

Gesa Köbberling und Vanessa Lux

## Evaluationsforschung zwischen „Ökonomisierung des Sozialen“ und Praxisreflexion

Evaluationen boomen seit einigen Jahren. Als wissenschaftliche Bewertungsverfahren mit dem Ziel der Analyse, Reflexion, Qualitätskontrolle und -entwicklung von „praktischen Maßnahmen“ (Wottawa/Thierau 2003, 14) sind sie zu einem festen Bestandteil psychosozialer Arbeit geworden. Es ist mittlerweile selbstverständlich, dass der eigene Arbeitsbereich einer Evaluation unterzogen wird, und für Psychologen<sup>1</sup> bietet die Tätigkeit in Evaluationsprojekten ein wachsendes Berufsfeld. Gleichzeitig gibt es vielfältige Widerstände gegen Evaluationen (vgl. Hirsland/Uhl o.J., 1), die als Mittel der Neoliberalisierung des Gesundheits-, Sozial- und Bildungsbereiches gesehen werden. Die Evaluationsforschung hat in diesem Prozess eine widersprüchliche Funktion, die zwischen dem Aufbau eines Konformitätsdrucks in Richtung auf die Ökonomisierung des Sozialen und dem Anspruch auf Praxisreflexion und „Qualitätsentwicklung“ changiert. Damit kommen nicht nur rhetorisch Fragen nach besserer Lebensqualität und emanzipatorischer Praxis zur Sprache. Diese Widersprüchlichkeit durchzieht den gesamten neoliberalen Umbau des Sozialstaats, der gerade in der Verbindung von repressiven und befreienden Elementen gründet (vgl. Candeias 2004, 231ff. u.ö.). Jedoch ist die Umstrittenheit von Evaluationen aus der oben genannten Definition allein nicht nachvollziehbar. Wer spricht sich schon für eine unreflektierte Praxis oder schlechte Qualität von psychosozialer Versorgung aus? Stattdessen ist zu diskutieren, unter welchen institutionellen, konzeptionellen und methodischen Voraussetzungen Evaluationsforschung den Anspruch verfolgt, zur Qualität von psychosozialer Versorgung und Reflexion psychosozialer Praxis beizutragen und inwieweit sie diesen realisieren kann.

### *Kontrolle oder Praxisreflexion? Die zwei Wurzeln der (deutschen) Evaluationsforschung*

Die Anfänge der Evaluationsforschung werden in den 1960er Jahren in den USA verortet. Erste Evaluationstheorien von Scriven (1980 u. 1983), Campbell (1969) oder Cook (1979) orientierten sich an dem Ziel, mittels experimentell-statistischer Methoden Effekte von Programmen zu messen, um eine rationale Entscheidungsgrundlage für die Beendigung oder Weiterführung von Maßnahmen zu schaffen. Vor dem Hintergrund eines

---

<sup>1</sup> Wenn nicht anders angegeben, sind Psychologinnen, Evaluatorinnen etc. trotz der männlichen Schreibweise explizit mitgemeint.

positivistischen Wissenschaftsverständnisses entwickelte sich die sog. *summative* Evaluation, die auf eine abschließende Erfolgsbewertung zielt und methodisch *quantitativ-experimentell* oder *-statistisch* orientiert ist. Ab Mitte der 70er Jahre rückte dann, die wissenschaftlichen Paradigmen der ersten Generation in Frage stellend, die politische Funktion von Evaluationen in den Blick. Im Mittelpunkt stand das spezifische Informationsbedürfnis der Zielgruppen, dem mit sog. *formativen* Evaluationsmodellen, die auf Praxisreflexion im Sinne konzeptioneller Weiterentwicklung durch Rückmeldung während des Evaluationsprozesses zielen, besser entsprochen werden sollte. Es wurden die unterschiedlichen Interessen der an Evaluationen Beteiligten thematisiert, und die Perspektiven möglichst vieler „Stakeholders“ – d.h. aller Personengruppen, die von der Evaluation und ihren Ergebnissen betroffen sein könnten – sollten in den Reflexionsprozess einbezogen werden (*partizipative Evaluation*). In dieser v.a. von Stake (1983), Guba/Lincoln (1989) und Patton (1987) vertretenen Orientierung gewinnen qualitative Erhebungsmethoden (z.B. narrative Interviews), die konkrete Handlungsebene sowie die Einbeziehung des Kontextes, in dem ein Programm umgesetzt wird, an Bedeutung. Gegenüber der Auffassung, dass wissenschaftliche Objektivität nur durch eine Beurteilung von außen gewonnen werden kann (*externe* und *Fremdevaluation*), wurden zunehmend Konzepte erprobt, in denen die Evaluatoren selbst Teil der evaluierten Einrichtung oder Organisation sind (*interne Evaluation*) oder Praktiker mit Methoden der *Selbstevaluation* ihre eigene Praxis reflektieren und bewerten.

Dieser Strang der Evaluationsforschung entstand in den USA parallel zu groß angelegten staatlichen Sozial-, Bildungs-, Gesundheits- und Infrastrukturprogrammen (Stockmann 2000, 21), die den Markt für die neue Forschung schufen. Die sich rasch einstellende Abhängigkeit drückt sich in der weiteren Entwicklung aus: Als in den 70er und 80er Jahren die Mittel für Reformvorhaben und Modellprojekte reduziert wurden, verlagerte sich der Schwerpunkt der Evaluationsforschung auf die Frage nach dem Kosten-Nutzen-Verhältnis, der Effizienz und Wirtschaftlichkeit der Programme. Eingebettet in die staatlichen Maßnahmen institutionalisierte sich die Evaluationsforschung in kurzer Zeit und wurde zu einer „Wachstumsindustrie“ (vgl. 22).

In der Bundesrepublik ist dies bislang nicht gleichermaßen der Fall. Die Bedeutung der Evaluationsforschung hat jedoch in den letzten Jahrzehnten stark zugenommen und eine Vielzahl von Veröffentlichungen bringt das Bestreben nach einem eigenständigen Berufsprofil mit Aus- und Weiterbildungen, Publikationsorganen und Forschungsressourcen zum Ausdruck (Wottawa/Thierau 2003, Stockmann 2000, Koch/Wittmann 1990). Erste Evaluationen gab es in der BRD Ende der 60er Jahre im Zuge sozialdemokratischer Reformbewegungen in den Bereichen Entwicklungszusammenarbeit, Bildung, Erziehung und Stadter-

neuerung. Sie dienen zum einen der Verbesserung der Programme und Maßnahmen und sollten zum anderen gegen neoliberale und neokonservative Kritik deren Effektivität nachweisen. Bereits mit der Haushaltsreform von 1970 wurden Erfolgskontrollen staatlicher Maßnahmen ausdrücklich vorgeschrieben. Dennoch erlebten Evaluationen erst seit den späten 80ern im Rahmen des „New Public Management“, mit ihrer Anwendung als Steuerungsinstrumente staatlicher und nicht-staatlicher Finanzierungspolitik, einen Aufschwung. Interessiert daran, den Einsatz finanzieller Mittel zu kontrollieren, geben heute Sozialversicherungsträger und Ministerien zunehmend Evaluationen in Auftrag, und die Entwicklung der Evaluationsforschung wird u.a. durch den Bundesrechnungshof forciert (Stockmann 2000, 30). Aber auch die wachsende Bedeutung von EU-Mitteln in der Sozialen Arbeit, deren Vergabe an Evaluationen gebunden ist, sowie die Finanzierung von Projekten durch Stiftungen und NGOs, die die Verwendung ihrer Gelder überprüfen wollen, tragen zum aktuellen Boom bei.

Im Einklang mit der Deregulierung sozialer Sicherungssysteme und deren Privatisierung und Ökonomisierung werden hierzulande über Evaluationen Kürzungsvorgaben durchgesetzt, Konkurrenzverhältnisse etabliert und die betriebswirtschaftliche Kosten-Nutzen-Rechnung für Bereiche eingeführt, die bisher davon ausgenommen waren (für das Bildungssystem vgl. Klausenitzer 2002). Die wachsende Bedeutung von Evaluationen ergibt sich aus dem Umstand, dass die angestrebte Kostenreduktion nicht mehr durch punktuelle Kürzungen zu leisten ist. Im neoliberalen Jargon wird dieser Politikwechsel als „Abschied vom Gießkannenprinzip“ gefeiert, an dessen Stelle eine gezielte Förderpolitik treten soll (so z.B. Jürgen Rüttgers in der *Zeit* v. 11.5.2005). Das bedarf der Entwicklung von Förderkriterien und Konzepten zur Kontrolle der geförderten Maßnahmen. Beides bieten auf wissenschaftlichem Niveau Evaluationen – so wirbt die Deutsche Gesellschaft für Evaluation (DeGEval 2005 u. 2002a; vgl. Wottawa/Thierau 2003, 21; Freytag 1998). Dass Kostenreduzierung und Kontrolle das vorrangige Ziel sind, wird offen benannt (z.B. Koch/Wittmann 1990, 3; Stockmann 2000; DeGEval 2002b).

Neben diesem Interesse an kostengünstiger Steuerung lässt sich für die Evaluationsforschung noch ein zweiter Bezugspunkt ausmachen: die Diskussion um die gesellschaftliche bzw. Praxisrelevanz von Wissenschaft und das Bemühen von Berufspraktikern in Pädagogik, Sozialer Arbeit und Psychologie, die eigene Tätigkeit in der Perspektive emanzipatorischer Neuausrichtung zu reflektieren. In den 70er und 80er Jahren entstanden, z.T. in direkter Anknüpfung an die Handlungs- und Aktionsforschung, in enger Kooperation zwischen Universitäten sowie pädagogischen und psychosozialen Einrichtungen zahlreiche Praxis- und Begleitforschungsprojekte. Forschungsgegenstand waren in der Regel aus der Antipsychiebewegung entstandene sozialpsychiatrische und

gemeindepsychologische Modellprojekte (Sieverding 1989, Bergold/Zaumseil 1989), Frauenhäuser, Jugendwohnen, Stadtteilprojekte, Beratungsstellen, Schülerläden etc. Ziel des kooperativen Forschungsprozesses war, aus der „Intuition“ (Beerlage/Fehre 1989) der Praktiker und der Rekonstruktion praktischer Prozesse zu lernen sowie in systematischer Reflexion die Praxis zu verbessern (a.a.O. 3f.; Heiner 1988a, 7) und gemeinsam an den politischen und fachlichen Fronten für gesellschaftliche Veränderung einzutreten. Hiltrud von Spiegel (1994), Carl Wolfgang Müller (1978 u. 1988) und Maja Heiner, die zahlreiche Sammelbände zu Praxisforschung (1988a) bzw. (Selbst-)Evaluation (1988b, 1994, 1996a, 1998a) in der Sozialen Arbeit herausgegeben hat, stellen im Gegensatz zur oben beschriebenen Tradition die Möglichkeiten von Evaluation zur Praxis- und Selbstreflexion sowie Qualifizierung im Interesse der Praktiker und Klienten in den Vordergrund. Sie machen partizipative (vgl. Ulrich/Wenzel 2003) und Selbstevaluation sowie formative Evaluationen stark, betonen die Notwendigkeit, die Praktiker direkt einzubeziehen, und setzen auf qualitative Methoden zur Datenerhebung. Ihre Vorgehensweise gilt in der Evaluationsforschung als einschlägig für bestimmte Fragestellungen. Allerdings werden quantitativ-statistisch orientierte Modelle viel häufiger angewendet, da sie weniger Geld und Zeit kosten – was auch bedeutet, keinen Mehrbedarf gegenüber Auftraggebern und beteiligten Praktikern legitimieren zu müssen.

Verstärkt durch das Wegbrechen sozialer Bewegungen lässt sich beobachten, dass die gesellschaftskritische Perspektive der Praxisforschungsprojekte schrittweise verschwindet. Schon in ihrem ersten Sammelband charakterisiert Heiner „Praxisforschung“ in Abgrenzung zur Handlungs- und Aktionsforschung durch eine „skeptischere Einschätzung der Veränderungsmöglichkeiten“, die von der „Zähigkeit der bestehenden Verhältnisse“ herrühre (1988a, 13). Die weitere Entwicklung lässt sich an den von Heiner gewählten Titeln ihrer Sammelbände verdeutlichen: Nach „Praxisforschung [...]“ (1988a) bzw. „Selbstevaluation in der Sozialen Arbeit“ (1988b) knüpft sie mit „Qualitätsentwicklung durch Evaluation“ (1996a) an die Qualitätsdebatte an und übernimmt in „Experimentierende Evaluation: Ansätze zur Entwicklung lernender Organisationen“ (1998a) aus der Betriebswirtschaft entlehnte Konzepte und die entsprechende Sprache. Dieser Band rückt nahe an Konzepte der aktivierenden Sozialpolitik heran, mit denen der Umbau der sozialen Sicherungssysteme seit den 90er Jahren legitimiert wird (vgl. den Beitrag zur Evaluation eines kommunalen Beschäftigungsprojekts von Buri/Schulze-Böing 1998); Heiner schreibt: „Organisationen sollten mit einem Minimum an Konsens, Zufriedenheit, Überfluss, Rationalität und Konsequenz auskommen, um beweglich zu bleiben und sparsam mit ihren Ressourcen umzugehen. [...] Zuviel Konsens, Zufriedenheit, Geld und Prinzipientreue lassen die Organisation träge werden. Selbst an die eigenen Ziele sollte man nicht zu sehr glauben, vielleicht wäre es ja besser,

sich neue zu setzen!“ (1998b, 49) Bleibt diese Entwicklung unreflektiert, kann sie dazu führen, dass auch die unter der Prämisse gesellschaftskritischer Reflexion entwickelten Ansätze der Evaluationsforschung dem neoliberalen Umbau des Sozialstaats und dem Kosten- und damit einhergehenden Konformitätsdruck nichts entgegenzusetzen haben.

Die Ausrichtung von Evaluationen an den Vorgaben neoliberaler Finanzpolitik hat jedoch nicht unmittelbar zur Konsequenz, dass emanzipatorisch-gesellschaftskritische Projekte besonders von Kürzungen betroffen sind. Häufig schneiden diese bei Evaluationen nach marktwirtschaftlichen Rationalisierungskriterien erfolgreich ab. Hintergrund ist, dass viele von ihnen als ehrenamtliche Projekte begonnen haben und erst später eine staatliche oder anderweitige kontinuierliche Finanzierung durchsetzen konnten, wobei stets ein Teil der Arbeit unbezahlt blieb. Da von Anfang an mit wenig Personal gearbeitet werden musste und zusätzliches, unbezahltes Engagement im Selbstverständnis der Beschäftigten fest verankert ist, sind diese Einrichtungen im Kosten-Nutzen-Vergleich häufig konkurrenzfähig. So liegen etwa die Kosten gemeindepsychiatrischer ambulanter oder teilstationärer Versorgungseinrichtungen unter denen der stationären Unterbringung in psychiatrischen Kliniken (vgl. Landratsamt Ludwigsburg 1999, 6). Zudem führte die ständig prekäre Finanzsituation der aus den Neuen Sozialen Bewegungen entstandenen Einrichtungen dazu, dass vielfältige Erfahrungen mit der kurzfristigen Akquirierung von Projektmitteln und mit Mischformen aus staatlicher und spendenbasierter, sog. Ko-Finanzierung gesammelt wurden – Modelle, die in anderen gesellschaftlichen Bereichen erst durch politische Kampagnen hoffähig gemacht werden mussten (zum Umbau des Hochschulsystems vgl. Bultmann 1997). Insofern lässt sich aus der Funktion von Evaluationen für die Umsetzung neoliberaler Bildungs-, Sozial- und Gesundheitspolitik nicht unmittelbar auf die Abschaffung emanzipatorisch-gesellschaftskritischer Ansätze schließen. Dennoch besteht, soweit die Qualitätskriterien, an denen eine Evaluation ausgerichtet ist, allein „Marktkriterien“ sind, die Gefahr, dass die „Qualität“ psychosozialer Arbeit auf diese reduziert wird und es zu einer entsprechend konzeptionell veränderten Praxis kommt, wie es bspw. bei Frauenhäusern und anderen feministischen Projekten im Zuge ihrer Professionalisierung zu beobachten ist (vgl. Hack 2004).

### *Qualitätsmanagement oder Lebensqualität?*

Die widersprüchliche Funktion der Evaluationsforschung zwischen (Kosten-)Kontrolle und Praxisreflexion zeigt sich in der seit Mitte der 90er Jahre geführten „Qualitätsdebatte“ in der psychosozialen Arbeit. Unter dem Druck von Sparvorgaben wird einerseits versucht, der Willkürlichkeit staatlicher Mittelstreichungen Kriterien entgegenzusetzen, die beschreiben, was Qualität in Sozialer Arbeit, Bildung, Psychothera-

pie etc. ausmacht und wie sich diese erheben lässt. Bei „stagnierenden finanziellen Ressourcen“ sei es „unverzichtbar, die Qualität der zu fördernden und der geförderten Projekte transparent zu machen“ (Scherrer 1996, 12); „gut gemeinte Intransparenz“ sei durch nichts mehr zu legitimieren (Heiner 1996b, 21). Dieser Versuch resultierte jedoch in einem Legitimierungszwang für die einzelnen Einrichtungen, nach dem Motto: Nur wer seine Qualität nachweisen kann, wird weiter bestehen. „Die Träger der Jugendhilfe sind daher gut beraten, selbst an der Qualitätsbewertung ihrer eigenen Arbeit zu arbeiten“ (Scherrer 1996, 12).

Andererseits wurde die „Qualitätsdebatte“ auch von den Geld gebenden Instanzen forciert. Von dieser Seite wird vielfach auf aus der Wirtschaft kommende Verfahren wie Total Quality Management (TQM), Controlling-Methoden und Zertifizierungsverfahren, etwa auf Grundlage der DIN EN ISO 9000ff., zurückgegriffen (vgl. Boeßenecker u.a. 2003). Diese zielen jedoch nicht auf eine inhaltliche Bestimmung von Qualität, sondern nur auf die formale Überprüfbarkeit zuvor festgelegter Abläufe. Im Fall der DIN EN ISO 9000ff., der Qualitätsnorm für Dienstleistungen, wird die „Rückverfolgbarkeit“ von Maßnahmen oder die „Abrufbarkeit des Fortschritts“ einer Intervention überprüft, nicht deren Angemessenheit für die vorliegende Problemstellung (Hack 2004, 546f.). Transparenz und Standardisierung der Abläufe werden gleichgesetzt mit „Qualität“. Dies geht soweit, dass unter dem Buchtitel „Qualitätskonzepte in der Sozialen Arbeit“ (Boeßenecker u.a. 2003) die Frage, was Qualität hier ausmacht, an keiner Stelle diskutiert wird. Stattdessen werden unterschiedliche Verfahren zur Zeiterfassung vorgestellt. Durch solche Standardisierungen soll möglich werden, fest umrissene Einzeldienstleistungen zu definieren, denen dann Zeiteinheiten und Preise zugewiesen werden (vgl. a. den Beitrag von Wulff in diesem Heft). Hierüber sollen die Einrichtungen untereinander vergleichbar werden, und das bildet die Grundlage für die unter dem Schlagwort „Ökonomisierung des Sozialen“ kritisierte Privatisierung z.B. der Altenpflege.

Die Übertragung der betriebswirtschaftlichen Kontrollverfahren auf die psychosoziale Versorgung ist nicht bruchlos möglich – und entsprechend ist Kritik an den Zertifizierungsverfahren geäußert worden: Nicht nur sind die Problemlagen von Fall zu Fall unterschiedlich, so dass z.B. der Ablauf eines Beratungsgesprächs in einer Familienberatungsstelle nur schwer vereinheitlicht werden kann, sollen die Standards nicht vollständig von den Beratungsinhalten abstrahieren. Auch die für die Festlegung von Qualitätsstandards im Produktionsbereich zentrale Kategorie der „Kundenzufriedenheit“ ist nicht einfach für die psychosoziale Versorgung zu übernehmen, in der der „Klient [...] weder Kunde noch König“, sondern „mehr oder minder freiwilliger Nutzer“ ist (Heiner 1996b, 29). So kann die Bewertung von Therapieerfolg durch die Beteiligten (Psychologin, Klient, Angehörige) sehr unterschiedlich ausfallen (bzgl. der Zertifizierung von Bildungsangeboten s. Becker 1996, 301).

Unter dem Primat des Kostendrucks ist jedoch nahe gelegt, dass die billigste Variante bevorzugt wird. „Vielleicht kann und will sich die Gesellschaft gegenwärtig nicht mehr als die billigsten Problemlösungsvarianten leisten. Aber auch ein solches Qualitätsverständnis, das entsprechende Qualitätsniveau sowie die damit verbundenen politischen Setzungen müssten offen gelegt werden, damit sie verhandelbar bleiben oder überhaupt erst verhandelbar werden.“ (Heiner 1996b, 28)

Die Evaluationsforschung verspricht, spezifische Bewertungsverfahren für die jeweiligen Praxiseinrichtungen zur Verfügung zu stellen und damit das Qualitätsproblem bearbeitbar zu machen. Irritierend bleibt, dass auch hier – bis auf wenige Ausnahmen, z.B. Heiner für den Bereich Sozialer Arbeit (1994, 9f.; 1996b) – die Frage nach einer inhaltlichen Bestimmung von Qualität psychosozialer Arbeit nicht gestellt wird. Stattdessen wird diese an die jeweils zuständigen Fachwissenschaften weiter delegiert. So bezieht sich die Evaluation von Kindertageseinrichtungen in Brandenburg mit dem programmatischen Namen „Qualität trotz Kostendruck“ (Schlecht/Tietze o.J.) auf „anerkannte Standardverfahren zur Erfassung pädagogischer Qualität“ wie die Kindergarten-Skala „KES“ (Tietze u.a. 1997, 2001a), die Krippen-Skala „KRIPS“ und die Hort-Skala „HOS“ (Tietze u.a. 2001b). Wessen Qualitätsansprüche diese entsprechen, ob von diesen die Praxis adäquat erfasst wird oder welche Vorstellungen von Erziehung und Betreuung von Kindern darin eingehen, ist dem Evaluationsprozess vorgelagert bzw. bleibt außen vor. Hier deutet sich bereits der formale Charakter der Evaluationsforschung an, deren mangelndes Gegenstandsverständnis nicht zuletzt methodische Probleme zur Folge hat (s.u.). Die aus der Perspektive Kritischer Psychologie für die Bewertung der Qualität psychosozialer Praxis zentrale Frage, inwieweit diese auf die Überwindung erniedrigender, ausbeuterischer und unterdrückender gesellschaftlicher Verhältnisse zielt und in diesem Sinne einer Verbesserung von Lebensqualität<sup>2</sup> zuarbeitet, oder aber dazu beiträgt, gesellschaftliche Behinderungen aufrecht zu erhalten und Unterdrückung „reibungloser“ zu gestalten, ist dabei nicht reflektiert. Es besteht die Gefahr, dass Qualitätsentwicklung auf Qualitätsmanagement mit beliebig wechselnden Kriterien reduziert wird, wobei die Frage nach der emanzipatorischen Perspektive im besten Fall offen bleibt.

---

<sup>2</sup> Zur Diskussion um die Schwierigkeiten einer „objektiven“ Beurteilung von Lebensqualität und der Möglichkeit der Instrumentalisierung des Begriffs für die Abwertung menschlichen Lebens als „unwert“ vgl. Schramme (2005). Wir verwenden den Begriff hier ausschließlich mit Bezug auf seinen utopischen Gehalt und den darin enthaltenen Verweis auf die Verbesserung der Lebensbedingungen. Vgl. a. Markard (2005).

### *Praxisrelevanz in der Evaluationsforschung und das Problem der Erhebung praxisrelevanter Daten*

Reflektiert und bewertet Evaluationsforschung mittels wissenschaftlicher Methoden Praxis, nimmt sie eine Vermittlungsposition zwischen dieser und der Wissenschaft ein. Zentraler Anspruch ist die Praxisrelevanz der Forschungsergebnisse. Darunter wird die Zweck- und Zielorientierung von Evaluationen (Wottawa/Thierau 2003, 14), die durch diese ermöglichte „Verhaltensoptimierung“ in Bezug auf komplexe Maßnahmen (21) sowie der „Nutzen“ für (23) bzw. die „Verwertung von Projektergebnissen durch den Auftraggeber“ verstanden (1990, 232). Entsprechend zielt sie auf die Steigerung der „Akzeptanz bei den potenziellen Auftraggebern“ (239). Dieses Verständnis von Praxisrelevanz entspricht dem, was Holzkamp in Anlehnung an Habermas (1965) als „technische Relevanz“ bezeichnet, die dann gegeben ist, wenn „wissenschaftliche Forschung durch die Angabe von Ausgangsbedingungen für das Auftreten bestimmter Effekte ‚erfolgskontrolliertes Handeln‘ [...] in ökonomischen, sozialen oder gesellschaftlichen Bereichen ermöglicht“ (1972, 18f.). Der technischen setzt Holzkamp die „emanzipatorische Relevanz“ (32) gegenüber. „Emanzipatorisch relevant wäre psychologische Forschung, sofern sie zur Selbstaufklärung des Menschen über seine gesellschaftlichen und sozialen Abhängigkeiten beiträgt und so die Voraussetzungen dafür schaffen hilft, dass der Mensch durch Lösung von diesen Abhängigkeiten seine Lage verbessern kann.“ (Ebd.) Unter dem Primat neoliberaler Gesundheits- und Sozialpolitik hat sich die Beschränkung auf „technische“ Praxisrelevanz durchgesetzt; das bedeutet den Rückzug auf eine Praxisentwicklung, die die institutionellen und gesellschaftlichen Rahmenbedingungen unangetastet lässt.

Doch selbst reduziert auf eine struktur- oder institutionenimmanente Optimierung ist „Praxisrelevanz“ für die Evaluationsforschung ein methodisches Problem, das mit dem Anspruch auf Wissenschaftlichkeit der Verfahren kollidiert. Ein Versuch, zu „praxisrelevanten“ Ergebnissen zu kommen, ist die Datenerhebung mit mehreren wissenschaftlichen Methoden (vgl. etwa Wottawa/Thierau 2003, 39, 73). Ein „multipler methodischer Zugang“ liefere „ein reichhaltiges und aussagekräftigeres Bild als ein monomethodischer Ansatz“ (Brandtstädter 1990, 219). Es sei ein Merkmal der Evaluationsforschung, dass sie methodisch „pluralistisch“ ist (BZgA 1999, 20), und auch nach Heiner lassen sich qualitative und quantitative Verfahren „ausgezeichnet [...] kombinieren“ (1998b, 31; vgl. 1988a, 10). Die Forderung nach methodischem Pluralismus ist gerade für Ansätze, die vom positivistisch-nomothetischen Mainstream an den Rand gedrängt werden, wissenschaftspolitisch wichtig, um nicht als unwissenschaftlich abqualifiziert zu werden. Doch die Vorstellung einer „gegenseitigen Kompensation qualitativer und quantitativer Methoden“ ist illusionär (Markard 1993, 116). Die Annahme, die

„Abstraktheit“ quantitativer Datenerhebung sei durch qualitative Einzelfallanalysen oder die geringe Fallzahl qualitativer durch quantitative Untersuchungen zu kompensieren, klammert „die begriffliche Klärung des Gegenstands“ aus (118). Ohne einen Begriff vom Gegenstand ist die Bestimmung der adäquaten Methode nicht möglich und damit die Praxisrelevanz der erhobenen Daten dem Zufall überlassen.

Dieses Problem soll umgangen werden, indem die unterschiedlichen Praxiszusammenhänge vermittelt über einen gemeinsamen Gegenstand erfasst werden: die „erfahrungsgestützten subjektiven Theorien des Auftraggebers“ (Wottawa/Thierau 1990, 231) oder der beteiligten Praktiker. Bei der Erhebung dieser „subjektiven Theorien“ sind die Evaluatoren jedoch damit konfrontiert, die Gütekriterien quantitativer Verfahren in der Praxis nicht einhalten zu können (vgl. Brandtstädter 1990, 217f.; Wottawa/Thierau 2003, 131f.). Wottawa/Thierau sprechen sich explizit für eine „stärkere Berücksichtigung von nicht im engsten Sinne wissenschaftlichen Techniken bei der Projektgestaltung und Durchführung“ aus (1990, 239). Statt die Gütekriterien auf ihre Angemessenheit hin zu hinterfragen, wird im Namen der Praxisrelevanz die Wissenschaftlichkeit der Erhebungsverfahren relativiert. Insofern in der Evaluationsforschung aber die Spezifik menschlicher Selbst- und Welterfahrung in ihrer gesamtgesellschaftlichen Vermitteltheit ungenügend theorisiert ist, steht sie bei der wissenschaftlichen Erfassung subjektiver Erfahrung vor ähnlichen Problemen, wie sie von der „Sozialpsychologie des Experiments“ thematisiert worden sind (Bungard 1984; vgl. Wottawa/Thierau 2003, 69). Statt den wissenschaftlichen Anspruch zu ermäßigen, wäre das methodische Vorgehen an den Grunddimensionen menschlicher Subjektivität zu orientieren (vgl. Holzkamp 1983, 510). Darüber hinaus ist die Aufschlüsselung personaler Erfahrung mit der Tatsache konfrontiert, dass das, was erfahrbar ist, „anschauliche und nicht anschauliche Aspekte“ hat (Markard 2000, 18). Erst durch die (gesellschafts-)theoretisch geleitete Rekonstruktion des Vermittlungszusammenhangs von individueller Erfahrung und gesellschaftlichen Bedingungen/Bedeutungen wird deutlich, welche Elemente der unmittelbaren Erfahrung – etwa aufgrund von beschränkten Einsichtsmöglichkeiten in die Struktur gesellschaftlicher Verhältnisse – verzerrt und ggf. zu problematisieren sind. Die Gleichsetzung der Berücksichtigung von subjektiven Theorien mit Praxisrelevanz beinhaltet die Gefahr, dass die unmittelbare Erfahrung der Praktiker oder gar Auftraggeber lediglich theoretisch verdoppelt, statt zum Ausgangspunkt wissenschaftlicher Analyse wird (vgl. Markard 2000).

### *Konformitätsdruck statt Praxisanalyse*

In der Evaluationsforschung wird versucht, hierauf mit der Entwicklung von Evaluationstheorien zu antworten, über die eine gegenstandsadäquate Methodik entwickelt werden soll. Das Bemühen steht jedoch den

eigenen Professionalisierungsbestrebungen entgegen. Denn Professionalisierung jenseits fachlicher Zersplitterung bedürfe allgemeingültiger Standards und einer „einzelne Evaluationsbereiche überschreitenden Vorgehensweise“ (Wottawa/Thierau 2003, 66; vgl. Stockmann 2000, 32). Die Vielfalt der Evaluationsgegenstände – von Personen über Umwelt-/Umgebungsfaktoren, Produkte, Techniken/Methoden bis hin zu Programmen und Forschungsergebnissen (vgl. Wottawa/Thierau 2003, 59) – erschwert es, eindeutige methodische Vorgaben für die Evaluationsforschung als Ganze zu formulieren. Stattdessen wird die Frage, „wann, wo und welche Methoden angewendet werden sollen“, zur „erst im Evaluationsprozess zu lösenden Aufgabe“ (Westermann 2002, 22).

Programmtheorien und Evaluationsmodelle werden entwickelt, die den allgemeinen Evaluationstheorien untergeordnet sind und die für einzelne Gegenstände als angemessen gelten (für eine Übersicht vgl. Wottawa/Thierau 2003, 58; Stufflebeam et al. 2000). Um zu entscheiden, welches Modell für die jeweilige Fragestellung bzw. den Gegenstand adäquat ist, bedarf es eines „Vorwissens“ über Entwicklung, administrativen und politischen Kontext und die Struktur des Evaluationsgegenstands (Westermann 2002, 8). Neben einer „guten Allgemeinbildung“ verweisen Wottawa/Thierau (2003, 51) auf verschiedene Techniken zur „Konkretisierung des Evaluationsobjektes“: „Literaturanalyse, Hospitationen, Experteninterviews“; „Gruppendiskussion“ sowie „Brain-Storming“-Techniken“ (87). Deren Zweckmäßigkeit ist jedoch umstritten, denn die „vagen, globalen, umgangssprachlich gefassten Zielformeln oder Slogans (,Verbesserung der beruflichen Qualifikation’, ,psychosoziale Kompetenz’, ,Verbesserung des Betriebsklimas’, ,Förderung des Gesundheitsverhaltens’, ,Chancengleichheit’, ,Steigerung des Lernerfolgs’)", mit denen Evaluatoren konfrontiert sind, lassen sich „nicht durch operationale Definitionen“ zu messbaren Zielen konkretisieren (Brandtstädter 1990, 221): „um die Angemessenheit von Operationalisierungsvorschlägen beurteilen zu können, bedarf es schon eines Vorverständnisses der fraglichen Konzepte“ (ebd.).

Zwar werden bei der Evaluation von Bildungsprogrammen Annahmen über menschliches Lernen im Rahmen von Programmtheorien expliziert. Auch in einigen Evaluationskonzepten für Soziale Arbeit (vgl. Spiegel 1994, Landgrebe 1994, Heiner 1996b u. 1998b) werden solche Vorannahmen thematisiert. Deren „Auswahl“ wird jedoch – wie die Bestimmung von Qualität – explizit nicht in der Evaluationsforschung vorgenommen, sondern bleibt dem theoretischen Standpunkt der Evaluatoren bzw. den Vorgaben der Auftraggeber überlassen. Der Umstand, dass Theorien wissenschaftlich umstritten sein und verschiedene Konzepte und Paradigmen im wissenschaftlichen Feld konkurrieren können, wird nicht reflektiert. Stattdessen wird die durch Theorienkonkurrenz aufgeworfene Frage des Gegenstandsverständnisses auf die Frage der Organisation von Kommunikationsprozessen verschoben, die zentrale

Voraussetzung für den Beruf des Evaluators sei (vgl. z.B. Wottawa/Thierau 2003, 51). Es sei Aufgabe der Evaluatoren, „die unterschiedlichen Beziehungs- und Rollenkonflikte zu bearbeiten und die dahinterstehenden Interessensdivergenzen auszubalancieren“, was ein „hohes Maß an persönlicher Ambiguitäts- und Frustrationstoleranz“ erfordere (Frank/Seifert 1998, 191). Was theoretisch zu klären wäre, wird den Evaluatoren als Problem der (Berufs-)Fähigkeit „in die Schuhe“ geschoben.

Zur Orientierung wird oberflächlich auf „die Kenntnis der entsprechenden Theorien und Befunde“ (Wottawa/Thierau 2003, 38) sowie die „Verantwortung des Wissenschaftlers, auch bezüglich der eigenen Informiertheit über den aktuellen Sachstand“, verwiesen (39), wobei die „jeweils erforderlichen, spezifischen Fachkenntnisse [...] sich, entsprechendes Verständnis und Unterstützung durch den Arbeitgeber vorausgesetzt, ‚on the job‘ erwerben“ lassen (51). Diese Aneignung wird schnell zur unkritischen Orientierung am wissenschaftlichen Mainstream – insbesondere, wenn die eigene „Ambiguitätstoleranz“ bereits ausgereizt ist. Um die eigenen Evaluationsergebnisse zu stärken, raten Wottawa/Thierau zur „Konsensbildung“ mit anderen Wissenschaftlern bzgl. der Auswertungsstrategie (159). Auseinandersetzungen unter Evaluatoren würden die Gefahr bergen, „negative Eindrücke über wissenschaftliche Evaluationen“ zu erzeugen, indem sie Vorwürfen von „Kaufbarkeit“, „fachwissenschaftlicher Inkompetenz“ und unnötigen „Kosten der Evaluationsprojekte“ Vorschub leisten (20) und damit dem Ruf der Evaluationsforschung schaden. Gesellschaftskritik, die unter den bestehenden Macht- und Herrschaftsverhältnissen potenziell umstritten ist, wird unter der Hand als subjektiv, interessengeleitet und inkompetent diffamiert. Der Konformitätsdruck innerhalb der Evaluationsforschung wird dadurch verschärft, dass die Umstrittenheit von Gegenstand oder Sachverhalten unter den an einer Evaluation Beteiligten durch Druck in Richtung Konsensbildung mit dem Auftraggeber „aufgelöst“ wird, also quasi durch dessen Setzung (vgl. Wottawa/Thierau 2003, 158).

### *Interessenwidersprüche im Evaluationsprozess und Personalisierung als Konfliktmanagement*

Zur Gewinnung valider und praxisrelevanter Daten ist die Evaluationsforschung auf die Kooperation der Praktiker vor Ort angewiesen, die die Fragebögen ausfüllen oder Zeit für Gruppengespräche und Interviews aufbringen müssen. In dem Maße, wie diese Kontrollfunktionen, Ökonomisierungstendenzen und Kostendruck wahrnehmen, die Evaluation in letzter Konsequenz etwa die Einstellung der Förderung oder sogar Schließung einer Einrichtung bedeuten kann, ergibt sich die Schwierigkeit, eine Beteiligung gegen deren unmittelbare Interessen – etwa Abwendung von Arbeitsplatzverlust – zu erreichen. Diese Konfliktsituation

konstituiert ein potenziell instrumentelles Verhältnis der Praktiker zu den Evaluatoren, insbesondere wenn sich diese an der Evaluation lediglich aus Gründen der Legitimation der Einrichtung gegenüber dem Geldgeber beteiligen (Boebenecker 2003, 9). Die Praktiker liefern dann keine oder nur sozial erwünschte Daten und stellen die Relevanz der Evaluationsergebnisse in Frage; Vorschläge für eine veränderte Praxis werden nicht „als handlungsrelevant angesehen und im konkreten Feld adaptiert“ (ebd.). Die Probleme, die aus den unterschiedlichen, sich z.T. widersprechenden Perspektiven und Interessen von Auftraggebern, Praktikern und „Nutzern“ ergeben, werden in der Evaluationsforschung viel diskutiert (Wottawa/Thierau 2003, 38). In den „Standards“ der DeGEval wird darauf hingewiesen, dass im Sinne der „Durchführbarkeit“ ein „diplomatisches Vorgehen“ erforderlich ist, um „eine möglichst hohe Akzeptanz der verschiedenen Beteiligten und Betroffenen in Bezug auf Vorgehen und Ergebnisse“ zu erreichen (2002a, 9). Aus Gründen der „Fairness“ sollen Evaluationen „unparteiisch“ durchgeführt werden und „unterschiedliche Sichtweisen von Beteiligten und Betroffenen auf Gegenstand und Ergebnisse der Evaluation in Rechnung stellen“ (10). Heiner schreibt daher den Evaluatoren die Rolle der neutralen Moderation „zwischen den fachlichen und politischen Fronten“ (1988a, 13) zu.

Eine Umgangsstrategie mit den Interessenkonflikten ist die Verwendung von partizipativen Evaluationsmodellen, die die Einbeziehung der Beteiligten in den Forschungsprozess vorsehen. Die Widersprüche lösen sich jedoch durch die „Partizipation“ der Interessengruppen nicht auf, solange Evaluationen auf einem Markt bestehen müssen. Vielmehr werden die Interessen der Auftraggeber faktisch zum Maßstab. Dies wird z.B. deutlich bei der Evaluation des Schulprojekts „Farbe Bekennen“. Hupfeld (2004) erläutert, über „kooperative Zielbestimmung und Absprache in Hinblick auf die einzusetzenden Methoden“ das „notwendige Engagement und Commitment bei allen direkt und indirekt Betroffenen fördern“ zu wollen (10). Trotzdem berichtet er von massiven Konflikten bei der Durchführung der Evaluation, die aus einer grundsätzlichen Befürchtung des Kollegiums, es handele sich um externe Kontrollmaßnahmen, resultierten (60). Die Auseinandersetzungen spitzten sich um die Frage des Nutzens der summativ ausgerichteten Evaluation für die Schule und der Aussagekraft des quantitativen Forschungszugangs (Einstellungsmessungen der Schüler zu Beginn und zum Ende der Projektlaufzeit) zu. Dem wurde von Seiten der Forscher entgegengehalten, dass eine formative Evaluation aufgrund des eng gesteckten Kostenrahmens nicht möglich gewesen sei.

Gitte Landgrebe machte dagegen mit ihrer zeit- und kostenintensiven „Kombination von Verfahren der Fremd- und Selbstevaluation in der Familienberatung und der Gruppenarbeit mit Frauen“ (1994) die Erfahrung, dass ihre Forschungen von den Beteiligten akzeptiert wurden. Sie arbeitete eng mit den Praktikerinnen vor Ort zusammen und stimmte die

Datenerhebungsmethoden (Zeiterfassungsbögen, Fallberichte, Protokolle der Gruppensitzungen, Befragungen der Pädagoginnen und Tiefeninterviews mit Klientinnen) mit ihnen ab. Doch auch auf diesem scheinbaren „Königsweg“ (ebd.) werden die grundsätzlichen Interessenwidersprüche nicht suspendiert. Partizipative und selbstevaluative Verfahren bieten weniger eine selbstbestimmte Forschung im Sinne der Erweiterung von Handlungsfähigkeit für die Betroffenen als vielmehr eine Form, in der organisiert wird, wie die jeweiligen Anforderungen (z.B. die Kosten zu senken) von den Praktikern zu übernehmen sind. Sie können so zum Mittel der Befriedung und Konfliktlinderung bzw. -vermeidung werden. Monika Sieverding (1989) schildert die Schwierigkeiten von Praxisforschung unterm Primat des Kostendrucks im Modellverbund „Ambulante psychiatrische und psychotherapeutische/psychosomatische Versorgung“. Sie zeigt, wie die Praktiker sie in internen Machtkämpfen zu instrumentalisieren, zur „Hofberichterstatte“ zu machen versuchten, und wie sie die Veröffentlichung kritischer Daten und damit eine öffentliche Auseinandersetzung mit ihrer Praxis zu verhindern suchten, indem sie ihre Qualifikation in Frage stellten.

Die Versuche, durch partizipative Evaluationen und die Berücksichtigung der verschiedenen Interessengruppen das Problem der Interessenwidersprüche zu lösen, sind unter solchen Bedingungen tendenziell zum Scheitern verurteilt. Statt die Probleme als Ausdruck der widersprüchlichen Funktion von Evaluation und der institutionellen Bedingungen zu reflektieren, werden sie durch eine doppelte Personalisierung „gelöst“. Zum einen wird eine personalisierende Sichtweise der auf die Interessenwidersprüche zurückgehenden Probleme und Konflikte zwischen Forschern und Praktikern nahegelegt. Theo Schneid leitet seinen Artikel zu Evaluation in der Jugendarbeit mit der Beobachtung ein, dass Evaluation „zu Unsicherheit und auch zu Abwehrreaktionen“ führe (2001, 10). Im gleichen Sammelband gibt Beywl Tipps, wie Evaluatoren bei den Praktikern „persönliche Ängste vor der empirischen Evaluation mindern“ könnten (2001, 21). Zum anderen werden Widersprüche und Konflikte, die sich aus der gleichzeitigen Verpflichtung dem Auftraggeber, den Praktikern und Nutzern sowie einer „wissenschaftlichen Redlichkeit“ gegenüber ergeben, in der Evaluationsforschung ebenfalls in personalisierender Weise verhandelt. Diese Konflikte, die besonders durch die Marktförmigkeit der Forschung und die spezifische Art der Bestechlichkeit durch den Auftraggeber geprägt sind, werden von Wottawa/Thierau zwar benannt, aber nicht problematisiert: „Jeder, der eine berufliche Tätigkeit im Bereich der Evaluation anstrebt, sollte sich bewusst sein, dass er damit eben nicht im engeren Sinne wissenschaftlich tätig ist. Er ist weniger frei, stärker an die Berücksichtigung von Interessen gebunden und gleichzeitig weniger folgenlos als viele Mitarbeiter im Grundlagenbereich“ (2003, 38). Konsequenterweise wird die Verantwortung dafür, dass diese Konflikte im Forschungsprozess nicht aufbrechen und

trotz der strukturellen Behinderungen valide Ergebnisse zustande kommen, den Evaluatoren zugeschrieben. „Persönliche Voraussetzungen“ für diesen Beruf seien neben „fachlichem Können“ Aspekte, die „Persönlichkeit, Interessen [und] Motivation“ betreffen: Kommunikationsfähigkeit und die „Fähigkeit zum role-taking“ (51), womit das Einfühlen in die verschiedenen Interessenlagen und deren Vermittlung gemeint ist.

### *Überwältigende Zwänge?*

Aus der Perspektive kritisch-psychologischer Praxisforschung (vgl. Markard/Holzkamp 1989, Fahl/Markard 1993, Markard & ASB 2000) stellt sich die Frage: Wie ist mit dem Feld der Evaluationsforschung umzugehen? Einige Fragen letzterer, genauer der formativen und partizipativen Evaluationsmodelle, ähneln zentralen Punkten ersterer: der Anspruch wissenschaftlicher Reflexion psychologischer Praxis, die Einbeziehung der Betroffenen in den Forschungsprozess, die Betonung des Zusammenhangs von Praxis- und Theorieentwicklung mit dem Versuch der Überwindung des Theorie-Praxis-Bruchs sowie die Bedeutung, die den subjektiven bzw. Praxistheorien der Berufspraktiker für die wissenschaftliche Forschung eingeräumt wird. Bei genauerem Hinsehen werden jedoch folgenreiche Unterschiede deutlich. Insbesondere fehlt der Evaluationsforschung die für die Kritische Psychologie zentrale Perspektive auf die Überwindung gesellschaftlicher Herrschaftsverhältnisse und auf eine Gesellschaft, in der kein Mensch „ein erniedrigtes, ein geknechtetes, ein verlassenes, ein verächtliches Wesen ist“ (Marx 1844, MEW 1, 385).

Stattdessen gibt sich die Evaluationsforschung gerade neutral (vgl. Wottawa/Thierau 2003, 15; DeGEval 2002a, 10) und verschweigt die Machtverhältnisse, in denen sie wirkt. Das hat unmittelbare Auswirkungen auf die Forschung. Ihre Funktionen der Kontrolle, der Durchsetzung des Kosten- und des Aufbaus eines Konformitätsdrucks bei der neoliberalen Umstrukturierung der psychosozialen Arbeit produziert ein Bedingungsgefüge, welches kritische Praxisreflexion erschwert. Besonders deutlich wird das an dem Versuch, die Praktiker auch potenziell gegen ihre Interessen zur Kooperation zu bewegen, während der Evaluator, der dem Auftraggeber verpflichtet ist, den Kooperationsrahmen mit den Praktikern durch seine Position permanent aufkündigt. Wird diese strukturelle Einbindung nicht begriffen, erscheinen Schwierigkeiten v.a. als methodisch bedingt und resultieren in einer Unübersichtlichkeit von Konzepten und Ansätzen sowie in der von den Evaluationsforschern selbst beobachteten Beliebigkeit und Folgenlosigkeit der Ergebnisse. Patton (1987) bezeichnet das als „politisches Erbe“ der Evaluationsforschung. Praxisreflexion und Qualitätsentwicklung verkommen zu Lippenbekenntnissen, die der Legitimation der „Ökonomisierung des Sozialen“ – sei es die Privatisierung sozialer Einrichtungen oder die Simulation

von Marktkonkurrenz innerhalb der öffentlichen Versorgungsstruktur – und dem Geschäft mit dem Evaluationsbedarf dienen.

Damit stellt sich die Frage, ob im Rahmen von Evaluationen überhaupt Praxisreflexion möglich ist. Wir sind der Ansicht, dass diese trotz aller Zwänge und Interessenkonflikte stattfindet, dass Evaluationen für das Nachdenken über die eigene Praxis zeitlich und gedanklich Räume eröffnen können. Sie produzieren u.a. Daten – z.B. eine Übersicht darüber, wie die insgesamt oder individuell verfügbare Arbeitszeit auf die einzelnen Aufgabenbereiche verteilt ist –, die die Einsicht in die eigene Arbeitsstruktur erhöhen (vgl. Landgrebe 1994). Gerade formative partizipative Evaluationen, die auf die Explikation des Konzepts einer Einrichtung oder sogar impliziter Praxistheorien drängen und womöglich auf Theorienkonkurrenz und Brüche in der Konzeptstringenz aufmerksam machen, können helfen, Einblick in die eigene, manchmal als blindes Umhertasten erscheinende Praxis zu ermöglichen. Die systematische Beteiligung der von einer Maßnahme Betroffenen kann den Anstoß für eine veränderte Praxis geben und der Evaluationsprozess zum organisatorischen Rahmen für die Auseinandersetzung mit deren Sichtweisen werden. Die Explikation der subjektiven Theorien der Praktiker, in die nicht nur akademische psychologische (bzw. pädagogische etc.) Theorien, sondern auch Alltagsvorstellungen, persönliche Lebens- und Berufserfahrungen usw. eingehen (vgl. Fahl/Markard 1993, 11), kann zudem als Ausgangspunkt einer kritischen Diskussion der Praxis dienen. Ein Bezug auf die Kategorien der Kritischen Psychologie könnte – auch wenn deren Konkretisierung für die meisten Bereiche psychologischer Praxis aktuell noch zu leisten ist – dabei ermöglichen, Praxiserfahrungen und implizite Theorien der Praktiker auf einem ausgewiesenen theoretischen Hintergrund zu analysieren sowie für die Frage nach Qualität inhaltlich begründete Kriterien zu entwickeln, statt affirmativ beliebige vorfindliche (Mainstream-)Theorien unkritisch zu übernehmen.

Selbst wenn im Rahmen formativer partizipativer Evaluationsmodelle psychologie- oder gesellschaftskritische Fragestellungen einfließen, heißt dies noch lange nicht, dass die Ergebnisse Eingang in die Evaluationsberichte finden. Auch wenn die Zensur von Berichten als wissenschaftlich unredlich angesehen wird, ist es allgemein üblich, dass nur die Daten veröffentlicht werden, die dem Auftraggeber passen (vgl. Wottawa/Thierau 2003, 37). Praxisreflexion und Qualitätsentwicklung im Sinne der Verbesserung der Lebensbedingungen verkommen dabei mal offener, mal unter der Hand zu einem (zu vermeidenden?) Desiderat von Evaluation, statt ihr Ziel zu sein. Den Evaluatoren bleibt hierbei wenig Spielraum. Nehmen sie die kritischen Stimmen in den Evaluationsbericht, laufen sie Gefahr, zukünftig schwerer an Aufträge zu kommen. Zugleich müssen sie davon ausgehen, dass die Berichte vor der Veröffentlichung umgeschrieben werden, im Zweifel also die eigene Redlichkeit in der Berichterlegung niemandem kenntlich wird. Sind die

Evaluatoren vertraglich zur Verschwiegenheit verpflichtet, können sie mit ihren anders lautenden Ergebnissen auch nicht an die Öffentlichkeit gehen. Um ihre Erkenntnisse doch für die Betroffenen nutzbar zu machen, bliebe ihnen nur eine „doppelte Buchführung“, d.h. einerseits die von ihnen untersuchte Praxis mit den Berufspraktikern und Einrichtungsnutzern zusammen zu reflektieren und andererseits einen passenden, eben angepassten Evaluationsbericht für den Auftraggeber zu verfassen. Das bedeutet jedoch eine zusätzliche Belastung, da das Auseinanderhalten der inoffiziellen und der offiziellen Version besonderer Anstrengung bedarf, zumal letztere mit der ganzen institutionellen Macht des Finanzgebers ausgestattet ist. Aus doppelter Buchführung ergibt sich zudem eine spezifische Form von Erpressbarkeit, d.h. sie bedarf eines vertrauensvollen Rahmens. Dieser ist jedoch durch die institutionelle Aufgabe von Evaluationen, Kontrolle und Bewertung von Einrichtungen u.ä. zu leisten, gerade nicht gegeben. Inwieweit das Bedingungsgefüge den Evaluatoren oder beteiligten Berufspraktikern Selbstzensur nahe legt, weitergehende Fragen bereits im Kopf abgeschnitten werden oder gerade auch bei den Praktikern kein Interesse an Problematisierungen besteht, die ihre eigene Praxis in Frage stellen könnten, ist im Einzelnen aktual-empirisch zu erforschen.

Punktuell besteht vielleicht die Möglichkeit, sich mit den Berufspraktikern gegen den Kostendruck zusammenzuschließen. Dann wird aber das Infragestellen der Praxistheorien und „Schienen“, die ggf. zur Perpetuierung problematischer Praxis in der Einrichtung beitragen, schwierig (vgl. Sieverding 1989). Grundsätzlich ist unter den Bedingungen der Kontrolle naheliegend, dass Praktiker ihre Vorgehensweisen und Konzepte rechtfertigen. Diese als Evaluator zu problematisieren, kann bei der unter restriktiver Sozialpolitik vorherrschenden repressiven Funktion von Evaluationen schnell als Kritik am Fortbestehen der Einrichtung und damit als Loyalitätsbruch verstanden und wirksam werden. Die Wahrung des intersubjektiven Verständigungsrahmens, nach Holzkamp notwendige Voraussetzung für subjektwissenschaftliche Aktualempirie (1983, 541; vgl. 540ff.) und somit Praxisforschung, ist unter diesen Bedingungen permanent gefährdet.

Letztlich wäre es vielleicht eine Perspektive für Evaluatoren, ihre Arbeit darauf auszurichten, mit den Berufspraktikern und den Klienten der Einrichtungen die Kompetenz zu entwickeln, die Zusammenhänge zu analysieren, in denen die Evaluation stattfindet, so dass Praktiker wie Klienten sich bewusster zu den Evaluationsvorgängen verhalten können. Die Einbeziehung der Perspektive der Klienten ist dabei eine Grundvoraussetzung. Die Evaluatoren selbst müssen sich die Marktförmigkeit des Feldes und den Beitrag, den ihre Tätigkeit zur Durchsetzung der Ökonomisierung und Konkurrenzförmigkeit von psychosozialer Arbeit leistet, bewusst machen. Hier eröffnet sich ein empirisches Feld für Praxisforschungsprojekte mit Evaluatoren, in denen spezifische

Problemkonstellationen herausgearbeitet und Denkformen kritisch reflektiert werden könnten<sup>3</sup>. Perspektive eines solchen Projekts müsste sein, Ökonomisierung, Kosten- und Konformitätsdruck als institutionelle Beschränkungen der Praxis selbst zu Gegenständen der Forschung zu machen und sie nicht blind durchzusetzen oder zu ihrer Verschärfung beizutragen (z.B. durch die Förderung von Konkurrenz zwischen verschiedenen Einrichtungen). Das würde auch bedeuten, die geforderte Neutralität der Evaluationsforschung gegenüber den Konzepten der zu evaluierenden Praxis aufzukündigen sowie die Kooperation mit dem Auftraggeber ggf. in Frage zu stellen und die institutionellen und gesellschaftlichen Bedingungen des beruflichen Handelns als Evaluator wie die Voraussetzungen und Vorgaben neoliberaler Sozial- und Gesundheitspolitik und ihre Bedeutung für die Handlungsmöglichkeiten der davon Betroffenen mitzureflektieren. Eine solche Wende vom Evaluationsgeschäft zur Evaluations- bzw. Praxisforschung würde nicht nur eine gesellschaftlich-emanzipatorische, sondern auch eine wissenschaftlich adäquatere Form der Praxisreflexion und Qualitätsentwicklung ermöglichen – angefangen bei der Überwindung der theoretischen Beliebigkeit, über die wissenschaftliche (objektive) Rekonstruktion subjektiver Erfahrung bis hin zur Problematisierung von Personalisierungs- und Psychologisierungstendenzen. Schließlich könnte sie helfen, einen Kritikstandpunkt zu entwickeln, von dem aus der neoliberalen Evaluationsoffensive praktisch etwas entgegengesetzt werden kann.

### Literatur

- Becker, A. (1996). Qualität durch Zertifizierung in der außerschulischen Jugend- und Erwachsenenbildung. In: Heiner 1996a, S. 291-320
- Beerlage, I. & Fehre, E.-M. (Hg.) (1989). *Praxisforschung zwischen Intuition und Institution*. Tübingen: DGVT
- Bergold, J. B. & Zaumseil, M. (1989). Forschungsdienst Wedding: Versuch der Entwicklung eines Gemeindepsychologischen Forschungszugangs. In: Beerlage/Fehre 1989, S. 137-151
- Beywl, W. (2001). Evaluation in der Jugendverbandsarbeit einführen – Hindernisse und Lösungsansätze. In: Ders. et al. (Hg.). *Evaluation im Alltag. Jugendverbände untersuchen ihre Wirkungen* (S. 17-27). Münster: Votum
- Boeßenecker, K.-H. et al. (Hg.) (2003). *Qualitätskonzepte in der Sozialen Arbeit. Eine Orientierung für Ausbildung, Studium und Praxis*. Weinheim: Beltz
- Brandtstädter, J. (1990). Evaluationsforschung: Probleme der wissenschaftlichen Bewertung von Interventions- und Reformprojekten. *Zeitschrift für Pädagogische Psychologie* 4, S. 215-227
- Bultmann, T. (1997). Die standortgerechte Dienstleistungshochschule. *Prokla* 104, S. 329-355
- Bundeszentrale für gesundheitliche Aufklärung (BzgA) (Hg.) (1999). *Evaluation – Ein Instrument zur Qualitätssicherung in der Gesundheitsförderung*. Köln: Selbstverlag

<sup>3</sup> Wenn mit Evaluation befasste Praktiker hieran Interesse haben, würden wir uns freuen, wenn sie sich bei der FKP-Redaktion melden.

- Bungard, W. (1984). *Sozialpsychologische Forschung im Labor. Ergebnisse, Konzeptualisierungen und Konsequenzen der sogenannten Artefaktforschung*. Göttingen: Hogrefe
- Buri, Ch. & Schulze-Böing, M. (1998). Evaluation als Programmoptimierung. Das Beispiel eines kommunalen Beschäftigungsprojektes. In: Heiner 1998a, S. 125-148
- Campbell, D. T. (1969). Reform als Experiment. *American Psychologist*, 24. Jg., H. 4, S. 409-439
- Candeias, M. (2004). *Neoliberalismus – Hochtechnologie – Hegemonie. Grundrisse einer transnationalen kapitalistischen Produktions- und Lebensweise. Eine Kritik*. Hamburg: Argument
- Cook, T.D. & Campbell, D.-T. (1979). *Quasi-Experimentation. Design and Analysis Issues for Field Settings*. Chicago: Rand McNally
- Deutsche Gesellschaft für Evaluation (DeGEval) (2002a). *Standards für Evaluation*. Köln: Selbstverlag
- Dies. (2002b). *Presseerklärung vom 19.02.2002 – Evaluationsboom in Wissenschaft und Politik – Fachgesellschaft verbreitet Qualitätsstandards*. [www.degeval.de/index.php?class=Calimero\\_Webpage&id=9014](http://www.degeval.de/index.php?class=Calimero_Webpage&id=9014) (17.5.2006)
- Dies., Arbeitskreis „Evaluation von Entwicklungspolitik“ (Hg.) (2005). *Lernen aus Evaluationen. Leitfaden für Geberinstitutionen und GutachterInnen*. O.O.
- Fahl, R. & Markard, M. (1993). Das Projekt „Analyse psychologischer Praxis“ oder: Der Versuch der Verbindung von Praxisforschung und Psychologiekritik. *Forum Kritische Psychologie* 32, S. 4-35
- Frank, K. & Seifert, B. (1998). Evaluation von Modellprogrammen oder: Der Spagat der Kröte. In: Heiner 1998a, S. 167-193
- Freytag, C. (1998). Erfolgskontrolle mit System. Evaluation wird in Firmen und Ämtern immer wichtiger. *Kölner Stadt-Anzeiger*, 1.10.1998
- Graumann, S. & Grüber, K. (Hg.) (2005). *Anerkennung, Ethik und Behinderung*. Münster: LIT
- Guba, E. G. & Lincoln, Y. S. (1989). *Fourth Generation Evaluation*. Newbury Park/London/New Delhi: Sage
- Habermas, J. (1965). Erkenntnis und Interesse. *Merkur. Deutsche Zeitschrift für europäisches Denken* 12, S. 1139-1153
- Hack, E.-K. (2004). Vom effektiven Widerstand zur effizienten Anpassung. *Das Argument* 256, 46. Jg., H. 4/5, S. 542-553
- Heiner, M. (Hg.) (1988a). *Praxisforschung in der Sozialen Arbeit*. Freiburg: Lambertus
- Dies. (Hg.) (1988b). *Selbstevaluation in der Sozialen Arbeit. Fallbeispiele zur Dokumentation und Reflexion beruflichen Handelns*. Freiburg: Lambertus
- Dies. (Hg.) (1994). *Selbstevaluation als Qualifizierung in der Sozialen Arbeit*. Freiburg: Lambertus
- Dies. (Hg.) (1996a). *Qualitätsentwicklung durch Evaluation*. Freiburg: Lambertus
- Dies. (1996b). Evaluation zwischen Qualifizierung, Qualitätsentwicklung und Qualitätssicherung. In: Dies. 1996a, S. 20-47
- Dies. (Hg.) (1998a). *Experimentierende Evaluation: Ansätze zur Entwicklung lernender Organisationen*. Weinheim/München: Juventa
- Dies. (Hg.) (1998b). Lernende Organisation und Experimentierende Evaluation. Verheißungen Lernender Organisationen. In: Dies. 1998a, S. 11-53
- Hirsland, A. & Uhl, K. (o.J.). *Evaluationsmethoden für Modelle zur Förderung von Zivilcourage in der politischen Bildung*. (Unveröff. Ms.)
- Holzamp, K. (1972). Zum Problem der Relevanz psychologischer Forschung für die Praxis. In: Ders., *Kritische Psychologie. Vorbereitende Arbeiten* (S. 9-34). Frankfurt a. M.: Fischer
- Ders. (1983). *Grundlegung der Psychologie*. Frankfurt a. M.: Campus
- Hupfeld, J. (2004). *Bericht der Begleitforschung zum Teilprojekt „Farbe bekennen: Für Toleranz und Verständigung (Wiesmoor)“ im Rahmen des Modellvor-*

- habens Jugendarbeit zur Intensivierung der „Prävention gegen Rechts“ des Landeskriminalamtes Niedersachsen.* Hannover: arpos Institut e.V.
- Klausenitzer, J. (2002). Altes und Neues. Anmerkungen zur Diskussion über die gegenwärtige Restrukturierung des deutschen Bildungswesens. *Widersprüche* 83. Zit.n. [www.widersprueche-zeitschrift.de/article98](http://www.widersprueche-zeitschrift.de/article98) (17.5.2006)
- Koch, U. & Wittmann, W. (Hg.) (1990). *Evaluationsforschung. Bewertungsgrundlage von Sozial- und Gesundheitsprogrammen.* Berlin/Heidelberg/New York: Springer
- Landgrebe, G. (1994). Kombination von Verfahren der Fremd- und Selbstevaluation in der Familienberatung und der Gruppenarbeit mit Frauen. In: Heiner 1994, S. 78-112
- Landratsamt Ludwigsburg – Amt für Soziale Aufgaben (Hg.) (1999). *Wegweiser und Handlungskonzept für die Versorgung psychisch kranker und behinderter Menschen im Landkreis Ludwigsburg.* Ludwigsburg: Selbstverlag
- Madaus, G. F., Scriven, M. & Stufflebeam, D. L. (Hg.) (1983). *Evaluation Models.* Boston: Kluwe-Nijhof
- Markard, M. (1993). *Methodik subjektwissenschaftlicher Forschung.* Hamburg/Berlin: Argument
- Ders. (2000). Einleitung: Praxisausbildung im Hauptstudium oder die Frage nach den Umständen, unter denen man aus Erfahrung klug werden kann. In: Ders. & Ausbildungsprojekt Subjektwiss. Berufspraxis 2000, S. 9-27
- Ders. (2005). Gesellschaftliche Bedingungen und die Lebensqualität psychisch Kranker. In: Graumann/Grüber 2005, S. 99-112
- Ders. & Ausbildungsprojekt Subjektwissenschaftliche Berufspraxis (Hg.) (2000). *Kritische Psychologie und studentische Praxisforschung.* Hamburg/Berlin: Argument
- Ders. & Holzkamp, K. (1989). Praxis-Portrait. Ein Leitfaden zur Analyse psychologischer Berufstätigkeit. *Forum Kritische Psychologie* 23, S. 5-49
- Marx, K. (1844). Zur Kritik der Hegelschen Rechtsphilosophie. Einleitung. Zit.n. *Marx-Engels Werke* (MEW), Bd. 1, S. 378-391
- Müller, C.W. (Hg.) (1978). *Begleitforschung in der Sozialpädagogik. Analysen und Berichte zur Evaluationsforschung in der Bundesrepublik.* Basel: Weinheim
- Ders. (1988). Achtbare Versuche. Zur Geschichte der Praxisforschung in der Sozialen Arbeit. In: Heiner 1988a, S. 17-33
- Patton, M. Q. (1987). Evaluation's political Inherency: Practical Implications for Design and Use. In: Palumbo, D. J. (Hg.). *The Politics of Program Theory* (S. 100-145). Thousand Oaks, CA: Sage
- Scherrer, W. (1996). Qualitätssicherung in der Kinder- und Jugendhilfe. In: Heiner 1996a, S. 9-19
- Schlecht, D. & Tietze, W. (o.J.). *Abschlussbericht über die Begleituntersuchung zum Modellprojekt „Qualität trotz Kostendruck“ des Landes Brandenburg.* Berlin: PädQUIS
- Schneid, Th. (2001). Evaluation – Spurensuche in der Geschichte der Jugendarbeit. In: Beywl, W. et al. (Hg.). *Evaluation im Alltag. Jugendverbände untersuchen ihre Wirkungen* (S. 10-16). Münster: Votum
- Schramme, Th. (2005). Ist eine objektive Beurteilung von Lebensqualität möglich? In: Graumann/Grüber 2005, S. 87-98
- Scriven, M. (1980). *The Logic of Evaluation.* Carlifornia: Edgepress
- Ders. (1983). Evaluation Ideologies. In: Madaus et al. 1983, S. 229-260
- Sieverding, M. (1989). Konfliktfeld einer wissenschaftlichen Beratung vor Ort – Am Beispiel sozialpsychiatrischer Modellprojekte. In: Beerlage/Fehre 1989, S. 127-135
- Spiegel, H. v. (1994). Selbstevaluation als Mittel beruflicher Qualifizierung. In: Heiner 1994, S. 11-54
- Stake, R. E. (1983). The Case Study Method in Social Inquiry. In: Madaus et al. 1983, S. 279-286

- Stockmann, R. (2000). Evaluation in Deutschland. In: Ders. (Hg.). *Evaluationsforschung. Grundlagen und ausgewählte Forschungsfelder* (S. 11-40). Opladen: Leske + Budrich
- Stufflebeam, D. L., Madaus, G. F. & Scriven, M. (Hg.) (2000). *Evaluation Models*. Boston: Kluwe-Nijhof
- Tietze, W., Schuster, K.-M. & Roßbach, H.-G. (1997). *Kindergarteneinschätzungsskala*. Dt. Fassung d. Early Childhood Environment Rating Scale v. Th. Harms u. R. M. Clifford. Neuwied: Luchterhand
- Tietze, W. et al. (2001a). *Kindergarten-Skala. Revidierte Fassung*. Dt. Fassung d. Early Childhood Environment Rating Scale Revised Edition v. Th. Harms, R. M. Clifford & D. Cryer. Neuwied: Luchterhand
- Tietze, W. et al. (2001b). *Hort-Skala*. Dt. Fassung d. School Age Care Environment Rating Scale v. Th. Harms, M. Jacobs & K.R. White. (Unveröff. Ms.)
- Ulrich, S. & Wenzel, F. M. (2003). *Partizipative Evaluation. Ein Konzept für die politische Bildung*. Gütersloh: Bertelsmann Stiftung
- Westermann, R. (2002). Merkmale und Varianten von Evaluationen: Überblick und Klassifikation. *Zeitschrift für Psychologie* 210, S. 4-26
- Wottawa, H. & Thierau, H. (1990). Evaluationsprojekte: Wissensbasis oder Entscheidungshilfe? *Zeitschrift für Pädagogische Psychologie* 4, S. 229-240
- Dies. (2003). *Lehrbuch Evaluation* (3., korr. A.). Bern: Hans Huber