

Systematisches Bewerten – ethische Herausforderung für Evaluation

Exposee für einen Beitrag zur Session D5 - Ethik in der Evaluation auf der 11. Jahrestagung der DeGEval 24.-26. September 2008 - Universität Klagenfurt

Evaluation ist mehr noch als Forschung oder auch andere Ansätze wissenschaftsbasierter Dienstleistungen – wie z. B. Organisationsberatung oder Supervision – mit ethischen Herausforderungen konfrontiert. Sie muss sich ihrer ethischen Grundlagen in ebenso starkem Mass vergewissern wie beispielsweise die Rechtsprechung, die die Psychotherapie oder die Humanmedizin, die sich ebenfalls mit grundlegenden Fragen des menschlichen Zusammenlebens oder der menschlichen Existenz beschäftigen, und die als Institutionen oder Professionen eine besondere gesellschaftliche Verantwortung tragen. Es ist besonders wichtig, dass in der Evaluation stets darüber reflektiert wird, wie möglichst gute Evaluationen erreicht werden können, und wo Grenzen schlechter oder gar schädlicher Evaluationen überschritten werden, und wie dem in der Gemeinschaft der Evaluatoren und Evaluatorinnen vorgebeugt werden kann.

Die starke ethische Herausforderung resultiert daraus, dass es in der Evaluation nicht lediglich um die – bereits anspruchsvolle – Aufgabe des sachlichen, nüchternen, neutralen, mehrperspektivischen *Beschreibens* von sozialen Tatsachen geht („*factual claims*“), sondern dass es auch um das *Bewerten* auf der Basis von Werten geht, also auf der Basis von „*value claims*“ (vgl. House/Howe 1999). In dem Maße, in dem die beschreibenden und bewertenden Ergebnisse von Evaluationen für die *Steuerung* von Programmen, Organisationen oder anderweitige strategische Entscheidungen mit großer Tragweite für große Kollektive oder die Gesellschaft genutzt oder nicht genutzt werden, ist es um so wichtiger, die ethischen Implikationen möglichst schon proaktiv zu reflektieren.

Evaluation als in der Tradition der Sozialwissenschaften stehend ist zur Bewältigung der Bewertungsaufgabe mit einem spezifischen begrifflichen System ausgestattet. So kommt der Terminus „Wert“ auf mehreren Dimensionen vor:

- auf der eher technischen des „Messwertes“ (synonym mit der Merkmalsausprägung oder des *variable value*)
- auf der des sozialen Wertes, aus dem die Bewertungsdimensionen und schliesslich die Bewertungskriterien für die Bewertungsakte der Evaluation gewonnen werden (*value claims*)
- auf der der Zuschreibung von „Wert“ für den Evaluationsgegenstand, also von „Güte“ und „Verwendbarkeit“, von „*merit*“ und „*worth*“, neuerlich um ein dritte Dimension „*significance*“ ergänzt, was übersetzt werden könnte mit aktueller und künftiger Bedeutsamkeit für die Gesellschaft, wo hinein auch die Dimension der „Nachhaltigkeit“ einbezogen werden kann.

Diese drei begrifflichen Dimensionen von „Wert“ gilt es im Sprechen über Evaluation auseinander zu halten, wobei allen, insbesondere den beiden letztgenannten, hohe Bedeutung zukommt.

Tabelle 1: Wert-Begriffe in der Evaluation

Wert, statistischer	Für die Auswertung mittels statistischer Verfahren vorbereitetes Datum: Werte (repräsentiert durch numerische und alphanumerische Zeichen) in Datenmatrizen (nach Konvention: Untersuchungseinheiten/Fälle in den Zeilen, ->'Merkmale' in den Spalten) ermöglichen die Verdichtung von Daten durch Auszählen der Häufigkeit Ihres Vorkommens und – je nach Qualität des statistisches Niveaus – weiter gehende mathematische Operationen gemäß der Konventionen der beschreibenden und der schließenden Statistik.
Wert, sozialer	Eine zunächst biographisch erworbene, kulturell basierte Disposition von Individuen, bestimmte Umstände/Lösungen/Handlungen anderen vorzuziehen. Deskriptiv beschreiben Werte die tatsächlichen Dispositionen von Menschen, normativ solche (sozialen) Dispositionen, die mit unterschiedlichen Begründungen als verbindlich gelten oder eingefordert werden. Werte sind stärker auf längerfristiges Handeln bezogen und auch emotional verankert. Sie können mit kurzfristigen und rational begründeten (z. B. materiellen) Interessen in Widerspruch geraten. Werte im sozialen Kontext werden von Personenmehrheiten (Gruppen, Aggregaten, Teilkulturen) getragen. Für sozialpolitische ->'Programme' sind sie relevant, da differierende Werte sozialer Gruppen sowohl zu differierenden ->'Zielen' (z. B. bezüglich eines wünschbaren Maßes der gesellschaftlichen Einkommensdifferenzierung), als auch zu differierenden ->'Interventionen' (z. B. zwischen Überredung, Anreiz oder Zwang) führen. Damit ergeben sich für die ->'Evaluation' von Programmen unterschiedliche, möglicherweise unvereinbare Bewertungskriterien, was die zentrale Stellung der Werte für Evaluationstheorie und -praxis unterstreicht.
Wert (eines Evaluationsgegenstandes)	Summe der Eigenschaften eines ->'Evaluationsgegenstandes', die zu dem Urteil führt, dass er mehr oder weniger gut oder schlecht ist. Dabei kann die über Zeit und Raum relativ stabile Wertdimension der ->'Güte' von der sehr instabilen der Verwendbarkeit unterschieden werden, die zwischen den ->'Beteiligten und Betroffenen' meist stark umstritten ist. Die Bestimmung des Wertes eines Evaluationsgegenstandes, das Bewerten, ist die zentrale und unverzichtbare Aufgabe jeder ->'Evaluation'.

Quelle: (Beywl/Niestroj 2008)

Die verschiedenen Modelle der Evaluation, die sich auch wissenschaftstheoretisch unterschiedlich verorten lassen, lassen sich danach systematisieren, wie sie mit sozialen Werten (*value claims*) im Rahmen der Planung und Steuerung von Evaluationen umgehen.

Tabelle 2: Typologie von Evaluationsmodellen nach ihrem Umgang mit Werten

	Name des Modells
I. wertedistanziert	a) Programmzielgesteuerte Evaluation
	b) Experimentaldesigngesteuerte Evaluation
	c) Quasiexperimentaldesigngesteuerte Evaluation
	d) Kosten/Nutzengesteuerte Evaluation
	e) Programmtheoriegesteuerte Evaluation
	f) Kontextmechanismusgesteuerte Evaluation
II. werterelativistisch	a) Spannungsthemengesteuerte Evaluation
	b) Dialoggesteuerte Evaluation
III. wertepriorisierend	a) Entscheidungsgesteuerte Evaluation
	b) Nutzungsgesteuerte Evaluation
	c) Stakeholderinteressengesteuerte Evaluation
IV. wertepositioniert	Selbstorganisationsgesteuerte Evaluation

Unterschieden werden grob wertedistanzierte und werterberücksichtigende Modelle, wobei letztgenannte noch einmal nach werterelativistischen, wertepriorisierenden und wertepositionierten unterschieden werden können. (vgl. Beywl 2006)

Die dem herkömmlichen rationalen Wissenschaftsverständnis nahe stehenden Evaluationsmodelle, aber auch andere, verfahren eher wertedistanziert und klammern die Bewertungsakte aus dem Evaluationsprozess aus, verschieben sie *vor* die evaluative Untersuchung (z. B. in die relativ mechanistisch aufgefassten Hypothesen- oder Alternativhypothesenformulierung) oder *hinter* die Evaluation, indem sie den Auftraggebenden oder politisch Entscheidenden die Bewertungsakte überlassen. Sie verbinden diese Haltung mit dem Neutralitätsgebot und der Verpflichtung von Wissenschaft auf Objektivität, mit Verweis auf den Werturteilsstreit in der deutschsprachigen Soziologie oder auf die Tradition des kritischen Rationalismus im Anschluss an Karl Popper. Helmut Kromrey hat diese Position kürzlich in einem Beitrag für die Zeitschrift für Evaluation einerseits verteidigt andererseits aufgezeigt, dass Evaluation mit dieser Position in die Nähe einer Identitätskrise kommt: Denn verweigert sie das Bewerten, worin liegt dann ihre *differentia specifica* gegenüber der Forschung? (Kromrey 2007)

Die andere Position ist verbunden mit einer paradigmatischen Wende in der Evaluation, die etwa Mitte der 70er Jahre vollzogen wurde, mit Namen wie Robert Stake, Lee Cronbach, Daniel Stufflebeam, Michael Scriven und anderen verbunden ist, und die bei aller Verschiedenheit letztlich eine konstruktivistische Grundlage hat: Das Bewerten ist danach gar nicht vermeidbar, ist eine grundlegende menschliche Tätigkeit und auch Kompetenz, die ein Überleben des Individuums, des Kollektivs und der Gesellschaft überhaupt erst ermöglicht.

Metaphorisch lässt sich an Grenzfällen illustrieren, z. B. aus der Psychologie, in letzter Zeit in der Literatur und in Filmen über Autisten mit Asperger Syndrom populär aufgegriffen z. B. von Mark Haddon oder Nick Balthazar: Es zeigt sich dass Menschen, die in Bezug auf das Bewerten grundlegend gehemmt sind, erhebliche Probleme bei der Orientierung in der Welt und im Umgang mit anderen Individuen haben. (Balthazar 2003; Haddon 2003).

Diese Erkenntnis lässt sich auf Organisationen und andere soziale Systeme übertragen. So kann als These formuliert werden: In einer weitgehend enttraditionalisierten, aufgeklärten, demokratischen und global kommunizierenden Gesellschaft – für die Mehrperspektivität und Multistakeholder-Settings selbstverständliche Gegebenheiten sind – kommt der Evaluation die Funktion des systematischen Bewertens auf der Basis empirisch gewonnener Informationen zu. *Evaluationsversagen* – sei es durch Bewertungsabstinz oder durch Fehlbewertungen, zurückgehend auf Kompetenzmängel, Kunstfehler oder uneingestandene Interessenkonflikte – würden zu schwerwiegenden ethischen Belastungen für die Evaluation als Profession und Institution führen.

Systematisches Bewerten – durch professionelle Evaluation – ist also notwendig und stets prekär.

Die Frage soll diskutiert werden, ob sich im Rahmen wissenschaftlicher Evaluationen systematische und nachvollziehbare Bewertungsprozesse entwerfen und realisieren lassen, die sich grundsätzlich von willkürlichen Bewertungsprozessen unterscheiden.

Es fragt sich ob zu Beginn des 21. Jahrhunderts, fast 100 Jahre nach dem Werturteilsstreit in der deutschsprachigen Sozialwissenschaft, qualitativ verbesserte sozial- und kommunikationstechnologische Bedingungen gegeben sind, um systematische Bewertungsprozesse zu organisieren; ob die Evaluationstheorie, die seit 30 Jahren intensiv entwickelt ist, über eine methodologische Grundlage verfügt, für die es einen intensiven theoretischen Diskurs auch über ethische Probleme, Grenzen systematischen Bewertens, Rollenprobleme von Evaluatoren und Evaluatorinnen, Kompe-

tenzanforderungen usw. gibt; ob Bedingungen gegeben sind, angesichts derer die Evaluationsgemeinschaft zur Auffassung gelangt: sollte man es wagen, muss man es wagen oder ist es eine ethische Verpflichtung zu wagen, sich der Frage des systematischen Bewertens im Rahmen wissenschaftlicher Evaluation zu stellen?; sich zu fragen welche Ansatzpunkte bieten Evaluationsmodelle, was bieten Disziplinen wie Psychologie, Philosophie, Soziologie, Erziehungswissenschaft, Politikwissenschaft usw.; welche Hilfsmittel lassen sich gewinnen aus Organisationstheorien, Organisationsentwicklung, Coaching; Supervision?

Hierbei drohen immer wieder Missverständnisse zur Identität von Evaluation und der Verwechslung mit moralischen Urteilen und dezisionistischen Entscheidungen. Es entstehen schnell phantastische Bilder: Evaluatoren und Evaluatorinnen werden als oberste Richter und Richterinnen, als Personen mit nahezu göttlicher Einsicht und Weisheit ausgestattet wahrgenommen oder ihnen wird unterstellt, dass sie dieses Selbstbild hätten – in Wirklichkeit sehen sie sich eher als Schieds- und Linienrichter, während es die Stakeholder sind, die unterschiedlich einflussreichen Finanziers, Programmverantwortlichen und Interessengruppen, die das Spiel machen, ihre *value claims* einbringen und schliesslich auch entscheiden.

Die mächtigen Projektionen und starken Widerstände indizieren einerseits, dass es um etwas Wichtiges geht, mahnen andererseits zu umsichtigem und reflektiertem Vorgehen – die Diskussion im Workshop kann hierzu einen wertvollen Beitrag leisten

Literatur

- Balthazar, Nic (2003): *Niets is alles wat hij zei (Ben-X - 2007 verfilmt)*. Antwerpen: Maklu.
- Beywl, Wolfgang (2006): "Evaluationsmodelle und der Stellenwert qualitativer Methoden". In: Flick, Uwe (Hg.) (2006): *Qualitative Evaluationsforschung Konzepte - Methoden - Umsetzung*. Reinbek: Rowohlt. 92-116.
- Beywl, Wolfgang/Niestroj, Melanie (2008): Glossar wirkungsorientierter Evaluation - 2. überarbeitete Auflage. Unveröffentlichtes Arbeitspapier: Univation - Institut für Evaluation.
- Haddon, Mark (2003): *Supergute Tage oder Die sonderbare Welt des Christopher Boone (The Curious Incident of the Dog in the Night-Time)*. 1 München: Blessing.
- House, Ernest R./Howe, Kenneth R. (1999): *Values in evaluation and social research*. Thousand Oaks: Sage Publications.
- Kromrey, Helmut (2007): "Wissenschaftstheoretische Anforderungen an empirische Forschung und die Problematik ihrer Beachtung in der Evaluation - Oder: Wie sich die Evaluationsforschung um das Evaluieren drückt". In: *Zeitschrift für Evaluation*, 113-123.

Kontakt Dr. Wolfgang Beywl

Univation – Institut für Evaluation - Hohenstaufenring 63, D-50674 Köln
Tel.: +49 (0) 221 / 4 24 80 71 E-Mail: wolfgang.beywl@univation.org – www.univation.org

Universität Bern, Koordinationsstelle für Weiterbildung, Schanzeneckstrasse 1 CH-3001 Bern
Tel. +41 (0)31 631 5349 . E-Mail: wolfgang.beywl@kwb.unibe.ch. - www.evaluationsstudium.ch