

Gerd Meyer, Ulrich Dovermann,  
Siegfried Frech, Günther Gugel (Hrsg.)

# Zivilcourage lernen

Analysen – Modelle – Arbeitshilfen

Bundeszentrale für politische Bildung, Bonn 2004  
ISBN 3-89331-537-3  
[www.bpb.de](http://www.bpb.de)

Buchhandelsausgabe:  
Institut für Friedenspädagogik Tübingen e. V.  
Tübingen 2004  
ISBN 3-932444-13-2  
[www.friedenspaedagogik.de](http://www.friedenspaedagogik.de)

## EVALUATIONSMETHODEN FÜR MODELLE ZUR FÖRDERUNG VON ZIVILCOURAGE IN DER POLITISCHEN BILDUNG

### EVALUATION – BEDROHUNG ODER CHANCE?

Evaluation hat sich in den letzten Jahren zu einem Modethema entwickelt. Trotz inflationären Gebrauchs ist der Begriff der Evaluation nicht eindeutig definiert. Im Zentrum steht die Erfassung von Daten über ein bestimmtes Evaluationsobjekt und die anschließende Bewertung anhand von Kriterien. Das übergeordnete Ziel ist dabei die Optimierung von Handlungen in komplexen Situationen (vgl. Kinast 1998, 23). Auch im Bereich der Bildungsarbeit werden zunehmend Evaluationen eingefordert. Hier rücken neben Rahmen- und Ausgangsbedingungen zunehmend auch die Beobachtung von Prozessen und die Frage nach Ergebnissen von Bildungsmaßnahmen ins Zentrum des Interesses.

Mittlerweile gibt es eine große Bandbreite von Programmen, Seminaren und Trainingsmaßnahmen, die bei ihren Teilnehmern Einstellungs- und Verhaltensänderungen bewirken wollen, etwa im Sinne toleranten Verhaltens oder friedlicher Konfliktlösung. Angesichts dieser Angebotsvielfalt kommt die systematische Analyse und Bewertung pädagogischer beziehungsweise andragogischer Interventionen im Alltag der politischen Bildung häufig zu kurz. Manche Projekte und Programme evaluieren ihre Arbeit gar nicht, andere beschränken sich auf eine Abfrage der Teilnehmer oder auf rein beschreibende Dokumentationen. **Ob und weshalb politische Bildung erfolgreich ist oder scheitert, steht jedoch nur selten im Zentrum von längerfristig angelegten Untersuchungen.**

Ein Grund hierfür liegt in Vorbehalten und Widerständen von Trägern und in der Praxis Tätigen gegenüber umfassenden Evaluationen. Oft existieren diffuse Ängste, nicht zuletzt vor dem Verlust von Fördermitteln oder Stellen. Bewertungsdruck oder die Angst vor Veränderung, vor Unterbrechung der eigentlichen Arbeit oder vor der Möglichkeit, eine Evaluation könne unangenehme Punkte zutage bringen, befördern ein negativ besetztes Verständnis von Evaluation. Aber auch weniger emotionale Gründe führen zu Widerständen: Häufig sind etwa



die personellen und finanziellen Ressourcen für eine umfassende Untersuchung nicht vorhanden. Mit der Notwendigkeit zur Evaluation konfrontiert, stehen die Betroffenen somit oft vor einer Herausforderung, die ihnen zugleich bedrohlich und unbekannt erscheint. Evaluation wird so leicht zur Irritation.

Evaluationen von Bildungsmaßnahmen können unterschiedliche Funktionen einnehmen. Insbesondere die von Evaluation betroffenen Praktiker der Bildungsarbeit assoziieren mit ihr in erster Linie häufig eine unerwünschte, fach- und praxisferne **Fremdkontrolle** ihrer Arbeit. Denn meistens erfolgt eine Evaluation auf Wunsch von Geldgebern, die Auskunft über den effizienten und zielführenden Einsatz ihrer Mittel erhalten wollen. Evaluation findet zu diesem Zweck in der Regel in Form einer so genannten bilanzierenden oder **summativen Evaluation** am Ende einer Maßnahme statt. Seltener thematisiert, aber nicht weniger bedeutsam und der konkreten Arbeit vor Ort förderlich, ist die Evaluation mit dem Ziel der **Selbstkontrolle**, die auf dem Interesse von Trägern, Maßnahmeleitern, Trai-



nen oder Programmdesignern beruht, mehr über die Reichweite, Praxisnähe und Wirkung der eigenen Bildungsprogramme zu erfahren. Da hier die Qualität und Weiterentwicklung der Maßnahme im Mittelpunkt steht, erfolgt Evaluation begleitend und **formativ**. Evaluationsergebnisse sollen hier zur kontinuierlichen Verbesserung und Weiterentwicklung der Arbeit herangezogen werden. Evaluation als Mittel zur **Legitimationsbeschaffung** schließlich, ist insbesondere dann von Interesse, wenn es um die bessere Darstellbarkeit der Wirkung von Projekten oder Programmen nach außen beziehungsweise um die Weiterbewilligung oder Akquisition von Fördermitteln geht.

**Ein Verzicht auf Evaluation, aus welchem Grund auch immer, bedeutet somit Verzicht auf Kontrolle über die Erfolge von Programmen und Projekten.** Evaluation mit dem Ziel der Selbstkontrolle – der Vergewisserung über die Ergebnisse der eigenen Arbeit nach innen – oder der Legitimationsbeschaffung – als Mittel der Demonstration der Ergebnisse nach außen – birgt in sich weniger Bedrohliches als Nützliches für Programmentwickler,

Träger und Trainer. Der Versuch, mehr über die Wirkung von Maßnahmen herauszufinden und diese zu bewerten, kann produktive Lernmöglichkeiten eröffnen und einen zentralen Beitrag zur Qualitätssicherung und -steigerung leisten, nicht nur für die Programme und Projekte selber, sondern auch für die durchführenden Organisationen und deren Mitarbeiter.

## WAS KANN EVALUATION LEISTEN?

Im Bereich der politischen Bildung können Evaluationen Antworten auf unterschiedliche Fragen geben. Die Bandbreite reicht von Aussagen über Teilnehmerprofile, dem Einholen von Feedback zu Veranstaltungen, der Bewertung von Moderatoren oder Lehrmaterialien bis zu Evaluationen der Wirkung von Trainingsprogrammen und Seminaren. Am Beginn der Beschäftigung mit Evaluation muss daher zunächst geklärt werden, welche Ziele die Untersuchung verfolgen soll, auf welche Fragen Antworten gefunden werden sollen.

**Grundsätzlich lassen sich folgende Bereiche unterscheiden, über die eine Evaluation Aufschluss geben kann:**

- ▶ Programmentwicklung (Didaktik/Zielsetzung);
- ▶ Durchführung (Qualität);
- ▶ Wirkung (Outcome);
- ▶ Effizienz (Kosten/Nutzen).

Diese Aspekte lassen sich einzeln oder im Rahmen breiter angelegter Untersuchungen umfassender untersuchen. Bei den Beteiligten kann eine sehr heterogene Interessenlage in Bezug darauf herrschen, welche Aspekte im Zentrum einer Evaluation stehen sollten. Für Geldgeber steht insbesondere die Untersuchung der Effizienz des Mitteleinsatzes im Vordergrund. Programmträger möchten daneben auch Aussagen zur Qualität der Durchführung erhalten. Ebenso wird für diese die Wirkung von Maßnahmen von Belang sein – ein Bereich, der insbesondere auch im Erkenntnisinteresse der Programmentwickler steht. Aussagen zu Wirkungen sind zudem für Trainer von Interesse, wobei für diese die von ihnen geleiteten Einzelveranstaltungen wahrscheinlich größere Bedeutung haben als eine Evaluation des Gesamtprogramms. Schließlich werden auch die jeweiligen Teilnehmer an Erkenntnissen über die von ihnen besuchte Veranstaltung inter-

essiert sein. Für potenzielle Interessenten können umfassende Aussagen über Gesamtprogramme zu einem Kriterium bei der Auswahl von Bildungsangeboten werden.

#### PROGRAMM- UND MAßNAHMENEVALUATION

**Zwei Ebenen der Evaluation politischer Bildung sind vor diesem Hintergrund denkbar: Evaluation von Programmen und Evaluation einzelner Maßnahmen.**

**Programmevaluationen** verfolgen das Ziel, Aussagen über Programme der politischen Bildung zu machen, was sowohl die Aspekte Programmentwicklung, Durchführung und insbesondere Wirkung sowie die Effizienz einschließen kann. Im Zusammenhang mit Wirkungsevaluation erweist sich die Erfassung von Einstellungs- und Verhaltensveränderungen häufig als hochkomplexer Vorgang. Obwohl verallgemeinerbare Schlussfolgerungen auch bei groß angelegten Programmevaluationen vielfach nur begrenzt möglich sind, können durchaus Wirkungen von Programmen identifiziert werden. Mit einer Kombination von Breiten- und Tiefenanalyse lassen sich umfassende Ergebnisse erzielen, die sowohl summativ die Erfolge eines Programms belegen, wie auch formativ zu dessen Qualitätsentwicklung beitragen können. Evaluation dieser Art ist ressourcenintensiv, sowohl was Kosten, Zeit aber auch die erforderliche Fachkompetenz betrifft. Häufig werden Programmevaluationen deshalb als **externe Evaluation** in Auftrag gegeben und nicht intern von Mitarbeitern der jeweiligen Organisation durchgeführt. Externe Evaluatoren verfügen eher über die nötige Objektivität gegenüber dem Evaluationsobjekt, jedoch fehlt ihnen zumeist der direkte Einblick in die Praxis. Zur Erarbeitung der zentralen Evaluationsfragen sollte deshalb gewährleistet sein, dass der externe Evaluator eng mit den Programmverantwortlichen zusammenarbeitet.

**Maßnahmenevaluationen** nehmen einzelne Veranstaltungen in den Blick, sind weniger ressourcenintensiv und können häufig als **interne Evaluation** beziehungsweise **Selbstevaluation** von den Durchführenden selbst (den Organisatoren oder Trainern) entwickelt und angewendet werden. Damit werden sie vor allem den Informationsbedürfnissen derer gerecht, die in erster Linie an Erkenntnissen über einzelne Programmbestandteile interessiert sind.

Dabei kann das praxisnahe Detailwissen, über das der interne Evaluator verfügt, bei der Entwicklung gegenstandsadäquater Evaluationsinstrumente von Vorteil sein. Unabdingbar sind jedoch auch Kenntnisse über Methoden und Instrumente, um zu gewährleisten, dass aussagekräftige Daten generiert werden. Es kann daher sinnvoll sein, in Kooperation mit Evaluationsexperten kleinere Evaluationsverfahren beziehungsweise -instrumente zu entwickeln und diese dann selber anzuwenden.

**Die Ziele einer Evaluation hängen eng mit dem Erkenntnisinteresse und den verfügbaren Ressourcen derer zusammen, die den Anstoß zu ihr geben.** Bevor die Frage beantwortet werden kann, was denn untersucht werden soll, muss zunächst geklärt sein, wer die Antwort auf diese Frage gibt beziehungsweise geben darf. Von dieser Entscheidung wird der Verlauf einer Evaluation maßgeblich beeinflusst. Werden Evaluationen allein aus einer ganz bestimmten Perspektive betrieben, etwa aus der des Geldgebers oder Trägers, und andere Beteiligte einer Bildungsmaßnahme nicht in die Definition der Evaluationsziele einbezogen, wird die Evaluation nur bedingt umfassende Aussagen machen können. Gerade im Bereich des sozialen Lernens ist es von Vorteil, im Sinne eines anzustrebenden multiperspektivischen Ansatzes schon in die Formulierung der Ausgangsfragen möglichst viele Beteiligte einzubeziehen. Nicht nur werden auf diese Weise blinde Flecken im Untersuchungsansatz vermieden, sondern darüber hinaus auch die Identifikation aller mit der Evaluation erhöht und eine hohe Transparenz des Vorgehens und der Ergebnisse ermöglicht. Fragestellung, Teilergebnisse und Endergebnisse werden nach einem **partizipativen Verständnis von Evaluation** kontinuierlich von den Evaluatoren mit möglichst vielen Beteiligtengruppen – den so genannten Stakeholdern – diskutiert und validiert (vgl. Ulrich/Wenzel 2003). Evaluation, die in dieser Weise Betroffene als Akteure einbezieht, zeigt sich sensibel für den Kontext des sozialen Lernens, in dem es keine einfachen Antworten auf die Frage nach dem Erfolg einer Maßnahme gibt. Eine so angelegte Evaluation versteht sich als integraler Bestandteil von Projekten politischer Bildung, bekommt dadurch einen neuen Stellenwert und kann für alle Beteiligten eine produktive Lernerfahrung sein.

Sowohl Programmevaluationen wie auch Maßnahmeevaluationen, externe wie interne Evaluationen, haben ihre Vor- und Nachteile. Oft werden ganz pragmatische Gründe für die eine oder andere Form der Evaluation sprechen (vgl. Thomas 2000, 8 f.). Zur Illustration werden im Folgenden (1) kurz unterschiedliche Evaluationsdesigns externer Evaluationen dargestellt, (2) grundlegende Probleme der Evaluation (nicht nur) von Trainingsprogrammen für Zivilcourage diskutiert und (3) einige Möglichkeiten zur Selbstevaluation von Zivilcourage-Trainings vorgestellt. Damit sollen Anregungen gegeben und praxisnahe Instrumente vermittelt werden, die im Seminar- und Trainingsalltag angewendet werden können. Eine umfassende Programmevaluation können diese allerdings nicht ersetzen. Sie stellen aber sinnvolle und anwenderfreundliche Optionen für Träger, Projektkoordinatoren und Trainer dar, Informationen über ihre Seminare und Trainings zu erhalten.

#### EVALUATION DER WIRKUNG VON PROGRAMMEN FÜR DEMOKRATIE- UND TOLERANZERZIEHUNG

Einen Eindruck davon, was Programmevaluationen leisten können, vermitteln Studien, welche die Bertelsmann Stiftung in den Jahren 2001 und 2002 bei externen Evaluatoren in Auftrag gegeben hat. Diese haben die in Kooperation mit dem Centrum für angewandte Politikforschung (C·A·P) an der Universität München entwickelten beziehungsweise für den Einsatz in Deutschland adaptierten Demokratie- und Toleranzerziehungsprogramme **„Achtung (+) Toleranz“**, **„Eine Welt der Vielfalt“** und **„Miteinander – Erfahrungen mit Betzavta“** untersucht. Diese Bildungsprogramme zielen darauf ab, Menschen für den Umgang mit Vielfalt zu sensibilisieren und zu gewaltfreiem Konflikt- und Kommunikationsverhalten zu befähigen. Im Mittelpunkt der Evaluation standen die Aspekte Programmakzeptanz, Lernerfolg – insbesondere die Wirkung der Programme auf Einstellungen und Verhalten –, Praxisbezug sowie die Nachhaltigkeit von Wirkungen. Auf der Grundlage eines partizipativen Verständnisses von Evaluation, das eine größtmögliche Einbeziehung aller Beteiligtegruppen anstrebt, wurden jeweils unterschiedliche Evaluationsansätze gewählt:

#### Evaluation mit dem Fokus auf erwachsenen Teilnehmern („Achtung (+) Toleranz“)

In einer methodisch vielfältigen Mehrebenenevaluation wurden mittels einer Dokumentenanalyse von Programmmaterialien und explorativen Expertengesprächen die im- und expliziten Vorannahmen der Programmkonzeption rekonstruiert. Mittels einer Umfrage in Form eines teilstandardisierten Fragebogens an ehemalige Seminarteilnehmer wurde die Breitenwirkung des Seminars untersucht. Bei einer daran anschließenden qualitativen Einzelfallstudie einer ausgewählten Seminargruppe wurde die Tiefenwirkung des Programms erforscht. Dazu führte man drei Monate vor und nach dem Seminar problemzentrierte Interviews. Innovativ war insbesondere die Durchführung eines videoprotokollierten Gruppenexperiments und dessen Analyse hinsichtlich Gruppendynamik und Konfliktverhalten. Aufbauend auf einem diskursiv-dialogischen Ansatz wurden theoretische Annahmen und Ergebnisse der Evaluation kontinuierlich mit den Programmwicklern diskutiert.

#### Evaluation mit dem Fokus auf jugendlichen Teilnehmern („Eine Welt der Vielfalt“)

Anhand ausgesuchter Schulklassen ging die Evaluationsstudie der Frage nach, in welchem Ausmaß der Einsatz des Programms in Schulen zur Entwicklung beziehungsweise Stabilisierung gewünschter Einstellungen, Absichten und Handlungen bei den Schülerinnen und Schülern führt. In einer ersten Erhebungsphase wurden mittels leitfadengestützter Telefoninterviews mit Trainern und Lehrern erwartete und beobachtete Wirkungen erhoben. Auf dieser Grundlage erfolgten in einer zweiten Erhebungsphase qualitative Fallstudien mit zwei zehnten Gymnasialklassen sowie einer Abschlussklasse einer Fachschule für Sozialpädagogik. Mit altersgerechten Erhebungsmethoden wurde die von den Schülern wahrgenommene Wirkung des Programms untersucht. Im Sinne eines partizipativen Ansatzes wurde die gesamte Untersuchung durch eine aus Vertretern aller Beteiligtegruppen gebildete Steuerungsgruppe begleitet.

#### Evaluation mit dem Fokus auf den Trainern („Miteinander“)

Hier stand die Wirkungswahrnehmung durch die Trainer als Programmmultiplikatoren im Mittelpunkt, die in mehrfacher Hinsicht Erfahrungen mit

dem Programm haben: sowohl zunächst als Teilnehmer und weiterhin als Experten in Bezug auf den Einsatz des Programms. In Kombination von quantitativer Fragebogenerhebung bei einem großen Sample ausgebildeter Trainer und leitfadengestützten Intensivinterviews mit ausgewählten Trainern wurde die Wirkung des Programms in zwei Bereichen erfasst: (1) im Arbeits- beziehungsweise gesellschaftlichen Umfeld der Trainer selbst sowie (2) anhand von durch die Trainer vorgenommenen Einschätzungen der Wirkungen auf die Teilnehmer. Begleitet wurde diese Evaluation durch eine Forschungsbegleitgruppe von Trainern.

Diese umfassenden Studien resultierten in Aussagen zu Reichweite, Akzeptanz und Praxisnähe der jeweiligen Programme sowie insbesondere darüber, inwiefern Teilnehmende ihre Einstellungen und ihr Verhalten – bewusst oder unbewusst – verändern und das vermittelte Wissen beziehungsweise die erworbenen Kompetenzen in ihrem Alltag umsetzen können. Empfehlungen zur Weiterentwicklung von Programmdesign, -organisation und -durchführung leiteten Überlegungen zu Qualitätssicherung und -entwicklung ein. Solche breit angelegten, ressourcenintensiven Evaluationen sind im Alltag der politischen Bildung sicherlich eher die Ausnahme. Der Zugewinn an Erkenntnissen und Möglichkeiten der Programmentwicklung lassen den Aufwand jedoch als gerechtfertigt erscheinen. Auch wenn externe Evaluationen nicht realisierbar sind beziehungsweise man sich selbst in der Lage befindet, eine Evaluation durchführen zu müssen, ohne selbst Experte auf diesem Gebiet zu sein, gibt es eine Reihe von Punkten, deren Beachtung die Qualität und Aussagekraft eigener Evaluationsaktivitäten erhöhen können.

## **EVALUATION VON TRAININGSMABNAHMEN FÜR ZIVILCOURAGE**

### **WAS IST ERFOLG?**

Schickt eine Firma ihre Mitarbeiter auf eine Softwareschulung, so liegt ein einfaches und vor allem objektives Kriterium zur Erfolgsfeststellung auf der Hand: nämlich ob und in welchem Umfang die Besucher des Kurses anschließend in der Lage sind, die entsprechende Software zu bedienen. Die Maßnahme wäre in diesem Sinne erfolgreich, wenn

überhaupt Veränderungen in der gewünschten Richtung – Beherrschung der neuen Software – feststellbar sind. In einem zweiten Schritt müsste der Auftraggeber festlegen, ob die erzielten Erfolge ihrem Umfang nach den eigenen Erwartungen entsprechen (wozu unter Umständen Vergleichsdaten vorliegen beziehungsweise erhoben werden müssten). Was jedoch, wenn der Inhalt einer Bildungsmaßnahme die Entwicklung und Förderung von Zivilcourage ist? Verglichen mit dem Erlernen eines Computerprogramms lässt sich behaupten, dass die Fähigkeit zur Zivilcourage schwieriger bestimmbar ist. Wie lässt sich angesichts dessen die Wirkung, der Erfolg eines Zivilcourage-Trainings bestimmen?

Aus Sicht der Geldgeber, die auf Abhilfe gegenüber zunehmenden Übergriffen sinnen und aus diesem Grunde Programme wie etwa Trainingsprogramme für Zivilcourage ins Leben rufen oder unterstützen, könnte deren „Erfolg“ analog zum Beispiel der Softwareschulung daran gemessen werden, unerwünschte Erscheinungen abzustellen oder in „nennenswertem“ Umfang abzumildern und so die ergriffenen Schritte gegenüber der Öffentlichkeit als „erfolgreich“ zu legitimieren. Anstelle des Bewertungskriteriums „Abnahme von Übergriffen in der Öffentlichkeit“ könnte sich der Erfolg von Trainingsmaßnahmen auch darin zeigen, dass die Teilnehmer gegenüber Nichtteilnehmern ein höheres Maß an Eingreifen in spezifisch-kritischen Situationen zeigen. Vorausgesetzt, es wäre möglich, derartiges überhaupt (wissenschaftlich exakt) zu erfassen und somit das Ausmaß der erwünschten Wirkungen festzustellen, könnte man dann uneingeschränkt den Erfolg des Programms behaupten?

Die Frage müsste mit „Jein“ beantwortet werden. Sicherlich stellt eine Zunahme zivilcouragierten Verhaltens ein „letztes Ziel“ der Programme und Maßnahmen dar. Aber selbst wenn dies empirisch feststellbar wäre – was methodisch kaum einzulösen ist –, so bliebe dennoch beispielsweise ungeklärt, was die entsprechenden zivilcouragiert Handelnden ohne Teilnahme an derartigen Trainings getan hätten und inwieweit das Ergebnis als Wirkung des Programms selbst oder anderer, programmexterner Bedingungen gelten muss. **Zudem käme es bei der Bewertung der gewünschten Befunde in hohem Maße darauf an, wer von diesen Trainings erreicht wurde und teilgenommen hat – die ohnehin Handlungsbe-**

**reiten und tendenziell „Mutigen“?** In diesem Fall wäre die Erfüllung des Kriteriums zumindest auf den ersten Blick nicht verwunderlich, im entgegengesetzten Fall von Teilnehmern, die aus diversen Gründen auf einem Standpunkt des „ohne mich“ stehen, eher weniger wahrscheinlich.

Bereits von daher sind, was den Bereich von Zivilcourage-Trainings betrifft, eindimensionale Festlegungen „externer“ Wirkungskriterien problematisch hinsichtlich der anschließenden Bewertung von Erfolgen. **Die Evaluation solcher Maßnahmen stellt sich sowohl wegen der Komplexität beziehungsweise Multidimensionalität von Zivilcourage, als auch wegen der zu berücksichtigenden Interessen unterschiedlicher „Stakeholder“ komplizierter dar.** Hinsichtlich der Ziele von Maßnahmen und deren Evaluation sind Interessen- und Zielkonflikte zwischen Auftrag- beziehungsweise Geldgebern, Trägern, Trainern, Teilnehmern und anderen von einem Programm beziehungsweise einer Maßnahme direkt und indirekt Betroffenen nicht unwahrscheinlich. Träger und Durchführende von Programmen und Maßnahmen verfolgen manifeste und latente Zielsetzungen, die jenen der Geldgeber nicht immer entsprechen müssen; Erfolge und Erfolgskriterien werden daher auch nicht per se identisch sein. Bei Maßnahmen der politischen (Erwachsenen-)Bildung, wie etwa Zivilcourage-Trainings, ist zudem zu berücksichtigen, dass die Teilnahme an derartigen Angeboten in der Regel auf freiwilliger Basis erfolgt, sie also auf ein hohes Maß an Zustimmung bei den Teilnehmenden angewiesen sind, die durchaus ihre eigenen und teils unterschiedlichen subjektiven Maßstäbe für das „Gelingen“ oder die „Wirksamkeit“ ins Spiel bringen.

Diese kurze und holzschnittartig geführte Diskussion zeigt, dass die Wahl von Kriterien nicht nur methodische **Fragen der Messbarkeit** und normative **Fragen der Bewertung** beziehungsweise Bewertbarkeit aufwirft, sondern dass eben diese Wahl einen selektiven Blick auf die Wirklichkeit erzeugt. Anders ausgedrückt: Kriterien stellen Wahrnehmungsraster dar, die festlegen, worauf zu achten ist und worauf nicht. Völlig unbeachtet bliebe in dem hier skizzierten Szenario, dass es durchaus wünschbare Wirkungen jenseits des hier diskutierten (Wirkungs-) Kriteriums geben könnte. Solche könnten zum Beispiel darin bestehen, mit einer

Maßnahme Teilnehmer aus Milieus zu erreichen, die sonst eher „bildungsfern“ beziehungsweise „absinent“ sind oder Sensibilisierungseffekte zu erzielen, die bislang nicht realisierte Reflexions- und Handlungsbereitschaft erzeugen. Beides bliebe außerhalb des „Wahrnehmungsfensters“ der hier thematisierten offen erkennbaren „Wirkung“ und damit „unsichtbar“.

Es gibt kein vorgegebenes Kriterium für den Erfolg und die als positiv zu wertenden Wirkungen. Kriterien für ein erfolgreiches Programm zur Förderung von Zivilcourage sind stets an die (unterschiedlichen) Perspektiven der Beteiligten, die darin enthaltenen Erwartungen, Interessen, Legitimations- und Handlungszwänge gebunden. Die Wahl von Kriterien und Beurteilungsmaßstäben beruht auf mehr oder weniger bewussten Wertentscheidungen, weshalb es für die Evaluation unabdingbar ist, diese nicht nur offen zu legen, sondern auch zwischen allen Beteiligten kommunikativ auszuhandeln. Es empfiehlt sich, diesen Prozess der Kriterienfindung durch einen erfahrenen Evaluator begleiten und gegebenenfalls moderieren zu lassen. Dies erhöht die Chance, dass alle relevanten Perspektiven in den Prozess einbezogen, „blinde Flecken“ aufgedeckt und vor allem die oft erforderlichen „Übersetzungsleistungen“ zwischen den verschiedenen Perspektiven geleistet werden.

Der Sinn einer derartigen (moderierten) „partizipativen“ Evaluation besteht nicht nur in dem Versuch, Verständigung zu erzielen, sondern stellt in genau diesem Prozess sicher, dass auch tatsächlich nach praxis- und alltagsrelevanten Wirkungen gesucht wird und deren Bewertung Kriterien unterliegt, die dem Gegenstand gegenüber angemessen sind. Darunter sind solche Kriterien zu verstehen, die sich aus dem Konzept und der didaktischen Ausrichtung einer Maßnahme beziehungsweise eines Trainingsprogramms selbst rekonstruieren lassen. Die Ableitung und Explikation derartiger „immanenter“ (Wirkungs-)Kriterien hat den Sinn, bei der Evaluation nicht Erwartungen beziehungsweise Maßstäbe anzulegen, die der Logik und Zielsetzung der zu evaluierenden Maßnahme nicht gerecht werden. Dies ist ein Gebot der Fairness (vgl. DeGEval, Standards) sowohl gegenüber denjenigen, die in einer bereits laufenden Bildungsmaßnahme tätig sind, als auch gegenüber den Teilnehmern.

EINE PERSÖNLICHKEITSORIENTIERTE THEORIE  
ZIVILCOURAGIERTEN VERHALTENS ALS  
GRUNDLAGE EINER GEGENSTANDSADÄQUATEN  
EVALUATION

**Inwiefern Evaluationskriterien gegenstandsadäquat sind, ist abhängig von der jeweiligen (häufig implizit) unterlegten (Alltags-)theorie des Bildungsprozesses und den darin enthaltenen Annahmen über die Wirkungsweise von Programmen und Maßnahmen.** Viele an den vermeintlich unmittelbar „sichtbaren“ Erfolgen ausgerichtete Betrachtungsweisen orientieren sich mehr oder weniger offen an einer behavioristischen Theorie des „Lernens“ im Rahmen derartiger Bildungsmaßnahmen, welche die Teilnehmer selbst als „black boxes“ betrachtet: Die Maßnahme wird als „Input“ angesehen, der am Ende einer idealer Weise homologen „Output“ zur Folge hat, wobei das Wie der Verarbeitung durch die Teilnehmer selbst nicht interessiert oder als nicht zugänglich betrachtet wird. Eine derartige Theorie geht aber an den vielfältigen und unterschiedlichen Prozessen vorbei, die mit dem subjektiven Erwerb von Reflexions- und Handlungskompetenzen verbunden sind. Bildungsprozesse sind eine höchst komplexe und – auf der Ebene der zu bildenden Subjekte – immer auch höchst persönliche und subjektive Angelegenheit.

**Dies betrifft zum einen den inneren Prozess des Erwerbs von Wissen.** Die durch die Maßnahme vermittelten Inhalte müssen individuell angeeignet, mit subjektiver Bedeutung versehen und in bereits bestehende Weltbilder und Relevanzstrukturen sinnvoll eingepasst werden. Daher ist eine häufig mehr oder weniger offen angestrebte – bildlich gesprochen – „1 : 1 Übertragung“ von programmatisch intendierten Inhalten und Zielen auf die Teilnehmer nicht möglich. Vielmehr erfolgt in der subjektiven Aneignung von Maßnahmeninhalten eine Veränderung sowohl bereits bestehender als auch der neu vermittelten Wissensbestände. Dieses so entstehende Amalgam von Altem und Neuem bildet die Grundlage unterschiedlicher Transfers: Vorhandene Erfahrungen können im Lichte der erweiterten Wissensbasis reflektiert werden, ohne dass dieses „innere Handeln“ unmittelbar sichtbare Konsequenzen haben muss. Dennoch ist dies die Grundlage der angestrebten Umsetzung auf der Ebene „sichtbarer“ Handlungen, die letztlich zur Ausbildung neuer, routinierter Handlungsmuster führen kann. Diese

Sicht von Bildungsprozessen erweist sich als gegenstandsadäquat, da sie einen engen methodologischen Bezug zu einer Theorie von Zivilcourage-Trainings aufweist.

Zivilcourage bezeichnet einen Handlungstypus der darauf abzielt, Freiheit, Würde und Unversehrtheit anderer gegen Übergriffe Dritter zu verteidigen. Die soziale und gesellschaftliche Funktion zivilcouragierten Handelns besteht darin, einen Beitrag zur Konfliktdeeskalation, Befriedung und Zivilisierung des alltäglichen Zusammenlebens zu leisten – etwa durch „beherztes“ Eingreifen im Falle fremdenfeindlicher Übergriffe. Zivilcourage kann, je nach Sichtweise, bis zur bewusst in Kauf genommenen Gesetzesübertretung reichen, wenn legales (staatliches) Handeln als illegitime Verletzung universal gültiger Werte wahrgenommen wird. Das verbindende Element der unterschiedlichen Ansätze von Zivilcourage-Trainings ist die Absicht einer systematischen Förderung entsprechender Handlungsdispositionen und -kompetenzen.

Eine Theorie zivilcouragierten Verhaltens lässt sich zum Beispiel aus einem in der Praxis viel gebräuchlichen Handbuch für Zivilcourage-Trainings rekonstruieren (Lünse/Rohwedder/Baisch 2001). Demzufolge setzt praktizierte Zivilcourage eine Persönlichkeit mit hoch differenzierten emotionalen, (sozial)kognitiven und aktionalen Dispositionen voraus. Auf der emotionalen Ebene ist zunächst „Mut“ gefordert, wobei Mut hier als eine subjektivrelationale Eigenschaft gefasst wird: Mut erweist sich in Situationen, in denen eigene Hemmungen und Ängste überwunden werden, wobei die Autoren zwischen „negativem“ Mut – das Überwinden eigener Grenzen und Hemmungen in fremdbestimmten Handlungssituationen, etwa im Befolgen militärischer Befehle – und „positivem“ Mut unterscheiden, wobei letzterer auf selbstbestimmtes Handeln zielt. In (sozial)kognitiver Hinsicht beruht Zivilcourage einerseits auf „Empathie“, der Fähigkeit sich jenseits von Sympathie oder Antipathie in die Situation und den Standpunkt eines anderen versetzen zu können. Damit ist eine Form der vorurteilsfreien inneren Annäherung an die Subjektivität des Anderen impliziert. Diese setzt ein ausgewogenes Urteilsvermögen voraus. Die Möglichkeit zivilcouragierten Handelns ist zudem an bestimmte ethisch-motivationale Haltungen ge-



knüpft, die weder hedonistisch (das heißt am eigenen Lustgewinn) noch utilitaristisch (das heißt am eigenen oder kollektiven Nutzen, ohne Berücksichtigung von damit verbundenen Benachteiligungen) orientiert sind.

Die zivilcouragierte Persönlichkeit wird von daher als selbst- und wertreflexiv und somit auch als „in-nengeleitet“ (das heißt in ihrem Handeln nicht von den Urteilen, der Zustimmung anderer und nicht von äußeren materiellen und immateriellen Belohnungen bestimmt) beschrieben. **Innere Autonomie und Selbstbewusstsein des Handelnden sind eine Voraussetzung für die zu zivilcouragiertem Handelnden erforderliche Nonkonformität, da das Eintreten für diskriminierte Personen vielfach in Konflikt mit den Überzeugungen und normativen Erwartungen des jeweiligen sozialen Umfelds gerät.** Daher erfordert aktives Eingreifen in öffentlichen Situationen, etwa bei Belästigungen in öffentlichen Verkehrsmitteln, häufig die Überwindung verinnerlichter, handlungshemmender Scham- und Peinlichkeitsschwellen, die sonst den Umgang zwischen Fremden im öffentlichen Raum bestimmen.

Im Sinne dieser **persönlichkeitsorientierten Theorie** zivilcouragierten Handelns zielt das von Lünse und anderen konzipierte Programm darauf, die Teilnehmer auf emotionaler, kognitiver und handlungsorientierter Ebene für die Notwendigkeit zivilcouragierten Handelns zu sensibilisieren und ihnen entsprechende eigene Handlungsmöglichkeiten aufzuzeigen. Das entsprechende didaktische Konzept umfasst daher Trainingseinheiten, die sich auf folgende Erfahrungs- beziehungsweise Lernfelder beziehen:

#### **Kognition und Selbsterleben:**

Sensibilisierung für Formen meist verdeckter struktureller und psychischer Gewalt, die sowohl analytisch erfasst als auch in ihren Konsequenzen für die Betroffenen emotional nachvollziehbar gemacht werden sollen. Auseinandersetzung mit Feindbildern, deren Funktion und Wirkung sowie mit den Affekten Wut und Aggression, wobei deren ambivalenter, das heißt positiver und negativer Charakter erfahrbar gemacht und Wege zu einem konstruktiven Umgang aufgezeigt werden sollen.

#### **Wertbewusstsein, Reflexionsvermögen und Empathiefähigkeit:**

Übungen zum Erkennen der Unvermeidbarkeit von Widersprüchen sowie der Existenz von Wertepluralismus sollen zu differenziertem individuellen Urteilsvermögen führen, das die Basis gesteigerter Empathiefähigkeit darstellt.

#### **Handeln:**

Vermittlung handlungspraktischer Kompetenzen zum Umgang mit Konfliktsituationen, in denen Zivilcourage gefordert ist.

„Wirkungen“ auf diesen Ebenen sind nicht unmittelbar beobachtbar, sondern lassen sich nur aus dem kommunikativen wie dem praktischen Handeln der Teilnehmer erschließen, sei es, dass man hierzu Selbstaussagen von Teilnehmern heranzieht oder Daten aus theoriegeleiteten Interpretationen von Handlungen gewinnt.

#### **DIE SUCHE NACH WIRKUNGSKRITERIEN: ANBIETER- UND TEILNEHMERPERSPEKTIVE**

**Im Folgenden wird eine Suchstrategie beschrieben, die Anbieter und Durchführende bei der Suche nach Kriterien unterstützen kann.** Für die Entwicklung von Fragerichtungen und Bewertungskriterien ist eine Systematisierung sinnvoll, die zwischen „Anbieter-“ und „Teilnehmerperspektive“ unterscheidet. Beide Perspektiven können dabei durchaus Übereinstimmungen aufweisen.

Für die **Anbieterperspektive** lassen sich Kriterien aus den expliziten Zielsetzungen entwickeln, wie sie etwa bei der Antragstellung, bei der Ankündigung der Maßnahme oder ähnlichen, auf Außenanwendung gerichteten Mitteilungen formuliert wurden. Oft jedoch decken diese nicht das gesamte Feld (erwünschter) Wirkungen ab. Zielführende „Wirkungen“ können auf verschiedenen Ebenen des Bildungsprozesses vorliegen, ohne auf den ersten Blick als solche wahrgenommen zu werden. Hier kann eine Rekonstruktion der dem eigenen Vorgehen zugrundeliegenden Theorien und impliziten Annahmen sowohl über die Wirkungsweise des spezifischen Maßnahmenangebots „Zivilcourage-Training“ als auch über das zu erreichende Ziel „Förderung von Zivilcourage“ selbst, die Aufmerksamkeit auf relevante Wirkungsbereiche lenken. Die folgende Fragenliste soll als Anregung

zu einer systematisierten, von Programmentwicklern und Trainern gemeinsam durchführbaren Rekonstruktion dienen, die zudem durch den Einsatz von Metaplantechniken unterstützt werden sollte.

► **(1) Wie sieht die ideale zivilcouragierte Persönlichkeit aus?** Wie beurteilt sie die Welt und andere? Welche Eigenschaften sollte sie haben? Wie verhält sie sich?

► **(2) Welche Hindernisse und Hemmnisse für zivilcouragiertes Handeln nehmen wir an?** Worin sehen wir die Ursachen und Gründe für das Bestehen dieser Hindernisse und Hemmnisse?

► **(3) Wie ist unsere Maßnahme aufgebaut?** Welche Übungs- und Trainingseinheiten umfasst sie? Aus welchem Grund haben wir diese in unsere Maßnahme aufgenommen? Auf welche der unter (1) und (2) herausgearbeiteten Dimensionen beziehen sich die einzelnen Einheiten? Haben wir Schwerpunkte gesetzt oder versuchen wir alle erarbeiteten Dimensionen gleichermaßen abzudecken?

► **(4) Wie sehen der idealtypische Teilnehmer und die idealtypische Teilnehmergruppe aus, die durch eine Maßnahme angesprochen werden sollen?** Wie würden sie auf die Maßnahme „ansprechen“? Woran könnten wir das erkennen?

► **(5) Wie sehen die realen Teilnehmergruppen aus?** Wodurch unterscheiden sie sich von der idealtypischen (4)? Welche Folgen erwarten wir aufgrund dieser Unterschiede? Welche der unter (2) erarbeiteten Hindernisse und Widerstände sind von besonderer Bedeutung? Woran würden wir dies erkennen? Ist das Konzept der realen Maßnahme darauf zugeschnitten? Welche Übungs- und Trainingsteile (3) müssten verstärkt praktiziert, welche eventuell weggelassen werden?

Durch die Anwendung dieser, oder entsprechend den eigenen Interessen modifizierter, Fragestrategien lassen sich sowohl die eigenen Erwartungen an die Wirkungen des politisch-bildenden Handelns als auch die ihm zugrundeliegenden Annahmen über Wirkungsprozesse explizieren und dokumentieren. Besondere Bedeutung kommt der Frage zu, woran man die erwünschten Wirkungen, Veränderungen erkennen könnte. Die so gefundenen Indikatoren für die Beobachtung gewünschter Veränderungsprozesse (Bildungsprozesse) bilden zugleich Kriterien für die Wirksamkeit der Maßnahme in die beabsichtigte Richtung. **Dieses Vorgehen zielt auf die**

**Feststellung qualitativer Veränderungen und nur indirekt auf die Erfassung des messbaren, graduellen Ausmaßes dieser Veränderungen.** Deren Erfassung würde die aufwändige Konstruktion wissenschaftlich zuverlässiger und gültiger Testbatterien voraussetzen, ein Aufwand, der nur unter Zugriff auf umfangreiche wissenschaftliche Ressourcen zu leisten wäre und daher für eine selbst durchgeführte Evaluation nicht leistbar (und erwartbar) ist. Der hier vorgeschlagene Prozess der Kriterienentwicklung kann dazu beitragen, die für eine Selbstevaluation erforderlichen Kompetenzen zu stärken; er fordert (und fördert) das Vorstellungsvermögen beziehungsweise den „Möglichkeitssinn“ und schärft zugleich die für die Beobachtung und Beurteilung von Praxisprozessen und später zur Evaluation der herangezogenen Daten erforderlichen „Deutungskompetenzen“.

**Ergänzend oder alternativ zu einer Evaluationsperspektive aus Anbietersicht kann die Teilnehmerperspektive in die Evaluation einbezogen werden.** Grundlage hierfür können einmal Selbsteinschätzungen von Teilnehmern sein. Diese lassen sich mit dem oben erarbeiteten Kriterienkatalog abgleichen und ermöglichen so eine Einschätzung, wo die Teilnehmer in Relation zu den Zielvorstellungen der Anbieter „stehen“. Die genannte Strategie ist folglich an Kriterien der Durchführenden orientiert. Zur Abfrage eignet sich hier ein Verfahren mittels einer auf einer Metaplanwand oder Flipchart angebrachten „Seminarzielscheibe“, auf der die Teilnehmer am Ende des Trainings ihre Einschätzung eintragen können. Im Folgenden wird ein Zielscheiben-Modell zur Bewertung von Methodik und Lernerfolg vorgeschlagen, dessen Zielsektoren je nach Erkenntnisinteresse umbenannt werden können (Anregungen lassen sich dem am Ende dieses Kapitels folgenden Fragenkatalog entnehmen).

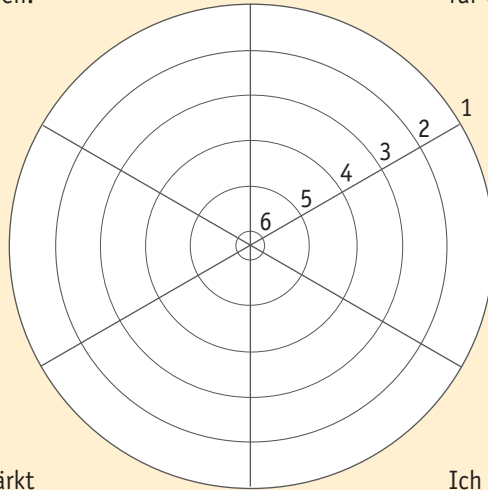
## METHODIK UND LERNERFOLG

Ich traue mir jetzt mehr zu,  
in Konflikte einzugreifen.

Ich empfand die Methode  
für das Thema als sinnvoll.

Ich habe hilfreiche  
Anleitungen für  
Konfliktsituationen  
erhalten.

Ich habe neue  
Kenntnisse vermittelt  
bekommen.



Ich denke jetzt verstärkt  
über meinen Umgang mit  
Konflikten im Alltag nach.

Ich erkenne Konflikte  
jetzt besser.

Bitte markieren Sie, inwieweit Sie den obenstehenden Aussagen zustimmen.

6 = höchste Zustimmung, 1 = keine Zustimmung

**Methodischer Hinweis:** Verteilen Sie an die Teilnehmer Klebepunkte und bitten Sie sie, vorab auf diesen ihre Bewertung einzutragen (6 = hohe Zustimmung, 1 = keine Zustimmung). Erst dann werden die Punkte auf der Zielscheibe angebracht.

Dieses Vorgehen verhindert einen Gruppeneffekt in dem Sinn, dass Teilnehmer sich an den Bewertungen anderer orientieren und – nach „oben“ oder „unten“ – abweichende Einschätzungen nicht mitgeteilt werden.

Die Fragen beziehen sich, im Uhrzeigersinn, auf folgende Bereiche: Methodik, Wissenserwerb, Kognition/Situationsanalyse, Selbstreflexion, Praxis- und Handlungsbezug sowie Kompetenzwahrnehmung/Handlungsbereitschaft. Die Antworten zu Methodik geben einen ersten, groben Aufschluss darüber, inwieweit die Teilnehmenden die eingesetzten Übungen und Trainingseinheiten aus ihrer Sicht als dem Thema angemessen einschätzen. Die nächsten Felder erfragen „Lernerfolge“ (Wirkungen) auf unterschiedlichen konkretisierten Ebenen. Zunächst werden Teilnehmereinschätzungen darüber erfragt, in welchem Maß die Teilnahme zu einem globalen, unspezifischen Wissenzuwachs geführt hat (wobei „Wissen“ hier sowohl „Faktenwissen“

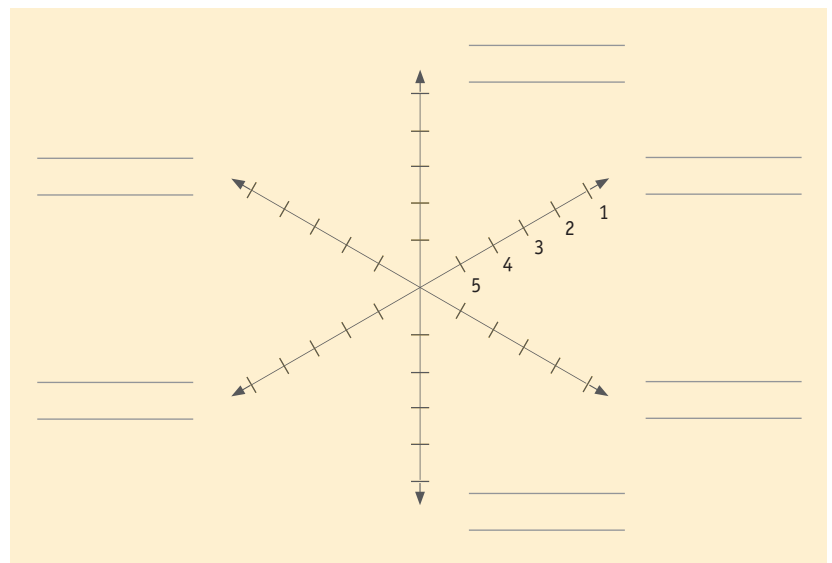
als auch deutungs- und handlungsrelevantes „Alltagswissen“ umfassen kann). Der wahrgenommene Wissenzuwachs im genannten Sinne ist ein Minieffekt einer Teilnahme. Die nächsten Themenbereiche erschließen Einzelaspekte subjektiv wahrgenommener Effekte im Sinne der oben ausgeführten Theorie: Kann das vermittelte „Wissen“ zur Interpretation konflikthaltiger Alltagssituationen herangezogen, die eigene Rolle in Konfliktsituationen reflektiert werden? Daran anschließend werden der subjektiv wahrgenommene Handlungsbezug der Maßnahme und der empfundene Zuwachs an Kompetenzgefühlen beziehungsweise der Bereitschaft zur praktischen Intervention in Konfliktsituationen erfragt.

Dieses Verfahren ermöglicht nicht nur eine Bewertung der einzelnen Aspekte durch die Visualisierung von Verteilungen und Nennhäufigkeiten, da die „Schwerpunkte“ der Einschätzungen „auf einen Blick“ erkennbar werden, sondern stellt zugleich auch die Streuung aufgrund eventueller „Ausreißer“, das heißt von der durchschnittlichen Gruppeneinschätzung abweichender Wertungen, dar. Zur Auswertung dieses Verfahrens ist es sinnvoll, eine Diskussion mit den Teilnehmern über die einzelnen abgefragten Kategorien anzuschließen, um Aufschluss über die individuellen Deutungen der Teilnehmer zu erhalten. Was aus Anbietersicht etwa als eine hilfreiche Anleitung für Konfliktsituationen empfunden wird, mag von Teilnehmerseite mit ganz anderem Deutungsgehalt versehen werden. Für die Teilnehmern können im Zusammenhang mit den einzelnen Themenfeldern andere Gesichtspunkte von Bedeutung sein. Wichtig ist in diesem Zusammenhang die Auseinandersetzung mit von der allgemeinen Tendenz abweichenden Einschätzungen.

Die Diskussion der beiden genannten Punkte kann dazu beitragen, neue, wenig beachtete, aber womöglich dennoch wichtige Aspekte aufzudecken.

Es kann außerdem durchaus sinnvoll sein, eine Evaluation auf Grundlage der von den Teilnehmern selbst an die Teilnahme angelegten Kriterien durchzuführen. Gerade im Bereich der politischen Bildung, im Speziellen von Zivilcourage-Trainingsmaßnahmen, erfolgt die Teilnahme zumeist aufgrund einer freiwilligen Entscheidung. Das Motiv besteht dann in der Regel in der Erwartung, eigenes Wissen und eigene Deutungs-, Urteils-, Handlungskompetenzen zu erweitern, die sich grob unter dem Aspekt einer angestrebten Persönlichkeitsentwicklung zusammenfassen lassen. Das Kriterium für die Beurteilung des Maßnahmen Erfolgs wäre also, in wie weit die Teilnahme an der Maßnahme einen Beitrag zur Realisierung der jeweiligen subjektiven Erwartungen beziehungsweise Zielvorstellung geleistet hat.

### MEINE PERSÖNLICHEN ERWARTUNGEN AN DAS ZIVILCOURAGE-TRAINING



Individuelle Teilnehmererwartungen vor und subjektive Einschätzungen zu deren Erfüllung nach Besuch eines Zivilcourage-Trainings lassen sich mit Hilfe eines so genannten Erwartungs- und Erfüllungsrasters visualisieren (in Anlehnung an transfer 1994, 2). Hierzu werden die Teilnehmer zu Beginn gebeten, jeweils auf einem, anhand der folgenden Grafik vorbereiteten Arbeitsblatt ihre Erwartungen an das Training zu formulieren (und eventuell nach ihrer Bedeutung für die subjektive Bewertung des Trainingserfolgs zu gewichten, wobei 5 einer hohen

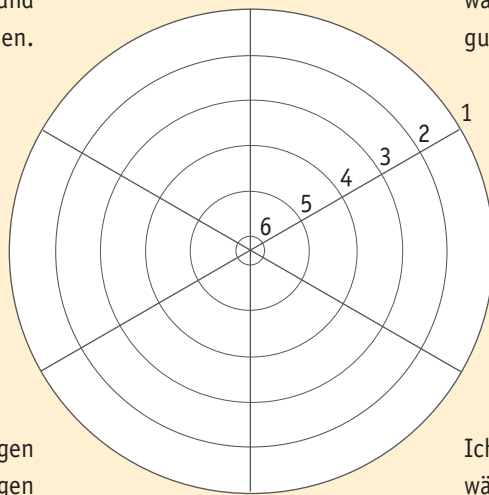
und 1 einer niedrigen Bedeutung für die Erfolgsbestimmung gleichkommt). Im Anschluss an das Training tragen die Teilnehmer auf demselben Arbeitsblatt ein, inwieweit die subjektiven Erwartungen ihrer Ansicht nach erfüllt wurden (wobei 5 für den höchsten und 1 für den niedrigsten Grad der Realisierung der subjektiven Ziele steht). Ein Vergleich der Teilnehmererwartungen und deren Erfüllung mit den Zielsetzungen des Trainings ermöglicht den Anbietern eine Optimierung des Angebots im Hinblick auf die Bedürfnisse der Zielgruppe.

## ORGANISATION UND DURCHFÜHRUNG

Im Training wurde auch über Alltagserfahrungen und -erlebnisse gesprochen.

Es gab neben den Übungen genug Zeit, im Training gemachte Erfahrungen zu verarbeiten.

Die Erläuterungen zu den Übungen durch den Trainer waren gut verständlich.



Der Veranstaltungsort war für das Training gut geeignet.

Die Dauer der Veranstaltung war dem Thema angemessen.

Ich fand die Atmosphäre während des Trainings angenehm.

Bitte markieren Sie, inwieweit Sie den obenstehenden Aussagen zustimmen.  
6 = höchste Zustimmung, 1 = keine Zustimmung

Da das Erreichen der Zielsetzungen – seien es jene der Anbieter oder der Teilnehmer – nicht nur von der abstrakten „Methode“, sondern auch von der jeweiligen konkreten Durchführung des Trainings beeinflusst wird, ist auch dieser Bereich zu evaluieren. Hierzu bietet sich folgende „Seminarzielscheibe“ zu Organisation und Durchführung an.

Im Uhrzeigersinn können die erhaltenen Antworten dazu beitragen, folgende evaluationsrelevante Felder zu erschließen: Eignung des räumlichen Umfeldes, das zeitliche Ablaufschema (zu „dichtes“ oder zu „lockeres“ Programm), das für den Bildungserfolg wichtige Gruppenklima, die Verständlichkeit der einzelnen Übungs- und Trainingseinheiten insgesamt, die Herstellung von Bezügen zu den Alltagswirklichkeiten der Teilnehmenden, das Vorhandensein ausreichender „Freiräume“ zur individuellen Verarbeitung und Reflexion von Erfahrungen während der Maßnahme. Es empfiehlt sich auch hier, die erhaltenen Zielscheibenergebnisse mit der Teilnehmergruppe zu besprechen, um Aufschluss über die erhaltenen Bewertungen zu erlangen.

Aufbauend auf den vorangegangenen Überlegungen bieten sich unter anderem folgende Aspekte und Fragestellungen für die Konstruktion von Fragebögen, aber auch zum Einsatz in „Zielscheiben“ an, um im Anschluss an ein Zivilcourage-Training Aussagen der Teilnehmer über die Durchführung der Veranstaltung, die Methode, Lernerfolg und Wirkung des Trainings zu erhalten. Selbsteinschätzungen der Teilnehmer zu ihrem Konfliktverhalten nach Besuch eines Zivilcourage-Trainings können eine umfassende Untersuchung der Wirkung auf die Bereitschaft und tatsächliches zivilcouragiertes Handeln nicht ersetzen, vermitteln aber durchaus erste Eindrücke.

**Als verhältnismäßig wenig ressourcenintensive Form der Evaluation von Trainings bietet sich darüber hinaus auch die Methode der Beobachtung an, bei der ein Beobachter den Trainingsprozess vor Ort dokumentiert.**

Als Beobachter kann sowohl ein Beteiligter – etwa ein Trainer – oder eine externe Person eingesetzt werden. Der Vorteil dieser Methode liegt darin, dass

sie es ermöglicht, Informationen über den Verlauf einer Bildungsmaßnahme, die Gruppendynamik und die Reaktionen der Teilnehmer auf einzelne Elemente des Trainings direkt und nicht erst rückbli-

ckend aus den Einschätzungen der Teilnehmer zu erhalten. Sie ist jedoch nicht voraussetzungslos: Beobachtung als Mittel der formativen Evaluation erfordert Erfahrung und Übung, um zu vermeiden, dass unerwünschte gegenüber erwünschten Ergebnissen „übersehen“ werden. Beobachtungen können unterschiedlich angelegt sein und von durch Co-delisten normierten Beobachtungen bis zu freien Beschreibungen reichen, bei denen sich die Leit-aspekte erst im Lauf der Beobachtung ergeben – in allen Fällen ist jedoch eine umfassende Dokumentation des Beobachteten notwendig. Für den Einsatz zur Evaluation von Zivilcourage-Trainings schlagen wir vor, eine Beobachtung unter anderem an folgenden Aspekten zu orientieren: dem Verlauf des Trainings, der Reaktion der Teilnehmer auf die Methode im Allgemeinen und die einzelnen Elemente im Besonderen, die Gruppendynamik und den Umgang mit Konflikten im Training. Die so gewonnenen Informationen ergeben Hinweise für die Prozessoptimierung und die zielgruppengerechte Gestaltung des Trainings. Zu bemerken ist jedoch, dass Beobachtung als Evaluationsmethode bei Bildungsmaßnahmen, die einen hohen Selbsterfahrungsanteil aufweisen, ethische Fragen aufwerfen können. Selbstverständlich ist deshalb Voraussetzung für den Einsatz eines Beobachters die Zustimmung aller Teilnehmer.

## FRAGENBEREICHE (AUSGEWÄHLTE FELDER UND BEISPIELE) ZU ...

### **Seminar und Gruppenprozess**

- ▶ Ausreichend Raum für Fragen und Reflexion während des Trainings
- ▶ Möglichkeit für Teilnehmer, eigene Erfahrungen in das Training einzubringen
- ▶ Ungelöste Konflikte im Trainingsverlauf
- ▶ Integration von Außenseitern in den Gruppenprozess

### **Methodik und einzelne Übungen**

- ▶ Angemessenheit der Methodik
- ▶ Übung, die besonders gefallen hat
- ▶ Übung, die ein besonders nützliches Handlungsrepertoire für Situationen gegeben hat, in denen Zivilcourage gefordert ist
- ▶ Übung, die als weniger geeignet für ein Zivilcourage-Training empfunden wurde
- ▶ Qualität der Materialien, die im Verlauf des Trainings ausgeteilt wurden

### **Trainer/Trainerin**

- ▶ Überzeugendes Auftreten des Trainers/der Trainerin
- ▶ Gute Vermittlung der Inhalte und Übungen durch den Trainer/die Trainerin

### **Praxisrelevanz des Programms, Lernerfolg und Transfer**

- ▶ Neues durch das Training erfahren (Stichworte abfragen)
- ▶ Situationen aus dem Alltag der Teilnehmer im Training angesprochen (Beispiele geben lassen)
- ▶ Hilfreiche Anregungen für Konfliktsituationen
- ▶ Verstärktes Nachdenken über ungerechte Situationen oder Konflikte im Alltag
- ▶ Selbstvertrauen nach dem Training, Konflikte anzusprechen beziehungsweise einzugreifen

### **Persönliche Erwartungen und Erfahrungen der Teilnehmer**

- ▶ Erwartungen an das Training (prä)
- ▶ Eingetretene Erwartungen (post)
- ▶ Schlüssel- beziehungsweise „Aha-Erlebnisse“ während des Trainings
- ▶ Selbsterfahrung(en)
- ▶ Reflexion eigenen Verhaltens in Konfliktsituationen
- ▶ Motivation, etwas im eigenen Leben zu ändern (post)

### **Bilanzierung**

- ▶ Zufriedenheit mit dem Verlauf des Trainings
- ▶ Zufriedenheit mit den Ergebnissen des Trainings

## DATENKRITIK UND -BEWERTUNG

**Die vorgeschlagenen Verfahren stellen bei weitem nicht alle Möglichkeiten von Evaluationsdesigns und -methoden dar, sondern verstehen sich als Anregung für mögliche Wege einer Selbstevaluation.** Das Erheben und Sammeln von Daten allein ist jedoch noch keine Evaluation. Bloße Punkte auf den Zielscheiben oder Kreuze auf Fragebögen sind im Sinne des Wortes „nichtssagend“. Damit sich mit den erhobenen Daten eine Evaluation vornehmen lässt, müssen diese zum Sprechen gebracht, das heißt interpretiert werden. Zunächst dienen Daten der Beschreibung eines Zustandes – sie dokumentieren gegebene Einschätzungen, Erwartungen, subjektiv erlebte oder fremdeingeschätzte Erfüllungserwartungen – kurz: eine bestimmte Situation zum Zeitpunkt ihrer Erhebung. Sie sind dadurch zustande gekommen, dass die Maßnahme in einer bestimmten Weise „befragt“ wurde, das heißt die Daten bilden nicht einfach eine gegebene Wirklichkeit ab, sondern stellen Reaktionen auf bestimmte Fragen dar. Andere Fragen könnten zu anderen „Bildern“ dieser Wirklichkeit geführt haben.

Hier schließt sich ein Kreis, der mit dem Vorschlag eröffnet wurde, die dem eigenen Handeln zugrundeliegenden „Theorien“ und „impliziten“ Annahmen zu rekonstruieren und zu explizieren: Die gewonnenen Daten sind Antworten auf die Fragen, die sich vor dem Hintergrund der je speziellen Annahmen als wichtig und interessant herausgestellt haben. Die Auswertung und Bewertung kann jeweils nur so gut sein, wie wir in der Lage sind, die gegebenen Antworten in ihrem „Sinn“ (was wird eigentlich gesagt) und in ihrer „Bedeutung“ (hinsichtlich der gestellten Frage) zu verstehen und zu bewerten. Dies trifft nicht nur auf die Einzelevaluation einer einzelnen Maßnahme zu. Vielmehr empfiehlt es sich, Daten zu dokumentieren und über mehrere Durchgänge hinweg zu vergleichen, um Veränderungen und Variationen feststellen zu können sowie einen umfassenderen Eindruck von der Bewertung und Wirkung einer Maßnahme beziehungsweise eines Programms zu erhalten.

Datenanalyse ist ein komplexer Prozess, auf dessen Einzelheiten und Methoden an dieser Stelle nur cursorisch eingegangen werden kann. Er umfasst das Dokumentieren, Sortieren, Interpretieren, Dis-

kutieren, Aus- und Bewerten von Informationen sowie letztendlich auch den verantwortungsbewussten Umgang mit den Ergebnissen einer Evaluation. Die Komplexität der Datenanalyse sollte jedoch nicht davon abschrecken, Bildungsmaßnahmen zu evaluieren. Die oben vorgestellten Möglichkeiten der Selbstevaluation liefern Daten, anhand derer Erwartungsmuster festgestellt und mit den Anbieterkategorien in Beziehung gesetzt werden können. Sie geben Antworten darauf, ob die Anbieterintentionen mit den Teilnehmerintentionen übereinstimmen. Einige dieser Antworten werden mehr, andere weniger leicht zu identifizieren sein. So müssen etwa die aus offenen Fragen in Fragebögen gewonnenen Antworten zunächst sortiert und im Hinblick auf wiederkehrende Aussagen und Kategorien ausgewertet werden (vgl. Glaser/Strauss, 1998), um aus ihnen Schlussfolgerungen über die Maßnahme oder das Programm zu ziehen. Eine detaillierte Dokumentation der Evaluationsergebnisse zu Präsentationszwecken ist auf jeden Fall zu empfehlen.

Evaluation ist ein Unterfangen, das nicht nur zur Befriedigung externer Interessen, sondern auch für die Praktiker der Bildungsarbeit von nicht zu unterschätzender Bedeutung ist. Wird sie unter möglichst großer Einbeziehung aller Betroffenen transparent und gegenstandsadäquat betrieben, kann sie die Grundlage für einen umfassenden Lernprozess bilden und Möglichkeiten der Programm- und Organisationsentwicklung eröffnen. Evaluation, so verstanden, wird langfristig kein „Fremdkörper“ bleiben, sondern Teil der pädagogischen Arbeit sein.

## LITERATUR

- Bommes, Michael u. a.: Eine Evaluation des Programms „Miteinander – Erfahrungen mit Betzavta“ – Abschlussbericht, Freiburg 2002.  
[www.cap.uni-muenchen.de/aktuell/news/2003/2003\\_04\\_toleranz.htm](http://www.cap.uni-muenchen.de/aktuell/news/2003/2003_04_toleranz.htm).
- Deutsche Gesellschaft für Evaluation (DeGEval): Standards für Evaluation der Deutschen Gesellschaft für Evaluation, Köln 2001.  
[www.degeval.de/standards/standards.pdf](http://www.degeval.de/standards/standards.pdf)
- Dimbath, Oliver u. a.: Evaluationsprojekt Achtung (+) Toleranz. Gesamtbericht, München 2002.  
[www.cap.uni-muenchen.de/aktuell/news/2003/2003\\_04\\_toleranz.htm](http://www.cap.uni-muenchen.de/aktuell/news/2003/2003_04_toleranz.htm).
- Glaser, Barney, und Anselm Strauss: Grounded Theory: Strategien qualitativer Forschung, Bern, Göttingen 1998.
- Guba, Egon G., und Yvonna S. Lincoln: Fourth Generation Evaluation, Newbury Park 1989.
- Heiner, Maja: Experimentierende Evaluation: Ansätze zur Entwicklung lernender Organisationen, Weinheim, München 1998.
- Kinast, Eva-Ulrike: Evaluation interkultureller Trainings, Lengerich 1998.
- Kirkpatrick, Donald L.: Evaluating Training Programs. The Four Levels, San Francisco 1998.
- Lünse, Dieter u. a.: Zivilcourage. Anleitung zum kreativen Umgang mit Konflikten und Gewalt, 3. überarb. Auflage, Hamburg 2001.
- Reischmann, Jost: Weiterbildungs-Evaluation. Lernerfolge messbar machen, Neuwied 2003.
- Sanders, James R. (Hg.): Handbuch der Evaluationsstandards. Die Standards des „Joint Committee on Standards for Educational Evaluation“. Aus dem Amerikanischen übersetzt von Wolfgang Beywl und Thomas Widmer, Opladen 1999.
- Thomas, Alain: Valuing Evaluation: A Practical Approach to Designing an Evaluation that Works for You. Bernhard van Leer Foundation Working Paper in Early Childhood Development No. 26, Den Haag, 2000.
- transfer e. V. (Hg.): Aus der Praxis – Für die Praxis 4/1994.
- Ulrich, Susanne, und Florian M. Wenzel: Partizipative Evaluation. Ein Konzept für die politische Bildung, in: Katrin Uhl u. a. (Hg.): Evaluation von politischer Bildung. Was können wir messen? Gütersloh 2004 (im Erscheinen).
- Univation – Institut für Evaluation und wissenschaftliche Weiterbildung e.V.: Abschlussbericht über die Evaluation des Programms „Eine Welt der Vielfalt“, Köln 2002. [www.cap.uni-muenchen.de/aktuell/news/2003/2003\\_04\\_toleranz.htm](http://www.cap.uni-muenchen.de/aktuell/news/2003/2003_04_toleranz.htm).
- Wottawa, Heinrich, und Heike Thierau: Lehrbuch Evaluation, Bern u. a. 1998.



