

Evaluation und Evaluationsforschung

Von Axel Groenemeyer und Holger Schmidt

Evaluation ist eine Idee von bestechender Einfachheit. Die Betrachtung und Bewertung von Ergebnissen und Folgen zielorientierten Handelns ist ein Vorgang, der im Alltag immer und überall routinemäßig und selbstverständlich eingesetzt wird, wenn Entscheidungen getroffen werden und Handeln mit gemachten Erfahrungen in Einklang gebracht werden. So weisen denn auch viele Autoren darauf hin, dass Evaluation ein Vorgang ist, der so alt wie die Menschheit sei (z. B. Stockmann 2004). Wenn es um politische oder professionelle Entscheidungsprozesse geht, ist die Forderung nach Evaluierung unwiderstehlich, sie weicht Entscheidungen, Maßnahmen, Programme oder Interventionen mit einer höheren Rationalität oder entzieht ihnen jeglichen Sinn und jegliche Legitimation. Evaluation ist zu einem modernen Zauberwort für Politik und professionelle Praxis geworden.

Sobald Evaluation mehr meint als Handeln im Kontext irgendeiner Art der Bewertung von Folgen und Erfahrungen und z. B. danach gefragt wird, wie Evaluation funktioniert, an welchen Kriterien eine „gute“ Evaluation ausgerichtet sein soll und welche Bedeutung sie hat, wird die Sache kompliziert und völlig unübersichtlich. So gibt es mittlerweile ausufernde Diskussionen darüber, was eigentlich Evaluation ist, ob und wie man sie von Evaluationsforschung abgrenzen kann und soll, welche Formen von Evaluation es gibt, welche methodischen Anforderungen an eine Evaluation zu stellen sind und welche Rolle ihr für politische Entscheidungen oder die professionelle Praxis zukommt oder kommen soll.

Modelle, Formen und Begriffe der Evaluation

Vor dem Hintergrund der steigenden politischen Bedeutung von Evaluation und Qualitätsmanagement, der allgemeinen Entwicklung von Methoden und Methodologien der empirischen Sozialforschung, noch stärker aber in Auseinandersetzung mit den Fragen der Verwendung wissenschaftlichen Wissens und der Politik- und Praxisrelevanz von Evaluationen hat sich das Feld in unübersichtlicher Weise ausdifferenziert und zu einer kaum zu systematisierenden Vielzahl an Evaluationsmodellen und zu einem „Begriffswirrwarr“ geführt, in dem nahezu jede Art der Reflexion über Politik und professionelle Praxis in irgendeiner Weise mit dem Etikett Evaluation belegt worden ist.

Dies beginnt bereits mit der Unterscheidung von *Evaluation* und *Evaluationsforschung*. Während in der Entwicklung der Evaluation nach dem 2. Weltkrieg zunächst in den USA, dann später auch in Deutschland Evaluation immer mit Evaluationsforschung gleichgesetzt wurde, melden sich in den letzten Jahren, insbesondere aus der professionellen Praxis vermehrt Stimmen, die gerade mit Hinblick auf die Praxisrelevanz der Evaluation eine explizite Abkehr von einer Orientierung an sozialwissenschaftlichen Methoden der empirischen Sozialforschung fordern und damit Evaluation von Evaluationsforschung deutlich abgrenzen. Parallel dazu und mit ähnlichen Argumenten haben sich auch die Methoden der Evaluationsforschung bzw. der Evaluation ausdifferenziert. Während zunächst Evaluationsforschung als quantitative sozialwissenschaftliche Kausalanalyse konzipiert war (Lange 1999), haben sich bereits seit den 1970er Jahren zunehmend Positionen entwickelt, die qualitative und interpretative Forschungsmethodologien in die Evaluationsforschung eingebracht und populär ge-

macht haben (Flick 2006). Zwar wird immer wieder darauf hingewiesen, dass Evaluationsforschung dem Bereich der angewandten Sozialforschung zugehört und in diesem Sinne eine Aufgabe der Politik- und Praxisberatung erfüllen soll, gleichwohl finden sich aber auch Ansätze, die der Evaluationsforschung einen wichtigen Beitrag zur Grundlagenforschung zutrauen (Kaufmann/Strohmeier 1981).

Weitere Differenzierungen des Feldes ergeben sich über die Differenzierung nach Akteuren der Evaluation (*interne vs. externe Evaluation; Fremd- vs. Selbstevaluation*) oder nach den Phasen im politischen Prozess, auf den sich die Evaluation bezieht. Stockmann (2004) grenzt z. B. die *preformative Evaluation*, die sich auf die Phase der Programmformulierung und -planung bezieht, von der *formativen Evaluation* ab, die sich auf die Umsetzung bzw. Implementation von Programmen bezieht. Diese Formen der wissenschaftlichen Politikberatung werden im deutschen Sprachbereich häufig auch als *wissenschaftliche Begleitforschung* bzw. auch international als *Implementationsforschung* bezeichnet. Dabei lassen sich diese Phasen durchaus weiter differenzieren und damit entsprechende Aufgaben der Evaluation auf dieser Ebene formulieren. So gliedern sich dann (pre-)formative Evaluationen in das *needs assessment*, mit dem die Bedürfnisse der am Programm beteiligten Akteure analysiert und bewertet werden, in das *evaluability assessment* zur Einschätzung der Möglichkeiten und Grenzen einer Evaluation, die *structured conceptualization* zur Auswahl und Begründung von Techniken und Methoden, der *implementation evaluation*, mit der die korrekte Umsetzung eines Programms bewertet wird und der *process evaluation* zur Analyse des Umsetzungsprozesses und Entwicklung möglicher Alternativen. Andere Autoren führen hier noch die Idee einer *Kontextevaluation* ein, mit der neben den Bedingungen der Programmdurchführung z. T. die ethischen Implikationen der Maßnahmen sowie Auswirkungen der Evaluation berücksichtigt werden sollen. Als weitere Begrifflichkeiten, die in Bezug auf die Planungsphase eines Programms verwendet werden und sich auf die Analyse der Durchführbarkeit von Programmen beziehen, lassen sich die Ideen einer *prospektiven* bzw. *antizipatorischen* oder *prognostischen* Evaluation anführen (Kury/Obergfell-Fuchs 2010).

Schließlich besteht die Aufgabe der Evaluation in einer dritten Phase in der Zusammenfassung, Bewertung und Bilanzierung der Ergebnisse, Effekte

und Folgen, die dann als *summative Evaluation* bezeichnet wird. Diese Form der Evaluation wird im deutschsprachigen Raum auch als *Wirkungsmessung*, *Wirkungsbewertung* oder, insbesondere in durch medizinische Perspektiven geprägte Felder, als *Katamnese* bezeichnet. Auch hierzu lassen sich weitere Differenzierungen anführen: So spricht man von *outcome evaluation* im Hinblick auf gewünschte Ziele bzw. von *impact evaluation*, wenn die Gesamtheit der Auswirkungen berücksichtigt werden soll. Bezieht man den Outcome einer Maßnahme oder eines Programms auf monetäre Größen, so kommen Verfahren der *cost-effectiveness* bzw. *cost-benefit-analysis* oder der *Kosten-Nutzen-Rechnung* zum Einsatz. Während es hierbei im Wesentlichen um die Messung und Bewertung von Effekten und Folgen der Maßnahmen und Programme geht, erhebt die *Wirkungsanalyse* darüber hinaus auch den Anspruch, das Zustandekommen der Effekte und Folgen wissenschaftlich zu erklären. Daneben finden sich zwei weitere Formen summativer Evaluationen, die keine eigenständige Datenerhebung beinhalten: die *Sekundäranalyse*, in deren Rahmen eine Re-Analyse vorhandener Daten vorgenommen wird, um noch nicht berücksichtigte Fragestellungen zu beantworten, bzw. andere Methoden eingesetzt werden, und die *Meta-Evaluation* bzw. *Meta-Analyse*, mit der die Ergebnisse verschiedener Studien unter dem Blickwinkel einer Evaluationsfragestellung vergleichend bewertet werden. Diese Liste der Begrifflichkeiten und Modelle zur Evaluation ist keineswegs abgeschlossen (vgl. z. B. Beywl 2006). Das Ausmaß ihrer Differenzierung verweist allerdings einerseits darauf, dass es die Evaluation bzw. die Evaluationsforschung mit komplexen Fragestellungen zu tun hat, andererseits aber auch darauf, dass die Verwendung des Evaluationsbegriffs mit Reputationsgewinnen verbunden ist.

Entwicklung der Evaluation und Evaluationsforschung in der Sozialen Arbeit

Tatsächlich wird in der Evaluationspraxis und insbesondere im Feld der Sozialen Arbeit zumeist von einer eingeschränkteren Perspektive ausgegangen und Evaluation im Wesentlichen auf die Frage nach der Messung, Analyse und Bewertung der Wirkungen