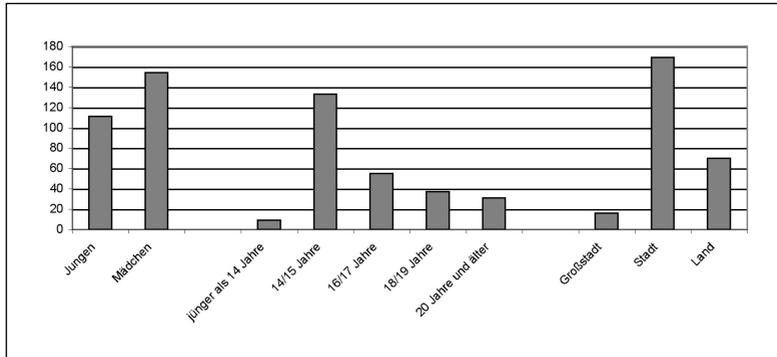
Erste Ergebnisse der wissenschaftlichen Begleitung

Die demographische Struktur der Teilnehmergruppe

An den beiden Trainingskursen in den Jahren 2002 und 2003 nahmen insgesamt 265 Schüler sowie Auszubildende im Alter von 13 bis 35 Jahren teil.

Es sind bisher im Projekt deutlich mehr Mädchen (N=154, 2002: 130) als Jungen (N=111, 2002: 91) beteiligt. Die Altersstruktur stellt sich folgendermaßen dar: jünger als 14 Jahre alt: 9 (2002: 9); 14/15 Jahre alt: 133 (2002: 111); 16/17 Jahre alt: 55 (2002: 45); 18/19 Jahre alt: 37 (2002: 31); 20 Jahre und älter: (2002: 25). Die meisten der Teilnehmer kommen aus mittleren und kleinen Städten: 179 (2002: 154); nur wenige aus einer Großstadt: 16 (2002: 13); etwa ein Drittel leben auf dem Land: 70 (2002: 54).¹

Abbildung 1: Geschlecht, Alter und Wohnort der Projektteilnehmer



Quelle: Befragungen 2002 und 2003 kumuliert

Die Teilnehmer kommen aus folgenden am Projekt beteiligten Schulen:

- zwei Regelschulen (ein Thüringer Schulmodell, in dem Haupt- und Realschulklassen in einer Schulform zusammengefasst sind) mit insgesamt fünf Realschulklassen (ab Klasse 8) aus Zella-Mehlis und Saalfeld
- einem Gymnasium mit zwei Klassen (ab Klasse 9) aus Eisenach
- zwei berufsbildende Schulen mit insgesamt drei Ausbildungsklassen (ab 1. Ausbildungsjahr) aus Jena und Weimar sowie eine Klasse aus dem Wirtschaftsgymnasium (ab Klasse 11) aus Jena
- einem überbetrieblichen Ausbildungsträger aus Merkers mit einer gemischten Gruppe Auszubildender ab dem ersten Ausbildungsjahr

1 Die statistischen Angaben zur Teilnehmerstruktur beziehen sich auf die Daten der beiden Befragungen, die jeweils vor Beginn der Seminare, an denen die Jugendlichen teilnahmen, stattgefunden haben. 2002 haben 221 Jugendliche mit dem Projekt begonnen, 2003 kamen 44 weitere hinzu. Bei den Zahlenangaben sind zunächst immer die Gesamtanzahl beider Jahre und in Klammern die Zahlen der Teilnehmer im Jahr 2002 angegeben. Teilnehmer, die älter sind als 20 Jahre, stammen aus den berufsbildenden Schulen.

Im Blick auf die Wohnsituation ergibt sich folgendes Ergebnis: Bei den Eltern leben 180, nur mit einem Elternteil 36; in anderen Wohnformen 43 Teilnehmer. Die meisten der Eltern sind berufstätig: Für den Vater geben 215 Jugendliche an, dass er berufstätig ist, für die Mutter geben das 208 Teilnehmer an.

Ergebnisse aus der Fragebogenuntersuchung

Die Bereitschaft der Auszubildenden und Schüler, sich auf den Fragebogen einzulassen, der relativ hohen Zeitaufwand und großes Maß an Konzentration erfordert, zeigt die Ernsthaftigkeit, mit der sowohl die Projektteilnehmer wie die Jugendlichen der Kontrollgruppe an die Sache herangegangen sind.

Wie bereits erläutert, haben wir Konflikte als zentralen Ansatzpunkt der Evaluation gewählt. In zahlreichen Fragen wird untersucht, wie die Schüler Konflikte beurteilen, einschätzen und bewerten, wie sie diese emotional und rational verarbeiten und welche Handlungsoptionen sie in Konfliktsituationen nutzen.

Welche Konflikte erleben die Schüler in der Schule, und woran entzünden sich diese Konflikte? Bei dieser Frage stellt sich heraus, dass Gewalt für die am Projekt beteiligten Schüler nicht das vordringliche Problem an ihrer Schule ist (7 Prozent). Weit häufiger sind die Konflikthanlässe Ungerechtigkeit (16 Prozent), Rivalitäten unter Schülern (11 Prozent), Vorurteile (13 Prozent), Demütigungen (15 Prozent) und Leistungsdruck/Stress (11 Prozent).

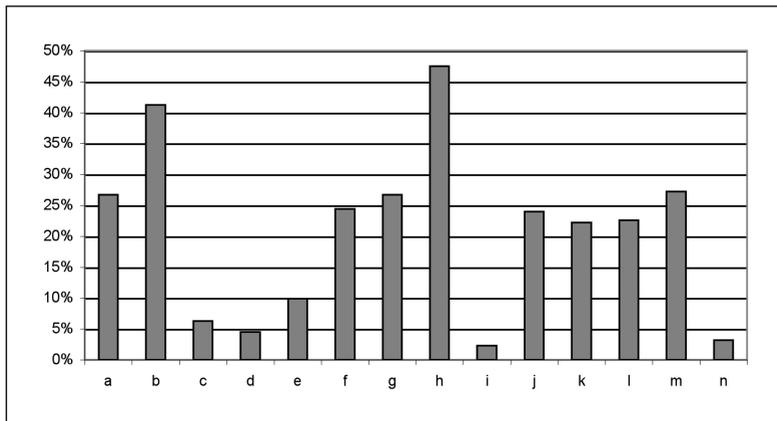
Interessant ist auch, dass die Zugehörigkeit zu Gruppen im Falle von Cliquen, Freundeskreisen oder Jugendkulturen (9 Prozent) eher als Konflikthanlass empfunden wird, als wenn es sich um religiöse, kulturelle oder nationale Gruppenzugehörigkeiten handelt (3 Prozent).

Die Schule ist eine Gesellschaft im Kleinen, eine »embryonic society«, wie es John Dewey (1986) formuliert hat. Hier manifestieren sich Normen und Werte unmittelbar erlebbar in den Handlungen der Akteure. Das kleine, überschaubare Gemeinwesen Schule ist der Testfall für die Einübung von Demokratie und Toleranz. Wie wird das Klima in der »Schulpolis«, um mit Hartmut von Hentig (1993) zu sprechen, eingeschätzt? In den Antworten zu den Fragen zum Schulklima steht das Thema Zusammenhalt im Mittelpunkt. Die Schüler haben den Wunsch, sich möglichst mit allen zu verstehen, ein positives Gruppenklima zu erreichen.

Konflikte scheinen für die Schüler nicht in erster Linie mit Gewalt verbunden zu sein, sie fühlen sich von ihnen mehr auf der emotionalen Ebene belastet. Das Bedrohliche an Konflikten sehen die Schüler in der unter-

schwelligen Wirkung von Konflikten, etwa zerstörten Freundschaften und einem schlechten Klassenklima.

Abbildung 2: Beurteilung von Konflikten



Quelle: Befragung 2002

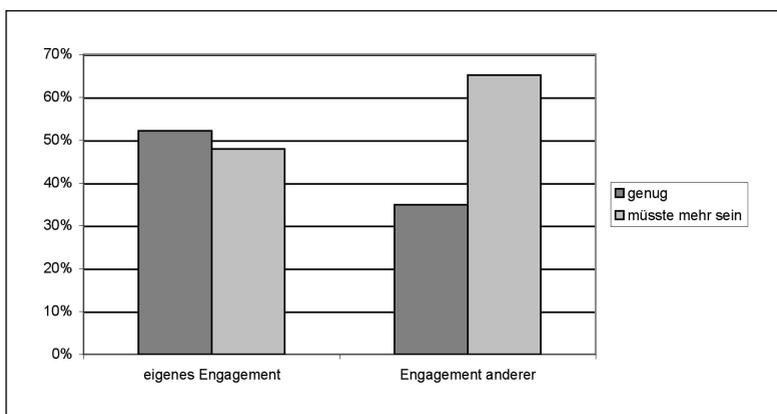
- a) Ich finde sie fürchterlich und versuche sie nach Möglichkeit zu vermeiden
- b) Ich sehe darin eine Chance (z. B. um zu einer echten Entscheidung zu kommen)
- c) Ich brauche sie (z.B. um meine geistige Überlegenheit zu demonstrieren)
- d) Ich suche sie (z.B. weil dann Leben in die Bude kommt)
- e) Sie sind für mich eine Möglichkeit, mit anderen in Kontakt zu kommen
- f) Ich sehe sie als Gewinn (z.B. weil man dabei viel über die Menschen lernen kann)
- g) Ich sehe sie als Herausforderung, der ich mich stellen möchte
- h) Ich rechne mit ihnen und möchte lernen, mit ihnen umzugehen
- i) Ich wünsche mir ein konfliktreiches Leben
- j) Ich hasse Konflikte, weil sie für mich gleichbedeutend sind mit Streit und ungunen Gefühlen
- k) Ich finde sie anstrengend und belastend
- l) Ich nehme sie als Ansporn, mich durchzusetzen
- m) Ich verstehe sie als Aufforderung, zu einer für alle befriedigenden Lösung beizutragen
- n) Ich weiß nicht, wie ich sie beurteilen soll

Sind Schüler in einen Konflikt verwickelt, sind ihre Verhaltensweisen primär dadurch gekennzeichnet, »mit allen gut auskommen« (22 Prozent) und »keine Freunde verlieren« (13 Prozent) zu wollen. Motive wie »Ich muss Stellung beziehen« (12 Prozent), »Ich möchte vor mir selber und vor meinem Gewissen bestehen können« (9 Prozent) und »Ich finde es feige, wenn ich ruhig

bleibe« (8 Prozent), die eher auf eine generelle Einstellung oder eine verinnerlichte Norm hindeuten, sind nicht ganz so hoch besetzt, aber dennoch deutlich erkennbar.

In der Beurteilung von Konflikten ergibt sich ein ambivalentes Bild. Die Schüler stimmen Aussagen, die Konflikte positiv bewerten und in ihnen eine Chance erkennen, zwar oft zu, in ähnlich hohem Maße bejahen sie jedoch auch Aussagen, die Konflikte negativ betrachten und ein Vermeidungsverhalten zum Ausdruck bringen. Viele Schüler schwanken in ihrer Konfliktbeurteilung zwischen diesen beiden Positionen: die Perspektive, einen Konflikt als etwas Positives wahrzunehmen, ist den Teilnehmern generell zwar nicht fremd. Es zeigt sich jedoch auch, dass Konflikte oft mit Ängsten besetzt sind und versucht wird, sie zu vermeiden.

Abbildung 3: Einschätzung des Engagements für Konfliktlösung



Quelle: Befragung 2002

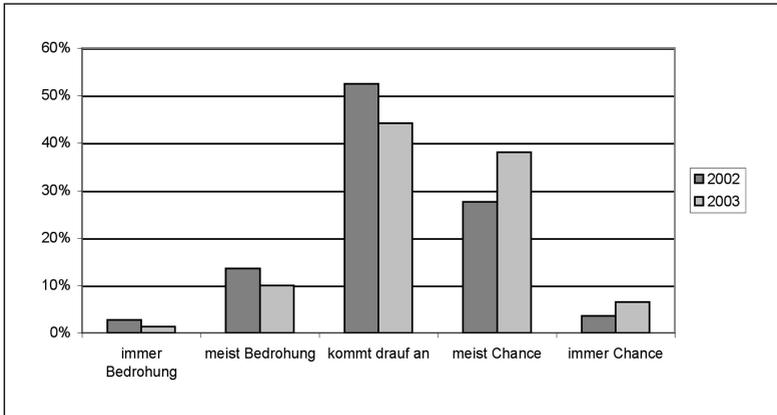
Es kann davon ausgegangen werden, dass die Schüler den Wunsch haben, mehr Kompetenzen im Umgang mit Konflikten zu erlernen. Knapp die Hälfte der Schüler denkt, dass sie selbst mehr tun müssten, um Konflikte zu lösen, von denen sie persönlich betroffen sind. Die Hauptverantwortung für die Lösung von Konflikten wird allerdings bei den anderen gesucht, die nach Meinung der Schüler hier mehr Engagement zeigen müssten.

Es besteht ein wechselseitig positiver Zusammenhang zwischen dem Blick auf Konflikte und der Streitkunst: Je positiver Konflikte beurteilt werden, je mehr sie als Chance denn als Bedrohung gesehen werden, desto mehr Kompetenzen sind subjektiv vorhanden, mit ihnen umzugehen. Und umgekehrt:

Jene Teilnehmer, die ihre Fähigkeiten zu streiten, positiv einschätzen, neigen eher dazu, Konflikte als Chance zu betrachten.

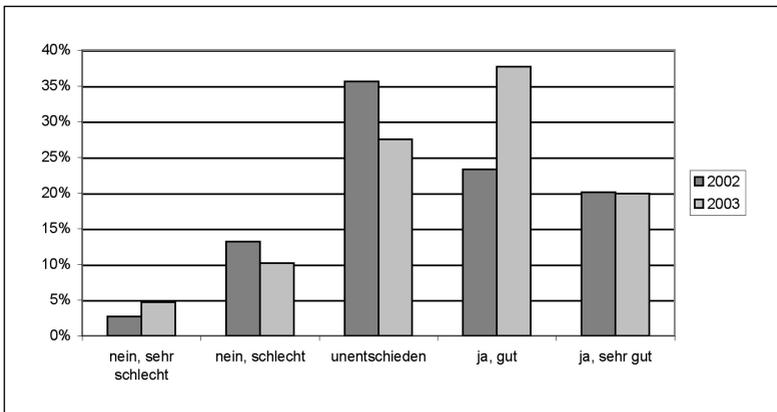
Im Vergleich der Erhebungen im Jahre 2002 und 2003 zeigt sich, dass sowohl die Zahl der Projektteilnehmenden, die Konflikte als Chance sehen, als auch derer, die angeben, sie könnten gut streiten, ansteigt.

Abbildung 4: Konflikt – Chance oder Bedrohung?



Quelle: Befragungen 2002 und 2003

Abbildung 5: »Kannst du gut streiten?«



Quelle: Befragungen 2002 und 2003

Die Kompetenzen im Umgang mit Konflikten sind aus der Sicht der Projektteilnehmenden eindeutig gestiegen. In den weiteren Untersuchungen wird zu überprüfen sein, ob sich diese Entwicklung fortsetzt und inwieweit sie auf das in den Seminaren Gelernte und Erlebte zurückzuführen ist.

Ergebnisse aus der Untersuchung der Reflexionsbögen

Wie bereits oben erwähnt, ist ein wichtiger Aspekt bei der Befragung der Teilnehmer die Veränderung des Verständnisses zu den vier Begriffen Konflikt, Gewalt, Toleranz und Demokratie, zu denen die Schüler jeweils nach jedem Seminar befragt werden. Die Fragen dazu sind eine Mischung aus offenen und standardisierten Fragen.

- In einer ersten offenen Frage werden die Teilnehmer nach Assoziationen zu den Begriffen gefragt: »Unter ... verstehe ich«.
- Eine zweite Frage geht den Veränderungen des Verständnisses nach: »Mein Verständnis zum Begriff ... hat sich deutlich vertieft/ist unverändert geblieben.«
- In einer dritten offenen Frage wird konkret nach den im Rahmen des Seminars erlebten deutlichsten Veränderungen des Verständnisses gefragt.

Diese drei Fragen sind im Zusammenhang zu betrachten. Eine umfassende Auswertung dieser Daten, besonders der offenen Fragen liegt derzeit aber noch nicht vor, so dass ein erster Blick zunächst auf die zweite Frage gerichtet werden soll.

Im Blick auf das Verständnis der vier zentralen Begriffe des Projektes lässt sich im Vergleich der Befragungen aus den ersten beiden Jahren eine deutliche Veränderung feststellen. Im Rahmen des Einführungsseminars hatte keiner der vier Begriffe eine besonders herausragende Bedeutung, alle Themenfelder wurden in verschiedenen Übungen und thematischen Zugängen in den Blick genommen. Die deutlichsten Veränderungen sind bei den Begriffen Toleranz (62 Prozent) und Konflikt (55 Prozent) zu erkennen, beim Begriff Gewalt (36 Prozent) gibt es die geringsten Veränderungen, etwa die Hälfte der Teilnehmer geben eine Veränderung beim Begriff Demokratie an.

Diese »Rangfolge« spiegelt auch den Seminarverlauf wider: Zu Toleranz wurde eine erste Annäherung in Form einer Übung versucht; Konflikt war ein immer präsent Thema in den Gruppen, da es viele gruppeninterne Konflikte im Verlauf des Seminars gab, die bearbeitet und von den Leitungsteams immer wieder auch in den Gesamtzusammenhang des Projektes gestellt wurden. Auch zum Stichwort Demokratie gab es erste inhaltliche Zu-

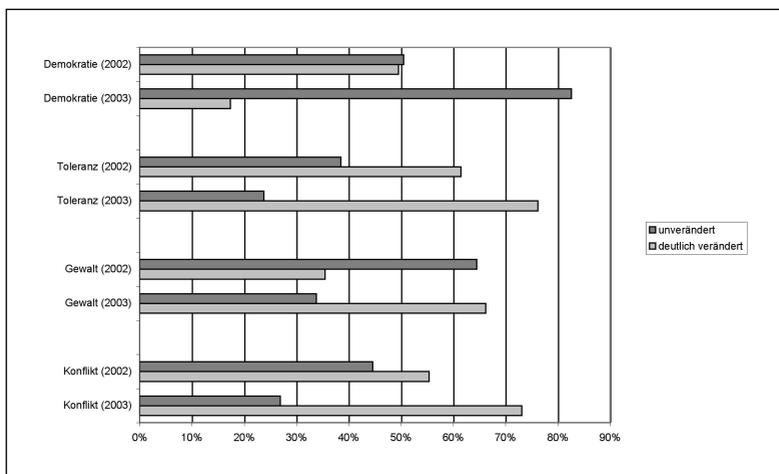
gänge, was sich in den Aussagen der Teilnehmer auch wiederfindet. Lediglich das Thema Gewalt wurde nicht explizit bearbeitet.

Während des zweiten Seminars fand dagegen eine ausführliche Beschäftigung mit drei der vier Begriffe statt: Konflikt, Gewalt und Toleranz wurden mit jeweils eigenen Definitionen theoretisch eingeführt und anhand von praktischen Übungen vertieft. In der zweiten Befragung wird deutlich, dass sich das Verständnis zu diesen drei Begriffen weiter vertieft hat. Am deutlichsten sind die Veränderungen bei den Begriffen Toleranz (76 Prozent) und Konflikt (73 Prozent), bei Gewalt (66 Prozent) sind es etwas weniger Teilnehmende, die eine Veränderung angeben.

Interessant ist auch der Blick auf den Begriff Demokratie: Hier geben in der zweiten Befragung nur 17 Prozent der Teilnehmer eine Verständnisveränderung durch das Seminar an. Zu diesem Thema hat es im zweiten Seminar auch keine ausdrückliche Arbeitseinheit gegeben, das Thema wird der Schwerpunkt des dritten und letzten Seminars sein.

Die geringe Angabe von Veränderungen ist trotzdem von Bedeutung, weil sie Rückschlüsse auf die anderen Antworten auf diese Fragen zulässt: Nach dem ersten Jahr waren wir uns in der Betrachtung der Antworten nicht ganz sicher, ob die Teilnehmer ehrlich angekreuzt hatten oder den Seminarteams und der Projektleitung »etwas Gutes tun wollten«, indem sie von ihnen erwartete oder vermeintlich »richtige« Antworten gegeben haben.

Abbildung 6: Veränderung des Verständnisses der Begriffe Konflikt, Gewalt, Toleranz und Demokratie



Quelle: Befragungen 2002 und 2003

Dieser Trend hätte sich aber in der zweiten Befragung fortsetzen müssen. Insofern stützt die geringe Anzahl der Verständnisveränderungen zu Demokratie alle anderen Antworten in diesem Frageblock. Diese zeigen, dass die Auseinandersetzung mit den zum Teil komplizierten und sehr theoretischen Definitionen im Rahmen des Seminarkonzeptes Sinn gemacht und offenbar auch zumindest kurzfristige Wirkungen hinterlassen hat.

Gestützt werden diese Hinweise auf Wirkungen zudem durch die Fragen nach dem Nutzen von einzelnen Übungen aus dem Seminarprogramm des zweiten Trainings. Dabei zeigt sich deutlich, dass einige theoretische Teile des Seminars in der Rangliste der Übungen, die den meisten Nutzen gebracht haben, erstaunlich weit vorn liegen.

Die Übung mit dem meisten Spielcharakter (»Die Firma«) liegt mit 82 Prozent erwartungsgemäß in der Bewertung ganz vorn, wobei hier zu fragen wäre, inwieweit Spaß und Nutzen sich beeinflussen. Weniger fraglich ist das bei anderen Übungen aus dem Seminarprogramm: Es folgen eine spielerische Übung zum Thema Toleranz (76 Prozent) und eine Übung zum Partnerschaftlichen Dialog (76 Prozent).

Unmittelbar danach folgt die erste rein theoretische Einheit: Konfliktdefinition (75 Prozent) und die Toleranzdefinition (73 Prozent). Nach einigen weiteren aktiveren Übungen wird die Auseinandersetzung mit dem Begriff Gewalt (65 Prozent) immer noch als sehr nutzbringend eingestuft. Auch andere Theorieteile (die aber nicht unmittelbar zu zentralen Begriffen gehören) werden in der Frage nach dem Nutzen sehr positiv bewertet.

Interessant werden auch die Auswertungen im Blick auf die Kombinationen von Demokratie- und Toleranztrainingsprogrammen mit erlebnispädagogischen Interaktionsübungen sein. Ein vorläufiger Rückblick auf das erste Projektjahr lässt die Vermutung zu, dass die Verknüpfung noch nicht vollständig gelungen ist: Die Interaktionsübungen (die weitgehend im Freien durchgeführt wurden) standen in der Bewertung der Teilnehmer eher neben den eigentlichen Seminarinhalten und sind von ihnen noch sehr wenig als integrale und wichtige Bestandteile des Gesamtkonzepts wahrgenommen worden.

In der ersten Befragung wurde lediglich danach gefragt: »Welche Übungen haben dir am meisten gebracht?«. Dabei wurden die Interaktionsübungen insgesamt und auch einzelne dieser Aktivitäten am häufigsten genannt. Werden daneben aber die (nicht sehr ausführlichen) Antworten auf die Frage »Warum gerade diese Übungen?« betrachtet, fällt auf, dass hier sehr oft benannt wird: »Weil sie Spaß gemacht und den Klassenzusammenhalt gefördert haben.« Kritisch bewertet werden die Übungen, bei denen es im Seminarverlauf die meisten Konflikte gegeben hat, die im Rahmen der Demokra-

tie- und Toleranzerziehung aber von Wichtigkeit sind und von den Leitungsteams gemeinsam mit den Gruppen bearbeitet wurden.

Im Rückblick der Teilnehmer entstand an dieser Stelle aber eine ambivalente Wahrnehmung des Seminars: auf der einen Seite die tollen Übungen, bei denen sie so gut zusammenhalten konnten – auf der anderen Seite die Teile des Seminars, in denen sie in Konflikte verstrickt waren und die Einigkeit in Gefahr geriet. Das war vom Konzept her aber nicht beabsichtigt. Die Interaktionsübungen waren nicht nur zur Förderung des Zusammenhalts gedacht, sondern verfolgten die gleichen Lernziele wie andere Übungen während der Woche auch.

Die Ergebnisse dieser Befragung mit dem Reflexionsbogen führten zu zwei wesentlichen Veränderungen: Zum einen wurde der Charakter der Interaktionsübungen für die beiden folgenden Seminarkonzepte deutlich verändert. Sie sind nun noch stärker mit den eigentlichen Seminarthemen und -inhalten verzahnt und sollen auf einer anderen als der sprachlichen Ebene einen Zugang zur Auseinandersetzung mit Toleranz und Demokratie für die Teilnehmer ermöglichen. Die Chance der Erlebnispädagogik liegt hierbei in ihrem vollständig anderen Ansatz – über das persönliche Erleben einer logistischen oder auch körperlichen Herausforderung das eigene Handeln und das Miteinander der Gruppe zu reflektieren.

Bereits im zweiten Trainingskurs wurde deutlich, dass diese Veränderung für die Akzeptanz des Gesamtkonzeptes durch die Teilnehmer sehr wichtig war. Zum anderen haben wir die Fragestellungen in den Reflexionsbögen anders gestaltet: Ab dem zweiten Jahr wird jede einzelne Übung nach ihrem Nutzen abgefragt. Hinzu kommen offene Fragen, bei denen die Teilnehmer ihre persönlichen Vorlieben und Abneigungen im Blick auf einzelne Übungen aufschreiben können: »Welche der Übungen haben dir am meisten/am wenigsten gebracht?« und »Welche der Übungen haben dir am meisten/am wenigsten Spaß gemacht?«.

Die Auswertungen dieser Sets an Fragen zu den Inhalten der Seminare lassen eine weitaus differenziertere Bewertung der Seminare durch die Teilnehmer zu: Bei der Frage, welche Übungen den meisten Spaß gemacht haben, werden die erlebnisorientierten Übungen zuerst genannt; die theoretischen Teile sind die Phasen, die nicht viel oder keinen Spaß gemacht haben.

Das ist nicht überraschend, sichert aber durchaus auch die Bewertung der Übungen, die etwas gebracht haben, deutlich ab, denn dort ist zu erkennen, dass sich die Bewertung des Nutzens nicht unbedingt am »Spaßfaktor« der Übungen orientiert.

Schlussbetrachtung

Diese ersten Einblicke in die Auswertung des umfangreichen Datenmaterials können nur andeuten, welch interessantes und spannendes Feld der Evaluation politischer Bildung sich vor uns auftut. Mit besonderer Spannung erwarten wir nun die qualitativen Interviews, die sowohl einen tieferen Einblick in die Akzeptanz der Trainings durch die Teilnehmer gewähren als auch die Konfliktwahrnehmung und den Konfliktumgang der Jugendlichen weiter untersuchen und damit Hinweise auf die Wirkungen der Trainings geben können.

Demokratie- und Toleranzerziehung mit den bereits mehrfach erwähnten Trainingsprogrammen, so viel ist bereits heute deutlich erkennbar, erweist sich als ein spannendes (durchaus auch spannungsreiches) Arbeitsfeld für die politische Jugendbildung. Und die Evaluation dieses Feldes steht dem in keiner Weise nach.

Literatur

- Bertelsmann Forschungsgruppe Politik (Hrsg.) (2001): *Toleranz. Grundlage für ein demokratisches Miteinander*. Gütersloh.
- Boock, Martin (2002): *Erlebnispädagogik und Demokratieerziehung. Reflexionen über die Möglichkeiten erfahrungsorientierter Lernmethoden in der außerschulischen politischen Bildung*. Diplomarbeit. Fachhochschule Jena.
- Bertelsmann Stiftung und Bertelsmann Forschungsgruppe Politik (Hrsg.) (2001): *Eine Welt der Vielfalt. Ein Trainingsprogramm des A WORLD OF DIFFERENCE®-Institute der Anti-Defamation-League, New York, in der Adaption für den Schulunterricht. Praxishandbuch für Lehrerinnen und Lehrer*. 2., überarbeitete Auflage. Gütersloh.
- Dewey, John (1986): *Erziehung durch und für Erfahrung*. Stuttgart.
- Hentig, Hartmut von (1993): *Die Schule neu denken*. München und Wien.
- Khanide, Marina, und Karl Giebeler (2003): *Ohne Angst verschieden sein. Ein Praxishandbuch für die interkulturelle Arbeit*. Gütersloh.
- Rudolf, Karsten, und Melanie Zeller (2000): »Der Bürger als Bildungspartner. Chancen und Möglichkeiten der Evaluation politischer Bildungsarbeit«. In: *kursiv – Journal für politische Bildung* 1 2000. 39–42.
- Ulrich, Susanne, und Florian M. Wenzel (2003): *Partizipative Evaluation. Ein Konzept für die politische Bildung*. Gütersloh.
- Ulrich, Susanne, Thomas R. Henschel und Eva Oswald (Hrsg.): *Miteinander – Erfahrungen mit Betzavta. Praxishandbuch für die politische Bildung. Ein*

*Praxishandbuch auf der Grundlage des Werks »Miteinander« von Uki Maro-
shek-Klarman, Adam Institute, Jerusalem. 3. Auflage. Gütersloh 2001.*

Ulrich, Susanne (2001): *Achtung (+) Toleranz. Wege demokratischer Konfliktre-
gelung. Praxishandbuch für die politische Bildung. Unter Mitarbeit von Jürgen
Heckel, Eva Oswald, Stefan Rappenglück und Florian Wenzel. 2. Auflage.
Gütersloh.*

Teil C

(Wirkungs-)Evaluation politischer Bildung: Methodenreflexion und Empfehlungen

Die Evaluation von »Achtung (+) Toleranz« – eine Methodenreflexion

Andreas Hirsland, Oliver Dimbath, Julia von Hayek, Werner Schneider

In diesem Beitrag möchten wir näher auf die im Zusammenhang mit dem Praxisbericht der Evaluation von »Achtung (+) Toleranz« vorgestellte Evaluationsmethode eingehen. Neben einer Beschreibung des Evaluationsprozesses wird es dabei auch um die methodologische Begründung des von uns gewählten Vorgehens, die Verfahren zur Datenerhebung, -auswertung und Ergebnisinterpretation gehen. Abschließend erfolgt eine kurze praxisbezogene Einschätzung der Eignung derartiger Verfahren für die Evaluation von Programmen der Demokratie- und Toleranzerziehung sowie eine Reflexion der erhaltenen Untersuchungsergebnisse als immer auch »methodische Artefakte«.

Im Zentrum des Evaluationsauftrages stand, neben dem Interesse an der Akzeptanz – »wie das Seminarangebot bei den Teilnehmern ankommt« –, die Frage nach den »Wirkungen« von »Achtung (+) Toleranz« und deren Nachhaltigkeit. Der Versuch, eine Antwort auf diese zunächst so klar und eindeutig erscheinende Fragestellung zu finden, erweist sich bei näherer Betrachtung als problematisches Unterfangen:

Eine einfache Kausalität zwischen Maßnahme und »output« oder »outcome« außerhalb experimenteller Untersuchungsdesigns ist kaum nachweisbar (vgl. hierzu Lynen von Berg und Hirsland in diesem Band und die dort angegebene Literatur). Gerade bei Programmen und Maßnahmen der politischen Bildung mit ihren komplexen, eher vage definierten und daher schwer operationalisierbaren Zielsetzungen – wie etwa »Toleranz« – ist nicht per se klar, worin genau die »Wirkung« bestehen solle und welche Wirkungen überhaupt auftreten. Es ist nicht per se klar, wie welche Wirkungszusammenhänge überhaupt feststellbar oder gar »messbar« und da es um Evaluation geht – schließlich bewertbar wären.¹

¹ Zur Unterscheidung von Feststellung und Messung mag folgendes Beispiel dienen: »Wetter« ist im meteorologischen Sinne durchaus anhand der Größen Luftdruck, Luftfeuchtigkeit, Temperatur und Sonnenscheindauer messbar und auf ei-

Derartige Fragestellungen lassen sich weder von den mit Evaluation beauftragten Wissenschaftlern noch von den Auftraggebern allein lösen, sondern bedürfen einer konsensuellen Klärung, d.h. eines Vorgehens, das als »partizipativ« begriffen werden kann (vgl. auch den Beitrag von Bommes et. al. in diesem Band).

Unser Verständnis von Partizipation ist ein »dialogisch-diskursives«, das darauf gerichtet war, in einem beständigen Austausch mit Auftraggebern und Programmentwicklern ein gemeinsames Verständnis der expliziten theoretischen Grundlagen des Programms ebenso wie der diesem unterlegten impliziten Vorannahmen etwa über seine Wirkungsweise zu gewinnen.

Auf diesen Grundlagen erst ließen sich sinnvolle und somit gegenstandsangemessene Kriterien und Forschungsfragestellungen definieren und Transparenz nicht nur des wissenschaftlichen Vorgehens, sondern auch der jeder Methode zugrunde liegenden erkenntnistheoretisch-methodologischen Voraussetzungen schaffen. Dieses Wissen ist Bedingung für die sachgerechte Einordnung und Beurteilung der erhobenen Befunde und ihrer konsensuellen Validierung durch Auftraggeber, Programmentwickler und Forschergruppe (für einen kurzen Abriss der in diesem Prozess zu bearbeitenden Probleme vgl. Wenzel 2003: 18–20).

Beginn der Evaluation: Exploration

Aus den genannten Gründen bestand die erste Stufe der Evaluation darin, zunächst (a) im Rahmen einer qualitativen Vorerhebung Programmmaterialien und andere Dokumente auszuwerten sowie (b) in explorativen Expertengesprächen mit Programmautoren und Seminarleitern eine Rekonstruktion des Programms vorzunehmen.

Bei dieser ersten Stufe ging es darum, das Programm und seine Übungsteile kennen zu lernen, d.h. allgemeines Wissen darüber anzusammeln, was das Programm »Achtung (+) Toleranz« überhaupt »ist«. Der zweite Schritt diente dazu, dieses Wissen zu vertiefen und am Selbstverständnis der Anbieter zu überprüfen, indem die dem Programm zugrunde liegenden Intentionen und Motive², Wirkungshypothesen und die daraus resultierenden Er-

ner »Mehr-Weniger-Skala« darstellbar. Subjektiv erfahrene Wirkungen des Wetters hingegen drücken sich in feststellbaren Qualitäten aus (schönes Wetter), die zu messen problematisch ist: War der eine Frühlingstag ein schönerer Frühlingstag als der andere, ist bei ansonsten gleichen klimatischen Bedingungen eines Sommers das Wetter in Italien nicht doch immer irgendwie ein wenig »besser« als am Starnberger See? (vgl. hierzu auch Hirsland in Wenzel 2003: 12–17)

wartungen und Ziele in Bezug auf Effekte einzeln durchgeführter Maßnahmen und Programmbausteine in Erfahrung gebracht, expliziert und zusammen mit Auftraggebern und Programmmentwicklern reflektiert wurden.

Auf dieser Grundlage wurden dann (c) gemeinsame Erwartungen an die Evaluation formuliert und ihre Zielsetzung definiert sowie (d) ein Evaluationsdesign entwickelt, das hinsichtlich seiner Gegenstandsangemessenheit und praktisch-organisatorischen Umsetzung an die besonderen Erfordernisse von »Achtung (+) Toleranz« angepasst war.

Ein Ergebnis dieser Auseinandersetzung war die Erarbeitung eines für die Evaluation maßgeblichen Grundverständnisses von Bildungsprozessen in der politischen Bildung, insbesondere im Zusammenhang mit »Achtung (+) Toleranz«. Bildungsprozesse in diesem Bereich funktionieren nicht nach einem behaviouristischen Modell, etwa dergestalt, dass das Seminarprogramm als Input in die (immer unterschiedlichen) Teilnehmer eingespeist wird und dann auf der anderen Seite ein für alle gleiches, beobachtbares Ergebnis als Wirkung herauskäme. In diesem Modell würden die Teilnehmer als »black boxes« begriffen, die durch die Programmteilnahme auf »programmierte« Weise geformt würden. Dies ist weder beabsichtigt, noch wäre es machbar.

Bildungsprogramme wie »Achtung (+) Toleranz« sind im Gegensatz dazu als relativ offene Angebote zu verstehen, die von unterschiedlichen Teilnehmern eben auch unterschiedlich aufgefasst und begriffen, d. h. mit je subjektiver Bedeutung versehen und dementsprechend auf die je eigenen Alltagswirklichkeiten bezogen werden.

Erwartbar ist daher, dass die Ergebnisse einer Teilnahme höchst unterschiedlich ausfallen, denn neues Wissen muss in subjektiven, »eigensinnigen« Aneignungsprozessen immer erst in bereits vorhandenes Wissen »eingebaut« werden, um praktische Bedeutung zu gewinnen. Handeln folgt nicht einfachen Reiz-Reaktion-Schemata, sondern ist stets in typischer Weise mit Intentionalität, mit einem auf andere bezogenen subjektiv gemeinten Sinn versehen.

- 2 Hier sind im Sinne einer phänomenologischen Soziologie so genannte »Weil-Motive« und »Um-zu-Motive« zu unterscheiden (vgl. Schütze und Luckmann 1994). Weil-Motive beziehen sich auf den Anlass zur Programmentwicklung, Um-zu-Motive auf die erwarteten, erhofften Effekte, die mit dem praktischen Einsatz des Programms verbunden sind.

Tiefenwirkung und Breitenwirkung: multimethodischer Mehrebenenansatz

Für die Evaluation bedeutet dies, dass Wirkungen zunächst dort auftreten, wo sie für den außenstehenden Beobachter unsichtbar sind: in den Veränderungen subjektiv bedeutsamer Wissensbestände. Hier liegt die Tiefenwirkung von durch die Seminareteiligung hervorgerufenen Bildungsprozessen. Da diese zunächst nur am einzelnen »Fall« festgestellt werden kann, beinhaltet das Evaluationsdesign (a) eine so genannte »Tiefenbohrung«, deren Ziel es war, am Beispiel einer Teilnehmergruppe eines »Achtung (+) Toleranz«-Trainings typische Formen und Muster der Veränderungen von Wissensbeständen und deren alltagspraktische Umsetzung – in veränderten Formen des Deutens und Handelns im Alltag – zu rekonstruieren.

Daneben bestand Interesse daran, einen Überblick über die Akzeptanz und Nachhaltigkeit von »Achtung (+) Toleranz« bei allen bisherigen Teilnehmern zu gewinnen, also gewissermaßen als Ergänzung zur fallbezogenen detaillierten Tiefenbohrung (b) die Fläche oder Breitenwirkung zu erfassen und zu dokumentieren. Hier ging es um die Einbeziehung möglichst großer Fallzahlen, was methodisch mit vertretbarem Aufwand nur durch eine postalische Befragung mittels eines teilstandardisierten Fragebogens möglich war.

Beide Teile zusammengenommen, ergaben das Evaluationsdesign eines multimethodischen Mehrebenenansatzes, mit dessen Hilfe wir eine möglichst umfassende Beschreibung und Bewertung relevanter Prozesse und Wirkungen erreichen wollten (vgl. Abb. 1 auf S. 63).

Teilstandardisierter Fragebogen

Der Einsatz standardisierter Verfahren eignet sich wegen der hohen Reliabilität und Objektivität des Verfahrens vorrangig zu einer Evaluation, die vorab definierte Reaktionen und Effekte mittels statistischer Kennwerte erfasst, und eröffnet darüber hinaus die Möglichkeit statistisch-repräsentativer Aussagen.

Als Nachteil eines derartigen Vorgehens kann sich hingegen erweisen, dass die Befragten eben nur auf das antworten können, was den Evaluatoren vorab als bedeutsam und fragenswert erschienen ist und dies lediglich in der durch die standardisierten Antwortmöglichkeiten vorgegebenen Form. Das ist insbesondere dann von Nachteil, wenn es um die Entdeckung unerwarteter, nicht-intendierter Wirkungen bei den Teilnehmern oder um die Erfassung komplexer Wissensbestände und Handlungskompetenzen gehen

soll. Durch Einbeziehung so genannter offener Fragen, die den Befragten die Möglichkeit selbst gewählter Antworten geben, können die genannten Probleme vermindert, aber letztlich nicht hinreichend gelöst werden, so dass standardisierte Befragungen keine Alternative zu dem erwähnten Einsatz qualitativer Fallstudien darstellen.

Wir haben das Instrument der postalischen Befragung eingesetzt, um in schriftlicher Form alle ehemaligen »Achtung (+) Toleranz«-Teilnehmer zu erfassen und auf diese Weise einen Überblick über allgemeine Einschätzungen des Seminarprogramms zu gewinnen und um anhand offener Fragen auch ausgewählte Felder subjektiver Bedeutungszuweisungen beschreibbar zu machen. Der hierzu entwickelte, sehr umfangreiche und daher eine hohe Motivation und Kooperationsbereitschaft der Respondenten voraussetzende Fragebogen deckte insgesamt sechs Fragebereiche ab:

- Rückblick auf »Achtung (+) Toleranz«
- Erwartungen an das Seminar »Achtung (+) Toleranz«
- Themen und Inhalte von »Achtung (+) Toleranz«
- Die Personen im Seminar
- Bewertungen und Urteile, Feedback
- Angaben zur Person

Die Auswertung der geschlossenen Fragen erfolgte mittels verschiedener statistischer Methoden. Verfahren der deskriptiven Statistik, etwa Häufigkeitsdarstellungen und Mittelwertberechnungen, ermöglichen die Darstellung der Verteilung von Merkmalsausprägungen der erhaltenen Antworten innerhalb des Sample. Auf diese Weise ließen sich etwa Kennzahlen für das Maß der Akzeptanz bestimmen (z. B. »Würden Sie die Veranstaltung weiterempfehlen?«).

In weiteren Untersuchungsschritten wurden diese Angaben auch mit Hilfe von Verfahren der korrelativen Statistik in Bezug gesetzt zu anderen untersuchungsrelevanten Merkmalen des Samples. Auf diese Weise konnten Fragen danach beantwortet werden, inwieweit das Programmkonzept etwa eine geschlechtsspezifische oder bildungsspezifische Ausrichtung hat: Kommt »Achtung (+) Toleranz« bei Frauen oder Männern, bei Teilnehmern mit höherer Bildung »besser« an?

Darüber hinaus haben wir, dort wo aufgrund theoretischer Vorüberlegungen ein »tieferer« Einblick erforderlich war, das Datenmaterial mit Hilfe einer explorativen Faktoranalyse untersucht. Unter inhaltlichen Gesichtspunkten bietet dieses Verfahren die Möglichkeit, latente, gewissermaßen »hinter« den Daten stehende Variablen aufzuspüren. Mit Hilfe dieses Vorgehens ließen sich die statistischen (quantitativen) Daten inhaltlich so interpretieren,

dass unter anderem für die Wirkungsweise von »Achtung (+) Toleranz« bedeutsame mehrdimensionale Typen der Konfliktwahrnehmung gebildet werden konnten (vgl. unseren Beitrag in diesem Band sowie Dimbath et al. 2003).

Dieses Beispiel zeigt, dass sich die Analysetiefe quantifizierender Verfahren in der Evaluation nicht in der bloßen Deskription erschöpft. Im Zusammenspiel von theoretischer Rekonstruktion des Programms und der ihm zugrunde liegenden Wirkungstheorie, deren Operationalisierung für die Verwendung im Fragebogen und dem Einsatz komplexerer statistischer Untersuchungsverfahren lassen sich auch komplexe Wirkungszusammenhänge entdecken.

Die Auswertung der offenen Fragen erfolgte mittels der aus der Grounded Theory (Strauss 1998) entlehnten Technik des Kodierens. Diese zielt darauf ab, die Daten »aufzubrechen« und zu einer gegenstandsbezogenen Theorie zu verdichten. Dazu werden die einzelnen erhaltenen Antworten in Codes »übersetzt« (Kodieren), d. h. ihre semantischen Felder und deren begriffliche Ordnung identifiziert und in sozialwissenschaftlich interpretierbare, begriffliche Kategorien transformiert (Dimensionalisieren und Kategorienbildung).

Kodieren als eine Analysetechnik, die ein Identifizieren, Dimensionalisieren und Kategorisieren der aus den Daten (re-)konstruierten begrifflichen Konzepte und ihrer wechselseitigen Bezüge beinhaltet, führt über den gesamten Analyseprozess hinweg zu einer fortschreitenden interpretativen Verdichtung und Abstraktion des Ausgangsmaterials.³

Um beim Kodieren längerer, komplexer Antworten (also solcher, die anstelle von Stichwortnennungen einen argumentativen Zusammenhang präsentierten) den Sinn des ursprünglichen Aussagenzusammenhangs nicht zu verlieren, wurden die betreffenden Textabschnitte zusätzlich einer hermeneutischen Sequenzanalyse unterzogen (vgl. einführung z. B. Reichertz 2000; Wernet 2000), um den Sinn des gesamten Aussagezusammenhangs zu erfassen und in das Kode-/Kategorienschema zu integrieren.

3 Die methodologische Grundlage für die von der Grounded Theory vorgeschlagene Kodiertechnik bildet ein Konzept-Indikator-Modell, nach dem so genannten empirische Indikatoren – d. h. in den Daten enthaltene Beschreibungen, Aussagen, Darstellungen von Verhaltensweisen, Ereignissen etc. – dem Forscher Möglichkeiten für eine Konzeptualisierung, ein Fassen dieser Indikatoren in sie bezeichnende Begriffe (»concepts«) anzeigen.

Indem der Forscher in seinem Analyseprozess eine Vielzahl solcher Indikatoren begrifflich bearbeitet (kodiert), erfolgt ein ständiger Vergleich, eine permanente, an den Daten arbeitende Systematisierung, Gruppierung, Klassifizierung und Dimensionalisierung dieser Indikatoren (zu Kodieren vgl. z. B. Böhm 2000: 475 ff.; Flick 1999: 196 ff.; Strauss und Corbin 1996; Strauss 1998).

Die auf diesem Wege gewonnenen Kategorien umreißen die Konturen der für die Befragten relevanten semantischen Felder und Rahmen, d.h. hier der von den Teilnehmern mit Blick auf das Seminar für relevant (mitteilenswert) erachteten Urteile und Bewertungen, erinnerten Erfahrungen usw.

Problemzentriert-diskursive Interviews

Um die individuellen, thematisch fokussierten Deutungen der Teilnehmer des Evaluationsseminars zu rekonstruieren, wurden mit den ihnen jeweils vor und – zur Feststellung der Nachhaltigkeit von Seminarwirkungen – einige Monate nach der Teilnahme an »Achtung (+) Toleranz« problemzentriert-diskursive Interviews geführt (vgl. Ullrich 1999; Witzel 2000).

Die Befragung sollte zum einen biographische Narrationen (Belegerzählungen) in Bezug auf Erfahrungen und Erlebnisse im Sinn- und Bedeutungskontext von Toleranz erzeugen, um darin die impliziten Wissensbestände und latenten Deutungsmuster der Befragten zu Toleranz, Achtung, Konflikten etc. rekonstruieren zu können. Zum anderen sollten die Interviews auch explizierbares Wissen in offener Form einholen (d.h. bisherige Erfahrungen mit ähnlichen Seminaren, Einstellungen und ihre Begründungen, Meinungsäußerungen etc.).

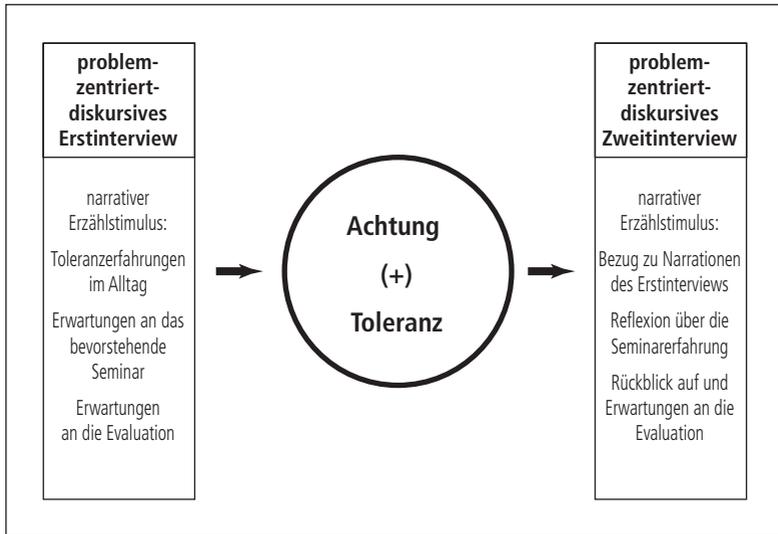
In den Nach-Interviews wurden die im ersten Interview berichteten, teilweise in Erzählungen gepackten konkreten Erfahrungen der einzelnen Seminarteilnehmer auf individuelle Weise in das zweite Interview mit einbezogen, um veränderte Interpretationszuschreibungen und Deutungen dieser Erfahrung aus der Sicht der Einzelnen zu ermitteln. Neben den problemzentrierten und diskursiven Abschnitten des Interviews wurde zum Schluss durch Nachfragen des Interviewers die Evaluation im Gesamtverlauf durch den Befragten reflektiert: Dies fand vor dem Hintergrund des »partizipativen« Interesses an der Perspektive der Teilnehmer statt und lotete nicht nur die subjektiven Einschätzungen hinsichtlich der Arbeit des Evaluationsprojekts, sondern auch mögliche Verbesserungsvorschläge aus.

Die Auswertung der Interviews erfolgte je nach erforderlicher (und damit theoretisch zu begründender) Analysetiefe gemäß unterschiedlicher Auswertungsstrategien. Durch inhaltsanalytische Vorgehensweisen konnten den Interviews wichtige Informationen ohne feinanalytische Textauslegung entnommen werden (z. B. biographische oder seminarbezogene Daten, Thematisierungen). Sequenzanalytische Verfahren der Interpretation wurden eingesetzt, um aus der sequenziellen Ordnung der in den Interviews enthaltenen Narrationen, Berichte und Reflexionen die subjektiven Bedeutungen zu re-

konstruieren, die der Interviewte dem Seminar und seinen Inhalten beimisst. Diese zeigen sich nicht nur auf der manifesten Ebene des inhaltlich Gesagten, sondern latent gerade in der Art und Weise des Erzählens, Darstellens, Begründens oder Bewertens.

Da Sequenzanalysen äußerst (zeit-)aufwändig sind, wurden sie nur an ausgewählten Schlüsselstellen eingesetzt. Die – dem eigenen Erkenntnisinteresse entsprechende und theoretisch begründbare – Auswahl dieser Stellen erfolgte auf der Grundlage einer umfassenden analytischen Bearbeitung jedes Interviews mittels eines an der Methodik der Grounded Theory (Strauss 1998) orientierten Kodierverfahrens (siehe hierzu auch die Auswertungsstrategie der offenen Fragen im Fragebogen).

Abbildung 1: Problemzentriert-diskursive Vorher- und Nachher-Befragung



Videoaufzeichnung und Gruppendiskussion

Der dritte Teil der Mehrebenenevaluation bestand in einem Video-Experiment – ein Weg, der in der qualitativen Evaluationsforschung bislang kaum beschritten wurde. Ziel war der Versuch, Effekte bzw. Wirkungen der Seminarteilnahme nicht nur aus Gesprächen mit Teilnehmern zu rekonstruieren, sondern durch die Beobachtung von Interaktionen. Damit wurde der Erwartung, die Inhalte von »Achtung (+) Toleranz« sollten sich über kognitive Veran-

kerungen und reflektierende Deutungen hinausgehend auch in konkret interaktivem Handeln niederschlagen, methodisch entsprochen.

Die Interaktionsbeobachtung per Video⁴ fand zu zwei Zeitpunkten statt – unmittelbar vor dem Seminar (t_1) und einige Monate danach (t_2). Die Seminargruppe erhielt beim ersten Durchgang t_1 den Auftrag, gemeinsam mit Hilfe einer Videokamera (Spot-Cam) einen ca. einminütigen »TV-Spot« zum Thema Toleranz/Intoleranz zu drehen; in t_2 sollte ein Werbespot für »Achtung (+) Toleranz« erstellt werden. Der Zeitrahmen für die Bewältigung der Aufgabe betrug jeweils 30 Minuten.

Zusätzlich wurde die Gruppe aufgefordert, den Gruppenprozess mit Hilfe einer weiteren Videokamera (Doku-Cam) zu dokumentieren. Aus der Art der Dokumentation erwarteten wir, Aufschluss über die Perspektive der Gruppe auf sich selbst und ihren gemeinsamen Arbeitsprozess zu gewinnen.⁵

Ferner wurde in einer Ecke des Raumes eine dritte Kamera (Raum-Cam) installiert, die das ganze Geschehen aus der Vogelperspektive aufzeichnete.⁶

Zum Zeitpunkt t_1 wurden der Gruppe unmittelbar im Anschluss an die Dreharbeiten beide Filme (Spotcam und Dokucam) vorgeführt und von der Gruppe diskutiert. So wurde uns die gemeinsame Reflexion der Gruppe über ihre Arbeit und ihre Erklärung dessen, was in den Filmen zu sehen ist, in einer für weitere Auswertungen protokollierbaren Form zugänglich (Bandaufzeichnung).

- 4 Gemäß Kai Brauer und Christopher Dehn liegt »der enorme Vorteil einer filmischen Aufzeichnung darin, dass der Ablauf bestimmter Ereignisse konserviert wird und unbegrenzt reproduziert werden kann. ... Es handelt sich hier sozusagen um visualisierte Beobachtungsprotokolle« (Brauer und Dehn 1995: 76). In unserem Fall wurden die Daten noch dazu z. T. in Eigenregie der Teilnehmer (als Selbstbeobachtung) erstellt.
- 5 Dies lässt sich beispielsweise an folgenden Fragen festmachen: Wie deutet die Gruppe den Dokumentationsauftrag? Wie organisiert sie die Durchführung? Wie gehen die Kameraführer mit den Gruppenbeschlüssen um? Zusammengefasst: Was wird von wem wie mit der DokuCam aufgezeichnet? Insofern dokumentiert das von der Doku-Cam Aufgezeichnete nicht nur den Interaktionsprozess, sondern ist selbst ein Dokument des Interaktionsprozesses, der dieser Aufzeichnung zugrunde liegt.
- 6 Eine solche »Überwachungskamera«, die, in einer entlegenen Ecke des Raums platziert, während der Gruppenübung völlig aus dem Aufmerksamkeitsbereich der Übungsteilnehmer verschwindet, bedarf einer sensiblen Einführung (vgl. hierzu Renne, Dowrick und Wasek 1983). Sehr schnell kann der Eindruck entstehen, man werde mit dem Dokumentationsauftrag nicht ernst genommen oder solle intransparenten »Psychotests« unterzogen werden. Für eine Teilnehmerin des Projekts war die Einführung durch die Evaluatoren offenbar nicht sensibel genug, was sie im Rahmen des nachbereitenden Gruppengesprächs zu kritischen Äußerungen in Bezug auf die Einrichtung einer Raum-Cam bewog.

In t_2 kehrte sich die Vorgehensweise um. Hier fand das Gespräch nebst Betrachtung der alten Filme (Spot und Doku) in Verbindung mit einer gemeinsamen Rückbesinnung auf die besuchte Seminarveranstaltung vor der Gruppenübung statt. Erst danach erhielt die Gruppe wieder einen Arbeitsauftrag – Erstellen eines Werbespots für »Achtung (+) Toleranz« – und die beiden Kameras (Spot-Cam und Doku-Cam). Aufgrund dieser Umkehrung der Untersuchungsanordnung erwarteten wir, die Gruppe würde durch das vorangehende Gespräch u. a. »Seminarwissen« aktualisieren und dieses in die anschließende Lösung der praktischen Aufgabe einfließen lassen.

Das hier beschriebene Untersuchungsarrangement war auf die Konfrontation unterschiedlicher Perspektiven angelegt: Die Selbstbeschreibung des Arbeitsprozesses mittels der Doku-Cam durch die Gruppe selbst, die »objektive« Beobachtung des gesamten Gruppenprozesses mittels der fest installierten Beobachtungskamera sowie die Reflexion des Handlungsprozesses und des individuellen Handelns im Rahmen der Gruppendiskussionen. Das hätte im Sinne einer »Triangulation« (vgl. Flick 1999) von Beobachtungs- und Untersuchungsperspektiven eine multiperspektivische, sozusagen »360-Grad-Analyse« des Gruppenprozesses ermöglicht.

So ist die Raum-Cam in der Lage, Prozesse zu erfassen, die dem intentionalen Blick der Doku-Cam verborgen bleiben, und damit überhaupt erst deren Selektionsprozesse transparent zu machen. Verdeutlicht wird so der Zusammenhang zwischen dem Handeln »vor der Kamera«, das durch das »Handeln hinter der Kamera« (Doku-Cam) stets nur ausschnitthaft dokumentiert werden kann (vgl. Reichartz 1994). Die Transkripte der Gruppengespräche enthalten zudem zusätzliche Informationen über Gruppen- und Einzelentscheidungen.

Eine systematische Bearbeitung sämtlicher sich hierdurch eröffnender Analyseebenen ist äußerst ressourcenintensiv und war nach Absprache mit den Auftraggebern der Evaluation auch nicht beabsichtigt. Vielmehr handelte es sich bei der Einbeziehung dieses methodischen Konzepts selbst um ein über den eigentlichen Auftrag hinausgehendes, exploratives Experiment.

Evaluatoren wie Auftraggebern war von Anfang an unklar, wie weit dieses im Rahmen der vertraglich vereinbarten Ressourcen ausgewertet werden könnte. Um den Aufwand vertretbar zu halten, mussten wir es bei der Auswertung der Protokolle der Gruppendiskussionen und der Aufzeichnungen durch die Doku-Cam bei einer Ersteindrucksanalyse belassen und haben lediglich die mittels der Raum-Cam erfolgten »neutralen« Aufzeichnungen des Interaktionsprozesses einer genaueren Analyse unterzogen.

Bei der Auswertung des filmischen Materials bietet sich die Arbeit mit Einstellungs- und/oder Sequenzprotokollen an (vgl. hierzu Maier 1995). Da

die Anfertigung eines partiturähnlichen Verlaufsprotokolls, in dem jedes Wort notiert und jede Kamerabewegung bzw. Einstellung erfasst und vor dem Hintergrund ihrer Kontextbedingungen kommentiert wird, enorm aufwändig ist, haben wir für die Raum-Cam auf die kürzeren Sequenzprotokolle zurückgegriffen, mit deren Hilfe die Filmhandlung in sinnvolle Abschnitte gegliedert werden kann.

Entlang der Fragestellung, wie die Gruppe ihren Schaffensprozess strukturiert, wurden auf der Grundlage dieser Sequenzprotokolle Aktionsmemos erstellt, die zu einem späteren Fallvergleich zwischen Doku-Cam und Raum-Cam ebenso wie zwischen den beiden Aufzeichnungszeitpunkten herangezogen wurden.

Beim Vergleich der beiden Gruppenübungen wurde der Fokus auf die Entwicklung der Kommunikation im Entscheidungsprozess gerichtet. Anhand einer Kurzbeschreibung der jeweils 30-minütigen Arbeitsphasen und der Konzentration auf die Sequenzprotokolle aus den beiden Raum-Cam-Filmen werden zunächst kommunikative und prozedurale Muster der Aufgabenbewältigung erstellt und anschließend miteinander kontrastiert.

Dieser Vergleich erfolgte vor dem Hintergrund einer Interpretationsfolie, die versuchte, sich auszumalen, welche Veränderungen im Sinne von Achtung (+) Toleranz idealerweise zu erwarten gewesen wäre. Das Ergebnis der Untersuchung ergab sich also durch eine trianguläre Gegenüberstellung der beiden Aufzeichnungszeitpunkte und des vermuteten Idealergebnisses.

Fazit

Den Kern des hier vorgestellten Evaluationskonzepts bildete die partizipative, dialogisch-diskursive Vorgehensweise, in der zusammen mit Auftraggebern und Programmentwicklern sowohl Zielsetzungen als auch Kriterien der Beurteilung von »Achtung (+) Toleranz« entwickelt wurden.

Im Zentrum dieses Prozesses stand die sorgfältige Analyse des Programms und die Rekonstruktion der ihm zugrunde liegenden Leitkonzepte (Welcher Toleranzbegriff wird unterlegt?) sowie der im didaktischen Aufbau enthaltenen Wirkungstheorie (Was soll welche Übung, welcher Trainingsteil erreichen, welches ist sein vermuteter Beitrag zur angestrebten Toleranzbildung?).

Auf der Grundlage dieser gemeinsamen Auseinandersetzung wurde als Rahmen für die spätere Analyse und Einschätzung eine gemeinsame Theorie des Toleranzbildungsprozesses entwickelt. Bereits vorab wurde deutlich, dass die erwartbaren Wirkungen einer Teilnahme nicht vorab eindeutig bestimm-

bar sein würden. Von daher erschien eine bloße Abfrage vorab operationalisierter Wirkungsdimensionen allein als nicht hinreichend. Dem wurde durch die Entwicklung des hier vorgestellten multimethodischen Mehrebenenansatzes Rechnung getragen.

Der Einsatz eines teilstandardisierten Fragebogens erbrachte Einsichten über die allgemeine Akzeptanz des Programms bei den Teilnehmern, über deren Erwartungen und deren Einschätzungen (Beurteilung) der im Seminar ablaufenden Gruppen- und Bildungsprozesse, die Qualität der Vermittlung von Seminarinhalten sowie Hinweise auf besonders gut erinnerte Seminarbestandteile und deren Nachhaltigkeit. Diese Fragen hängen inhaltlich miteinander zusammen: So ist die Akzeptanz der didaktischen Form sicherlich mit dafür ausschlaggebend, ob vermittelte Inhalte auch auf positive Resonanz stoßen und »angenommen« werden; die persönliche Zufriedenheit mit dem Ablauf insgesamt beeinflusst über positive Erinnerungen die Nachhaltigkeit etc. Derartige Erkenntnisse dienen dazu, eine erste Bewertung des Programms zu ermöglichen.

Dennoch erfassen die mit Hilfe derartiger Erhebungs- und Auswertungsverfahren gewonnenen Einsichten eher die »Oberfläche« rückblickend erteilter manifester Meinungsäußerungen und Zustimmungsbekundungen. Ein Einblick in die Wirkungsweise des Programms selbst lässt sich auf diesem Wege nicht gewinnen.

Hier setzte die Fallstudie an. Anhand der Interviewprotokolle konnten Prozesse der durch »Achtung (+) Toleranz« angestrebten Sensibilisierung für Vielfalt und Relationalität von Werturteilen, für das Vorhandensein stillschweigender Annahmen, die der je eigenen Wirklichkeitsdeutung unterlegt sind, und für eine Veränderung des subjektiven Toleranzverständnisses aufgezeigt werden.

Bedeutsam an dieser Form der Rekonstruktion war, dass sie über die Klassifizierung einzelner Veränderungen hinauswies und die Erstellung von Typologien der Wissensaneignung und des Wissenstransfers ermöglichte. Diese machen die Logik des Bildungsprozesses transparent und erlauben eine Einschätzung von Seminarerfolgen auch dort, wo manifeste Verhaltens- oder Einstellungsänderungen nicht oder noch nicht beobachtbar sind.

Wie weit die Teilnahme an einem »Achtung (+) Toleranz« Seminar wirkt, ist demzufolge nicht nur vom Programm und seiner Durchführung selbst, sondern ebenso von den reflexiven und sozialen Kompetenzen der daran Teilnehmenden abhängig. Wie man die erreichten Ergebnisse beurteilt, ist dann nicht nur eine Frage objektivierter Wertungsskalen, sondern bemisst sich an den (aufgrund der Typologien beschreibbar gemachten) Entwicklungsprozessen, die bei den Teilnehmern ausgelöst wurden.

Das Video-Experiment brachte demgegenüber noch eine weitere Dimension zum Vorschein, die mit den beiden genannten Verfahren nicht hinreichend erfasst werden kann, die jedoch gerade für das Feld der Toleranz von ausschlaggebender Bedeutung ist: die Eigendynamik des sozialen Prozesses bzw. sozialer Prozesse. Toleranz ist weniger in Form einer Einstellung bedeutsam, sondern letztlich dort, wo es situativ um das praktische Handeln in sozialen Zusammenhängen geht.

Welche Eigenlogik hier zum Tragen kommt, deuten die auf den ersten Blick paradox anmutenden Ergebnisse der Videobeobachtung an (vgl. unseren Praxisbericht in diesem Band). Soziologisch ist dies nicht weiter verwunderlich. Unserer Auffassung nach hat dieses Ergebnis jedoch weiter reichende methodisch-methodologische Implikationen für die Evaluation von auf soziale Kompetenzsteigerung gerichteten Programmen und Maßnahmen der politischen Bildung.

Es zeigt, dass das Soziale eine eigenständige Wirklichkeitsebene darstellt, die nicht als Summe ihrer einzelnen Teile begriffen werden kann. Von daher sind in diesem Zusammenhang Verfahren, die lediglich individuelle Meinungen, Auffassungen und Beurteilungen erfassen, ungeeignet, das für das Zusammenleben letztlich ausschlaggebende Soziale methodisch adäquat in den Blick zu nehmen – denn: Denken und Handeln sind nicht (immer) ein und dasselbe.

Literatur

- Böhm, Andreas (2000): »Theoretisches Codieren. Textanalyse in der Grounded Theory«. In: Uwe Flick, Ernst von Kardoff und Ines Steinke (Hrsg.): *Qualitative Forschung. Ein Handbuch*. Reinbek. 475–484.
- Brauer, Kai, und Christoph Dehn (1995): »Videotechnik. Ein Leitfaden zur Orientierung«. In: Kai Brauer und Christoph Dehn: *Sozialwissenschaftliche Videographie. Erfahrungsberichte, Anwendungsmöglichkeiten und technisches Grundwissen*. Berlin. 57–84.
- Dimbath, Oliver, Julia von Hayek, Andreas Hirseland und Werner Schneider (2003): »Toleranz, Konflikte und Wirkungen. Zur multimethodischen Mehrebenen-Evaluation von Toleranzbildungsprogrammen am Beispiel von Achtung (+) Toleranz«. In: Heinz Lynen von Berg und Roland Roth (Hrsg.): *Maßnahmen und Programme gegen Rechtsextremismus*. Opladen 193–227.
- Flick, Uwe (1999): *Qualitative Forschung. Theorie, Methoden, Anwendung in Psychologie und Sozialwissenschaften*. Reinbek.

- Maier, Birgit (1995): »Zur Methodik der Filmanalyse von ethnographischen Filmen«. In: Edmund Ballhaus und Beate Engelbrecht: *Der ethnographische Film: eine Einführung in Methoden und Praxis*. Berlin. 223–267.
- Reichertz, Jo (2000): »Objektive Hermeneutik und hermeneutische Wissenssoziologie«. In: Uwe Flick, Ernst von Kardoff und Ines Steinke (Hrsg.): *Qualitative Forschung. Ein Handbuch*. Reinbek. 514–523.
- Reichertz, Jo (1994): »Selbstgefälliges zum Anziehen. Benetton äußert sich zu Zeichen der Zeit«. In: Norbert Schröer (Hrsg.): *Interpretative Sozialforschung. Auf dem Weg zu einer hermeneutischen Wissenssoziologie*. Opladen. 253–280.
- Renne, C. M., Peter W. Dowrick und G. Wasek (1983): »Considerations of the Participant in Video Recording«. In: Peter W. Dowrick und Simon J. Biggs: *Using Video. Psychological and Social Applications*. Chichester. 23–32.
- Schütze, Alfred, und Thomas Luckmann (1994): *Strukturen der Lebenswelten*. Band 1. Frankfurt am Main.
- Shotter, John (1983): »On Viewing Videotape Records of Oneself and Others. A Hermeneutical Analysis.« In: Peter W. Dowrick und Simon J. Biggs: *Using Video. Psychological and Social Applications*. Chichester. 199– 210.
- Strauss, Anselm L. (1998): *Grundlagen qualitativer Sozialforschung. Datenanalyse und Theoriebildung in der empirischen und soziologischen Forschung*. München.
- Ullrich, Carsten G. (1999): »Deutungsmusteranalyse und diskursives Interview. Leitfadenkonstruktion«. In: *Interviewführung und Typenbildung*. Mannheimer Zentrum für Europäische Sozialforschung, Arbeitspapier 3.
- Wenzel, Florian (2003): »Ist politische Bildung messbar? Zur internen Evaluation und Qualitätssicherung in der außerschulischen Jugendarbeit«. Unveröffentlichte Workshop-Dokumentation. München.
- Wernet, Andreas (2000): *Einführung in die Interpretationstechnik der Objektiven Hermeneutik*. Opladen.
- Witzel, Andreas (2000): »Das problemzentrierte Interview (26 Absätze)«. In: *Forum Qualitative Sozialforschung/Forum Qualitative Social Research* (Online-Journal) (1) 1 2000, <http://qualitative-research.net/fqs>.
- Ziller, Robert C. (1999): *Photographing the Self. Methods for Observing Personal Orientations*. Newbury Park, Cal.

Datenerhebungsmethoden der Evaluation

»Eine Welt der Vielfalt«

Jochen Kehr

Die in der Evaluation verwendeten Datenerhebungsmethoden wurden für das Programm »Eine Welt der Vielfalt« eigens konzipiert und konnten so den Besonderheiten und Bedingungen des Programms angepasst werden. Das Gesamtdesign der Evaluation wurde mit seinen zwei Phasen so angelegt, dass die gewonnenen Informationen aus einer Fremd- sowie Selbsteinschätzung der »outcomes« des Programms auf die Schüler bestanden. Die erste Phase der Fremdeinschätzung durch die Lehrer und Koordinatoren diente dabei als explorierende Vorbereitung, um die Schüler passgenauer befragen zu können.

Telefoninterviews als Methode der ersten Untersuchungsphase

In der ersten Untersuchungsphase wurde die Methode des Telefoninterviews zur Ermittlung der Fremdeinschätzung eingesetzt. Telefoninterviews haben den Vorteil, dass Personen aus unterschiedlichen Regionen ohne großen zeitlichen Aufwand fokussiert befragt werden können. Allerdings erfordert das Durchführen von Telefoninterviews ein hohes Befragungsgeschick und eine gute Vorbereitung.

Auf der Grundlage der vereinbarten Fragestellungen wurde ein strukturierter Interviewleitfaden entwickelt, der geschlossene Fragen (z. B. »Welche der fünf Lektionen setzen Sie komplett ein?« mit Antwortvorgaben eins bis fünf) mit offenen Fragen (z. B. »Welche Wirkungen, die Ihnen sehr wichtig wären, werden Ihrer Meinung nach wahrscheinlich nicht erreicht?«) kombiniert.

Diese Kombination von geschlossenen und offenen Fragen fordert die Befragten entweder auf, sich auf eine Antwortvorgabe festzulegen oder eine subjektive eigene Beschreibung/Einschätzung mit eigenen Worten zu geben. Dabei wurden in den Interviewleitfaden Anregungen von Auftraggeberseite (Bertelsmann Stiftung und Centrum für angewandte Politikforschung) auf-

genommen. Der Leitfaden wurde vor dem Echteinsatz intern und extern Pretests unterzogen.

Durch die Antworten der Befragten sollten Trends bzw. offene Fragen (Hoffnungen, Befürchtungen) zur Wirkung des Programms ermittelt werden, die für die zweite Phase der Schüler-Befragung relevante Ansatzpunkte darstellen. Dabei sollten erwünschte und tatsächliche »outcomes« auf Schüler-Ebene identifiziert werden. Ebenfalls sollte geklärt werden, in welchem Umfang das Programm an den Schulen der Befragten tatsächlich umgesetzt wird.

Befragt wurden 39 Lehrer, die die Voraussetzungen erfüllten, an einem Lehrertraining teilgenommen zu haben und das Programm mit hoher Wahrscheinlichkeit in ihrer Schule einzusetzen. Die Interviewpartner wurden entweder durch die Länderkoordinatoren in einem Gespräch über das anstehende Interview informiert oder erhielten durch die Evaluatorinnen ein Schreiben, in welchem kurz der Hintergrund des geplanten Interviews erläutert und ein Anruf zur Terminvereinbarung angekündigt wurde. Mit den Interviewpartnern wurden individuell Interviewtermine für 30-minütige Interviews vereinbart.

Der zeitliche Aufwand, die Interviewpartner zu erreichen, war unerwartet hoch. In einigen Fällen wurden individuell vereinbarte Interviewtermine mehrfach nicht eingehalten, Nachfragen telefonischer und schriftlicher Art wurden teilweise nicht oder sehr spät beantwortet. Zudem war es eine Schwierigkeit, dass von vielen Interviewteilnehmern nur die dienstlichen Nummern der Schulen vorlagen. Der Schulalltag und die Bedingungen in der Schule (nur ein Telefon o.Ä.) sind bestimmend dafür, dass es enorm schwierig war, die Lehrer zu erreichen. Die Interviews hatten eine durchschnittliche Länge von 30 Minuten.

Im Anschluss an die Lehrer-Interviews wurden Interviews mit den Koordinatoren durchgeführt. Die Koordinatoren hatten hier die Rolle als Experten, da sie intensiv mit dem Programm auf verschiedenen Ebenen vertraut sind (Koordination, Multiplikation, Fortbildung). Die zehn mit Koordinatoren durchgeführten Interviews hatten eine durchschnittliche Länge von 45 Minuten. Auch hier wurde mit einem zuvor abgestimmten und in einem Pretest geprüften Interviewleitfaden gearbeitet. Alle Interviews wurden mit Headset geführt und die Antworten mittels Textverarbeitung direkt protokollarisch festgehalten.

Die Antworten zu den geschlossenen Fragen wurde mit dem Computerprogramm SPSS ausgewertet, die Antworten auf die offenen Fragen mit der Text-Sortier-Technik (TST) (Beywl und Schepp-Winter 2000). Die TST-Technik ist eine Auswertungstechnik für qualitative Daten auf der Basis eines

konventionellen Textverarbeitungsprogramms (z. B. Word). Die Methodik basiert genau wie bei anderen qualitativen Auswertungsmethoden darauf, die Aussagen in sinnverwandte Kategorien zu ordnen und diesen Prozess nachvollziehbar zu machen.

Rückblickend hat sich das Telefoninterview als geeignete Methode bewährt, da damit bundesweit eine relativ große Zahl an Lehrern und Experten befragt und wichtige Informationen zur Umsetzung sowie Wirkung des Programms gesammelt werden konnten. Mithilfe dieser Informationen aus der Fremdperspektive konnte die Befragung der Schüler zielgenau konzipiert werden. Allerdings zeigte sich, dass der Aufwand, Telefoninterviews durchzuführen, bei dieser Zielgruppe nicht unterschätzt werden darf. Nur gut vorbereitete Telefoninterviews geben gehaltvolle Informationen, die für die Evaluation solch eines Programms notwendig sind.

Fallstudien als Methode der zweiten Untersuchungsphase

Um die »outcomes« bei den Schülern fokussiert zu untersuchen, wurden in der zweiten Untersuchungsphase Fallstudien in zwei ausgewählten Schulen durchgeführt, um die Akzeptanz des Programms und seine Wirkung bei Schülern zu untersuchen.

Ursprünglich war geplant, die Schulen nach theoretischen Annahmen (z. B. verschiedene Schultypen, Regionen) auszusuchen. Dies erwies sich jedoch als nicht leistbar. Die Anzahl der Schulen, die im vorgegebenen Erhebungszeitraum bereit war, teilzunehmen, war sehr gering. Teilweise wurde Zeitmangel durch die Lehrer, teilweise wurden andere Begründungen durch die Schulleitungen angegeben.

Dies verweist auf die Notwendigkeit, bereits bei einem Programmstart auf die geplanten Evaluationsaktivitäten hinzuweisen und zu versuchen, bei den Beteiligten eine Selbstverpflichtung zur Teilnahme und aktiven Unterstützung der Evaluation zu erzeugen. Damit dies möglich wird, müssen Evaluationen nachvollziehbar nützlich für den angestrebten Zweck (hier: Entscheidungsbasierung) sein und die einzusetzenden Instrumente hohe augenscheinliche Validität besitzen (Beywl und Joas 2000: 83–100).

Die Fallstudien wurden an den beiden Schulen durchgeführt, die sich zur Teilnahme bereit erklärt hatten und in denen viel mit dem Programm gearbeitet wird, viele Übungen des Materialordners zum Einsatz kamen, und die Lehrer im Interview deutlich machten, dass sie Wirkungen bei ihren Schülern beobachten konnten (zwei 10. Klassen des Rouanet-Gymnasiums in Beskow, Brandenburg, insgesamt 28 Schüler; sowie eine Abschlussklasse der

Fachschule für Sozialpädagogik in Osterholz-Scharmbeck, Niedersachsen, insgesamt 17 Schülerinnen). Die Schüler, die an der Befragung teilnahmen, taten dies auf freiwilliger Basis. Sie hatten zuvor schriftlich der Befragung zugestimmt.

In den Fallstudien wurden die Perspektive und die diesbezüglichen Wahrnehmungen und Meinungen der Schüler abgefragt und die Annahmen zu »outcomes« aus Phase I vertieft überprüft: Die Hauptfragestellungen lauteten: Welche von den Lehrern wahrgenommenen »outcomes« werden von den ausgewählten Schülern der Fallstudien bestätigt? Welche werden nicht oder in geringerem Maße bestätigt? Gibt es weitere »outcomes«, die sich feststellen lassen?

Die zweistündige Untersuchung war methodisch so angelegt, dass sie einerseits treffgenau die wesentlichen Fragestellungen bearbeitet, andererseits die Zielgruppe Schüler altersgemäß aktiviert und auch Spaß mit sich bringt. Wichtig war den Evaluatorinnen, dass es sowohl Phasen des aktiven Austauschs gibt (Brainstorming, Kleingruppenarbeit) als auch Phasen der Einzelarbeit bei sehr persönlichen Fragestellungen (z. B. Fragekarten zum Bereich Persönlichkeit/Selbstbewusstsein). Der unten stehende Ausschnitt des Untersuchungsdesigns verdeutlicht die methodische Anlage.

Tabelle 1: Ausschnitt des Untersuchungsdesigns der Fallstudien

Wirkungsbereich	Fragen (Beispiele)	Methode
Eröffnung	»Wenn ihr noch einmal an das Programm EWDV zurückerdenkt, was fällt euch spontan ein?«	Brainstorming
Reflexion/ Sensibilisierung	»Während wir mit Eine Welt der Vielfalt gearbeitet haben, ist es mir öfter als vorher passiert, dass ich im Nachhinein über Situationen nachgedacht habe, in denen diskriminierende Handlungen passiert sind. Trifft das für dich zu?«	Moderatorin gibt Fragen ins Plenum. Schülerinnen notieren auf Matrix.
Persönlichkeit/ Selbstbewusstsein	»Hat dir das Programm die Möglichkeit gegeben, Dinge von dir zu erzählen, die du dich vorher nicht getraut hast zu sagen oder zu zeigen?«	Fragekarte in Umschlag wird ausgegeben. Schüler antworten schriftlich.
Wissenszuwachs/ Information	»Weißt du jetzt besser, wie Vorurteile entstehen?«	Stellung einnehmen, Abstand zum Mittelpunkt auf großem Papier markieren

<i>Umgang miteinander</i>	»Hat sich der Umgang miteinander und das Klima in eurer Klasse verändert, seit ihr mit dem Programm arbeitet?«	Sammeln auf Moderationskarten
<i>Akzeptanz und Gesamtbewertung des Programms</i>	»Stellt euch vor, die Schülersprecher einer Schule in Berlin laden euch ein, über eure Erfahrungen mit Eine Welt der Vielfalt zu sprechen. Sie überlegen, das Programm in ihrer Schule zu nutzen und möchten – mal ganz unabhängig von irgendeiner Lehrmeinung – von euch wissen: Lohnt sich das Programm? Worauf müssen diejenigen unbedingt achten, die damit arbeiten? Was haltet ihr für schwierig oder diskussionsbedürftig?«	Kleingruppenarbeit und Präsentation im Plenum

Die Methoden der Fallstudie wurden auf das Programm Eine Welt der Vielfalt zugeschnitten. So werden bei folgender Fragemethode schon während des Informationsprozesses Unterschiede innerhalb der Gruppe und dadurch auch den Beteiligten deutlich. Die Schüler sollten Fragen wie »Hast du durch das Programm neue Aspekte über die Auswirkungen von Vorurteilen erfahren?« beantworten, indem sie auf einem großen kreisrunden Plakat, das im Raum auf dem Boden ausgelegt war, im wörtlichen Sinne Stellung bezogen.

Je nach dem Grad der Zustimmung zu der Aussage sollten sie sich relativ zum Mittelpunkt (=o) positionieren. Je weiter sich die Schüler dabei vom Mittelpunkt wegstellten, desto weniger, je näher sie am Mittelpunkt standen, desto mehr stimmten sie der Frage zu. Ihre Position haben die Schüler mit einer Farbe auf dem Plakat verewigt, die Abstände der Füße zum Mittelpunkt wurden in cm gemessen.

In der Evaluation wird damit das Ziel von »Eine Welt der Vielfalt« aufgegriffen, Unterschiede in der Gruppe deutlich zu machen und gegenseitigen Respekt zu fördern. Der Lernprozess, den das Programm angestoßen hat, wird also durch diese Evaluationsmethoden weitergeführt.

Rückblickend lässt sich sagen, dass die Schüler mit großer Ernsthaftigkeit und hohem Engagement mitgearbeitet haben. Die Antworten der schriftlichen Abfragen sind sehr ausführlich und von beeindruckender Differenziertheit.

Der Methodenmix der Fallstudie scheint den Schülern ein Setting gegeben zu haben, in dem sie sich mit ihren Meinungen und Erfahrungen ernst genommen fühlten und dementsprechend gerne mitgearbeitet haben. Vor allem der Wechsel aus Einzelantworten und Gruppenmethoden ist für Schüler

ein geeigneter Weg der Befragung, der eine eigene Stellung innerhalb einer Gruppe zulässt und Tendenzen der Anpassung wegen wahrgenommenen Gruppendrucks mindert.

Literatur

- Beywl, Wolfgang, und Steffen Joas (2000): »Evaluation ist unnatürlich! Eine Einführung in die nutzenfokussierte Evaluation entlang eines Seminars von Michael Q. Patton«. In: Jörg Strübing u. a. (Hrsg.): *Empirische Sozialforschung und gesellschaftliche Praxis. Bedingungen und Formen angewandter Forschung in den Sozialwissenschaften. Festschrift für Helmut Kromrey*. Opladen. 83–100.
- Beywl, Wolfgang, und Ellen Schepp-Winter (2000): »QS 29. Zielgeführte Evaluation von Programmen. Ein Leitfaden.« In: Bundesministerium für Familie, Senioren, Frauen und Jugend (Hrsg.): *Materialien zur Qualitätssicherung in der Kinder- und Jugendhilfe*. Berlin.

»Miteinander – Erfahrungen mit Betzavta« – Methode der Evaluation

Michael Bommes, Jochen Guter, Ulrike Wolff-Jontofsohn

Methodischer Rahmen

In Ergänzung zu unserem Praxisbericht wird im Folgenden der methodische Rahmen der Evaluation des Programms »Miteinander – Erfahrungen mit Betzavta« genauer erläutert. Unsere Ergebnisse stützen sich im Wesentlichen auf eine quantitative Erhebung mittels postalisch versandter Fragebögen sowie eine qualitative Erhebung, in der ausgesuchte Multiplikatoren des Programms intensiver befragt wurden.

Darüber hinaus wurden die Berufserfahrungen einer aus Praktikern des Programms bestehenden Forschungsbegleitgruppe genutzt, um den zur Erforschung des Programms gewählten methodischen Rahmen zu verbessern und für die Interpretation der gewonnenen Daten zusätzliche Perspektiven zu gewinnen. Die Arbeit mit einer Praktikergruppe gilt in der Evaluation als innovatives Element. Wir gehen daher in der nachfolgenden Diskussion unseres Vorgehens ausführlicher auf die Chancen, Grenzen und Verbesserungsmöglichkeiten im Umgang mit Begleitgruppen ein.

Quantitative Erhebung

Der für den Zweck der quantitativen Erhebung erstellte Fragebogen ist mit 18 Seiten sehr umfangreich und enthält zudem überwiegend so genannte offene Fragen (im Internet unter: www.cap.uni-muenchen.de/2003/2003_toleranz_anhang_betzavta.pdf). Der Umfang ergab sich daraus, dass wir entsprechend der Zielsetzung des Projektes von den Multiplikatoren möglichst ausführliche Auskünfte zu Themenbereichen wie Einsatz des Programms, Zielgruppen, Implementierungsmöglichkeiten in unterschiedlichen Organisationen, Wahrnehmung und Einschätzung der Grenzen und Möglichkeiten politischer Bildung, Kriterien des Erkennens von Erfolg pädagogischen Handelns u. a. erhalten wollten.

Dabei sollten die Befragten durch die offenen Fragestellungen die Möglichkeit erhalten, ihre Perspektiven zur Geltung zu bringen, auch im Wissen darum, dass dies den Auswertungsaufwand erheblich steigern würde.

Wir gingen davon aus, dass Umfang und Fragemodus die Bereitschaft zur Teilnahme und Beantwortung des Fragebogens im Fall der befragten Probandengruppe nicht absenken würden, denn wir nahmen an, dass die Befragten hoch motiviert seien, möglichst umfassend Auskunft zu geben und daher aufgrund des Designs des Fragebogens eine entsprechende Möglichkeit dazu geboten bekommen würden. Man darf davon ausgehen, dass das Programm »Miteinander« nur von Personen verwendet wird, die von seiner Zielsetzung überzeugt sind. Selbst wenn negative Arbeitserfahrungen mit dem Programm gemacht worden sein sollten, ist davon auszugehen, dass angesichts der geteilten Zielsetzung umso mehr die Bereitschaft besteht, gegebenenfalls solche Erfahrungen mitzuteilen. Diese Annahme wurde weitgehend bestätigt.

In den beantworteten Fragebögen merken einige Befragte in der Tat die Länge des Fragebogens kritisch an, andere heben aber gerade dies als sein Positivum hervor.

Qualitative Erhebung

Die im Rahmen der Untersuchung befragten Interviewpartner wurden nach einer ersten Durchsicht der Fragebögen, jedoch noch vor einer endgültigen statistischen Auswertung ausgewählt. Im Fragebogen wurden die »Miteinander«-Multiplikatoren gebeten, im Falle einer Bereitschaft zur Teilnahme an einem Tiefeninterview ihre Kontaktadresse zu notieren. 26 von 69 Befragten (38 Prozent) erklärten ihre Einwilligung zu einem solchen Interview.

Dabei ist anzumerken, dass manche der restlichen Befragten notierten, in ihrer eigenen Einschätzung über zu wenig Erfahrung mit dem Programm zu verfügen und somit eine Intensivbefragung von sich aus ausschlossen.

Für die Auswahl der Befragten wurden in Absprache mit einer Forschungsbegleitgruppe folgende Gesichtspunkte zugrunde gelegt:

- Bezüglich der Erfahrung im Umgang mit dem Programm sollte die Gruppe der Befragten möglichst heterogen zusammengesetzt sein und somit sowohl langjährige Trainer als auch Neulinge umfassen.
- Der Fragebogen sollte in einem Maße vollständig ausgefüllt sein, dass Beziehungen zwischen quantitativen und qualitativen Daten hergestellt und analysiert werden können.
- Berücksichtigung sollte des Weiteren die Herkunft der Multiplikatoren finden. Ziel war es, Teilnehmer aus möglichst vielen Bundesländern zu inter-

vieren sowie gleichermaßen Befragungen in den alten und neuen Bundesländern durchzuführen.

- Die Arbeitsfelder, in denen die Teilnehmer des Interviews das Programm zum Einsatz bringen, sollten unterschiedlich sein, sodass Rückschlüsse auf mögliche Wirkungen in Schulen, auf dem freien Bildungsmarkt und in anderen Bildungseinrichtungen gezogen werden können.
- Es sollte darauf geachtet werden, dass den Verhältnissen der quantitativen Befragung hinsichtlich Geschlecht und Alter bei der Auswahl der Befragten Rechnung getragen wurde.

Aufgrund der formulierten Bedingungen zur Interviewteilnahme mussten vier potenzielle Kandidaten von vornherein ausgeschlossen werden, da sie den Fragebogen bis auf einige wenige Fragen nicht ausgefüllt hatten oder nach eigener Auskunft über keine Erfahrung im selbständigen Einsatz des Programms verfügten.

Bei der verbleibenden Gruppe der Befragten, die ihre Bereitschaft zur Teilnahme an einem Intensivinterview erklärten, fällt auf, dass es sich – nach den Kriterien der quantitativen Analyse – vorwiegend um »erfahrene« Multiplikatoren handelt (12 von 22). Dies wurde insofern zum Problem, da der Bedingung einer heterogenen Zusammensetzung der Interviewgruppe hinsichtlich Erfahrungsgrad sowie einer Widerspiegelung der Verhältnisse der quantitativen Befragung schwer Rechnung zu tragen war.

Unter Berücksichtigung aller übrigen Bedingungen ergab es sich, dass die Interviews zu einem großen Teil mit Multiplikatoren, die über viel Erfahrung verfügten, geführt wurden. Insgesamt wurden elf Interviews zwischen März und Juni 2002 durchgeführt, von denen sieben mit »erfahrenen« Befragten stattfanden. Diese Tatsache gilt es im Auge zu behalten, da – wie die statistische Analyse gezeigt hat – der tatsächliche Anteil dieser Gruppe weitaus geringer ist.

Hinsichtlich der Merkmale Geschlecht und Alter gibt es annähernd die gleichen Verhältnisse in der Interviewgruppe wie in der Fragebogenerhebung (sieben Frauen, vier Männer; Durchschnittsalter 43 Jahre). Unter den Interviewten finden sich vier Lehrer, fünf andere Pädagogen und ein Polizist, womit die Berufsgruppe der Lehrer etwas über- und die Berufsgruppe der »Sonstigen« unterrepräsentiert ist. Letzteres erklärt sich aus der Tatsache, dass zum einen die vier Befragten, die von vornherein aufgrund ihres Antwortverhaltens bzw. ihrer fehlenden Erfahrung aus der Gruppe der potenziellen Interviewpartner ausgeschlossen wurden, zum Teil aus der Gruppe der »Sonstigen« stammen, zum anderen finden sich in der Gruppe Mitarbeiter des Centrums für angewandte Politikforschung, die aufgrund ihrer Nähe

zum Auftraggeber für die regulären Interviews nicht ausgewählt werden konnten, da dies den gängigen Forschungsstandards widersprochen hätte (vgl. Deutsche Gesellschaft für Evaluation 2002).

Eine Einschränkung musste hinsichtlich der regionalen Herkunft gemacht werden, da lediglich drei Befragte aus den neuen Bundesländern ihre Bereitschaft zum Tiefeninterview erklärten und aufgrund oben genannter Kriterien nur zwei dieser Personen ausgewählt werden konnten.

In Zusammenarbeit mit der Forschungsbegleitgruppe wurde ein Gesprächsleitfaden mit sechs Themenkomplexen entwickelt, der einige Fragestellungen des Fragebogens wieder aufnimmt und vertieft. Die verschiedenen Themen sind im Folgenden unter Angabe einiger Schwerpunktfragestellungen dargestellt:

- Werdegang (Haupt- und nebenberufliche Tätigkeit; Wege zur politischen Bildung; Kontakt zum Programm; Einstieg ins Programm)
- Organisatorische Rahmenbedingungen (erster Einsatz von »Miteinander«; strukturelles Umfeld; Auftragsakquirierung)
- Zielgruppen (welche Zielgruppen werden erreicht; mögliche Unterschiede zwischen Gruppen; welche Beobachtungsweisen entwickeln sich bei Multiplikatoren)
- Praktische Durchführung (eigene Erfahrungen mit dem Programm; Materialeinsatz; Handhabung der Reflexionsphase; Rolle des Programms in der Schule)
- Rolle und Person (Selbstwahrnehmung im Seminar hinsichtlich Rolle und Person)
- Lernerfolg (Wahrnehmung und Interpretation von Lernerfolg)

Zusätzlich zu den Leitfadeninterviews wurden im Laufe des Evaluationsvorhabens weitere Gespräche geführt, die dem Zweck dienten, spezifische Fragestellungen, die während des Forschungsprozesses entstanden, näher zu beleuchten und gegebenenfalls zu klären. In einem Fall war dies ein Gespräch mit einem Mitglied der Adaptionierungsgruppe des Programms. Ziel war es hierbei, die Intentionen und konkreten Abläufe zu ergründen, die bei der Adaption im Vordergrund standen, und die gewonnenen Informationen zu den quantitativen und qualitativen Daten in Beziehung zu setzen.

Des Weiteren wurde im Rahmen des »Miteinander«-Jahrestreffens 2002 eine Gruppendiskussion durchgeführt, an der sich die Teilnehmer dieses Treffens auf freiwilliger Basis beteiligen konnten. Von besonderer Bedeutung war dieses Gruppengespräch deshalb, weil die Teilnehmer zum Großteil über sehr wenig bzw. keine Erfahrung im Umgang mit »Miteinander« verfügten und ihre speziellen Perspektiven und Problemstellungen hinsichtlich des

Programms zur Geltung bringen konnten. Dadurch konnte die zuvor formulierte Problematik des Übergewichts erfahrener Multiplikatoren im Interviewsample etwas korrigiert werden.

Eine weitere Informationsquelle stellten zudem zahlreiche Rückmeldungen vor allem unerfahrener »Multiplikatoren« und Teilnehmer einer Schulung dar, die aus eigenem Antrieb schriftlich oder telefonisch das Evaluationsteam kontaktierten, um eigene Anliegen, Wünsche oder Kritik vorzubringen.

Die ausgewählten Gesprächspartner waren auf Anfrage ausnahmslos zu einem leitfadengestützten Interview bereit. Die Durchführung der Interviews verlief in allen Fällen problemlos. Die Dauer der Interviews betrug zwischen 45 und 105 Minuten. Zur Wahrung der Anonymität der Intensivbefragten werden hier Detailergebnisse aus den Interviews nur in allgemeiner Form referiert.

Aufbau und Arbeit mit einer Forschungsbegleitgruppe

In Absprache mit den Auftraggebern war vereinbart worden, das Evaluationsdesign partizipativ und interaktiv zu gestalten. Generell lässt sich sagen, dass eine gegenstandsangemessene Evaluation von Demokratie- und Toleranzerziehungsprogrammen noch in den Anfängen steckt und Formen der Beteiligung von Praktikern erst nach und nach ausprobiert werden. »Partizipative Evaluation« wird in der Tradition der Aktionsforschung (Altrichter 1997; Bargal 1992) in der Regel mit mehreren oder allen der folgenden Charakteristika in Verbindung gebracht:

- Die untersuchten Personen sind an der Evaluation zu beteiligen.
- Die Auffassungen der untersuchten Personen im Hinblick auf die Zielbestimmung, die Durchführung und die Berichterstattung der Untersuchung sollen berücksichtigt werden.
- Die Evaluation selbst ist als bewusstseinsbildender interaktiver Prozess anzulegen.
- Das neu erworbene Reflexionswissen soll den Praktikern wieder bereitgestellt werden.
- Die Untersuchung hat die Verbesserung der beruflichen Möglichkeiten und Tätigkeiten der untersuchten Personengruppe zum Ziel.
- Die Evaluation kann gegebenenfalls Impulse für politisches Handeln und Veränderung geben.

Um diese Forderungen zu gewährleisten, wird im Rahmen der Aktionsfor-

schung üblicherweise auf eine Gruppe von Praktikern zurückgegriffen, die den Forschungsprozess begleiten. Ziel der Arbeit mit einer solchen »Forschungsbegleitgruppe« ist in der Regel die Bereitstellung relevanter Informationen im Anschluss an die praktischen Kenntnisse dieser Gruppe sowie die Gewährleistung des Teilnehmerbezugs während eines Forschungsvorhabens mit der Absicht, die gewonnenen Erkenntnisse in die pädagogische Praxis einzuspeisen.

Darüber hinaus bietet eine Begleitgruppe die Chance, Interpretationen zu diskutieren und gegebenenfalls zu korrigieren, zusätzliche Daten und Informationen zu erhalten und daran den laufenden Forschungsprozess zu kontrollieren. Bei der Einrichtung von solchen Begleitgruppen bieten sich zwei unterschiedliche Vorgehensweisen an:

- gleichberechtigte Beteiligung und Teilhabe an Entscheidungen betreffend die gesamte Durchführung der Evaluation verbunden mit einer (Mit-)Verantwortung der Begleitgruppe für die Ausgestaltung und Durchführung des Forschungsprozesses sowie die dabei erzielten Ergebnisse
- Begleitung des gesamten Forschungsprozesses durch Diskussion und Information, mit letzter Verantwortung für den Forschungsprozess und seine Ergebnisse ausschließlich bei den Evaluatoren

In diesem Zusammenhang wird in der einschlägigen Literatur wiederholt auf wiederkehrende Fragestellungen und typisch auftretende Problematiken hingewiesen, wie sie mit einem solchen Konzept verbunden sein können:

- In welchem Ausmaß ist eine Forschungsbegleitgruppe zu beteiligen? Ist dies permanent neu auszuhandeln oder zu Anfang des Prozesses definitiv festzulegen?
- Wer trägt die Verantwortung für die Ergebnisse?
- Welche Rolle spielen verschieden geartete Loyalitäten der teilnehmenden Gruppen (Evaluierter, Evaluatoren, Auftraggeber) gegenüber einem zu evaluierenden Programm, einer Trägerorganisation, der Evaluationseinrichtung etc.? Welche Rolle spielen hier finanzielle Subventionen oder Beschäftigungsverhältnisse?
- Wie ist mit vertraulichen Informationen oder der Verwertung von Zwischenergebnissen zu verfahren? Wer darf was bei welcher Gelegenheit verwenden?
- Wie lassen sich Arbeitsfähigkeit und Stabilität einer solchen Begleitgruppe angesichts von Zeitproblemen, Freiwilligkeit, Fluktuation von Mitarbeitern etc. gewährleisten?

Vor dem Hintergrund der Tradition der Aktionsforschung und einiger damit

verbundener und bereits im Verlauf der 80er Jahre in den Sozialwissenschaften ausführlich diskutierter methodologischer Probleme und angesichts der Erwartungen der Auftraggeber ist an dieser Stelle eine kurze Erläuterung zu dem Konzept der »partizipativen und interaktiven Evaluation« einzuschleiben.

Partizipation meint das Vorsehen von Möglichkeiten der Äußerung von Einschätzungen, Wahrnehmungen und Perspektiven durch die Evaluierten hinsichtlich aller Aspekte einer laufenden Evaluation. Interaktiv ist eine Evaluation, sofern die Evaluatoren ihr Vorgehen betreffend Datenerhebung und Auswertung sowie ihre Einschätzungen im Laufe des Prozesses und nach seinem Abschluss in Auseinandersetzung mit den Evaluierten erläutern und sich deren Kritik aussetzen.

Beides meint nicht, dass am Ende ein gemeinsam verantwortetes Ergebnis steht. Partizipative Evaluation operiert wie jede empirische Sozialwissenschaft unter der Bedingung der Reflexivität der Forschungsverhältnisse: Alle Individuen, die in Forschungsprozesse einbezogen werden, artikulieren in ihren Auskünften ihre Beobachtung der Art und Weise, in der sie durch die empirische Forschung beobachtet werden und die Erwartungen, die sie damit verbinden. Dies ist selbst ein auch für Evaluationen aufschlussreicher Sachverhalt.

Vor diesem Hintergrund ist von der Differenz der Perspektiven der an einer partizipativen Evaluation teilnehmenden Individuen (und Gruppen, sofern sie als solche agieren) auszugehen, die in dieser Form der Evaluation aneinander gebrochen werden können, aber nicht ineinander aufgehen. Dies ist nicht nur nicht zu erwarten, sondern wohl auch nicht wünschenswert, wenn man nämlich davon ausgeht, dass Reflexionsgewinn nicht aus Einheit, sondern aus Differenz und damit aus Variationen möglicher Beobachtungs- und Beschreibungsformen resultiert.

In der Evaluationsuntersuchung, über die hier berichtet wird, entschieden wir uns aufgrund der Anlage der Evaluation und der Zielgruppe der Multiplikatoren für eine Forschungsbegleitgruppe, die aus erfahrenen Multiplikatoren des Programms »Miteinander« bestehen sollte. Bei der Bildung dieser Gruppe wurde darauf geachtet, dass die Mitglieder unterschiedliche Berufsbilder und Zielgruppen des Programms präsentieren und zugleich regional möglichst weit gestreut residieren.

Beim ersten Treffen mit Multiplikatoren bekundeten neun Personen ihr Interesse an der Mitarbeit in einer Forschungsbegleitgruppe. Bald deutete sich jedoch an, dass sich die feste Kerngruppe aus weniger Personen zusammensetzen würde. Insgesamt fanden mit dieser Gruppe drei Treffen statt, wobei sich die Gruppe aus verschiedenen Gründen als wenig stabil erwies (z. B. aufgrund beruflicher oder familiärer Belastungen, fehlender Ver-

gütung und zu überbrückender großer regionaler Distanzen). Bei keinem der Treffen nahmen alle ursprünglichen Mitglieder teil.

Abschließende Bemerkung zur zukünftigen Arbeit mit Forschungsbegleitgruppen

Aus unserer Sicht und nach Auffassung der von uns befragten Multiplikatoren erwiesen sich die Anlage des Fragebogens und die Konzeption und Durchführung der Interviews als ertragreich und erlaubten uns genauere Einblicke in den Verlauf und die Resultate der Implementation des Programms »Miteinander – Erfahrungen mit Betzavta«, wie sie in dem Abschlussbericht und dem Beitrag in diesem Band genauer dargelegt sind.

Da die Arbeit mit einer Forschungsbegleitgruppe als ein innovatives und aber zugleich nicht einfach zu handhabendes Element der Evaluation betrachtet werden kann, möchten wir hier abschließend noch einmal näher auf die positiven und negativen Erfahrungen mit einer solchen Gruppe eingehen:

- Als gewinnbringend und ertragreich für den Forschungsprozess erwiesen sich die Kommentare zum Fragebogen, die gemeinsame Auswahl der Interviewpartner, die Gestaltung des Leitfadens für die Intensivbefragungen und die Diskussion des Zwischenberichts. Besonders hilfreich waren darüber hinaus die Zusatzinformationen zur Implementierung und Praxis des Programms und die kritische Reflexion des Begriffs »Wirkung«.
- Zusätzlich haben die Mitglieder der Gruppe die Evaluation unter den Multiplikatoren des Programms bekannt gemacht. Dies eröffnete den Zugang zu weiteren relevanten Informationen.
- Als problematisch erwies sich die sukzessive Schrumpfung der Gruppe mit negativen Auswirkungen auf den Informations- und Austauschprozess. Bei der Implementierung wurde darüber hinaus ein unterschiedliches Verständnis der Funktion und der verfolgten Zielsetzungen der Evaluation deutlich. Dies äußerte sich einerseits in einer offenen Kritik an der Transparenz des Forschungsdesigns und andererseits im Verlassen der Gruppe durch ein Mitglied.
- Aus Sicht der Evaluatoren spielten auch Loyalitäts- und Rollenkonflikte der Gruppenmitglieder bei den auftretenden Divergenzen eine gewisse Rolle.
- Der Zeitaufwand, der durch die Begleitgruppentreffen und den Versuch, den Austausch aufrechtzuerhalten, erforderlich war, ist durch das Evaluationsteam unterschätzt worden.
- Vor diesem Hintergrund ist kritisch anzumerken, dass es nicht immer gelungen ist, die Begleitgruppe in ausreichendem Maße über den For-

schungsverlauf zu informieren und in relevante Entscheidungen einzubeziehen.

Das Evaluationsteam hat einige der dargestellten Schwierigkeiten u. a. durch die Hinzuziehung von weiteren Experten auf diesem Gebiet zu kompensieren versucht. In einer Diskussion mit einem ausgewiesenen Experten für Evaluation und Organisationsentwicklung wurde deutlich, dass die aufgetretenen Probleme und insbesondere die erwähnte Fluktuation der Begleitgruppe eher als der »Normalfall« eines solchen Forschungsdesigns zu betrachten sind (s. auch Altrichter et al. 1997).

Eine Alternative wäre, sich statt auf eine feste Gruppe auf Informationen aus einem losen »Pool« von Praktikern mit unterschiedlicher Erfahrung zu stützen. Vor diesem Hintergrund haben wir uns um eine Vielzahl von Kommentaren bemüht und auf diese Weise im Laufe der Evaluation eine Reihe zusätzlicher Einsichten gewonnen.

Als Konsequenz aus diesen Erfahrungen schlagen wir für den Aufbau zukünftiger Begleitgruppen die Beachtung folgender Gesichtspunkte vor:

- Der Aufbau einer Gruppe sollte frühzeitig gewährleistet werden. Dabei sollten Verantwortungsstrukturen, Zuständigkeiten und Zielsetzungen klar formuliert sowie die meist durchaus differenten Motivationen zur Teilnahme an einer solchen Gruppe geklärt und offen gelegt werden.
- Bereits in der Frühphase des Forschungsprozesses sollte gewährleistet werden, dass in der Gruppe alle relevanten Forschungsschritte dargelegt und diskutiert werden, um die Evaluation möglichst transparent zu gestalten.
- Bei der Konzeption einer Forschungsbegleitgruppe ist zu überlegen, ob die Beteiligung einer regional konzentrierten Gruppe oder einer überregional gestreuten Gruppe für die Zielsetzungen der Forschung präferiert wird. Überregionale Streuung erhöht den Organisations-, Kosten- und Zeitaufwand, Regionalität schränkt die Reichweite der zugänglichen Perspektiven ein.
- Um einen gewissen Grad an Stabilität der Begleitgruppe zu gewährleisten, sind auch finanzielle Mittel (Vergütung) bereitzustellen.

Literatur

Altrichter, Herbert u.a. (1997): »PraktikerInnen als ForscherInnen, Forschung und Entwicklung durch Aktionsforschung«. In: Barbara Frie-

- bertshäuser und Annedore Prengel (Hrsg.): *Handbuch qualitative Forschungsmethoden in der Erziehungswissenschaft*. Weinheim und München. 640–660.
- Bargal, David, Martin Gold und Miriam Lewin (1992): »The Heritage of Kurt Lewin. Theory, Research, and Practice«. In: *Journal of Social Issues* (48) 2 1992.
- Creswell, John W. (1998): *Qualitative Inquiry and Research Design. Choosing Among Five Traditions*. London.
- Deutsche Gesellschaft für Evaluation e.V. (2002): *Standards für Evaluation*. Köln.
- Eser Davolio, Myriam (1999): *Prävention und Bekämpfung von fremdenfeindlichen, rassistischen und gewaltaffinen Einstellungen bei BerufsschülerInnen*. Edition Soziothek. Bern.
- Friebertshäuser, Barbara, und Annedore Prengel (Hrsg.) (1997): *Handbuch qualitative Forschungsmethoden in der Erziehungswissenschaft*. Weinheim und München.
- Kvale, Steinar (1996): *InterViews. An Introduction to Qualitative Research Interviewing*. London.
- Ulrich, Susanne, und Florian M. Wenzel (2003): *Partizipative Evaluation. Ein Konzept für die politische Bildung*. Gütersloh.

Selbstevaluation wertschätzend gestalten – methodisches Vorgehen in sechs Schritten

Florian M. Wenzel

Einleitung: Stellenwert von Selbstevaluation

Evaluationen, die von außen an Bildungsprogramme und -projekte herangebracht werden, stoßen häufig auf großen Widerstand. Strukturelle und organisatorische Kontexte machen es für Evaluatoren schwierig, Daten zu erheben. Wenn Rollenverteilungen, Hierarchieebenen und Beziehungen untereinander unklar oder konfliktgeladen sind, ist es schwer, Evaluation als Bewertung aus einer externen Perspektive durchzuführen. Auch ist häufig zu beobachten, dass keine klaren Zielkataloge oder Konzepte für die pädagogische Arbeit vorliegen und es somit schwer fassbar ist, was und nach welchen Kriterien evaluiert werden soll. Ergebnisse solcher Evaluationen sind deshalb innerhalb von Organisationen oft geradezu »umkämpft«, da sie nicht auf einer gemeinsamen Basis stehen, sondern im Resultat unterschiedlichen Zielvorstellungen zugeordnet werden.

Zunehmend wird klarer, dass Beteiligte und Betroffene von Evaluationen mit in die Planung und Durchführung von Evaluationsvorhaben einbezogen werden sollen. Vor allem vor dem Hintergrund mangelnden praktischen Nutzens und fehlender Konsequenzen vieler Evaluationen soll auf diesem Wege erreicht werden, dass Evaluation einen konstruktiveren Stellenwert erhält.

Eine Möglichkeit, die Verantwortung für die Erhebung und Durchführung von Evaluationen umfassend bei den Beteiligten und Betroffenen zu belassen, ist die Selbstevaluation. In diesem Verfahren untersuchen Projektverantwortliche ihr eigenes professionelles Handeln innerhalb der eigenen Organisation selbst. Sie geben Rechenschaft über ihr eigenes Tun und dokumentieren Potenziale der Optimierung.

Selbstevaluation wird aber nicht nur aus eigener Initiative ergriffen, sondern inzwischen systematisch für die Durchführung besonders von Modellprojekten gefordert – auch dies eine Reaktion auf die starke Abwehr gegenüber vielen Fremdevaluationen. Problematisch bleibt dabei die Tatsache, dass

auch im Falle von Selbstevaluation viele Unklarheiten bestehen: Was ist die Stellung von Selbstevaluation im Kontext meines Bildungsprojektes? Kann ich mich selbst bewerten und dies als gültige Evaluation ansehen? Was geschieht mit den Ergebnissen und welche Konsequenzen haben sie? Wie kann ich methodisch vorgehen?

Im Folgenden wird ein Vorgehen für Selbstevaluation vorgeschlagen, das auf diese Fragen antwortet. Ausgangspunkt ist dabei, Selbstevaluation mit »Appreciative Inquiry«, einem wertschätzenden Verfahren der Organisationsentwicklung, zu verknüpfen, das mit einer speziellen Interviewtechnik arbeitet. Selbstevaluation kann so zum Motor von Qualitätsverbesserung der eigenen Arbeit werden und diese zielgerichteter organisieren.

Selbstevaluation nimmt die Strukturen ernst, innerhalb derer sie arbeitet und organisiert einen gleichberechtigten Prozess, an dem alle relevanten Personen beteiligt sind. Gleichzeitig werden »en passant« Daten erhoben, die es ermöglichen, auch nach außen darzustellen, was innerhalb eines Bildungsprojektes geleistet wird.

Selbstevaluation ist also im besten Falle eine bewusste systemische Intervention, die sich sinnvoll in einen Bildungskontext einfügt und diesen weiterträgt. So hat Selbstevaluation immer den Charakter einer »formativen Evaluation«, die nicht wie eine »summative (bilanzierende) Evaluation« einen vergleichenden Schlussstrich zieht, sondern auf fortlaufende Optimierung ausgerichtet ist und selbst zum Motor eines Bildungsprojektes werden kann. Selbstevaluation hat damit einen doppelten Auftrag:

- Praxisveränderung: die eigene Praxis soll zielgerichtet verändert oder stabilisiert werden, um besser fachlich handeln zu können
- Erkenntnisgewinn: über die eigene Praxis sollen durch eine methodische Untersuchung Erkenntnisse gewonnen oder abgesichert werden

Ausgangspunkt für Selbstevaluation: Problemzentrierung vs. Wertschätzung

Die Grundhaltung gegenüber einem Evaluationsvorhaben ist entscheidend für dessen Erfolg. Evaluation wie auch Bildungsprojekte selbst richten häufig den Blick auf Defizite, die es zu beseitigen gilt. Große Teile des gesellschaftlichen Denkens gründen auf dem Defizitansatz, sehen auf das, was nicht funktioniert, was mangelhaft und zu kritisieren ist. Dementsprechend werden Negativbotschaften vermittelt und innovative positive Ansätze verdeckt.

Dies kann insbesondere in Programmen und Projekten, die sich ohnehin mit »Problemen« der Gesellschaft wie Intoleranz oder sozialer Benachteiligung

gung beschäftigen, zu einer Demotivationsspirale führen. Evaluation kann unter dieser Sichtweise als destruktive Intervention wirken.

»Appreciative Inquiry« ist eine Methode, die versucht, jenseits von problemorientiertem Denken die innovativen Ressourcen der Menschen als Grundlage jeglicher Interaktion miteinander zu sehen. Es wird mit den Ressourcen, Qualitäten und Motivationen gearbeitet, die in der Vergangenheit erfolgreich waren. Die Visionen und der Wille der Einzelnen werden als Ausgangsbasis für eine Weiterentwicklung einer Gruppe, Organisation oder eines Projektes genommen.

Philosophische Ursprünge finden sich in der Positiven Theologie (Matthew Fox: Betonung eines »Ursegens« anstelle einer »Ursünde«), Positiven Psychologie (in 30 Jahren psychologischer Forschung gab es 45 000 Studien zu Depression, aber nur 300 Studien über Glück und Freude), Positiven Abweichung (anstelle von Studien etwa über die Gründe von krankhafter Fehlernährung Fokussierung auf Beispiele erfolgreicher und innovativer Abweichung vom Krankheitsmuster) und dem Positiven Management (anstelle von zu erfüllenden Kompetenzkatalogen Stärkung der individuellen Fähigkeiten von Mitarbeitern).

Der Ansatz der Appreciative Inquiry entstand in den 80er Jahren in den USA und wurde hauptsächlich von David Cooperrider und Diana Whitney entwickelt (Cooperrider und Whitney 1999). Appreciative Inquiry ist eine Methode und Haltung, um in Organisationen Veränderungsprozesse anzustoßen und zu entwickeln. Wörtlich übersetzen könnte man »Appreciative Inquiry« mit »wertschätzender Untersuchung«. Dabei kann Inquiry als »Befragung« im Sinne von neugierigem Erforschen mit der Offenheit für neue Potenziale und Möglichkeiten bezeichnet werden, appreciative als »wertschätzend«, im Sinne eines wertgebundenen Aktes, der das Beste in den Menschen oder der Welt um uns betont.

Der Veränderungsprozess bezieht sich bei Appreciative Inquiry auf den »Organismus«, auf die menschlichen Aspekte und das soziale System der Organisation. Der Fokus wird dabei nicht auf die Verdeutlichung der Diskrepanz zwischen Soll und Ist gelegt, vielmehr werden die Potenziale der Menschen und der Organisation freigelegt und gefördert. Statt Veränderung im Glauben an Fakten und Daten zu verorten, werden die »weichen« Faktoren als entscheidend für Veränderung von Wirklichkeit angesehen.

Besonders für Evaluation ist dies ein sinnvoller Ansatz, da er schon im Ausgangspunkt von der Achtung gegenüber dem, was untersucht werden soll, getragen wird. In den neuen Qualitätsstandards für Selbstevaluation der Deutschen Gesellschaft für Evaluation wird dies unter dem Gesichtspunkt der Fairness ebenfalls betont: »Wertschätzend gestaltete Interaktion. Die Zu-

sammenarbeit in der Selbstevaluation soll so angelegt werden, dass die Kontakte zwischen den Beteiligten von gegenseitiger Achtung und Wertschätzung geprägt sind.« (Beywl und Müller-Kohlenberg 2003: 91)

Die Formulierung dieses Standards weist auf einen weiteren Aspekt hin, der sich mit Appreciative Inquiry verbindet: Im Mittelpunkt einer Selbstevaluation stehen Menschen, ihr Handeln und ihre Interaktionen und nicht zuerst Daten und Fakten oder statische Zustände. Dies wiederum schlägt eine Brücke zur Arbeit in Bildungsprogrammen, die sich nicht in der Vermittlung von Wissen erschöpfen, sondern prozessorientiert und immer mit Blick auf die Freiheit und Autonomie ihrer Zielgruppen arbeiten (Wenzel und Seberich 2002). Appreciative Inquiry ist damit ein angemessenes methodisches Verfahren, um Selbstevaluation wertschätzend zu gestalten. Wertschätzende Selbstevaluation arbeitet mit einer Reihe der für die Selbstevaluation modifizierten Grundprinzipien von Cooperrider und Whitney (1999), die verdeutlichen, dass der Schwerpunkt einer Selbstevaluation nach diesem Ansatz nicht darauf liegt, die »Wahrheit« über ein Programm zu eruieren und zu bewerten, was falsch und was richtig gemacht wurde, sondern zu analysieren, was hilfreich für ein Projekt ist und weiter gefördert werden kann.

Das konstruktivistische Prinzip

Projekte und Organisationen müssen bewusst als menschliche Konstruktionen wahrgenommen werden. Die Bilder, die über eigene Projekte und eigenes Handeln in Bildungskontexten geformt werden, bestimmen die eigene Wirklichkeit, das Fühlen und Handeln. Es gibt keine objektive Wahrheit, die evaluierbar wäre, sondern eine Vielzahl von Perspektiven, die für sich als sinnvolle Konstruktionen ernst zu nehmen sind. Wertschätzende Selbstevaluation analysiert, welche dieser Perspektiven hilfreich und unterstützend für die weitere Qualitätsverbesserung eines Bildungskontextes sein können.

Das Prinzip der Gleichzeitigkeit

Untersuchung, Befragung, Analyse und Planung sind nicht trennbar von tatsächlicher Veränderung, sondern geschehen gleichzeitig. Die Art der Perspektive, die Art und Weise, wie Fragen gestellt werden, bilden die Grundlage für den zukünftigen Umgang miteinander. Die Frage nach Defiziten und Problemen wird innerhalb eines Projektes sehr schnell zu einem »Frustdiskurs« führen, der sich verfestigt.

Wertschätzende Befragung stellt keine »neutralen« Fragen, um Wahrheiten zu evaluieren, sondern stellt mit ihrer Interviewtechnik selbst eine Intervention dar, die unterstützend und motivierend wirkt. Bereits mit der ersten Frage beginnt eine Handlungs- und Organisationsveränderung. Dieser systemisch-vernetzte oder auch »holografische« Ansatz steht im Gegensatz zu einer linear-zerlegenden Analyse, die erst Ursachen analysiert und dann Maßnahmen plant.

Das vorwegnehmende Prinzip

Das Bild, das Menschen von der zukünftigen Form ihrer Projekte und Organisationen haben, beeinflusst entscheidend, wie sie sich bereits in der Gegenwart in diesen Kontexten verhalten. Der Erwartungshorizont bringt die Zukunft in die Gegenwart. Jede künstlerische Gestaltung von positiven Metaphern für die Zukunft hat deshalb innovationsfördernde Wirkungen in der Gegenwart. Wertschätzende Selbstevaluation arbeitet mit Visionen, aus denen Zielformulierungen abgeleitet werden können, die nicht abstrakt zu erfüllen, sondern aus den eigenen Erfahrungen und Motivationen entstanden sind.

Das poetische Prinzip

Bildungskontexte bieten, offenen Büchern gleich, unendliche Interpretations- und Auslegungsmöglichkeiten. Zudem wird ihre Geschichte von den beteiligten Autoren immer weiter geschrieben. Wie ein gutes Gedicht sind zahlreiche Möglichkeiten der Inspiration vorhanden, es gibt nicht nur ein Thema, auf das man sich konzentrieren müsste.

In der wertschätzenden Befragung wird Wert gelegt auf »literarische« Formulierungen, auf die genaue Wahl der Worte. Packende Metaphern motivieren Menschen und bilden Ankerpunkte, um die herum sich Tun ereignen kann. Wertschätzende Selbstevaluation orientiert sich deshalb weniger an Daten und Fakten, als an den Bildern und Emotionen, die Menschen zu Handlungen bewegen.

Das positive Prinzip

Die Welt ist kein Problem, das gelöst werden muss. Der positive Antrieb, etwas gemeinsam zu gestalten, ist der stärkste Antrieb für dauerhafte Veränderung.

In der Selbstevaluation von Gruppen und Organisationen ist die positive Art des Fragens entscheidend für die tatsächliche weitere Veränderung in ihnen. Die Umsetzung von Evaluationsergebnissen wird so nicht zur vermeidenden Pflichtaufgabe, sondern entspringt dem eigenen gestalterischen Antrieb.

Die nachfolgende Gegenüberstellung verdeutlicht zentrale Unterschiede zwischen problemorientierter und wertschätzender Herangehensweise an Evaluation.

Tabelle 1: Problemzentrierte vs. wertschätzende Evaluation

Problemzentrierung		wertschätzende Orientierung
	Ausgangspunkt	
Mangel		Funktionsfähigkeit
Defizit		Wertschätzung
Problem		Motivation
<i>nachsorgend</i>		<i>zukunftsorientiert</i>
	Herangehensweise	
Ursachen suchen		Motivationen kennen lernen
Lösungen finden		Visionen ausdrücken
Aktionen planen		Verwirklichen
<i>Schwächen abschwächen</i>		<i>Stärken bestärken</i>
	Weltbild	
linear-zerlegend		systemisch-vernetzt
wahr vs. falsch		hilfreich vs. nicht hilfreich
<i>tatsachenorientiert</i>		<i>personenorientiert</i>

Methodisches Vorgehen

Im Folgenden wird beschrieben, wie der Ansatz der Selbstevaluation nach der Appreciative Inquiry methodisch umgesetzt werden kann. In den USA wird »Appreciative Evaluation« bereits auf einigen Kongressen thematisiert.

Es existieren größere Projektvorhaben und Erfahrungsberichte sowie die erste umfassendere Publikation zu diesem Thema.¹

In Deutschland gibt es noch wenig konkrete Erfahrungen, auch wenn Evaluationsvorhaben sich zunehmend der notwendigen Achtung und Wertschätzung gegenüber ihren Klienten bewusst sind.

Wertschätzende Evaluation lässt sich in sechs Teilschritte gliedern, die nachfolgend ausführlich beschrieben werden. Zu den einzelnen Schritten werden zur Verdeutlichung Beispiele angeführt, die mit diesem Ansatz arbeiteten.²

Das methodische Vorgehen einer wertschätzenden Selbstevaluation umfasst sechs Schritte:

1. Einbeziehung der Beteiligten und Betroffenen und Definition des Evaluationsthemas: Wer soll den Prozess der wertschätzenden Selbstevaluation tragen und durchführen? Was soll konkret untersucht werden?
2. Erhebung vorhandener Stärken und Ressourcen: Was funktioniert bereits innerhalb unseres Projektes? Was motiviert uns im professionellen Handeln? Welche Indikatoren und Erfolgsfaktoren haben wir für unsere Arbeit?
3. Entwerfen gemeinsamer Visionen: Wie können wir uns eine umfassende Verwirklichung von Erfolg in unserem Projekt vorstellen? Wie sieht die optimale Zukunft unserer professionellen Arbeit aus?
4. Formulierung von Zielen: Was wollen wir konkret erreichen? Welche Leit- und Teilziele haben wir, die fassbar und umsetzbar sind?
5. Planung, Durchführung und Bewertung von Interventionen: Mit welchen Mitteln erreichen wir unsere Ziele? Wie können wir deren Erreichung untersuchen?
6. Erstellung eines Evaluationsberichtes: Was haben wir aus dem Evaluationsprozess gelernt? Was ist der größte Nutzen der Untersuchung für unsere Zukunft? Wie können wir das für andere darstellen?

Es ist sinnvoll, einen solchen Prozess durch einen externen Moderator begleiten zu lassen. Dieser hat nicht die Funktion eines Evaluators, sondern eines

1 Für die Planung und einen Erfahrungsbericht vgl. »The Synapse Group« (1997) und Mohr, Smith und Watkins (1999). Ein umfassendes Werk haben Mohr und Jacobsgaard (2003) vorgelegt. Fettermann (2001) betont ebenfalls die Notwendigkeit, innerhalb von Evaluation die Beteiligten zu ermächtigen, ihre eigenen Belange in die Hand zu nehmen, ohne dabei näher auf Appreciative Inquiry einzugehen.

2 Diese sind ausführlicher in den entsprechenden Dokumentationen dargestellt, die in der Bibliographie aufgeführt sind. Die Manuskripte sind beim Autor erhältlich.

Evaluationsberaters, der eine Selbstevaluation initiiert und das methodische Vorgehen in seinen Teilschritten steuert, um Rollenkonflikte zu vermeiden. Er gewährt einen sicheren Rahmen, innerhalb dessen die Selbstevaluation stattfinden kann.

Zusätzlich steht er während des Prozesses beratend zur Verfügung, um zu klären, ob bestimmte Fragestellungen auftauchen, für deren Evaluation eine Auslagerung an einen externen Evaluator sinnvoll ist. In Selbstevaluationen nehmen die Beteiligten ihre Evaluation also selbst in die Hand, gerade um auch selbst entscheiden zu können, wann sie Unterstützung von außen benötigen.

1. Schritt: Einbeziehung der Beteiligten und Betroffenen und Definition des Evaluationsthemas

In diesem ersten Schritt werden diejenigen zusammengebracht, die ihr eigenes professionelles Handeln untersuchen möchten. Dies kann ein Koordinationsteam oder ein Kollegium sein, eine Fachabteilung, ein Pool von Trainern. Wichtig ist, dass diese Personen ihr eigenes professionelles Handeln thematisieren und beleuchten und nicht das anderer Menschen. Durch ein vorbereitendes Team werden sie eingeladen, an der Selbstevaluation mitzuwirken. Zusammen mit dem Evaluationsberater sollte genau überlegt werden, wer die Beteiligten und Betroffenen sind und wie sie in die Selbstevaluation mit einbezogen werden können.

Gerade der Beginn einer Selbstevaluation entscheidet über Akzeptanz für oder Widerstand gegen ein solches Vorhaben und erfordert genaue Planung. Vorbereitend wird auch das Thema der Evaluation definiert, das in seiner Formulierung besondere Relevanz für die entsprechende Gruppe haben muss. Da es sich bei der Selbstevaluation um eine formative, also qualitätsverbessernde Evaluation handelt, sollte die Formulierung des Evaluationszieles auch diesem Optimierungsgedanken folgen.

Zudem muss es in seiner Formulierung ergebnisoffen sein und sollte auf einer Ebene liegen, die mit Mitteln der Selbstevaluation gut zu bewältigen ist. Evaluationen tendieren im Allgemeinen dazu, in ihrer Themenbestimmung Vorgaben zu treffen, die mit bestehenden Mitteln kaum realisierbar sind. Gerade Selbstevaluationen sollten sich ein klar umrissenes und konkretes Thema setzen, das gut in die eigene Arbeit integrierbar und mit den vorhandenen Ressourcen zu bewältigen ist.

Beispiel

Im Bezirk Schwaben in Bayern fand durch die Stiftungsfachhochschule Benediktbeuern eine Erhebung statt, die analysierte, in welchen Institutionen politische Bildung angeboten wird und welcher Bedarf vorhanden ist. Aus dem Rücklauf dieser Analyse wurde deutlich, dass sehr umfassend der Wunsch nach Vernetzung und Kommunikation im Bezirk gegeben war. Deshalb wurden die an der Erhebung beteiligten Institutionen und Personen zu einem Präsenzworkshop eingeladen, in dem sie einerseits ihr eigenes Handeln systematisch selbst evaluierten und zusammenführten und zum anderen gemeinsam Ziele für die weitere Arbeit formulierten.

Das Vorbereitungsteam des Workshops arbeitete in diesem Fall mit Metaphern, um ein dieser Selbstevaluation angemessenes Thema und Titel zu finden, das Zustand und Optimierung von politischer Bildung in Schwaben benennen sollte und Lust machen wollte auf den Workshop. Dies war essenziell, da die Fremd- und Selbsteinschätzung zu Umfang und Stellenwert politischer Bildung häufig negativ und problemzentriert ist. Als Titel wurde schließlich gewählt: »Wegen Renovierung geöffnet – Baustelle Politische Bildung in Schwaben« (Wenzel und Boeser 2003a).

2. Schritt: Erhebung vorhandener Stärken und Ressourcen

In diesem Schritt wird erarbeitet, wann sich die Beteiligten der wertschätzenden Selbstevaluation in ihrem Arbeitszusammenhang lebendig und engagiert gefühlt haben, wann sie in ihrem professionellen Handeln motiviert waren und viel Energie verspürten. In »wertschätzenden Interviews«, die die Basis für diese Form der Selbstevaluation bilden, befragen sich die Beteiligten in Paaren.

Dies geschieht optimalerweise in einem Präsenzworkshop, um eine möglichst hohe Personenzentrierung und Interaktion zu ermöglichen. Wenn dies nicht möglich ist, können vorbereitete Interviewfragen auch versandt oder in kleineren lokalen Teams beantwortet werden. Dies mindert und bremst allerdings die Dynamik des Selbstevaluationsprozesses (siehe hierzu Wenzel 2001 und Grieger 2001).

Die Interviews werden anhand des Evaluationsthemas vom Evaluationsberater zusammen mit einem Vorbereitungsteam entsprechend der Methodik der Appreciative Inquiry entwickelt. Der besondere Fokus während dieser Interviews wird dabei auf die Wahrnehmung des Positiven gelegt. Die Interviewfragen bieten zahlreiche unterschiedliche Ankerpunkte, um eine lineare

Datenabfrage zu verhindern und die Interaktion während der Interviewsituation zu fördern.

Die Fragenden versetzen sich in die Lage eines interessierten Forschers, um bei den Befragten Geschichten und Erlebnisse aus der Vergangenheit wieder neu zu beleben und zu entdecken und ihre Visionen zu erfahren, die sich daraus ableiten. Die Fragen enthalten wertschätzende, bejahende und bestärkende Elemente. Persönliche Erfolge und Stärken kommen zur Sprache, Situationen und Ereignisse im Arbeitszusammenhang werden in einem anderen Licht gesehen.

Diese konkreten Geschichten erzählen von der erfolgreichen Gestaltung der Arbeit und vom Engagement, das die Beteiligten in ihrem professionellen Handeln gezeigt haben. Sie spiegeln energiegeladene Momente und Momente der Freude wider. Es wird erkundet, verstanden und wertgeschätzt, welcher konstruktive und positive Kern bereits in der eigenen Arbeit vorhanden ist.

Die besten Geschichten, Metaphern und konkreten Erlebnisse werden anschließend in Kleingruppen vorgestellt und gemeinsam zu Kriterien verdichtet, die darstellen, was als Erfolgsfaktoren innerhalb des entsprechenden Bildungsprojektes gelten kann. Dabei stellen jeweils die Interviewer ihre Ergebnisse den anderen vor. Durch diese wechselseitige Perspektivübernahme wird eine Gesprächskultur ermöglicht, die schon zu Beginn der Selbstevaluation einlädt, die jeweiligen Sinnkonstruktionen der anderen empathisch nachzuvollziehen.

In der Formulierung der Kriterien wird auf Sprache besonders geachtet. Packende Formulierungen und ansprechende Bilder können den gemeinsamen Anker für weitere Veränderungen bilden.

Darauf aufbauend wird erkundet, was zu den Erfolgsfaktoren geführt hat, was sie ermöglicht hat. Rahmenbedingungen und Zusammenhänge innerhalb des Systems, Handlungen und Interventionen, die dies möglich machten und gefördert haben, werden sichtbar gemacht. Die eigenen Ressourcen und die Vergangenheit der Beteiligten werden dabei als eine Quelle gesehen, aus der in Zukunft geschöpft werden kann.

Beispiel

Im Rahmen der ersten Fachtagung zur Vernetzung von ca. 200 XENOS-Projekten (www.xenos-de.de) gegen Rechtsextremismus in Deutschland wurde thematisiert, dass bisher unklar sei, was die vielen unterschiedlichen Projekte innerhalb des Rahmenprogramms bezüglich ihrer Erfolgsfaktoren und Ziele zusammenführt. Dies wurde in einer modellhaften Selbstevaluation im

Rahmen eines Workshops mit folgendem wertschätzenden Interview erarbeitet, das die Anlage und Grundstruktur dieses Selbstevaluationsschrittes verdeutlicht (Wenzel 2003).

Das ideale XENOS-Projekt

Es tut sich etwas. Die seit einiger Zeit laufenden und neuen Projekte des Bundesprogramms »XENOS – Leben und Arbeiten in Vielfalt« haben deutschlandweit zu einer erhöhten und auch umfassenderen Wahrnehmung der Themen Demokratie und Toleranz beigetragen. Als Beteiligte eines solchen Projektes tragen wir alle mit unserer Arbeit dazu bei. Unsere persönliche wie professionelle Herangehensweise ist sehr unterschiedlich, und wir haben verschiedene Perspektiven und Visionen hinsichtlich unseres Projektes. Dies wollen wir heute betrachten, vernetzen und gemeinsam weiter tragen.

Bitte befragen Sie Ihr Gegenüber zu den positiven Erfahrungen innerhalb ihres/seines XENOS-Projektes, den Höhepunkten und Erlebnissen, die sich dort ereignet haben. Stellen Sie sich vor, Sie besuchen eines der Projekte, die sehr erfolgreich laufen. Fragen Sie dabei nicht wie ein Analytiker, der Fakten sammelt. Sie wollen vielmehr die besten Geschichten erfahren. Achten Sie auf spannende Formulierungen und Metaphern und halten Sie diese fest. Sie sind neugierig auf die positiven Erfahrungen und Motivationen Ihres Gegenübers! Lassen Sie Ihrem Interviewpartner Zeit, in sich zu gehen, und unterstützen Sie ihn durch hilfreiches Nachfragen.

- Erzählen Sie mir bitte, wie Sie mit dem XENOS-Programm in Kontakt gekommen sind. Was hat Sie motiviert, ein Projekt zu starten bzw. sich daran zu beteiligen, was fanden Sie daran attraktiv? Was ist es, was Sie noch heute in Ihrem Projekt engagiert sein lässt?
- Können Sie sich an ein Erlebnis oder eine Situation erinnern, die ein echter Höhepunkt in Ihrem Projekt war? Etwas, wo Sie sich richtig lebendig und wohl gefühlt haben und wussten, so soll es sein? Was ist dort geschehen? Wer war daran beteiligt? Welche Rolle haben Sie und andere Menschen gespielt?
- Was sind Ihrer Meinung nach Faktoren, die einem XENOS-Projekt Lebendigkeit und Stärke geben? Was sind wichtige Einflussfaktoren?
- Jetzt seien Sie mal nicht so bescheiden! Erzählen Sie mir, was Sie an sich selbst gut finden. Welche Talente oder Fähigkeiten haben Sie, was können Sie besonders gut? Was davon könnten Sie in Ihr

Projekt künftig einbringen? Worauf hätten Sie so richtig Lust? Was würden Sie mit viel Energie anpacken?

- Stellen Sie sich nun vor, zehn Jahre sind vergangen. Ihre kühnsten Träume sind Wirklichkeit geworden. Die Ziele und Visionen Ihres XENOS-Projekts, die an einzelnen Stellen aufblitzten, sind nun umfassend vorhanden. Welches Bild taucht vor Ihrem inneren Auge auf? Wodurch lässt sich Ihr erfolgreiches Projekt nun symbolisieren?

Dieser Schritt der Selbstevaluation verdeutlicht, wie bereits von Anfang an die Beteiligten und Betroffenen ihre Evaluation selbst in die Hand nehmen. Durch die Art der Fragestellung und die gemeinsame Vernetzung findet oft zum ersten Mal eine systematische Erfassung von Kriterien und Erfolgsfaktoren für ein Bildungsprojekt statt. Dieser Schritt ist die notwendige Basis, um später umsetzbare Ziele für eine Veränderung der Praxis entwerfen zu können.

Die gegenseitige Befragung mit wertschätzenden Interviews unterscheidet sich grundlegend von einem Vorgehen, das in den Sozialwissenschaften und zahlreichen Fremdevaluationen Verwendung findet: Durch umfassende Fragenkataloge und zahlreiche »Prüffragen« soll indirekt die »Wahrheit« über einen bestimmten Sachverhalt erhoben werden.

Auch narrative Interviews versuchen mit einem sehr offenen Vorgehen, aus umfangreichen Erzählungen der Interviewten später interpretierte Aussagen zu erstellen. Im Gegensatz dazu ist beim wertschätzenden Interview der Sinnzusammenhang von Fragen und Datenerhebung unmittelbar erkennbar. Die gesammelten Aussagen werden transparent für alle verdichtet und zusammengeführt.

Nicht die Objektivität der Daten, sondern der Nutzen dieser Intervention steht bei wertschätzender Selbstevaluation im Vordergrund. Durch die Integration der verschiedenen Perspektiven zu Kriterien und Erfolgsfaktoren wird eine gemeinsame und für alle tragfähige Basis geschaffen. Dies ist ein Prozess, in dem zunächst nicht Sachverhalte oder gegebene Daten im Mittelpunkt stehen, sondern die befragten Personen und ihre konkreten und handlungsleitenden Erlebnisse und Fähigkeiten. Diese sind es, die als Motor der Selbstevaluation für Motivation sorgen und schon hier einen Bogen zur weiteren Umsetzung der Ergebnisse bilden können.

3. Schritt: Entwerfen gemeinsamer Visionen

Auf dem Fundament der Vergangenheit, die sich in den gehörten Geschichten widerspiegelt, werden in Kleingruppen Zukunftsvisionen entworfen, die in ihrer Formulierung als Gegenwart bestimmt sind. Was könnte sein? Wie sollen sich der Arbeitskontext und seine Bildungsangebote entwickeln? Welche Träume und Wünsche haben die Beteiligten für die Zukunft ihres Projektes? Diese Visionen gründen in den Geschichten und Motivationen der Vergangenheit und bleiben somit mit der Realität verwurzelt. Sie sind quasi Verlängerungen der Erfolgsfaktoren.

Die Visionen, die als Gegenwart formuliert sind, können sich auf sofortige Veränderungen oder auf die nächsten fünf bis 25 Jahre beziehen. Das Visionieren geschieht als kreativer und ästhetischer Prozess, in dem unterschiedliche Methoden wie z. B. Malen, Collagen, Theater, Zeitungsartikel oder Brief zum Einsatz kommen. Während im konventionellen Ansatz Visionen oft durch die Problemorientierung unterdrückt werden bzw. sich im Utopischen verlieren, wird bei der Appreciative Inquiry ganz bewusst auf kreative und ästhetische Elemente gesetzt, die Phantasie fördern und dadurch auch neue und manchmal ungewohnte Wege aufzeigen können, die gangbar sein können.

Die Formulierung von Visionen ist ein Schritt auf dem Weg zur Formulierung von Zielkriterien, die konkret und realistisch sind. Im Gegensatz dazu stehen Utopien. Der Unterschied von Vision und Utopie ist entscheidend, um den Stellenwert dieses Schrittes richtig einzuschätzen:

Vision	Utopie
– verwurzelt in Erfahrungen der Vergangenheit	– Verlängerung der Gegenwart
– unabhängig vom »Geworden-Sein« der Dinge	– »Nirgendwo-Land«
– in Wertschätzung verhaftet	– von Problemen ausgehend
– Möglichkeiten: reformatorisch	– Umsturz: revolutionär
– Handlungsrelevanz für die Gegenwart	– Wirklichkeitsverlust der Gegenwart

Visionen bieten immer die Möglichkeit der weiteren Anknüpfung an vorhandene Ressourcen und Wirklichkeiten. Evaluation dient hier nicht dazu, den Beschluss zu unternehmen, etwas völlig anderes zu versuchen und mit Bestehendem völlig zu brechen. Diese »utopische Annahme« steckt hinter zahlreichen Befürchtungen gegenüber Evaluationen, die der Legitimation zur Ab-

schaffung oder Umstrukturierung von Bildungsprojekten dienen, ohne dass auf vorhandene Ressourcen aufgebaut wird.

4. Schritt: Formulierung von Zielen

Die Selbstevaluationsschritte zwei und drei dienen der Verankerung und Zusammenführung dessen, was bereits existiert und dessen, was möglich sein könnte. Auf dieser Basis können nun Ziele formuliert werden. In Bildungsprojekten existieren häufig keine ausgesprochenen Ziele oder sehr allgemeine Ziele wie »Qualifizierung fördern und Beschäftigung schaffen«, die nicht operationalisierbar sind und eine Vielzahl von Möglichkeiten der Intervention offen lassen. Selbstevaluation ist auch hier eine Intervention, die zur Präzisierung eines Bildungskontextes beiträgt.

In den Zielen wird formuliert, wie die gemeinsame Zukunft aussehen soll, welche Möglichkeiten der Weiterentwicklung des Arbeitskontextes gegeben sind. Dabei kann es sich z. B. um Leitlinien der Organisation oder ihrer Abteilungen ebenso handeln wie um einen Katalog dazu, wie sich die Organisation versteht. Die Formulierungen sollen dabei das bislang Verwirklichte herausfordern, provokativ sein, gleichzeitig aber auf dem Boden der Tatsachen bleiben, also mit Beispielen aus der Vergangenheit belegen, dass sie möglich sind.

Die Aussagen sollen bejahend, ausdrucksstark, konkret, motivierend und in der Gegenwartsform formuliert sein, damit sie greifbar sind und einen erreichten Status darstellen. Durch die provokative Formulierung der Vorschläge wird zum einen der Status quo der Organisation herausgefordert, zum anderen werden alte Verhaltensweisen und Muster in Frage gestellt und neue, kreative Wege erwogen.

Die Formulierung von Zielen, die diesen Anforderungen entsprechen, ist eine umfassende und intensive Arbeit. Für die Formulierung muss viel Zeit und Arbeit investiert werden. Als Überprüfung kann im Rahmen von Selbstevaluationen der SMART-Check (Beywl und Schepp-Winter 1999: 63) angewandt werden. Danach sollen Ziele spezifisch, messbar, angemessen, realistisch und terminiert sein, um einerseits davon abgeleitete Interventionen auf der Handlungsebene gestalten und zum anderen im Rahmen von Selbstevaluation erheben zu können, ob sie erreicht wurden. Gut formulierte Ziele stellen eine handlungsleitende Orientierung dar, die zur Optimierung eines bestehenden Bildungsprojektes anregt.

Beispiel

Im Rahmen der Selbstevaluation und geplanten Optimierung eines Tutoriensystems im Diplomstudiengang Pädagogik der Universität Augsburg wurden zu diesem Schritt ausführliche Leitziele sowie ihnen zugeordnete Teilziele erstellt, die zum Teil in den bereits bestehenden Angeboten und Erfahrungen mit den Tutorien gründen und sich zum Teil aus gemeinsamen Visionspräsentationen entwickelten. Beispielhaft sind hier zwei Leitziele und die dazugehörigen Teilziele dargestellt (Wenzel und Boeser 2003). Sie stellen eine systematische Basis dar, auf deren Grundlage entsprechende Interventionen geplant wurden und nachfolgend in ihrer Erreichung bewertet werden konnten.

Leitziel 1: Es gibt einen »Raum« für Tutoren

- Wir haben an der Fakultät ein Zimmer
 - Wenn es offen ist, ist es für alle zugänglich.
 - Es gibt eine regelmäßige Sprechstunde.
 - Uns stehen vielfältige Materialien zur Verfügung.
- Wir haben einen angemessenen Arbeitsraum für die Tutorien
 - Licht und Fenster
 - Grundausstattung wie Beamer, Video, Overhead
- Flexibilität als Freiraum
 - Es besteht die Möglichkeit, auch nach außen zu gehen oder Teilnehmer nach Hause einzuladen.
 - Es gibt ein Treffen einmal pro Monat außerhalb der Uni (alle drei Monate mit der Fachschaft).
 - Einmal pro Monat findet ein runder Tisch mit Dozenten statt.

Leitziel 2: Training und Support sind umfassend vorhanden

- Es gibt Trainingsangebote für Tutoren von Tutoren/Dozenten zu:
 - gruppendynamischen Prozessen (z. B. Kontaktspiele, TZI etc.)
 - Kommunikation (z. B. nonverbale Kommunikation, Rhetorik etc.)
 - Präsentationstechniken
 - Lerntechniken (z. B. Umgang mit Texten, Mindmapping etc.)
 - neuen Methoden
- Wir recherchieren auch selber außerhalb der Universität entsprechende Angebote und Referenten für die verschiedenen Themen und bringen diese ein.

5. Schritt: Planung, Durchführung und Bewertung von Interventionen

Auf Basis der formulierten Ziele werden nun Interventionen und Maßnahmen geplant. Diese können im Rahmen einer wertschätzenden Selbstevaluation zwei unterschiedliche Richtungen annehmen:

- Interventionen und Maßnahmen, um die formulierten Ziele zu erreichen: In diesem Falle hat Selbstevaluation vor allem die in der Einleitung beschriebene Aufgabe, fachliche Veränderung herbeizuführen. Aufgrund des in der Wertschätzung und den Ressourcen der Beteiligten verankerten Prozesses wird eine hohe Bereitschaft entstehen, Optimierungsmöglichkeiten zu nutzen.
- Ausdehnung des Evaluationsprozesses: Innerhalb dieses Schrittes ist es möglich, für Leit- und Teilziele, zu denen bereits Interventionen und Maßnahmen bestehen, den Selbstevaluationsprozess zu erweitern. Hier dient die Selbstevaluation dazu, weitere Daten zu erheben, um darstellen zu können, ob die jeweiligen Ziele erreicht wurden. Gemeinsam mit dem Selbstevaluationsteam werden für ein oder mehrere Teilziele entsprechend spezifische Evaluationsthemen festgelegt, der Prozess beginnt wieder bei Schritt 1.

Mit den Mitteln der wertschätzenden Befragung ist es möglich, auf effiziente Art und Weise eine Vollerhebung für eine bestimmte Zielgruppe durchzuführen, indem Interviewer geschult werden und relativ schnell Ergebnisse der Interviews zusammentragen und verdichten. Beispielsweise hat The Synapse Group auf diese Art und Weise mit 60 Interviewern 480 wertschätzende Interviews durchgeführt, deren Ergebnisse in einem anschließenden Workshop zu Indikatoren, Erfolgsfaktoren und Empfehlungen verdichtet wurden.

Beispiel

Im Rahmen des XENOS-Workshops, dessen Interviewleitfaden weiter oben dargestellt ist, erarbeitete eine Teilgruppe aufgrund der gemeinsamen Zielformulierungen modellhaft ein spezifisches wertschätzendes Interview für ein Projekt für benachteiligte junge Frauen, um ihren eigenen Evaluationsprozess auszudehnen und zu überprüfen, ob von ihnen formulierte Ziele erreicht wurden. Vor der eigentlichen Befragung, die von der Zielgruppe selbst zu zweit und dann in Vernetzung stattfinden soll, wird ein biografischer Einstieg gewählt, um die jungen Frauen für den anschließenden Schritt

zu motivieren. Nach den Interviews sollen Themen, Erfolgsfaktoren und Visionen analog der Selbstevaluationschritte 2-4 gebündelt werden.

Fragen nach der Vorgehensweise wertschätzender Interviews zum Thema »Der Frauencomputerkurs und meine Zukunft«
Wie hast du von dem Kurs erfahren? Warum nimmst du gerade an diesem Kurs teil? Was hast du bis heute gemacht?
Welche Situation war für dich ein echter Höhepunkt im Kurs? Etwas, bei dem du dich so richtig wohl gefühlt hast? Was ist geschehen, wer war dabei?
Was kannst du richtig gut? Was ist deine persönliche Stärke? Was macht dir viel Spaß? Was davon könntest du in einen Beruf einbringen?
Weißt du (nun), was du beruflich machen willst? Was hat dir der Umgang mit EDV gebracht? Wie war das Lernen mit Frauen?
Stell dir vor, ein Jahr ist vergangen. Was machst du? Viele deiner Wünsche sind in Erfüllung gegangen. Wie sieht dein Alltag aus?

Das Beispiel verdeutlicht, wie im Rahmen einer Selbstevaluation die Beteiligten ihre Untersuchung in die Hand nehmen. In der Ausdehnung des Vorhabens gestalten sie aufgrund der eigenen Erfahrungen selbst wertschätzende Fragen und können unterschiedliche Zielgruppen dazu anregen, ebenfalls ihr eigenes Tun und Handeln zu betrachten und zu optimieren. Selbstevaluation wird auch hier als eine passende bildungsbegleitende Intervention gestaltet, die Lerncharakter für die entsprechende Zielgruppe hat.

6. Schritt: Erstellung eines Evaluationsberichtes

Der gesamte wertschätzende Selbstevaluationsprozess sollte möglichst ausführlich dokumentiert werden. Abschließend wird von einem Kernteam der Beteiligten und Betroffenen ein Evaluationsbericht in schriftlicher Form erstellt. Dies sollte möglichst im Rahmen eines Präsenzworkshops gemeinsam geschehen – so kann dieser Bericht systematisierter Ausgangspunkt für die weitere Steuerung des Bildungsprojektes insgesamt werden.

Hauptziel eines Evaluationsberichtes ist es, die Nutzung, den Nutzen und damit die Umsetzung der Evaluationsergebnisse zu fördern. Im Rahmen einer wertschätzenden Selbstevaluation ist ein Bericht keine objektive Darstellung einer wissenschaftlichen Wahrheit, sondern das Produkt der Vernet-

zung unterschiedlicher Perspektiven auf eigenes professionelles Handeln und der Darstellung von Möglichkeiten der Optimierung.

Im besten Falle ist ein Evaluationsbericht ein gut nutzbares Handbuch, das selbst wieder dazu beiträgt, ein Bildungsprojekt in seiner Qualität zu fördern. Entsprechend didaktisch ansprechbar sollte er gestaltet sein. In den Standards für Selbstevaluation wird dies wie folgt definiert: »Vollständigkeit und Klarheit der Berichterstattung. Selbstevaluationsberichte sollen den Gegenstand einschließlich seines Kontextes ebenso wie die Ziele, die Fragestellungen, die Verfahren und Befunde der Evaluation beschreiben, damit die wesentlichen Informationen zur Verfügung stehen, leicht verstanden werden und allen Beteiligten und Betroffenen zugänglich sind.« (Beywl und Müller-Kohlenberg 2003: 90)

Selbstevaluation ist immer formative Evaluation. Um die Ergebnisse der wertschätzenden Selbstevaluation wieder zu thematisieren, sollte ein konkreter Termin vereinbart werden, an dem aus dem Abstand von einigen Monaten mit Hilfe des Berichtes der Prozess insgesamt nochmals reflektiert wird. Häufig wird erst mit diesem zeitlich distanzierten Blick deutlich, welche umfangreiche Arbeit geleistet wurde.

Schlussbetrachtung: Wertschätzende Selbstevaluation und lernende Organisationen

Mit den methodischen Schritten einer wertschätzenden Selbstevaluation wurde deutlich, wie sich Evaluation und Organisationsentwicklung verbinden. Der Stellenwert des Evaluators verschiebt sich vom Experten der Sozialwissenschaften zum Moderator und Berater, der einen eigenständigen Prozess begleitet. Strukturen und Systeme, innerhalb derer evaluiert wird, werden auf gleichberechtigte Art und Weise mit einbezogen. Wenn sich ein integriertes Verständnis von Evaluation durchsetzt, so gewinnt der Begriff der »lernenden Organisation« seine eigentliche Bedeutung. Eine lernende Organisation würde Evaluation von Anfang an als Bestandteil ihrer Zielbestimmung sehen, der ebenso spannend und tragend ist wie die inhaltliche Arbeit.

Literatur

Beywl, Wolfgang, und Ellen Schepp-Winter (1999): »Zielfindung und Zielklärung«. In: Bundesministerium für Familie, Senioren, Frauen und Jugend

- (Hrsg.): *QS Nr. 21. Materialien zur Qualitätssicherung in der Kinder- und Jugendhilfe*. Berlin.
- Beywl, Wolfgang, und Hildegard Müller-Kohlenberg (2003): »Standards der Selbstevaluation«. In: *Zeitschrift für Evaluation* (2) 1 2003. 65–75.
- Zur Bonsen, Matthias, und Carole Maleh (2001): *Appreciative Inquiry (AI). Der Weg zu Spitzenleistungen*. Weinheim und Basel.
- Bundesvereinigung Kulturelle Jugendbildung (1998): »Leitfaden für Selbstevaluation und Qualitätssicherung«. In: Bundesministerium für Familie, Senioren, Frauen und Jugend (Hrsg.): *QS Nr. 19. Materialien zur Qualitätssicherung in der Kinder- und Jugendhilfe*. Berlin.
- Cooperrider, David L., und Diana Whitney (1999): *Appreciative Inquiry*. San Francisco.
- Fetterman, David M. (2001): *Foundations of Empowerment Evaluation*. London.
- Grieger, Gunnar (2001): *Appreciative Inquiry. Wertschätzende Organisationsentwicklung*. Paderborn 2001 (als E-Book erhältlich unter www.active-books.de).
- Heiner, Maja (Hrsg.) (1988): *Selbstevaluation in der sozialen Arbeit*. Freiburg i. Br.
- Mohr, Bernhard und Mette Jacobsgaard (2003): *Appreciative Evaluation*. Taos, N.M.
- Spiegel, Hiltrud von (1993): *Aus Erfahrung lernen. Qualifizierung durch Selbstevaluation*. Münster.
- Ulrich, Susanne, und Florian M. Wenzel (2003): *Partizipative Evaluation. Ein Konzept für die politische Bildung*. Gütersloh.
- Wenzel, Florian M. (2001): *Erinnern, Erzählen, Darstellen. Evaluation und Dokumentation des Projektes ESKA (Europäische SeniorenKulturArbeit)*. München.
- Wenzel, Florian M., und Michael Seberich (2002): »Handeln reflektieren – Demokratie erleben«. In: Gotthard Breit und Siegfried Schiele (Hrsg.): *Demokratie Lernen als Aufgabe der politischen Bildung*. Bonn.
- Wenzel, Florian M., und Christian Boeser (2003a): »Wegen Renovierung geöffnet. Baustelle politische Bildung in Schwaben. Dokumentation des Bezirks und des Bezirksjugendringes Schwaben«. Manuskript. Augsburg.
- Wenzel, Florian M., und Christian Boeser (2003b): »Intensivseminar für Tutoren. Das ideale Tutorensystem im Studiengang Diplom-Pädagogik der Universität Augsburg. Dokumentation der Ergebnisse, Prozesse, Methoden«. Manuskript. Augsburg.
- Wenzel, Florian M. (2003): »Dokumentation des Facharbeitskreises Partizipative Evaluation«. XENOS-Fachtagung »Vielfalt und Chancengleichheit am Arbeitsmarkt und in der Gesellschaft«. Manuskript. Freising.

Die Autoren

Michael Bommers, Prof. Dr. phil., Professor für Soziologie/Methodologie interkultureller und interdisziplinärer Migrationsforschung am Institut für Migrationsforschung und interkulturelle Studien (IMIS) der Universität Osnabrück. Durchführung und Leitung deutscher und internationaler Forschungsprojekte im Bereich Migration. Autor und Herausgeber von Büchern und Beiträgen zu den Themenbereichen und Problemfeldern Migration – Nationalstaat – Wohlfahrtsstaat, Folgen der Migration in Kommunen, Inklusionschancen ausländischer Jugendlicher in Erziehung und Ausbildung, methodologische Probleme der Migrationsforschung.

Oliver Dimbath, Dr. rer. pol., Studium der Soziologie, Sozialpsychologie und Sozial- und Wirtschaftsgeschichte in München. Mitarbeiter am Lehrstuhl für Soziologie an der Universität Augsburg. Arbeitsgebiete: Soziologie der Arbeit und der Berufe, Kirchen- und Religionssoziologie, Jugendsoziologie, Methoden qualitativer Sozialforschung, Evaluation und Evaluationsforschung.

Stephan Eschler, Dipl. Sozialpädagoge/Sozialarbeiter (FH); pädagogischer Mitarbeiter der Europäischen Jugendbildungs- und Jugendbegegnungsstätte Weimar im Bereich politisch-soziale Bildung; Projektleiter »Schule und Ausbildung für Toleranz und Demokratie«; Trainer für Demokratie- und Toleranzerziehungsprogramme.

Jochen Guter, Student der Soziologie, Anglistik und Ethnologie an der Albert-Ludwigs-Universität Freiburg. Projektassistent bei der Evaluation des Programms »Miteinander – Erfahrungen mit Betzavta« 2001 bis 2002. Zur Zeit Mitarbeit in einem Projekt des Instituts für Migrationsforschung und interkulturelle Studien (IMIS), Osnabrück, über die zweite Generation von Ausländern in verschiedenen europäischen Staaten.

Julia von Hayek, Dipl.-Soz., Studium der Soziologie, Volkswirtschaft und Psychologie in München. Lehrbeauftragte am Institut für Soziologie der Ludwig-Maximilians-Universität München. Arbeitsgebiete: Körpersoziologie, Thanatosoziologie, Evaluation und Evaluationsforschung.

Andreas Hirseland, Dr. rer. pol., Mitarbeiter am Institut für Soziologie der Ludwig-Maximilians-Universität München. Arbeitsschwerpunkte: Wissens- und Kultursoziologie, Diskurstheorie und -analyse, Methoden qualitativer Sozialforschung, Biografienforschung, Geld- und Konsumsoziologie, modernes Alltagsleben, Kommunikationstheorie, Evaluationsforschung, Evaluator in den Bereichen Jugendhilfe und politische Bildung.

Jochen Kehr, Studium der evangelischen Theologie und Psychologie (Diplom) in Köln und Bonn. Trainer für Evaluation, Bildungscontrolling und körperorientierte Verhaltens- und Kommunikationstrainings, freiberuflich als Evaluator und Fortbildungsreferent bei Univation, Institut für Evaluation Dr. Beywl & Associates GmbH in Köln. Schwerpunkte: Evaluation von Bildungsmaßnahmen sowie die proaktive und klärende Evaluation.

Heinz Lynen von Berg, Dr. phil, von 1992 bis 1999 wissenschaftlicher Mitarbeiter und Lehrbeauftragter am Institut für Sozialwissenschaften der TU Berlin; von 1999 bis 2002 Geschäftsführer von »Miteinander – Netzwerk für Demokratie und Weltoffenheit in Sachsen-Anhalt e.V.«; seit 2002 wissenschaftliche Begleitforschung des CIVITAS-Programms am Institut für interdisziplinäre Konflikt- und Gewaltforschung (IKG) der Universität Bielefeld. Arbeitsschwerpunkte: Evaluationsforschung, Zivilgesellschaft, Rechtsextremismus und minderheitenfeindliche Gewalt, politische Kommunikation und Diskursanalyse.

Erich Schäfer, Dipl.-Soz., Dr. phil.; Professor für Methoden der Erwachsenenbildung am Fachbereich Sozialwesen der Fachhochschule Jena; Forschungsgebiete: Medienpädagogik, wissenschaftliche Weiterbildung, Methoden der Jugend-, Erwachsenen- und Weiterbildung.

Werner Schneider, Dr. phil., Professor für Soziologie an der Universität Augsburg. Arbeitsgebiete: Familiensoziologie, Soziologie der Lebensalter, Kultur- und Wissenssoziologie, Thanatosoziologie, Postmoderne/Poststrukturalismus sowie diskursanalytische Perspektiven und qualitative Sozialforschung, Evaluationsforschung.

Katrin Uhl, M.A., Studium der Politikwissenschaft, Japanologie und interkulturellen Kommunikation in Tübingen, Seattle, Kyoto und Fulda. Von 2000 bis 2003 Projektmanagerin der Bertelsmann Stiftung im Themenfeld Demokratie und Bürgergesellschaft. Arbeitsschwerpunkte: Demokratie- und Toleranzerziehung, Konfliktprävention, interkulturelle Kommunikation/interkulturelle Pädagogik, Zivilcourage, Migration und Integration, Evaluation politischer Bildung.

Sandra Uhl, Dipl. Sozialpädagogin/Sozialarbeiterin (FH), seit 2003 wissenschaftliche Mitarbeiterin in der Evaluation des Projekts »Schule und Ausbildung für Toleranz und Demokratie«.

Susanne Ulrich, Dipl.-Verw. Wiss., Studium der Verwaltungswissenschaften in Konstanz; 1987 bis 1990 Referentin für Internationale Jugendbegegnungen und entwicklungspolitische Bildungsarbeit beim Kolpingwerk. 1992 bis 1995 freiberufliche Seminar- und Kulturarbeit; seit 1995 wissenschaftliche Mitarbeiterin des Projektes »Erziehung zu Demokratie und Toleranz« am Centrum für angewandte Politikforschung München. Schwerpunkte: Entwicklung, bzw. Adaption und Implementierung der Programme: Betzavta – Miteinander, Eine Welt der Vielfalt, »Achtung (+) Toleranz«, Power of Language. 2003 Gründung der »Akademie Führung und Kompetenz am Centrum für angewandte Politikforschung«.

Florian M. Wenzel, M.A., Studium der politischen Theorie, Philosophie und Pädagogik in Edmonton (Kanada), Essex (Großbritannien) und München. Von 1999 bis 2003 wissenschaftlicher Mitarbeiter des Projektes »Erziehung zu Demokratie und Toleranz« am Centrum für angewandte Politikforschung München. 2003 Gründung der »Akademie Führung und Kompetenz am Centrum für angewandte Politikforschung«. Prozessbegleiter für Open Space und Appreciative Inquiry.

Ulrike Wolff-Jontofsohn, Dr. phil., Erziehungswissenschaftlerin, seit 2003 Leiterin der Abteilung wissenschaftliche Weiterbildung der Pädagogischen Hochschule Freiburg in der Akademie für wissenschaftliche Weiterbildung von Universität und PH Freiburg, Expertin im BLK-Projekt »Demokratie lernen und leben«. Arbeitsschwerpunkte: internationale Konzepte von Menschenrechtspädagogik, Demokratie- und Toleranzerziehung, Interkulturelle Bildung, Erziehungs- und Bildungswesen in Israel, Schulentwicklung im internationalen Vergleich, Evaluation von politischer Bildung, Weiterbildungsmanagement.