



Manfred Moldaschl, Christina Schwarz

Die Farben der Evaluierung

Eine Verteidigung der Evaluierung gegen ihre Befürworter

Wir sind alle so flexibel. Und sollen immer noch flexibler werden. Und weil das nicht nur für uns, sondern auch für alles andere gilt, wird alles ziemlich kompliziert. Da aber irgendwo Komplexitätsreduktion stattfinden muss, erfreut sich Schwarzweißdenken allergrößter Beliebtheit, und Vereinfachungsangebote finden reißenden Absatz. Das gilt auch für die Debatte über Sinn und Funktion der Ergebnis- und Leistungsbewertung an Universitäten. Die einen drängen auf Evaluierung, in der erhebenden Selbstgewissheit, notwendige Modernisierungsattacken auf ein verkrustetes System zu reiten, in großen Lettern "Effizienz " und "Qualität " in ihrem Schilde führend. Wo Marktmechanismen im Hochschulsystem und damit Selektionsdruck fehlten, müsse die *visible hand* der Evaluierung das kompensieren. Zumindest dort, wo Marktförmigkeit selbst (noch) nicht realisierbar sei. Andere verteidigen das Humboldtsche Ideal der Universität, das sich mit kleinlicher Rechenhaftigkeit nicht verträgt. Sie warnen vor der Unterordnung von Lehre und Forschung unter die Herrschaft der Krümelzähler - und vor der Kommodifizierung unserer höheren Bildung.

Natürlich ist auch diese Darstellung schwarzweiß - es gibt mehr als diese zwei Positionen, in denen zudem manches zusammenwächst, was nicht notwendig zusammengehört. Aber es sind dominante Positionen, wie wir sie auch aus der Debatte über die Modernisierung der Gesellschaft insgesamt kennen, speziell im Hinblick auf den Staat und die Sozialsysteme. Dass diese Debatte heute ideologisch so aufgeladen ist, verdanken wir nicht zuletzt einer Ideologie, die sich als das Ende aller Ideologien versteht: dem *ökonomischen Fundamentalismus*. Indem dieser alles angreift, was in das "freie Spiel der Kräfte" eingreift, sei es ein staatlicher Schutz der Berufe, Regulierungen der Medikamentenzulassung, Tarifverträge mit der organisierten Arbeitnehmerschaft, öffentliche Hochschulen oder Polizei, oder Sozial- und Umweltstandards: eine sachliche Auseinandersetzung mit den Nebenfolgen von Sozial-, Umwelt-, Bildungs- und generell Modernisierungspolitik wird erschwert, Lagerdenken geschürt.

Das ist ein wenig wie im Krieg, zu dessen ersten Opfern nicht nur die Wahrheit, sondern auch das Differenzierungsvermögen zählt: Freund und Feind, gut und böse, das Dazwischen entfällt. Mit der Evaluierungskritik ist es ähnlich wie mit der Staatskritik, nur umgekehrt. Kam letztere früher "von links", als Kritik an der Herrschaft von Eliten, die sich durch den Staat die Reproduktion ihrer Machtquellen sichern, so kommt sie heute bevorzugt "von rechts"¹, als Kritik an Mittelverschwendung für soziale Zwecke und an mangelnder Wirtschafts- und Herrschaftsdienlichkeit. In ihrer Einseitigkeit erzeugt diese Kritik bei den Kritisierten quasi unbeabsichtigt konservative Verteidigungsreflexe. Kam die Forderung nach Evaluierung an Hochschulen früher von Kritikern, welche die Herrschaft der Ordinarien monierten und mittels Lehrevaluation deren undemokratischen, die Bedürfnisse der Studierenden missachtenden Lehrstil aufzudecken (und zu verändern) hofften, so kommt sie heute eher von jenen, die nach Legitimation für eine zunehmende Ungleichverteilung der immer knapperen Mittel für Hochschulbildung suchen. Freilich unter der ungleich legitimeren Chiffre "Wettbewerb". Das bringt nun diejenigen, die für eine zumindest relative Autonomie der Bildung gegenüber der Wirtschaft eintreten, in die seltsame Lage, vor allem jene Argumente sammeln zu müssen, die den Unsinn gängiger Evaluierungsstrategien, -praktiken und -erwartungen belegen. Die gibt es leider überreichlich. Mit der Evaluierung ist es wie mit Zielvereinbarungen: Es könnten so sinnvolle Sachen damit gemacht werden, doch geschieht meist das Gegenteil - eine Verschwendung von Zeit und anderen nicht rückholbaren Ressourcen.

Von den genannten Promotoren wollen wir uns die Art und die Inhalte des Diskurses über Evaluierung aber nicht aufzwingen lassen. Man muss Evaluation vielmehr gegen die unkritische Befürwortung verteidigen, als ein Instrument *wissenschaftlicher Selbststeuerung, verantwortlicher Selbstbestimmung* und *institutioneller Selbstkritik*. Das Ziel unseres Beitrags ist es, Verständnisse von Evaluierung zu skizzieren (Abschnitt 1), den grassierenden Taylorismus der Hochschulevaluierung und seine Folgen unter die Lupe zu nehmen (Abschnitt 2), und eine Perspektive zu umreißen, wie man den aktuellen Evaluierungsboom für die drei genannten Funktionen konstruktiv aufgreifen könnte (Abschnitt 3).

1. Farben, Formen und Funktionen der Evaluierung

Die Frage, was Evaluierung ist, kann derart breit beantwortet werden, dass sich gleich die Frage anschließt, was dann nicht Evaluierung sei. Denn 'Evaluierung' ist eines jener Zauberworte der technischen Moderne, die eine rationalere Einrichtung der Welt verheißen. Vom Lateinischen kommend bedeutet es Wertbestimmung oder *Auswertung*, und nach gegenwärtiger Konvention versteht man darunter - vor allem im öffentlichen Sektor - die Auswertung von Maßnahmen oder die Bewer-

¹ Bündiger hat Matthias Beltz das formuliert: "Der Staatsfeind trägt heute Brillanten".

tung von Programmen im Hinblick auf ihre *Wirkung*. Definiert man Evaluierung so allgemein, ist sie im Grunde unabdingbarer Teil jedes zweckrationalen Handelns, das sich stets mit seinen Effekten auseinandersetzen muss. Dagegen kann auch niemand etwas haben, geht es doch um eine nüchterne Bewertung der Verhältnisse von Aufwand und Ertrag (*Effizienz*) sowie von Ziel und Ergebnis (*Effektivität*). Das braucht jede Praxis. Wer würde nicht wissen wollen, ob mit den eingesetzten Mitteln die angestrebten Ziele erreicht werden, und ob sie ökonomisch (effizient) erreicht werden? Umgekehrt zieht, wer heute etwas gegen Evaluation sagt, sogleich den Verdacht auf sich, etwas zu verbergen zu haben.

Definition²

Evaluierung ist die systematische Nutzenbewertung eines Prozesses, Programms, Produktes oder einer Institution, die der Information und Urteilsbildung von Entscheidungsträgern in Unternehmen, Politik, öffentlichen Einrichtungen und Gemeinden dienen soll. Systematisch heißt, sie erfolgt in der Regel mit den Mitteln der empirischen Forschung.

Wir kennen die Funktion der Evaluierung und ihre Maßstäbe unter vielen anderen Namen, etwa aus dem betriebswirtschaftlichen Kontext als *Controlling*. Dieses wird ebenfalls verstanden als eine Institution und als Set von Verfahren zur "Sicherstellung der Rationalität" (Weber, Schäffer 1999). Natürlich steckt in Worten wie *Auswertung*, *Bewertung* und *Maßstäben* überall der *Wert* (valus, valor, value). Insofern gilt für Evaluierungsverfahren generell, was Ekkehard Kappler für das Controlling als betriebliche Evaluierungsinstanz feststellt:

"Dabei erweist sich die Behauptung eines wertfreien, rationalen, neutralen, objektiven, apriori transparenten, ahistorischen und apolitischen Rechnungswesens wie Informations- und Wissensmanagements als nicht haltbar bzw. idealistisch bis ideologisch. Es kann gezeigt werden, dass das Rechnungswesen systematisch zur Erzeugung disziplinärer Macht beiträgt, wenn seine Normen und Standards nicht als Konventionen erkannt, sondern als vermeintliche rationale Muster und Sachzusammenhänge individuell verinnerlicht werden, weil es interessenbezogen auffällbare Worthülsen, wie z.B. Kosten, Leistung, Effizienz, Rationalität, Wirtschaftlichkeit, Ziele, Mittel, Entscheidung, Information, Prognose, Abweichung, Anreiz, Sanktion, institutionell verankert und als gesellschaftliches Steuerungsmuster unkritisch zu übernehmen gestattet" (Kappler 2000, 237).

Ein zentrales Schlüsselwort des Evaluierungsdiskurses hat Kappler hier noch nicht einmal erwähnt: die *Qualität*. Sie eignet sich in besonderer Weise zur Legitimation

² Das ist unsere eigene Definition. Wir könnten uns aber auch der von Peter Baumgartner (1999, 71) anschließen: "Unter Evaluation sind alle Aktivitäten und/oder Ergebnisse zu verstehen, die die Bedeutung, Verwendbarkeit, (Geld-)Wert, Wichtigkeit, Zweckmäßigkeit, ... einer Sache beurteilen bzw. bewerten. Nur dieses weit gefaßte Verständnis von Evaluation kann sowohl die Charakteristika besonderer Evaluationsfelder berücksichtigen als auch einen adäquaten Beitrag zur Theoriebildung leisten".

von Bewertungen und Maßnahmen, die unter anderem Label weniger Akzeptanz fänden. Für Wirtschaftsunternehmen hatte der Kaizen-Guru Mazaaki Imai (1992) das bereits mustergültig erkannt:

"Während das Hauptanliegen des Managements die Produktivitätssteigerung sein mag, ist Qualität vielmehr ein gemeinsames Anliegen von Belegschaft und Management. Wenn ein Management die Produktivität im Unternehmen steigern will, wird der Betriebsrat fragen: 'Warum denn eigentlich? Das heißt doch nur, dass wir im Grunde härter arbeiten müssen. Was ist für uns dabei drin?' Gegen Qualität kann aber niemand etwas haben" (Imai 1992, 130).

Dass trojanische Qualitätsstrategien auch die Hochschulreform durchziehen, haben Laske/Weiskopf (1996) hervorgehoben (vgl. auch Laske/Habersam/Kappler 2000). Natürlich ist auch umstritten, was unter 'Qualität' zu verstehen ist. In einem von Hartmut Wächter (1997) herausgegebenen Band werden in einer Weise verschiedene Organisationstheorien zur Erklärung der teils sektenartige Züge annehmenden Qualitätsbewegung in der Wirtschaft durchgespielt, die man ohne Mühe auf die Qualitäts- und Evaluierungsdebatte in der Wissenschaft übertragen könnte. Da wird sie als soziale Bewegung mit primär identitätsstiftender Funktion analysiert, als Konstruktion von Legitimationsfassaden zur Vermeidung wirklicher Veränderungen, oder als mikropolitische Spiel, mit dem verschiedene Akteursgruppen jeweils ihre Ziele durchzusetzen versuchen. Jeder, der schon mal eine Universität von innen gesehen hat, kennt die oft gleichzeitige Plausibilität solcher Erklärungen.

Statt nun aber bezogen auf unseren Gegenstand, Forschung und Universitäten, zu beklagen, dass "Qualität" oder "Effizienz" den universitären und externen Akteuren nur die mehr oder minder anerkannten Legitimationen für die Verfolgung ihrer Partikularinteressen bieten, sollte man dies zum *Ausgangspunkt der ganzen Betrachtung* machen. Insofern stellen sich für jedes Verfahren der Evaluierung - sei es unter diesem Namen oder unter *Audit*, *Review*, *Controlling*, *Akkreditierung* - stets dieselben Grundfragen:

- Wessen Ziele und welche Zwecke sind Gegenstand der Bewertung?
- Wer bewertet wen?
- Mit welchen Maßstäben wird bewertet?

Wenn man so fragt, kommt man zur Metafrage, welchem Ziel die Evaluierung selbst dient. Legt man ferner eine (ggf. mikro-)politische Akteursperspektive zugrunde, wird man weiters fragen, welchen Zweck eine Evaluierung für welche der beteiligten Akteursgruppen hat. Man kann davon ausgehen, dass sich jede Gruppe des Verfahrens im eigenen Interesse zu bedienen versucht. Von *der* Evaluierung kann auch deshalb nicht die Rede sein, weil es vielfältige "Farben" und Formen, Anwendungsbedingungen und Zwecke von Evaluierung gibt.

Zunächst zu den *Funktionen von Evaluierungsverfahren*, wie sie üblicherweise beschrieben werden, z.B. bei Stockmann (2004, 18f.). Hervorzuheben sind insbesondere folgende vier: Eine Funktion der *Erkenntnisgewinnung*, bei der das Lernen im Vor-

dergrund steht; sie richtet sich darauf, etwas über die Wirkung von Eingriffen in Praxiszusammenhänge zu erfahren, und damit über diese selbst (welche Kausalbeziehungen bestehen; erreicht ein Programm seine Zielgruppe; etc.). Emphatisch könnte man hier auch von einer Aufklärungsfunktion sprechen - der Glaube über die Wirkung von Maßnahmen wird durch Wissen ersetzt. Zweitens eine *Kontrollfunktion*, die sich u.a. auf Kontrolle der Mittelverwendung, der Aufgabenerfüllung, der Wirksamkeit von Maßnahmen, und Ähnliches bezieht. Drittens eine *Dialogfunktion*, die darin besteht, beteiligte Akteure zum Dialog über Ziele, notwendige Maßnahmen und geeignete Maßstäbe zu animieren. Schließlich eine Funktion der *Legitimierung*, in der beispielsweise gefragt wird: Sind die Ziele öffentlicher Investitionen erreicht worden? Sind die Verantwortlichen verantwortlich mit den Mitteln umgegangen? Dies sind gewissermaßen die "rationalen" Funktionen, in denen wir uns alle die Anwendung von Evaluierung erhoffen, besonders, wenn es um den Einsatz unserer Steuermittel geht. Mit der Wahrscheinlichkeit dieser Funktionalität befassen wir uns später im Text. Die beiden folgenden Tabellen (Tab. 1 und 2) geben einen Überblick über wichtige Formen und Konzepte der Evaluierung, und damit bereits einige methodische Antworten auf unsere obigen drei bzw. vier Grundfragen.

	<i>Ort der Steuerung</i>	<i>Instanz der Durchführung</i>	<i>Ort der Nutzung</i>
<i>Innerhalb des Projekts/ der Institution</i>	Selbstevaluation	interne Evaluation	formative Evaluation
<i>Außerhalb des Projekts/ der Institution</i>	Fremdevaluation	externe Evaluation	summative Evaluation

Tabelle 1: Formen der Evaluierung (nach Widmer 2004, 85)

	<i>Analyseperspektive</i>	<i>Erkenntnisinteresse</i>	<i>Evaluationskonzepte</i>
<i>Programmformulierung Planungsphase</i>	ex ante	analysis for policy; science for action	präformativ, formativ: aktiv gestaltend
<i>Implementation</i>	on-going	interne Evaluation	formativ/summativ (beides möglich)
<i>Wirkungsphase</i>	ex post	analysis for policy; science for knowledge	summativ: bilanzierend, ergebnisorientiert

Tabelle 2: Konzepte der Programmevaluierung nach Phasen des Prozesses (Stockmann 2004, 17)

Mit *interner* Evaluation ist die Auswertung von Maßnahmen oder Programmen durch die jeweilige Handlungsinstanz selbst gemeint. Das heißt aber nicht notwendig, dass z.B. eine Universität oder Fakultät sich selbst bewertet. Sie kann das auch an externe Instanzen delegieren, hat aber dabei die Initiative und behält die Hoheit über das Verfahren. Bei der *externen* Evaluierung bestimmen andere Instanzen die Ziele und setzen die Maßstäbe. Mit *summativer* Evaluierung wird in der Regel eine

bilanzierende, expertengestützte, ergebnisorientierte Bewertung von Endergebnissen bezeichnet. Bei der *formativen* Evaluation wird auch der Verlauf der Intervention bzw. des Programms analysiert und ggf. beeinflusst (z.B. Bortz/Döring 2002, 576).

Bei Literatur- und Internetrecherchen stellten wir allerdings eine weitgehende Beliebtheit und vor allem Anspruchslosigkeit dieser Terminologie fest. So hat sich etwa leider als Minimalkonsens durchgesetzt, formative Evaluierung lediglich als *prozeßbegleitend* zu verstehen, wobei der Gestaltungsaspekt meist ausgeblendet wird. Das ist eine gravierende und symptomatische Verkürzung, insbesondere wenn man an einen Ursprung der Evaluierung in den Arbeiten von Kurt Lewin (z.B. 1953) und seiner Gruppe aus den 40er Jahren anknüpft. Hier ging es mit der Survey-Feedback-Methode - zugleich Kernstück der *Organisationsentwicklung* (OE) - nicht nur um eine Veränderung der Maßnahmen und Ziele nach Auswertung des Erreichten. Vielmehr bestand das Ziel dieses humanistischen Ansatzes in der Berücksichtigung der Interessen von Arbeitskräften oder generell derjenigen, die evaluiert und "verändert" werden sollten. Dies wiederum beschränkte sich nicht auf Partizipation, sondern zielte auf *Demokratisierung*, also gleichberechtigte Beteiligung an der Gestaltung sozialer Praxis, und auf *Bildung*, verstanden als Emanzipation der Person(en). Deshalb wird der Ansatz auch *reduktiv* genannt.

Symptomatisch ist der Bedeutungswandel, weil sich auch der OE-Ansatz in der Praxis zunehmend zum reinen Herrschafts- und Verfügungsinstrument "verfärbt" hat, in dem Partizipation nur noch instrumentell als Technik zur reibungslose(re)n Durchsetzung von Managementzielen verstanden wird (vgl. Hanft 2003; Moldaschl 2004a).³ Dasselbe "Schicksal" erfuhren die Leitbilder der Hochschulreform in den letzten drei Dekaden, deren Farbe ausgehend von der Demokratisierung in den 1970er Jahren über die Dezentralisierung und Professionalisierung immer mehr ins Kalt-bläuliche tendierte, bis hin zur aktuellen "Entlassung in die Autonomie" (Laske/Zauner 2000, 459) im Geiste der Vermarktlichung.

2. Der Doppelcharakter der Evaluierung

Das demokratische Versprechen, welches der Evaluierung einmal innewohnte und heute noch in der Utopie einer wissensbasierten Politik fortlebt⁴, scheint gleichwohl

³ Im Lewinschen Geist entfaltet wird der (im englischen Wortsinn eigentliche) Sinn formativer, *bildnerisch-emanzipativer Evaluierung* z.B. im Ansatz von Lincoln und Guba (1985). An den politischen Entstehungskontext knüpft u.a. Charles Lindblom (1990) an. Evaluation begann sich in den Vereinigten Staaten der 30er Jahre mit dem New Deal zu etablieren, ausgehend von der Wirtschaftskrise und dem darauf reagierenden kooperativen Föderalismus in der Roosevelt-Ära. Sie sollte zur gerechteren und effektiveren Verteilung von Finanzen im expandierenden Wohlfahrtsstaat dienen.

⁴ Diese Utopie bzw. dieses Leitbild, wird besonders vom Blair-Berater Anthony Giddens weitergetragen (z.B. 1990). Die Entscheidungen des Einzelnen, auch in der Politik, würden mit der Enttraditionalisierung und wachsender Optionsvielfalt zunehmend legitimationspflichtig,

immer mehr auf die *Kontrollfunktion* einzuschnurren: auf den Zweck einer mehr oder weniger demokratischen Kontrolle der Verwendung von Gemeinwesenressourcen. Insbesondere dann, wenn es um die Umverteilung knapperer Mittel geht, also um Selektionsentscheidungen, soweit nicht nach dem Rasenmäherprinzip verfahren werden soll. Daneben und darunter vollzieht sich im Namen so legitimer Ziele wie "Transparenz" und "Qualität" der Übergriff des *new managerialism* auf alles, was sich nicht (kostendeckend) rechnet oder schlecht rechnen lässt.

Das verweist auf andere Modi der *Legitimationsfunktion* als diejenigen, die in Lehrbüchern und Legitimierungen der Evaluierung beschrieben werden. Dabei kennt jeder diese Modi, wie sie etwa von Suchman (1972) beschrieben wurden. Interessengruppen können ein schlechtes Programm durch Schönfärberei retten (*eye-wash*), und Programmfehlschläge kaschieren, wenn sie geeignete Indikatoren bilden (*white-wash*). Mit umgekehrter Indikatorenwahl lässt sich ein Programm sabotieren, indem ausschließlich negative Resultate gesucht werden (*submarine*). Handlungsaufschub lässt sich bestens begründen mit der Notwendigkeit, weitere Daten zu beschaffen (*postponement*). Und Evaluation lässt sich als Ritual der Unparteilichkeit betreiben, während man die Ergebnisse im Entscheidungsprozess ignoriert (*posture*). Über den größten Evaluierungsschub in der Geschichte des deutschen Wissenschaftssystems bei der deutschen Wiedervereinigung wäre ebenfalls zu sprechen - aber nicht in diesem Beitrag.

Dieser *Doppelcharakter* von Evaluation als Instrument von Herrschaft und Befreiung, Aufklärung und Verschleierung, kennzeichnet auch die anderen Verfahren *sozialer Rationalisierung*, wie Elias, Foucault oder zuvor Max Weber herausgestellt hatten. Unter sozialer Rationalisierung versteht Weber die sozialhistorische Ausdehnung und Durchsetzung der Warentausch-Beziehungen und der Zweck-Mittel-Rationalität, der Rechnungs- und Betriebsführung sowie der bürokratischen Verwaltung, und nicht zuletzt der entsprechenden Wirtschaftsgesinnung (vgl. Schwarz 2004). Diese Durchsetzung beschreibt er als Koevolution, in der sich die private Lebensführung an die kapitalistische Wirtschaftsverfassung anpasst, indem die ihr gemäße Vorstellung vom richtigen Leben von den Subjekten verinnerlicht wird, was ihr wiederum den weiteren Weg ebnet.

Tatsächlich finden wir diese Art von Subjektivität und Rationalität idealtypisch in der von Kappler eingangs kritisierten Objektivitätsfiktion des Controlling, ferner im Formalisierungswahn der Qualitäts-Zertifizierung, oder im beflissenen Meß-Eifer größerer Teile der Hochschulpolitik. Einige BeobachterInnen wie Michael Power (1997) oder Marilyn Strathern (2000) sehen in der *evaluation explosion* eine pathologische Entwicklung der Gesellschaft, in der quasi jeder jeden bewertet: Rating-Agenturen die Banken und ihre Kunden, ja ganze Staaten, Akkreditierer die Universitäten und Studiengänge, Auditierer die Zulieferer und Abnehmer, Controller

und Evaluierung schaffe das zur Legitimation nötige Wissen. Die Gutachtensflut und die Kommissionitis scheinen ihm Recht zu geben.

die Mittelverwender, und Berater die Controller. Powers *audit society* ist eine Misstrauensgesellschaft, in der zwanghaft möglichst viele Daten registriert werden, um sich Handlungsoptionen offen zu halten, in der aber immer seltener wissend gehandelt wird. Obwohl sich doch Management und Regierung gerade über den Anspruch eines *wissensbasierten Handelns* legitimieren.

Komplementär dazu wird die Wissenschaft bzw. die Universität seitens der Wirtschaft und der Politik immer offensiver mit dem Anspruch eines immer *handlungsorientierteren Wissens* konfrontiert. Wissenschaftliche Wissensproduktion und Ausbildung verliert damit zusehends Leitbegriffe wie 'Erkenntnis' oder 'Bildung' als Legitimationsquellen. Gemessen wird sie stattdessen an ihrem ökonomischen Beitrag zur Standortsicherung und an ihrer Lieferfähigkeit für politisch umweglos verwendbares Steuerungswissen. In einem solchen Szenario wird Evaluierung vollends zur Fremdbestimmung. Unter den Klängen sich entfernender Demokratisierungsmelodie mutiert sie zur Drohung einer *Evaluierungsdemokratie*, in der die Sachwalter des Gemeininteresses keinen Eigensinn in der Begründung des evaluierten Handelns mehr dulden und Evaluierte nur noch in die Umsetzung involvieren wollen.

Stimmt dieses Szenario? Wie weit geht die Vereinnahmung durch Forderungen wirtschaftlicher und politischer Verwertbarkeit? Und wie ausgeprägt ist die *Rationalität* der Kandare, an die die Universität genommen werden soll? Diesen Fragen und diesem wichtigen Strang der Debatte wollen wir hier nicht weiter nachgehen (vgl. dazu etwa Wings 1998; Willmott 2000; Prisching 2001 oder Greenwood 2004). Stattdessen interessiert uns die Frage, wo der Messwahn ins Zeremonielle umschlägt, in reines Ritual, dessen Geltung mit Effizienz oder Qualität nicht erklärt werden kann. Die Frage also, wo der Legitimationsbedarf verschiedener Akteursgruppen jene *Fassadenrationalität* schafft, die dem soziologischen Neoinstitutionalismus zufolge in der wirtschaftlichen Praxis ebenso bedeutsam ist (z.B. Meyer/Rowan 1977; 1978). Letztlich zielt genau darauf auch die Kritik von Power und Strathern ab: dass sich Bewertungsroutinen bürokratisch verselbstständigen und der Mangel an inhaltlichem *Sinn* durch den Triumph der Verwaltungsmittel kompensiert wird. Wer dürfte dabei nicht an Mark Twains zwingende Charakterisierung jedes Formalismus: "Als sie das Ziel aus den Augen verloren hatten, verdoppelten sie ihre Anstrengungen".

2.1. Kennzahlen-Konstruktivismus in der akademischen Praxis

Natürlich ist Evaluierung als Prüfprozedur für Qualität und Leistung nichts Neues in der akademischen Praxis. Wissenschaft ist durchdrungen von Verfahren der Selbstevaluierung, wobei in der *Forschung* das *Gutachtersystem* bzw. die Peer-Review an erster Stelle steht ("TÜV der Wissenschaften", Fröhlich 2003). Es wird angewandt in wissenschaftlichen Zeitschriften und Berufungsverfahren, bei Vergabe von Drittmitteln und Preisen, bei der Kongresszulassung und vielen weiteren Selektionssituationen. Zu den wichtigsten Indikatoren dieser akademischen Selbstbewertung zählen die Menge der Veröffentlichungen, ihr Rang nach Reputation der Pub-

likationsorgane, Zitations-Indices und die Drittmittelwerbung durch die Forscher, Forschergruppen oder Fakultäten. Die Einrichtung von Hochschulbeiräten mit Vertretern anderer gesellschaftlicher Funktionsbereiche ist ein neueres Verfahren, ebenso die geplante Forschungsmittelverlosung der amtierenden deutschen Bildungsministerin (unter der Überschrift *Eliteum*), die mit neuen Ranking-Prozeduren einhergehen wird. Die Vorbehalte der scientific communities richten sich also weniger gegen Evaluierung an sich als gegen die zunehmende Fremdbestimmung des Verfahrens und der Kriterien.

In der *Lehre* ist die Evaluierung im deutschsprachigen Raum traditionell schwach ausgeprägt: Bislang gelten vorrangig Kriterien der Tonnenideologie, wonach die Mittelzuweisung entsprechend der berechneten "Belastungskennziffern" (Lehrbelastung) erfolgt und Forschung allenfalls am Rande zählt. Die Lehrevaluation, die in den 1990er Jahren eingeführt und nun im Zuge des "Bologna-Prozesses" flächendeckend praktiziert werden soll, gehört ebenso zu den neueren Verfahren wie die Akkreditierung (durch Private!) bei Einrichtung von (öffentlichen) BA-MA-Studiengängen und Weiterbildungsstudiengängen oder die Zeitschriften-Rankings bzw. Uni-Rankings nach "Qualität" der Absolventen durch Personalmanager.

Die Wissenschaftsforschung weist auf die systematischen Grenzen und Nebenfolgen schon der bisherigen Evaluierung hin, ganz im Gesite Thomas Kuhns, der auf die Irrtümer in der Selbstbeschreibung von Wissenschaft aufmerksam machte. Viele ihrer Praktiken der Selbstevaluierung legitimieren den Studien (und unserer eigenen Erfahrung) zufolge lediglich partikuläre Interessenverfolgung mit höheren Werten; oder sie legitimieren ihre Selbstreferenzialität allein mit der Regelmäßigkeit ihrer Praktiken.⁵ Die Erklärungen hierfür sind unterschiedlich: neben systemtheoretischen im Sinne von Luhmanns *Legitimation durch Verfahren* (1969) und soziologisch-neoinstitutionalistischen, die auf Konformität mit Umwelterwartungen setzen, hat sich eine eigene "Ökonomik der Wissenschaft" etabliert (Übersichten z.B. bei Stephan 1996 und Peter 2004). Mit den altbekannten Mitteln des ökonomischen Ansatzes analysiert sie die Anreizwirkungen des akademischen Systems und deren Kompatibilität mit den Leistungserwartungen an dieses System.

Alle genannten Ansätze dürften aber zum gleichen Schluss kommen, und das ist die gute Nachricht für alle Evaluierungsgläubigen: Kennzahlen produzieren das, was sie messen. Zumindest als Artefakt. Man könnte das auch als *Kennzahlen-Konstruktivismus* bezeichnen. Kennzahlen fördern konstruktive (kreative?) Anpassungsleistungen der Akteure und ihrer explizierten Wertsysteme. Es handelt sich eben nicht um nonreaktive Verfahren. Britische Kollegen, die nicht genannt werden möchten, berichten uns beispielsweise, dass ihre Fakultät vor jedem großen Evaluierungstermin (alle vier Jahre) "Auswechselspieler" einkauft, die "golden eggs" legen und in den betreffenden high-rated refereed journals (oder sollten wir sagen:

⁵ Ein Vorwurf, der auch gegen die DIN/ISO-Normung von Qualitätssicherung in der Wirtschaft immer wieder vorgebracht wird.

deep impact journals?) besonders publikationsschwanger sind. Das hübscht ihre faculty scores auf und zahlt sich ökonomisch in Form höherer Mittelzuweisung aus (ähnlich Willmott 2000, 107). Zu diesem System der Peer Review und dem von ihm produzierten "Prestigeschmuck" hat Fröhlich (2002; 2003) u.a. folgende Befunde zusammengetragen:

- Den Mythos, die wissenschaftliche Qualität eines Journals steige mit seiner Abweisungsrate, bedienten manche Journals, indem sie schon Kürzungsforderungen oder eine Verlegung in die Forschungsnotizen als Abweisung deklarierten;
- vielfach sei unklar, welche Zeitschriften überhaupt extern begutachtet werden und welche nicht, und Lesern würden bewusst Angaben dazu vorenthalten, wie die Peer-Review praktiziert wird;
- nicht wenige *behaupteten*, auf Basis eines Peer-Review-Verfahrens zu editieren, praktizieren es aber kaum oder gewährten großzügige Ausnahmen;
- auch in sauber begutachteten Journals würden oft nur Bruchteile extern qualitativ evaluiert; gleichwohl schmücke sich alles dort Veröffentlichte mit diesem Nimbus;
- Journals schmückten sich mit einem großen Pool an Gutachtern, während faktisch "nachweislich nur ein winziger Bruchteil überhaupt eingesetzt" werde;
- Gutachten über Beiträge anderer Theorierichtungen würden bestimmten Gutachtern zugeschanzt, die - teils für mehrere Journals tätig - mitunter eine ganze Disziplin beherrschen könnten.

Zugleich würden NaturwissenschaftlerInnen "kaum offiziell Publiziertes lesen, bzw. Sozial- und KulturwissenschaftlerInnen die (paradigmatische) Konkurrenz möglichst nicht zur Kenntnis nehmen. Das liegt (a) an der Publikationsüberflutung (diese wiederum ein Effekt der Evaluationsemsigkeit, die sich vor allem an formalen Kriterien wie Zahl der Publikationen festmacht" (Fröhlich 2003, 29). Das fördert die Salamipublikationstaktik und das Selbstzitat, denn dieses geht ungehindert in den Zitierindex ein, und macht nach Fröhlich in Disziplinen wie der Medizin bis zu 60% aller "erwirtschafteten Zitate" aus (*Inzucht-Faktor*). Wie das Absurde rekursiv stabilisiert wird, lässt sich leicht weiter illustrieren. Durch Zitierkartelle und individuelles "strategisches Zitieren" dienen sich akademische Glattlinge in high rated Journals. In der *Berufungspraxis* wird die "wissenschaftliche Exzellenz" der Bewerber immer mehr über *impact factors* operationalisiert. Das favorisiert "Glasperlenspieler", wie der Leiter des ZEW Werner Franz (2003, 8) beklagt. Die Bewerberauswahl verkomme ...

"... zu einem bloßen Zählappell von wissenschaftlichen Beiträgen in international höchst angesehenen Zeitschriften. Alle anderen Aspekte bleiben unberücksichtigt, wie beispielsweise weitere wissenschaftliche Arbeiten, das didaktische Können, die Teamfähigkeit und die Kooperationsbereitschaft. Folgerichtig werden die Bewerber

lediglich nach der Anzahl ihrer Publikationen in den internationalen Top-Journalen gereiht".⁶

Wer daraus nur folgert, das Problem liege in der Selbstevaluierung, weshalb Fremdevaluierung zu bevorzugen sei, dem hat sich die Funktionsweise des Kennzahlen-Konstruktivismus nicht ganz erschlossen. Die Versuchungen und die Möglichkeiten des Fassadenbaus sind 'nach außen' keineswegs kleiner. Zumal sich externe Instanzen ja ebenfalls auf diese oder ähnliche Indikatoren stützen. Umgekehrt können sich die Evaluierten nicht auf die Rationalität oder Neutralität der Evaluierer verlassen, ebenso wenig auf die Konsistenz ihrer Maßstäbe. Im Frühjahr 2004 wurde z.B. das Forschungsinstitut ISO in Köln evaluiert und nach negativer Bewertung zur Schließung empfohlen. Ihm wurde zum Verhängnis, was man jahrelang von ihm gefordert hatte, nämlich handlungsorientiertes Steuerungswissen für die regionalen Politik- und Wirtschaftsakteure zu produzieren. Die Evaluierer hatten hingegen eine nicht ausreichende Produktivität von Artikeln für "deep impact" Journals moniert. Dass dies kein Einzelfall ist, macht Friedhelm Neidhard (2001) im Vorwort zu einer Studie über Evaluierungsprozesse von Instituten der Wissenschaftsgemeinschaft Leibnitz deutlich - also einer Studie zur Evaluierung der Evaluierung:

"In mehreren Instituten wurde berichtet, dass ihre Anstrengungen und Leistungen für außerakademische Nutzergruppen weder hinreichend wahrgenommen noch angemessen eingeschätzt wurden. ... Wissenschaftsratsevaluationen ... besitzen hier offenkundig einen akademischen Bias". Besonders wichtig seien daher "Konventionen zur Bestimmung und Messung auch der außerwissenschaftlichen Leistungen wissenschaftlicher Einrichtungen" bzw. zur "Balancierung von akademischen und außerakademischen Verpflichtungen" (Neidhard 2001, 8f.).

Hier könnte man anmerken, dass externe Evaluation eben nicht allein auf Angehörige des Wissenschaftssystems setzen dürfe, weil dies a priori die akademische Selbstreferenzialität stärkt. Doch es ist weder gesichert, wenn nicht unwahrscheinlich, dass an der Evaluierung alle relevanten Gruppen beteiligt werden (Parker/Jary 1995), noch dass außerakademische Beobachter wie Kultusministerien andere als formale und leicht handhabbare Kennzahlen (z.B. Auslastungskennziffern) berücksichtigen. Ebenso wenig ist gesichert, dass sie neben finanzrelevanten Zahlen auch solche zur Qualität verwenden, oder dass sie die Evaluierung überhaupt verwenden. So stellt z.B. Berthold (2002, 160) zur Lehrevaluation fest, dass "Evaluations-

⁶ Vgl. dazu auch Laske/u.a. (2000a, 190). Abgesehen davon können natürlich nur englischsprachige Journals die große weite Welt der wissenschaftlichen Exzellenz repräsentieren. Und von den Vertretern einer *Punktesammelwissenschaft* kann nicht erwartet werden, dass sie dem widersprechen. Man mag sich damit trösten, dass es sich für sie selbst nicht unbedingt auszahlt. Wie Franz (2003, 8) weiter berichtet: "Vor einigen Jahren beantwortete ein amerikanischer Wissenschaftler die Frage, warum er nicht den in einer renommierten, englischsprachigen Fachzeitschrift abgedruckten Originalbeitrag deutscher Wissenschaftler, sondern lieber dessen einleitende Zusammenfassung seitens eines anderen amerikanischen Wissenschaftlers zitiert habe, mit 'It does not pay off to quote a German'".

berichte relativ folgenlos blieben und mittlerweile einige zig-tausend Lehrevaluationsbögen unausgewertet in Aktenschränken auf den Tag ihrer Vernichtung warten". Wo die Mittelzuweisung (bzw. -einwerbung) an die Zahl der Absolventen gekoppelt ist, wird deren Zahl eben steigen, die Anforderungen an sie werden sinken, Notendurchschnitte werden sich "bessern". Kieser (2003, 10) berichtet von der Erfahrung in angelsächsischen Ländern, wo Rankings einen hohen Anreiz zur (Schön-)Färbung von Evaluierungsergebnissen induziert wird:

"Ein Abrutschen in den Rankings ist eine Katastrophe: Weniger Stiftungsgelder, weniger gute Bewerber, geringere Wertschätzung durch die einstellenden Unternehmen ... Also versuchen Dekane alles, um ein solches Abrutschen zu verhindern. ... Sie instruieren ihre Studierenden nachdrücklich, auch in ihrem ureigensten Interesse ... anzukreuzen, dass sie einen hervorragenden Unterricht genießen. Dies führt wiederum dazu, dass die Studierenden die Macht, die ihnen als Beurteiler von Studienprogrammen zukommt, zu missbrauchen, um von ihren Professoren vorwiegend unterhaltsamen und weniger fordernden Unterricht und gute Noten zu erpressen".

Wettbewerb an sich ist eine gute Sache. Die Frage ist nur, über welche Rahmenbedingungen man ihn worauf lenkt. Insofern kann man das alles nicht unter schlechter Anwendung einer guten Sache abtun, nach dem rationalistischen Credo: *Das Prinzip ist gut, nur die Realität ist schlecht* (Moldaschl 1990). Vielmehr muss nach systematischen Gründen gefragt werden. Und die liegen, unserer Ausgangsannahme zufolge, zu einem erheblichen Teil im politischen Charakter der Evaluierung. Sie sind begründet in der Tatsache, dass die internen und externen Akteure sich die Evaluierung interessenbestimmt zu Nutze machen.

"Evaluationsergebnisse werden in diesem Prozess instrumentell nur dann und dort genutzt, wo sie den Akteuren, egal auf welcher Ebene, für die Durchsetzung ihrer Interessen dienlich sind. Fächer, die sich bislang um die Umsetzung von Zielvereinbarungen ... nicht kümmerten, ziehen diese nun plötzlich zur Legitimation ihrer eigenen Handlungsstrategien heran. Die Landeregierung wiederum nutzt Evaluationen, um die Durchsetzungsmacht ihrer bereits getroffenen Entscheidungen zu erhöhen. Und Hochschulleitungen entscheiden sich für oder gegen den Erhalt von Fächern, völlig unabhängig von deren Evaluationen, weil die Altersstruktur des Personals eine schnelle Erfüllung der Einsparauflagen gewährleistet. Evaluationen werden negiert und verschwiegen oder aber in den Vordergrund gerückt, wenn sich dies im Machtpoker als nützlich erweist" (Hanft 2003, 9f.).

Heißt das, man solle die Evaluierung besser sein lassen? Soll man die immer knapperen Ressourcen also besser in den Ausbau der Universitäten oder wenigstens in die Bremsung ihres Abbaus stecken als in eine leer laufende Legitimationsmaschine, einen Konformismus-Generator? Die Frage greifen wir im dritten Abschnitt nochmals auf. Zuvor ist noch ein Wort zum gegenwärtigen ideellen Kontext der Hochschulmodernisierung zu sagen.

2.2. Taylorismus der Hochschulreform und Taylorismus der Seele

Was man mit Evaluierung macht, bzw. was wahrscheinlich gemacht werden wird, ist nicht nur abhängig von den jeweiligen Partikularinteressen, sondern auch von übergreifenden Diskursen, hegemonialen Leitbildern und eingefahrenen Praktiken. Das eigentümliche an der laufenden Hochschulmodernisierung mindestens im deutschsprachigen Raum ist nun die kuriose Mischung von Mustern 'alter' und 'neuer' Moderne, von bürokratischer und managerieller Steuerungsrationalität.

"Die Herrschaft erster Ordnung durch das Ministerium beginnt zwar formal durch mehr Autonomie abgelöst zu werden; gleichzeitig wird - über managementbezogene Effizienzforderungen und Controllingsysteme - sehr nachdrücklich am Aufbau einer Herrschaft zweiter Ordnung gearbeitet" (Laske/u.a. 2000a, 181).

Die gegenwärtige Hochschulreform weist frappierend viele Züge des Taylorismus auf. Frappierend deshalb, weil sie ihre Vorbilder und Legitimationsressourcen eben nicht mehr von der Demokratisierungsidee bezieht, sondern vom manageriellen Modernisierungsdiskurs (in der Verwaltungsmodernisierung generell, Moldaschl/u.a. 2004). Und der definiert sich gerade über die Abgrenzung vom Taylorismus. Mehr noch: er bezieht seine Identität aus dessen "Überwindung". Tatsächlich finden sich viele "post-tayloristische" Begründungen, Modellvorstellungen und Maßnahmvorschläge in den Reformkonzepten: größere finanzielle Autonomie der Hochschulen wie bei Profit Centers; Dezentralisierung der Vergabe von Studienplätzen und Auswahlrecht der Hochschulen; Kunden- und Dienstleistungsorientierung; größere Selbständigkeit der Mitarbeiter zumindest in Gestalt der Juniorprofessoren, und so fort.

Auf der anderen Seite aber leben die alten Leitbilder der Verwissenschaftlichung des Produktionsprozesses in der *Massenproduktion der Bildung* (besser: von Qualifikation) fort, und damit die Zentralisierung der Verwaltung, funktionale und hierarchische Differenzierung, die penible Kontrolle der Ausführung (*Misstrauenskultur*), die Einführung von Leistungslohnsystemen,⁷ und natürlich die Priorität quantitativer Leistungsmaße (*Tonnenideologie*). Selbstverständlich geschieht und geschah das nie vollständig. Die Logik bürokratischer Rationalisierung bricht sich stets an den Rea-

⁷ Was sich die Konstrukteure der W-Besoldung und anderer Anreizmechanismen in Deutschland von ihren Lenkungsinstrumenten versprechen, ist ulkig. Professoren mit kleinen Geldbeträgen lenken zu wollen wie den Herrn Smith vom Herrn Taylor, wird allenfalls bei den Unselbständigen und Unmotivierten ansatzweise aufgehen können. Und bei denen stellt sich eher die Frage, aufgrund welcher Fehlsteuerung sie berufen wurden, siehe oben. Zumal uns die ökonomische Forschung lehrt, wie intrinsische Motivation durch extrinsischen Anreiz zerstört werden kann (z.B. Frey 1997). Aber es geht auch hier nicht um rationale, sondern um *symbolische* Funktionen, wie Stephan Laske (1977) schon am Thema "Anforderungsgerechtigkeit" der Arbeitsbewertung gezeigt hatte. Wissenschaftlicher Wettbewerb jedenfalls läuft im Großen und Ganzen nicht über den Code Geld, sondern über Codes wie Anerkennung, Selbstbestimmung und Erkenntnis. Schließlich hat man sich zu früheren Zeitpunkten entschieden, nicht in den Wettbewerb um möglichst hohe Einkommen einzutreten.

litäten, und die Realität ist stets "schlecht" aus der Sicht des Verfügungsinteresses und Berechenbarkeitsglaubens. Entscheidend ist somit nicht der Grad der Umsetzung dieser Logik, sondern ihre *Gültigkeit als Leitlinie* der guten Gestaltung - im Sinne des Thomas-Theorems.⁸

Insofern findet an den Universitäten eine Art nachholende Modernisierung statt - im Modus einer mittlerweile *veralteten Moderne*.⁹ Die Systemtheorie würde hier von paradoxer Intervention sprechen: "Neue Steuerungsinstrumente" wie die Evaluation sollen den Rückzug staatlicher Verantwortung einleiten, aber auf traditionell regulierende Weise. Wo die wirtschaftliche Praxis unter gewaltigen Geburtswehen neue, flexible Organisations- und Koordinationsformen hervorbringt, suchen viele Universitäts- und Ministerialverwaltungen den Weg im Bewährten, das sich schon lange nicht mehr bewährt. Aber das teilen sie immerhin mit manchen Wissenschaftsdisziplinen, die sich (wie etwa Psychologie oder Volkswirtschaftslehre) in ihrem Mainstream noch immer am kybernetischen Ideal einer Naturwissenschaft orientieren, die sich davon ihrerseits schon vor Jahrzehnten entfernt hatte.

Man kann also nicht sagen, dass Tonnenideologie, Zählappelle und kreativer Stillstand vorrangig von außen in die unschuldig nach Wahrheit suchende Wissenschaft induziert würden. Der schreiende Widerspruch zwischen erklärtem Ziel und beschrittenem Weg, zwischen dem Prinzip kreativer Wissens- und Wertschöpfung und der Form seiner Verfolgung, ist in beiden Welten alltäglich. Verwunderlich ist allenfalls der Eifer, in dem manche Befürworter des Messens das Schlechteste beider Welten zu verbinden trachten.

Nun vollzieht sich die Vermessung der Universität aber nicht bruchlos, und das nicht nur wegen deren innerer Widersprüche. Nach vormals verbreiteter Ableitungsseligkeit haben soziologische Beobachter gelernt, dass sich die mit Taylorismus Überzogenen wehren können, sei es durch explizite Gegenmachtstrategien oder (verdeckte) mikropolitische Taktiken - wie schon oben beschrieben. Allerdings findet dieses Korrekturpotential seinerseits Grenzen durch das, was Georg Zilian (2004) "Taylorismus der Seele" nennt, und was schon Max Weber als Formung einer Wirtschaftsgesinnung beschrieben hatte, die sich mit der sozialen Rationalisierung entfaltet und sie zugleich fördert. Die neuen Verhältnisse bringen also jene Subjektivität mit hervor, auf die sie sich stützen. Sie produzieren in ausreichender Zahl jene Subjekte, die das alles gut finden und sich damit arrangieren: Glasperlenspieler und Kennzahlen-Gaukler, akademische Eichhörnchen und admi-

⁸ "If men define situations as real, they are real in their consequences". Auf diese Einsicht von William Thomas und Florian Znaniecki (1920) stützt sich der Sozialkonstruktivismus, den man vor Verwechslung mit dem erkenntnistheoretischen Relativismus offenbar nicht schützen kann.

⁹ Zur Gegenüberstellung und Diskussion von Prinzipien klassischer und reflexiver Modernisierung - oder von "erster" und "zweiter" Moderne (vgl. Beck/u.a. 1996).

nistrative Erbsenzähler, Blockwarte der Effizienzkontrolle und all jene, für die Bildung nichts weiter ist als eine prozessoptimiert herzustellende Ware.

Die um sich greifenden Verfahren des Transparentmachens, des wechselseitigen Bewertens und der Selbstbewertung legen auch den Rückgriff auf Foucault nahe, auf seine materialreichen Analysen der Individualisierungspraktiken in Form der *Geständnistheorie*, der *Untersuchungsverfahren* bzw. *Prüfungen*, mit denen Menschen zu Subjekten gemacht werden (dazu Moldaschl 2003), bzw. zur "regierbaren Person". Darauf stützen sich z.B. Kappler (2000) und Vormbusch (2004), wenn sie versuchen zu erklären, warum die Kontrollierten mit den Prämissen und Maßstäben des Controllings einverstanden sind.

3. Reflexive Evaluierung als soziales Lernen

Damit erweitert sich nochmals die Frage, wie die Bedingungen beschaffen sein müssen, damit eine sinnvolle Aneignung von Evaluierungsverfahren an den Universitäten möglich und wahrscheinlich ist. Das können wir freilich nur anreißen, und fragen zunächst noch einmal nach, warum denn Evaluierung boomt.

3.1. Warum Evaluierung?

Dass sich jede Interessengruppe der Evaluierung ihrem jeweiligen Partikularinteresse gemäß bedienen kann, ist klar. Ebenso, dass die Evaluierenden bessere Chancen als die Evaluierten haben, ihre Interessen durchzusetzen. Jenseits dieser Interessendienlichkeit gibt es aber doch strukturelle Bedingungen, welche die *evaluation explosion* (Power 1997) fördern, also die Verbreitung von Evaluierung, Audits, Ratings, und Ähnlichem. Modernisierungstheorien haben die Vervielfältigung von Handlungsmöglichkeiten und Entscheidungserfordernissen als Generaltrend der Moderne beschrieben, begrifflich codiert mit "funktionaler Differenzierung" (Luhmann), "Multioptionsgesellschaft" (Gross) oder "reflexiver Modernisierung" (Beck; Giddens). Individuell, organisational und gesamtgesellschaftlich nimmt die Zahl der praktisch und der kognitiv verfügbaren Gestaltungslösungen und Rationalisierungsstrategien zu, womit Entscheidungen "entselbstverständlicht" werden. Wächst die Zahl der Optionen, so wächst auch die Kontingenz, d.h. die wechselseitige Bedingtheit von Gestaltungsentscheidungen. Das bedeutet: (a) Die Ergebnisse der Entscheidungen betreffen immer mehr andere Akteure, direkt oder indirekt, gegenüber denen sie vertreten werden müssen. Und sie lassen sich (b) zugleich schlechter vorhersagen.

Statt sich also entlang tradierter Praktiken mehr oder weniger "naturwüchsig" zu entfalten, wird der Rationalisierungsprozess zunehmend begründungspflichtig oder *reflexiv*, wie besonders Ulrich Beck (1986) mit dem Begriff der *Nebenfolgen* hervorhob. Zumindest im Prinzip ist nichts von der Begründungspflicht und Revisionsmöglichkeit ausgenommen. Es ist jedoch für Individuen, Organisationen und Insti-

tutionen essentiell, aus jenen Nebenfolgen bzw. nichtintendierten Folgen des Handelns zu lernen, die im zweckrationalen Handeln und im rationalistischen Weltbild meist ausgeblendet werden. Den Handel etwa mit impact factors (siehe obiges Beispiel zu akademischen "Auswechselspielern") hat sich niemand gewünscht. Er ist aber nicht die Perversion, sondern der Effekt eines Systems, welches man in guter Absicht (Auswahl der Besten) eingeführt hat, und das selbst neue Probleme schafft. Schlechte Systeme oder Maßnahmen sind solche, deren Anwendung mehr oder gravierendere Probleme schaffen, als sie lösen.

Die ganze westliche Wissenschaft ist schon unter diesen Verdacht geraten, und mit ihr die Aufklärung. Auf manche der Evaluierungsverfahren und -praktiken an Hochschulen trifft er sicherlich zu, wie *Evaluierungen der Evaluierung* gezeigt haben (z.B. Hanft 2001; Röbbcke/Simon 2001). Wenn Evaluierung aber eine unsinnige Praxis aufdecken kann, wenn sie Ressourcenverschwendung (und sei es: durch Evaluierung) nachweisen kann, dann hat sie ja einen Nutzen. Insofern ist die Evaluierung der Evaluierung - *reflexive Evaluierung* - eine der notwendigen Bedingung jeder dauerhaften Anwendung von Evaluierungsverfahren. Man muss analysieren, welchen Nutzen und welche Nebenfolgen sie haben, und sie dann entweder ändern oder abschaffen. Das Prinzip der Reflexivität, welches Verfahren der Evaluierung in einer Praxis realisieren, muss auch für diese Verfahren selbst gelten. Ein infinites Regress? Nein, schlicht die Anwendung der Regel, dass eine Regel stets auf Angemessenheit zu überprüfen ist. Das gilt für Handeln auf jeder Aggregationsebene, wie im Ebenenmodell des organisationalen Lernens von Argyris/Schön (1978).

Evaluierungsverfahren von Entscheidungen, Maßnahmen und Programmen sind somit als Teil einer umfassenden sozialen Rationalisierung zu verstehen, in all ihrer angedeuteten Ambivalenz. Die regelhafte, auf Dauer gestellte Anwendung entsprechender Prüfverfahren kann man (mit oder gegen Giddens, z.B. 1990) als *institutionelle Reflexivität* bezeichnen (vgl. zu Begriff und Operationalisierung Moldaschl 2004b). Zu den von (fast) allen akzeptierten Verfahren institutioneller Reflexivität auf der Ebene von Wissenschaft und Gesellschaft zählen die *Technikfolgenabschätzung*, die sich zur Technikgeneseforschung erweitert hat, und das *science assessment* (z.B. Gill 1994), das in der Wissenschaftsforschung aufging. Schließlich wollen (fast) alle wissen, was hinten rauskommt: Steuerzahler, Kreditgeber, Umweltschützer, sogar Wissenschaftler. Jedenfalls wenn wir erwarten können, dabei nicht allzu schlecht auszusehen. Auf organisationaler Ebene haben intelligentere Ansätze des *Organisationslernens* Methoden beschrieben oder entwickelt, mit denen sich Organisationen kontinuierlich selbst prüfen.¹⁰ Auf individueller und Gruppenebene gibt es

¹⁰ Der Kontinuierliche Verbesserungsprozess (KVP) ist ein Beispiel für entsprechende Methodensets - auch nicht per se, aber bei entsprechender Gestaltung. Auch reflexive Fassungen des Organisationslernens bieten solche Instrumentarien; freilich gibt es eine Mehrheit anderer Fassungen, die lediglich auf eine (zweckrationale) Objektivierung, ein Verfügbarmachen des in den Köpfen von Organisationsmitgliedern befindlichen Wissens hinauslaufen.

Institutionalisierungen wie *Supervision* für Berater, Therapeuten und andere Professionelle.

Man könnte in ähnlicher Weise wie oben bei der Evaluierung auch die Nebenfolgen des Controllings bilanzieren und die enormen Schäden beziffern, die es in den Unternehmen verursacht. Das tun, wenn auch selten quantifizierend, mittlerweile einige britische Journals.¹¹ Auch das ist eine Evaluierung gesellschaftlicher Praxis, die nur nicht unter diesem Namen firmiert. Dass Controlling auch einen Nutzen hat - oder haben kann, bezweifelt deshalb niemand. In Deutschland wurde 1997 mit der DeGEval ein Fachverband der Evaluierenden gegründet, der auch eine Zeitschrift für Evaluationsforschung herausgibt; lange nach entsprechende Aktivitäten in angelsächsischen und skandinavischen Ländern.

Neben solchen Indizien der 'evaluation explosion' gibt es allerdings auch andere, bei denen man fragen muss, ob wir nicht lediglich eine Aufblähung der Diskurse als der (anderen) Praktiken erleben. In Deutschland werden z.B. nur etwa 0,06%, also ein halbes Promille, der gesamten Forschungsmittel für Technikfolgenforschung und -bewertung aufgewandt (Kreibich 2004, 10). Und Unternehmen setzen meist nur einen winzigen Bruchteil jener Mittel für die Evaluierung jener Maßnahmen und Programme ein, die sie mit hohem Aufwand planen und implementieren (Moldaschl 2004b).

3.2. Was, wie und wen evaluieren?

Zu den Mysterien und Zweck-Mittel-Verkehrungen der Hochschulevaluierung zählt, dass die wissenschaftlichen Instanzen der Universität gegenüber ihrer Verwaltung keine Mittel und Maßstäbe der Evaluierung besitzen, schon gar keine Sanktionsmittel, obwohl ihre Arbeit doch in hohem Maße von diesen Verwaltungen - sagen wir: beeinflusst wird. Warum entwickeln wir keinen Bürokratiefaktor oder Behinderungskoeffizienten, keine Maße für Durchlaufzeit von Anträgen, Reiseabrechnungen; für das Commitment zwischen Unileitung und Kollegium; für den Grad der institutionellen Förderung von Drittmittelwerbung und -abrechnung; für den Aufwand und Erfolg von Weiterbildung für Universitätsmitarbeiter, und so fort? Warum sprechen wir nur auf den Fluren über die neuesten Erfindungen kulturministerieller Evaluierungsbürokratie, und kehren den Spieß nicht um? Wenn in der *audit society* jeder jeden auditiert, warum dann nicht wir die von Ministerien geschaffenen Rahmenbedingungen für die Qualität von Forschung und Lehre?¹²

¹¹ Z.B. "Critical Perspectives on Accounting": <http://www.academicpress.com/cpa>; "Accounting, Organizations and Society" oder "Advances in Public Interest Accounting". Weitere Links unter <http://accfinweb.account.strath.ac.uk/df/c5.html>.

¹² Auch das gibt es - wie sollte es anders sein - im Vorbildreich der Wirtschaft. Und dort soll das 360-Grad-Feedback (vgl. Neuberger 2000; Bröckling 2003) ja nicht nur für wechselseitiges Lernen sorgen, sondern mitunter auch für den Verzicht auf Mitarbeitergespräche, wenn sich die Führungskräfte damit ihre eigene Evaluierung ersparen.

Verstehen wir Evaluierung eben doch nur als Kontrollinstrument höherer politischer Entscheidungsinstanzen?

Wie gehen wir ferner mit dem wohl fortdauernden Abbau öffentlicher Mittel für die Universitäten um? Weigern wir uns im Interesse interner Konfliktvermeidung, selbst Kriterien der Verteilung zu entwickeln, werden die Entscheidungen entweder 'fremd' bestimmt, oder der Rasenmäher geht über uns hinweg. Oder der Zufall, wie es wohl jeder in seiner Fakultät schon erfuhr und wie Anke Hanft oben erklärte. Zum Ausgangspunkt der seit 2002 laufenden Hochschulreform in Österreich erklären Burtscher/Pasqualoni (2004, 25): "Der Universitätsreform leistete unter anderem eine Ressourcenverteilung nach dem Gießkannenprinzip Vorschub, da darin defensive Routinen sichtbar wurden, die den Interessensaushandlungsprozess und das Austragen notwendiger Konflikte sehr erschwerten". Sie sprechen damit die unseres Erachtens wichtigste Funktion der Evaluierung an, nämlich ihre *reflexive Funktion als Instrument des Diskurses* zwischen Akteuren über Ziele und Wege.

Das ist es, was im anspruchsvollen Verständnis von *formativer Evaluierung* als *bildendem* Prozess sozialer Verständigung zwischen verschiedenen Stakeholdern stets maßgeblich war. Eine erste Antwort auf die Wie-Frage. Praktisch muss sie immer konkret beantwortet werden. Es kann z.B. durchaus sinnvoll sein, internen Verteilungskämpfen aus dem Weg zu gehen und sich dem Rasenmäher zu unterziehen, oder der Willkür einer übergeordneten Instanz, wenn man damit gemeinsam einen Schuldigen hat und sich intern solidarisiert. Es sind Situationen denkbar, in denen damit mehr soziales Kapital gerettet werden kann als Einzelnen an finanziellem oder symbolischem Kapital verloren geht (vgl. zur Struktur und Dynamik universitärer Verhandlungssysteme Laske/Zauner 2000). Das Ziel kann also nicht in "brutalstmöglicher Aufklärung"¹³ liegen, nicht in der generellen Maximierung von Reflexivität.

3.3. Mit welchen Maßstäben?

In der Einleitung dieses Beitrags haben wir en passant drei Sinnbestimmungen genannt, unter denen wir uns eine konstruktive Aneignung von Evaluierungsverfahren an den Universitäten und in öffentlichen Forschungseinrichtungen vorstellen können.

Das erste ist das Prinzip *wissenschaftlicher Selbststeuerung*. Wissenschaft und Universitäten brauchen Autonomie, um ihre Aufgaben zu erfüllen, wie jedes andere soziale Teilsystem auch (auch wenn man ihre Autonomiefähigkeit skeptisch beurteilt, wie z.B. Laske/Hammer 1996). Wenn es unsere Aufgabe ist, neues Wissen und Erkenntnis zu generieren, gesellschaftliche (Fehl-)Entwicklungen zu diagnostizieren,

¹³ Dieser Begriff stammt von einem anerkannten Spezialisten für politische Unwahrheit, dem hessischen Ministerpräsidenten Roland Koch, ge- und versprochen anlässlich der CDU-Spendenaffäre in Hessen 2002.

zu prognostizieren und ggf. zu kritisieren, jüngere Menschen mit Urteilsvermögen auszustatten, Basisqualifikationen für eine selbsttätige Fortbildung zu schaffen, und künftig hoffentlich in vermehrtem Maß den im Berufsleben stehenden Menschen Reflexions- und Weiterbildungsmöglichkeiten geben zu können, dann ist nichts gegen Evaluierung einzuwenden, die uns Rückmeldungen darüber gibt, wie gut wir das hinbekommen und wie weit wir damit kommen. Evaluierung sollte dann aber auch *von uns ausgehen* und nachweisen, was uns dabei behindert.

Das zweite Prinzip, mit dem wir uns identifizieren können, ist das der verantwortlichen Autonomie bzw. der *verantwortlichen Selbstbestimmung*. Mit verantwortlich ist gemeint, dass unsere wissenschaftlichen Einrichtungen und wir selbst nicht autistisch auf jene "reine" Autonomie pochen, wie es im Vorwurf "Elfenbeinturm" impliziert ist. Die Gesellschaft finanziert unsere Institution und hat legitime Ansprüche an sie, an uns. "Die Gesellschaft" als Abstraktum kann aber keine Ansprüche an uns stellen. Das tun andere ihrer Institutionen, konkrete Anspruchsgruppen und ihre Vertreter. Und deren Ansprüche an Forschung und Lehre konkurrieren bzw. konfliktieren nicht nur oft untereinander, sondern mitunter auch mit den Erfordernissen guter Forschung und Lehre aus Sicht der Wissenschaft. Jeder externen Gruppe steht es frei, summative Evaluierungen unserer Arbeit vorzunehmen (z.B. die Universitäts-Rankings seitens der Wirtschaft oder der Medien), und damit auch Kritik.¹⁴ Manche haben die Macht, auf dieser Basis über das Schicksal der evaluierten Einrichtung zu entscheiden. Evaluierung aber, die für uns handlungsrelevant und subjektiv verbindlich werden soll, setzt jedoch unsere Beteiligung voraus, und *Diskurs* mit den gesellschaftlichen Gruppen. Dafür kommen nur anspruchsvolle Varianten formativer Evaluierung in Betracht. Sie haben die Funktion, diesen Diskurs mit zu organisieren, und die Verständigung über gesellschaftliche Ziele und Prioritäten zu fördern, über legitime und illegitime Ansprüche, dominant vorgetragene und gesellschaftlich vernachlässigte. Zum Beispiel über die Frage, wie man das Verhältnis von "Spitze" und "Breite" in der Bildung gestalten will. Evaluierung, die diesen Verständigungs- und Lernprozess fördert, nennen wir *reflexive Evaluierung*.

¹⁴ "Insgesamt genießen die staatlichen Hochschulen übrigens", so schreibt der Spiegel 3/2004, "bei potenziellen Arbeitgebern einen erstaunlich guten Ruf. Nur jeder dritte Personalchef meint, Absolventen privater Business Schools seien besser ausgebildet als ihre Konkurrenten von staatlichen Unis. ... Die Frage, ob das häufig sehr teure Studium an einer Privatuniversität besser auf Führungsaufgaben vorbereite, beantworten 70 Prozent mit 'Nein'. Übervolle Hörsäle, schlecht ausgestattete Bibliotheken und blätternder Putz an den Wänden sind also doch zu etwas gut, meinen die Personalchefs: Sich an einer Mammut-Universität durchboxen zu müssen, stähle fürs Berufsleben. Offenbar lässt sich ein guter Ruf auch nicht so schnell wieder ruinieren. Seit Jahren tummeln sich in Personaler-Rankings stets die gleichen Namen in bestimmten Fächern an der Spitze. Und allein das Image zählt - harte Faktoren wie die Absolventen- und Doktorandenquote, die Zahl der wissenschaftlichen Veröffentlichungen und der eingeworbenen Drittmittel spielen keine Rolle". Bedarf es noch anderer Beweise, dass man die "materiale Rationalität" der Evaluierung nicht überschätzen sollte?

Institutionelle Selbstkritik ist eine Notwendigkeit auch in der Wissenschaft, wie die obige Kritik am System der Peer Review und anderen Mechanismen innerer Konformitätsproduktion belegt. Ferner müssen wir uns mit den Folgen und Nebenfolgen unseres Handelns befassen, mit den "Externalitäten". Das kann man nicht einer kleinen Spezialdisziplin wie dem science assessment überlassen. Das ist die zweite Bedeutung *reflexiver Evaluierung*. Über die bloß rekursive Definition (Evaluierung der Evaluierung, vgl. Röbecke/Simon 2001) geht das weit hinaus. Die Selbstkritik wiederum ist eine unabdingbare Voraussetzung von Autonomie. Gegenwärtig ist freilich eher mangelnde Souveränität wissenschaftlicher Institutionen im Umgang mit Evaluation festzustellen, wie die eben zitierte Studie zeigt. Dazu noch einmal Neidhard (2001, 9f.):

"Wir haben bei der WZB-Recherche den Eindruck gewonnen, dass in den Instituten die Kompetenz, mit eigenen Erfahrungen die Evaluationsprozesse zu qualifizieren, bislang nur gering entwickelt ist. Dies wird damit zusammenhängen, dass in den meisten Instituten dauerhafte Erfahrungen mit Selbstevaluationen, in denen Leistungsmaßstäbe definiert und darauf bezogene Leistungskriterien experimentiert werden, kaum vorliegen, also auch nicht diskutiert und rasoniert werden". Die WZB-ForscherInnen wunderten sich ferner darüber, "... wie unbesehen in den WGL-Instituten gegenüber den Empfehlungen des Wissenschaftsrats Fügsamkeit und Folgebereitschaft entwickelt sind. Wir gehen davon aus, dass die Qualität der Evaluationen steigen würde, wenn die Institute sich selber jenseits aller Routinen stärker reflektieren und dann auch zu einem Selbstbewusstsein kommen würden, das aus der Rolle bloßer Evaluationsobjekte befreit".

Quis custodes ipsos custodiet?

"Wir mußten diese Stadt zerstören, um sie zu retten". Dieser berühmte Satz eines US-Offiziers im Vietnamkrieg, der kürzlich auf ein ganzes Land ausgedehnt wurde, scheint irgendwie auch auf die Hochschulen zuzutreffen. Wie andere Organisationen "verharren [sie] in ihren Routinen und zeigen erst dann Bereitschaft zu echten Reformen, wenn sie in ihrem Bestand gefährdet sind". So versucht Anke Hanft (2003, 10) der aktuellen Lage Positives abzugewinnen. Und echte Reformen sind nötig, denn "entstanden sind Massenausbildungsstätten, die den Namen Universität nicht mehr verdienen", so Jeanne Rubner (in der SZ vom 29./30.5.2004). Wenn Rubner dann allerdings die Lösung in "Spitzenforschung", "Eliteuniversität", "Eliteförderung", etc. sieht, dann reproduziert sie den ideellen Rahmen, in dem sich auch die Evaluierungsdebatte bewegt. Und dieser lässt keinen Enthusiasmus aufkommen über die Chancen der von uns favorisierten Leitbilder der Evaluierung. Vieles, womit man an deutschen Universitäten die Aufgaben der Massenbildung zu bewältigen sucht, erinnert vielmehr an Strategien der Massenproduktion vergangener Tage: Taylorisierung des Universitätsbetriebs und Verbetrieblichung nach einem Muster, das an den Lehrstühlen eben dieser Universitäten für längst überholt erklärt wird. Das heißt: Einführung von analytischer Arbeitsbewertung und Leistungslohn, Priorität quantitativer Leistungsmaße (Tonnenideologie) und Kontroll-

systeme, damit verbundene Misstrauenskultur, Bestenauslese statt breiter Investition in den Humankapitalstock, und so fort. Das wirft unter anderem eine Frage auf, die schon die Römer beschäftigt hatte: Wer kontrolliert eigentlich die Kontrolleure? Auf der anderen Seite heißt es zu recht: Wer kämpft, kann verlieren. Wer nicht kämpft, hat schon verloren.

Literatur

- Argyris, Ch./Schön, D.A. (1978): *Organizational Learning. A Theory of Action Perspective*. Reading, Mass.
- Baumgartner, P. (1998): Evaluation mediengestützten Lernens. In: Kindt, M. (Hg.): *Projektevaluation in der Lehre. Multimedia an Hochschulen zeigt Profil(e)*. Münster/u.a., 71.
- Beck, U. (1986): *Risikogesellschaft - auf dem Weg in eine andere Moderne*. Frankfurt/M.
- Beck, U./Giddens, A./Lash, S. (1996): *Reflexive Modernisierung. Eine Kontroverse*. Frankfurt/M.
- Berthold, Ch. (2002): Von der Evaluation zur strategischen Hochschulentwicklung - 16 Thesen. In: Reil, Th./Winter, M. (Hg.): *Qualitätssicherung an Hochschulen: Theorie und Praxis*. Bielefeld, 160-165.
- Bortz, J./Döring, N. (2002): *Forschungsmethoden und Evaluation für Human- und Sozialwissenschaftler*. Berlin/u.a.
- Bröckling, U. (2003): Das demokratisierte Panopticon. Subjektivierung und Kontrolle im 360°-Feedback. In: Honneth, A./Saar M. (Hg.): *Zwischenbilanz einer Rezeption*. Frankfurt/M., 77-93.
- Burtscher, Ch./Pasqualoni, P.-P. (2004): Demokratie im Universitätsbetrieb? In: Weber, W. G./Pasqualoni, P.-P./Burtscher, Ch. (Hg.): *Wirtschaft, Demokratie und soziale Verantwortung - Kontinuitäten und Brüche*. Göttingen, 347-376.
- Franz, W. (2003): Berufungen. In: ZEW news 11, 8.
- Frey, B. S. (1997): *Markt und Motivation. Wie ökonomische Anreize die (Arbeits-)Moral verdrängen*. München.
- Fröhlich, G. (2002): Anonyme Kritik. Peer Review auf dem Prüfstand der empirisch-theoretischen Wissenschaftsforschung. In: Pipp, E. (Hg.): *Drehscheibe E-Mitteleuropa*. Wien, 129-146.
- Fröhlich, G. (2003): Evaluation wissenschaftlicher Leistungen. In: *medizin-bibliothek-information* 3 (2), 29-32.
- Giddens, A. (1990): *The Consequences of Modernity*. Oxford.
- Gill, B. (1994): Folgenerkenntnis. Science Assessment als Selbstreflexion der Wissenschaft. In: *Soziale Welt* 45, 430-453.
- Greenwood, D. (2004): University Reforms and the Future of the Social Sciences. In: Weber, W. G./u.a. (Hg.): *Demokratie und soziale Verantwortung in Unternehmen*. Göttingen (im Erscheinen).
- Hanft, A. (Hg.) (2001): *Grundbegriffe des Hochschulmanagements*. Neuwied.

- Hanft, A. (2003): Evaluation und Organisationsentwicklung. Eröffnungsvortrag zur 6. Jahrestagung der Deutschen Gesellschaft für Evaluation e.V. (DeGEval). 8.-10. Oktober 2003. Hamburg.
- Imai, M. (1992): Kaizen. Der Schlüssel zum Erfolg der Japaner im Wettbewerb. Landsberg.
- Kappler, E. (2000): Die Produktion der regierbaren Person. In: Witt, F. H. (Hg.): Unternehmung und Informationsgesellschaft. Wiesbaden, 237-278.
- Kieser, A. (2003): Entwicklungsperspektiven der wirtschaftswissenschaftlichen Ausbildung in Deutschland. In: Neue Ökonomie der Arbeit. Ergänzungsband zur Festschrift zum Anlass des 10-jährigen Gründungsjubiläums der Fakultät für Wirtschaftswissenschaften der TU Chemnitz, 3-18.
- Kreibich, R. (2004): Zur Organisation von Verantwortung im Dialog von Wissenschaft und Politik. IZT ArbeitsBerichte Nr. 1. Berlin.
- Laske, St. (1977). Die "Anforderungsgerechtigkeit" in der Arbeitsbewertung oder die Funktion von Fiktionen. In: Gohl, J. (Hg.): Arbeit im Konflikt. Probleme der Humanisierungsdebatte. München, 142-162.
- Laske, St./Zauner, A. (2000): Architektur und Design universitärer Verhandlungssysteme. In: Laske, St./u.a. (Hg.): Universität im 21. Jahrhundert. München/Mering, 447-487.
- Laske, St./Scheytt, T./Meister-Scheytt, C./Scharmer, C. (Hg.) (2000): Universität im 21. Jahrhundert. Zur Interdependenz von Begriff und Organisation der Wissenschaft. München/Mering.
- Laske, St./Habersam, M./Kappler, E. (Hg.) (2000): Qualitätsentwicklung in Universitäten. Konzepte, Prozesse, Wirkungen. München/Mering.
- Laske, St./Meister-Scheytt, C./Weiskopf, R. (2000a): Qualitäten der Qualität in Universitäten. In: Laske, St./Habersam, M./Kappler, E. (Hg.): Qualitätsentwicklung in Universitäten. München/Mering, 177-221.
- Laske, St./Hammer, R. (1996): Zur Autonomiefähigkeit der Universität - eine eher skeptische Nabelschau. In: Altrichter, H./Schatz, M. (Hg.): Qualitäten von Universitäten. Innsbruck.
- Laske, St./Weiskopf, R. (1996): Quality - A linguistic Trojan Horse? Reflections on the Quality Discussion in Higher Education. Unveröff. Manuskript. Innsbruck.
- Lewin, K. (1953): Die Lösung sozialer Konflikte. Bad Nauheim.
- Lincoln, Y. S./Guba, E. (1985): Naturalistic Inquiry. Beverly Hills, CA.
- Lindblom, Ch. (1990): Inquiry and Change: The Troubled Attempt to Understand and Sharpe Society. New Haven, Conn.
- Luhmann, N. (1969): Legitimation durch Verfahren. Neuwied.
- Mendelsohn, E./Nelkin, D./Weingart, P. (Hg.) (1978): The Social Assessment of Science. Wissenschaftsforschung. 13 Report. Universität Bielefeld.
- Meyer, J. W./Rowan, B. (1977): "Institutionalized Organizations: Formal Structure as Myth and Ceremony". In: American Journal of Sociology 83 (2), 340-363.
- Meyer, J. W./Rowan, B. (1978): The structure of educational organizations. In: Meyer, M. W./Freeman, J. H./et al. (Eds.): Environments and Organizations. San Francisco, 78-109.
- Moldaschl, M. (1990): Das Modell ist gut, nur die Realität ist schlecht - Expertensysteme in der Arbeitswelt. In: Technische Rundschau 82 (49), 104-112.

- Moldaschl, M. (2003): Foucaults Brille. Eine Möglichkeit, die Subjektivierung von Arbeit zu verstehen? In: Moldaschl, M./Voß, G. G. (Hg.): Subjektivierung von Arbeit. München/Mering, 149-192.
- Moldaschl, M. (2004a): Partizipation und/als/statt Demokratie. Zum Entwicklungsverhältnis von gesellschaftlicher Demokratisierung und organisationaler Partizipation In: Weber W. G./u.a. (Hg.): Demokratie und soziale Verantwortung in Unternehmen. Göttingen (im Erscheinen).
- Moldaschl, M. (2004b): Institutionelle Reflexivität. Papers und Preprints des Lehrstuhls für Innovationsforschung und nachhaltiges Ressourcenmanagement der TU Chemnitz Nr. 1, 20.
- Moldaschl, M./Hinz, A./Wex, T. (Hg.) (2004): Reorganisation im Dritten Sektor. München/Mering.
- Neidhard, F. (2001): Vorwort zu: Röbbcke M./Simon D.: Reflexive Evaluation. Ziele, Verfahren und Instrumente der Bewertung von Forschungsinstituten. Berlin, 7-12.
- Neuberger, O. (2000): 360°-Feedback: Alle fragen? Alles Sehen? Alles Sagen? München/Mering.
- Parker, M./Jary, D. (1995): The McUniversity. Organizations, management and academic subjectivity. In: Organization Science 2 (2), 319-338.
- Peter, F. (2004): Legitimierung des Wissens - Zur Rolle des impliziten Wissens. In: Held M./Kubon-Gilke, G./Sturn, R. (Hg.): Ökonomik des Wissens. Marburg, 151-172.
- Power, M. (1997): The Audit Society. Rituals of Verification. Oxford.
- Prisching, M. (2001): Die Vermarktlichung der Universität. Beitrag für die Österreichische Forschungsgemeinschaft Oktober 01, <http://members.eunet.at/oefg/text/universitaet/universi.pdf>. [Zugriff 10.9.02].
- Röbbcke, M./Simon, D. (2001): Reflexive Evaluation. Ziele, Verfahren und Instrumente der Bewertung von Forschungsinstituten. Berlin.
- Schwarz, C. (2004a): Evaluation als modernes Ritual. Zur Ambivalenz gesellschaftlicher Rationalisierung am Beispiel Virtueller Universitäten. Unveröff. Dissertation.
- Schwarz, C. (2004b): Evaluation im Spannungsfeld: zwischen Projektmanagement und Politikberatung, in: Tergan, S. O./Meister D./Zentel, P. (Hg.): Evaluation von E-learning. Münster.
- Stephan, P. E. (1996): The Economics of Science. In: Journal of Economic Literature 34 (3), 1199-1235.
- Stockmann, R. (2004): Evaluation in Deutschland. In: Stockmann, R. (Hg.): Evaluationsforschung. Opladen, 13-43.
- Strathern, M. (2000): Audit Cultures. Anthropological studies in accountability, ethics and the academy. London/New York.
- Suchman, E. (1970): Action for what? A Critique of Evaluative Research. In Toole, R. O. (Ed.): The Organisation, Management, and Tactics of Social Research. Cambridge, MA.
- Thomas, W./Znaniecki, F. (1920): The Polish Peasant in Europe and America. Boston.
- Wächter, H. (Hg.) (2001): Qualitätsmanagement auf dem Prüfstand. Reflexionen über DIN ISO 9000 und TQM. Wiesbaden.
- Weber, J./Schäffer, U. (1999): Sicherstellung der Rationalität von Führung als Aufgabe des Controlling? In: Die Betriebswirtschaft 99, 731-747.

- Widmer, Th. (2004): Qualität der Evaluation - wenn Wissenschaft zur praktischen Kunst wird. In: Stockmann, R. (Hg.): Evaluationsforschung. Opladen, 83-110.
- Willmott, H. (2000): On measuring and Commodifying Research Quality. In: Laske, St./Habersam, M./Kappler, E. (Hg.): Qualitätsentwicklung in Universitäten. München/Mering, 101-118.
- Wingens, M. (1998): Wissensgesellschaft und Industrialisierung der Wissenschaft. Wiesbaden.
- Vormbusch, U. (2004): Management by Foucault? Betriebliches Controlling und Subjektivierung. In: Moldaschl, M. (Hg.): Selbstmanager und Intrapreneure. München/Mering (im Erscheinen).
- Zilian, H. G. (2004): Die Seele des Menschen im Turbokapitalismus. In: Moldaschl, M. (Hg.): Selbstmanager und Intrapreneure. Subjektivierung von Arbeit II. München/Mering (im Erscheinen).

AutorInnen:

Prof. Dr. phil. habil. Dr. rer. pol. Manfred Moldaschl, Professur für Betriebswirtschaftslehre an der TU-Chemnitz, Lehrstuhl für Innovationsforschung und nachhaltiges Ressourcenmanagement, und Direktor des Instituts für Innovationsmanagement und Personalentwicklung (IfIP), Chemnitz

Arbeitsschwerpunkte: Organisation und Arbeit, Innovation und Strategie, Intervention/Beratung, Netzwerke, Subjektivität, Reflexivität, Human- und Sozialkapital

Dr. des. Christina Schwarz, Soziologin und wissenschaftliche Mitarbeiterin am Weiterbildungsstudium Arbeitswissenschaft der Universität Hannover

Arbeitsschwerpunkte: Evaluation, e-learning, Medien- und Organisationsentwicklung sowie Technik- und Wissenschaftssoziologie