

## **Forschungsevaluation als Organisationslernen**

### **Eigenerfahrungen**

Während mich gelegentlich Forschungen im Umkreis des Topos Organisationslernen (OL) beschäftigten (aber nur anhand von Fragestellungen der Politischen Soziologie), blieb mein Wissen über Evaluationsprozesse auf die eigenen Erfahrungen beschränkt. Doch schon die Primärerfahrungen liefern etwas Material, um Hypothesen über den Zusammenhang von Evaluation und OL zu formulieren.

Immerhin bin ich in den letzten 15 Jahren ungefähr zehn Mal evaluiert worden, dreimal im Rahmen der Mitgliedschaft bei einer größeren Organisationseinheit (Fachbereich bzw. Institut) und sieben Mal als Mitglied oder Leiter einer selbständigen kleineren Forschungseinheit. Im Nachhinein kann ich mit Bestimmtheit sagen, dass allein letztere, also die „internen“ Evaluationen durch einen Beirat, deutliche Auswirkungen auf die eigene Praxis hatten. Dagegen blieben die als evaluatorische Großereignisse durchgeführten „externen“ Begutachtungen der größeren Organisationseinheit durchweg ohne Nachhall auf die Praxis der Forschungsgruppe, der ich angehörte.

Die ungleiche Effektivität der beiden Verfahren ist leicht zu erklären. Die als „intern“ klassifizierte Evaluation durch Beiräte profitierte von den evaluationsunabhängig existierenden und i.d.R. längerfristigen Beziehungen zwischen Evaluatoren und Evaluierten. Das ermöglichte eine relativ offene und hinsichtlich der eigenen Stärken und Schwächen wohlinformierte Kommunikation. Gleichzeitig waren die Gegenstände der Beurteilung, nämlich die Leistungen der „kleinen“ (fünf- bis 20-köpfigen) Forschungseinheit präzise beschreibbar, nämlich anhand von Objekten, Zielen und Instrumenten. Die Gutachtenden waren oft schon mit Publikationen der Gruppe vertraut. Jedenfalls schützte die Kontinuität der Beziehung stets vor Indifferenz und Ignoranz. Gleichwohl war zu keiner Zeit ein Abgleiten der Kommunikation in freundliche Unverbindlichkeit zu befürchten. Beide Seiten vermieden es sorgfältig, die professionellen Standards aufzuweichen oder gar zu ignorieren.

Die „großen“ Verfahren, in denen ein ganzer Fachbereich bzw. ein aus mehreren Forschungseinheiten bestehendes Institut evaluiert wurden, waren von anderer Natur. Im Vordergrund standen Forschungsprogramme, die in recht allgemeiner Form dargestellt und abgehandelt wurden. Außerdem zählten quantitative Indikatoren, d.h. Stellenausstattung, Publikationslisten und Drittmittelvolumina. Für die Erörterung der Qualität einzelner Forschungsleistungen fehlte es an Zeit und hinreichender Informiertheit der Evaluatoren. Die Ergebnisse der Begutachtung waren deshalb nur von geringer Relevanz für die in einzelnen Projektzusammenhängen Arbeitenden. Da Urteile und Empfehlungen ausschließlich der „großen“ Organisationseinheit galten, gab es keine unmittelbar „Betroffenen“. Jedes

Organisationsmitglied blieb frei, positive Urteile mit eigenen Leistungen und negative mit den Leistungen anderer in Verbindung zu bringen. Die wichtigsten „Lerneffekte“ der Evaluation betrafen die Art und Weise der Selbstdarstellung.

Soviel zu den Eigenerfahrungen als Evaluierter. Die folgenden Bemerkungen beziehen sich nicht direkt auf sie, aber sind zweifellos durch die Typik dieser Erfahrungen konditioniert. Wenn nun von Wissenschafts- bzw. Forschungsevaluation die Rede ist, sind damit in erster Linie extern beauftragte fachspezifische peer reviews gemeint, d.h. weder die anhand quantifizierter Daten aufgestellten Rankings noch die oben erwähnten und nach Dafürhalten des Autors recht effektiven („internen“) Beirats-Evaluationen.

### **Aspekte des Organisationslernens**

Bevor wir uns wieder dem Gegenstand ‚Evaluation von Forschungsorganisationen‘ zuwenden, seien ein paar Aspekte des Topos ‚Organisationslernen‘ in Erinnerung gerufen. So gilt es zu beachten, dass es in Sachen OL keineswegs ein konsistentes Theoriegebäude zu besichtigen gibt. Denn die einschlägigen Theorieansätze müssen sich mit dem Umstand herumschlagen, dass die existierenden Organisationen zwar eine enorme Strukturdiversität aufweisen, aber *per definitionem* auf zwei scheinbar widersprüchlichen Funktionsprinzipien festgelegt sind. Organisationen müssen – als auf Dauer gestellte, komplex strukturierte Interaktionsordnungen – erstens ein Set von Strukturmerkmalen und Sinnprämissen gegen alle unmittelbaren Umwelteinflüsse immunisieren, um sich in veränderlicher Umwelt als berechenbar und beständig zu behaupten, während sie sich, zweitens, keine selbstschädigende Umweltignoranz leisten dürfen, die ihren Zugriff auf externe Ressourcen (Kapital, Kunden, Subventionen und ganz allgemein: Funktionsreputation) gefährden könnte. Angesichts dieser zwei komplementären Grundfunktionen ist schon der ‚normale‘ Organisationsprozess eine Gratwanderung zwischen zu weit getriebener Routinisierung und dem Risiko des Identitätsverlustes bzw. der Selbstauflösung in die Umwelt.

Es liegt auf der Hand und wird vom Augenschein bestätigt, dass diese Gratwanderung umso eher gelingt, je leichter es der Organisation fällt, ihre operativen Einheiten den sich wandelnden Umweltbedingungen anzupassen und dabei einen Kern von Identitätsmerkmalen konstant zu halten. Um ihren Fortbestand zu sichern, müssen sich Organisationen gleichzeitig wandeln und treu bleiben. (Das gelingt bekanntlich nicht allen Exemplaren so gut wie z.B. der römisch-katholischen Kirche.) Organisationswandel ist also in der Welt der realen Organisationen endemisch. In Anbetracht der Ubiquität von Organisationswandel ist es nicht einfach, einen distinkten Begriff des Organisationslernens zu begründen.<sup>1</sup> Müsste dieser doch wegen der Referenzebene ‚Organisation‘ weiter gefasst sein als ein Lernbegriff, der lediglich für solche Fälle gilt, in denen sich Individuen komplexeres Wissen aneignen und zur Anwendung bringen. Hier wird der Einfachheit halber

---

<sup>1</sup> Es verwundert auch nicht, wenn die Mehrzahl der Beiträge zu dem derzeit umfassendsten Sammelwerk zum Thema (Dierkes et al. 2001) nahe legt, die Begriffe Organisationswandel und Organisationslernen als Synonyme und alle Differenzierungsanstrengungen als nutzlos zu betrachten. Eine Ausnahme bilden Gherardi und Niccolini (2001), die allein fünf soziologische Konzepte (i.S. theoretischer Zugänge) des Organisationslernens unterscheiden.

für den mikro-interaktionistischen Zugang zum Thema (i.S. von Gherardi und Niccolini 2001), also eine handlungsanalytische Perspektive optiert. Sie unterstellt Akteure als Autoren von Entscheidungen und erlaubt es, alle möglichen Phänomene als Zuschreibungsobjekte zu behandeln. In dieser Perspektive lässt sich immerhin sinnvoll von Lernenden (und Lehrenden), Lernanreizen und Lernanlässen, ja sogar Lernzielen und Lernergebnissen sprechen.

### **Evaluation als Lernanlass**

Auf den ersten Blick scheint die Begutachtung von Organisationsleistungen durch peers als kundige Nichtmitglieder einen unproblematischen Fall von OL darzustellen. Die zugrunde liegenden Wertmaßstäbe und angewendeten Kriterien sind vorgeblich dieselben, die auch innerhalb der Organisation Geltung beanspruchen. Und sie sind außerdem recht stabil und zählen folglich nicht zu den dynamischen, d.h. immer wieder aufs Neue zu dechiffrierenden Umwelтанforderungen. Das macht die von peer reviews repräsentierten Anforderungen prinzipiell antizipierbar: Ihnen sollte nichts Unerwartetes und Überraschendes anhaften. Die Organisation könnte sich auf sie problemlos einstellen – im Rahmen des ohnehin mehr oder weniger dynamischen Organisationswandels.

Auf einen zweiten und dann mehr empirisch als normativ informierten Blick erkennen wir zwei Besonderheiten, die das Bild kongruenter Erwartungen diesseits und jenseits der Organisationsgrenze trüben. So stellt sich bei näherer Betrachtung heraus, dass es mit der Uniformität von Wertmaßstäben und Qualitätskriterien nicht weit her ist. Wissenschaft ist eine Unternehmung, an die eine Vielzahl von Erwartungen und Wertbezügen herangetragen werden. Die Differenziertheit der wissenschaftlichen Gegenstände und Praxen widersetzt sich der Idee ihrer Beurteilung anhand einer überschaubaren Wertehierarchie. Die stärker axiomatisierten und anhand des Forschungsstands besser identifizierbaren Naturwissenschaften sind dabei nur wenig besser dran als die Sozial- und Geisteswissenschaften.<sup>2</sup> Weil die These gleicher Werte und Kriterien in der Praxis ebenso gut falsifiziert wie bestätigt werden mag, ist in der Beziehung von Evaluierten und Evaluatoren systematisch Unsicherheit zu vermuten. Würde es diese nicht geben, könnten die Evaluatoren auch schwerlich ihr Handeln für relevant halten. Also ist den Evaluierten eine gewisse Unsicherheit über die Kriterien zuzugestehen, nach denen ihr Tun tatsächlich beurteilt wird.

Dieser unvermeidliche Rest an Unsicherheit akzentuiert eine zweite Besonderheit, die es nahe legt, Evaluation als einen Fall von OL zu betrachten: *die Differenz von Auftraggeber und Objekt* der Evaluation. M.a.W., es sind in aller Regel nicht die Evaluierten selbst, welche eine Evaluation in Auftrag geben. Vielmehr wird der Auftrag i.d.R. von der Organisationsleitung, einem Finanzierungsträger bzw. einem Träger der fachliche oder politischen Verantwortung (für die Organisation) erteilt. Durch diesen Modus der Beauftragung werden die allemal schon bestehenden *Differenzen von*

---

<sup>2</sup> Das belegt u.a. die kontroverse Diskussion über die Finalisierung von scheinbar erfolgreicher Disziplinen (Böhme et al. 1973).

*Informationsverfügung und Entscheidungsrechten* in der Organisation aktualisiert und – so platt es auch klingen mag – Organisationsmitglieder in die ihnen unangenehme Rolle von Entscheidungs*nehmern* gezwängt. Das ist insofern relevant, als es für alle Beteiligten einen starken Anreiz impliziert, ihre Aktionen im Verfahren dezidiert *strategisch* zu kalkulieren.

Strategisches Handeln der Evaluierten äußert sich in angestrengtem „window dressing“ bei gleichzeitig demonstrierter Auskunftsbereitschaft und Kooperationsbereitschaft, in Bemühungen, Kohärenz vorzuspiegeln, wo in Wirklichkeit unverbundene Diversität herrscht, in schier endlosen Listen vermeintlicher Kooperationspartner oder in der Verwischung des Unterschieds zwischen wissenschaftlichen und politischen Publikationen bzw. Forschungsaufenthalten und Urlaubsreisen.

Der Anreiz zu strategischen Ausweichaktionen dieser Art ist besonders stark, wenn mit ungünstigen Ergebnissen der Evaluation gerechnet wird. Die Evaluation besteht – in Lichte dieser Erwartung – überwiegend in Akten der Normkontrolle: nicht in dem Sinne, dass nach der Angemessenheit der empirischen Normenorientierung gefragt würde, sondern in der schlichten Überprüfung, ob als fraglos gültig unterstellte Normen auch ausreichend Beachtung finden. Wo sich Evaluatoren auf ein solches Abprüfen von Normgeltung beschränken, haben wir es folglich mit dem Modus des *einfachen* Organisationslernens zu tun, in dem das Lernsubjekt eine Bekräftigung der geltenden Regeln und Routinen erfährt.

Dieser Modus des „single-loop learning“ (Argyris 1976) erscheint in einer organisationssoziologischen Perspektive auf die Voraussetzungen flexibler Anpassung an kontingente Umwelten als hochgradig riskant. Der Fokus auf Regelmäßigkeit verleitet die Organisation dazu, *pathologisch* zu lernen und alle Aufmerksamkeit auf die Optimierung vermeintlicher Performanzindikatoren zu richten. Genau daraus resultieren die vielen aufgeblähten Publikationsverzeichnisse, in denen die naturgemäß begrenzte Zahl bemerkenswerter Fachaufsätze in einem Wust von Schriften populärwissenschaftlichen, praxisberatenden und meinungsbekundenden Inhalts untergeht oder die Vielsprachigkeit der Veröffentlichungen als Beleg für internationale Reputation ausgegeben wird.

Aus systemtheoretischer Sicht ist ferner zu berücksichtigen, dass die im Rahmen der Evaluation ergehenden Urteile und Empfehlungen mit hoher Wahrscheinlichkeit am „Sinnfilter“ des sozialen Systems „Organisation“ auflaufen oder durch diesen erheblich modifiziert werden. Innerhalb der Organisation können sie nur insoweit sinnkonforme Wirkungen entfalten, wie sie dort auf „Resonanz“ stoßen, d.h. Anschlussstellen in der internen Kommunikation finden (Luhmann 1986). Was die Organisationsmitglieder nicht schon selbst für sinnvoll halten, kann ihnen nur schwer organisationsextern vermittelt werden. Gelingender Sinntransfer, wie ihn die Abgabe von Empfehlungen voraussetzt, beruht allemal auf der Anschließbarkeit von vorhandenen Einschätzungen, Interessen und Planungen bzw. der Eignung für innerorganisatorische Wettbewerbs- und Konfliktstrategien. Jedenfalls reagiert die Organisation – als interdependentes Kommunikations- und Entscheidungssystem – autonom auf die an sie adressierten Signale. Dass Organisationsführungen und -mitglieder regelmäßig gute Gründe haben, die Evaluationsurteile nicht zu ignorieren, versteht sich von selbst. Was sie damit anfangen, ist jedoch eine andere Sache. Ob auf die eventuell resultierende Revision eingefahrener

Routinen – der Forschungsplanung, des Publizierens, der Kooperation usw. – auch eine Steigerung der Forschungsqualität folgt, ist eine ganz andere Frage.

Welche Möglichkeiten bestehen, die Chance für „evaluationsadäquate Resonanz“ zu erhöhen? Das ließe sich womöglich dadurch erreichen, dass sich die Evaluatoren – natürlich vor dem Hintergrund des üblichen Gerüsts quantitativer Daten – über die von den Evaluierten verwendeten Ziele und Kriterien informieren ließen. Dazu gehört die Kenntnis der tradeoffs zwischen den diversen (in der Literatur finden sich zehn und mehr<sup>3</sup>) Bewertungsstandards, die von den Forschenden in Kauf genommen oder absichtlich vermieden wurden und die folglich ihrer Arbeit ein individuelles Gepräge geben. Dazu gehört ferner die Zurkenntnisgabe und Kenntnisaufnahme etwaiger besonderer Probleme, welche die Erfüllung des einen oder anderen Standardkriteriums bereitet.

Auf diese Weise könnten die Evaluatoren den Evaluierten die logische Selbstverständlichkeit konzederen, dass sich der Gesamtheit aller in Frage kommenden Bewertungsgesichtspunkte nicht einmal unter den denkbar glücklichsten Bedingungen Rechnung tragen lässt. Denn einem komplexen Kriterienset lässt sich nicht auf mechanisch maximierende Weise gerecht werden, sondern allenfalls durch „satisficing behavior“ im Sinne von Herbert A. Simon. Die Besonderheit dieser einzig praktikablen Weise des Umgangs mit Unsicherheit liegt in der Rückverlagerung des Rationalitätskriteriums aus der für Maximierungsstrategien unzugänglichen Umwelt in die „Innenwelt“ des Akteurs. Nur er bzw. sie kann entscheiden, was im konkreten Fall Geltungsvorrang beanspruchen soll und was nicht.

### **Innovation – ein blinder Fleck im Evaluationsverfahren?**

In Theorien des Organisationslernens wird neben der Bekräftigung von Regelkonformität ein weiterer, und zwar anspruchsvollerer Modus thematisiert, der jedoch im Hinblick auf die Praxis der Wissenschaftsevaluation nur von geringerer Relevanz zu sein scheint. Gemeint sind komplexe (sog. *double-loop*) Lernprozesse, in denen die Organisation ihre Entscheidungsprämissen und -prozeduren modifiziert und dabei u.U. auch *tiefergreifende* Änderungen ihrer maßgeblichen Selbst- und Umweltdeutungen vornimmt. Das ist die organisationssoziologisch interessantere Variante des OL, da sie sich deutlich abhebt vom alltäglichen und insofern trivialen Vorgang des umweltinduzierten Organisationswandels. In der mikro-interaktionistischen Perspektive ist „komplexes“ OL demnach als intern *intendierter und gesteuerter* Wandel struktureller und kognitiver Organisationsvariablen zu charakterisieren – ein Prozess, der keiner äußeren Veranlassung bedarf, sondern den Resultaten interner Selbstbeobachtung und –evaluation entspringt. Die Voraussetzungen und Prozesscharakteristika eines derart pro-aktiv inszenierten Gestaltwandels sind folglich andere als die Bedingungen der fortlaufenden Sicherung eines bestandsförderlichen

---

<sup>3</sup> Beispielsweise listet Andreas Stucke (1998) 13 Kriterien für wissenschaftliche Qualität auf: Einbindung in Forschungszusammenhänge, Kohärenz des Forschungsprogramms, qualifizierte Publikationstätigkeit, angeworbene Drittmittel, regelmäßige Beiratsevaluation, Qualifikation des Personals, Kooperationsbezüge mit anderen Forschungseinrichtungen, organisationsübergreifende Berufungsverfahren, Beteiligung der Forschenden an der Lehre, Nachwuchsförderung, eingehende Konferenzeinladungen, Forschungsaufenthalte im Ausland und eingeladene Gastwissenschaftler.

Korrespondenzverhältnisses zwischen Organisation und Umwelt, also dessen, was besser dem Begriff des Organisationswandels zu subsumieren wäre und sich überwiegend in Tests auf Regelkonformität manifestiert.

Man würde wohl fehlgehen, Belege für das entsprechende Selbstgestaltungsvermögen ausgerechnet bei bürokratisch verfassten Organisationen zu suchen, um die es sich auch bei den Trägerorganisationen von Wissenschaft handelt (Universitäten, Forschungsgesellschaften, Instituten usw.). Für die wissenschaftlich erfolgreicherer Bürokrationen scheint vielmehr das Muster eines Strukturdualismus typisch: auf der einen Seite die forschenden Organisationsmitglieder mit Freiraum für die Gestaltung ihres Arbeitsprogramms bzw. flexible, projektorientierte Arbeitszusammenhänge zwischen Organisationsmitgliedern; auf der anderen Seite eine fixe Struktur von Dienstleistungsfunktionen, die durch ihre bürokratische Verfasstheit den Forschenden eine verlässliche Infrastruktur und Abschirmung vor unwillkommenen Außenereignissen bietet. „Komplexes“ OL beschränkt sich hier auf die Tätigkeit der forschenden Wissenschaftler. Es manifestiert sich in kognitiver Kreativität und einer flexibel handhabbaren Kopplung mit dem Wissensbestand. Es ist genau besehen nicht von den Erscheinungsformen individuellen (aber allemal interaktiven) Lernens zu unterscheiden. Letztere scheint für die Qualität wissenschaftlicher Arbeit häufig entscheidend zu sein, was sich einerseits an der oft geringen Organisationsloyalität der Forscher und andererseits an der transorganisatorischen Struktur der Netzwerke für Forschungsk Kooperation und Forscherreputation ablesen lässt.

Die Evaluation von Wissenschaftsorganisationen müsste also zweigleisig verfahren: Regelkonformität und zweckadäquate Leistungsindikatoren mögen die Qualität des organisatorischen Rahmens ‚messen‘, Produktivitäts- und Kreativitätsindikatoren das Forschungspotential. Tatsächlich tragen die verbreiteten Kriterienkataloge den beiden Bewertungsebenen sehr ungleichgewichtig Rechnung. Von den „13 Kriterien für wissenschaftliche Qualität“, die der Wissenschaftsrat einst für peer reviews empfahl (Stucke 1998), scheinen lediglich zwei (das Kohärenzkriterium und „qualifizierte Publikationstätigkeit“) geeignet, Aufschluss über die Qualität der wissenschaftlichen Arbeit zu geben; drei weitere (Einbindung in externe Forschungszusammenhänge, Drittmittelvolumen und Beiratsevaluationen) rekurrieren auf evaluative Beiträge Dritter und können als indirekte, aber weniger valide Qualitätsindikatoren gelten. Die restlichen acht Kriterien betreffen Leistungen der *Wissenschaftsorganisation* i.e.S. und bezeichnen bestenfalls nützliche Voraussetzungen von Wissenschaftsqualität, aber sind keine Qualitätsindikatoren. Dabei scheint es noch an Indikatoren für die wichtigste Infrastrukturleistung zu mangeln: die Herstellung einer innovationsförderlichen Gelegenheitsstruktur einschließlich adäquater Leistungsanreize.

Was vor allem anderen bedenklich stimmt, ist die völlige Abwesenheit eines Indikators für Originalität, Kreativität und Innovativität. Auch und gerade in den oben erwähnten zwei direkt qualitätsbezogenen Kriterien wird letztenendes auf Konformität mit dem Status quo Bezug genommen. Wo Forschungen einen dezidiert innovativen Charakter haben und bewusst gegen Wissenstraditionen verstoßen, hängt es allein von der Aufgeschlossenheit der Redakteure und Gutachter der Publikationsorgane ab, ob innovative Leistungen eine Chance

erhalten, in den auf sie abhebenden Bewertungsroutinen als positiv wahrgenommen zu werden. Häufiger werden die Innovatoren als Regelverletzer identifiziert. Diesen Sachverhalt monierte Friedhelm Neidhardt, als er feststellte, dass „der Selbststeuerungszirkel der DFG für 'kleine Fortschritte' auch im ungünstigsten Fall durchaus tauglich“ sei, aber kaum für „Wissenschaftsrevolutionen, also 'große Durchbrüche' und ‚schöpferische Zerstörung‘ der herrschenden Standards“ (Neidhardt 1988: 136). Der großen Bedeutung, welche in Evaluationsverfahren dem Drittmittelaufkommen beigelegt wird, entspricht folglich ein Tunnelblick auf das Spektrum der qualitätsverdächtigen Arbeiten. Das bedeutet letztenendes „eine erhebliche Einschränkung der Zahl möglicher Innovationen und eine Prämierung von standardisierter Normalwissenschaft“ (Münch 2006 mit weiteren Literaturverweisen).

### **Evaluation als Impuls zur Pfadkorrektur**

Will man die bescheidenen konzeptionellen Ansätze der Forschungen zum Organisationslernen für die Organisation von Qualitätskontrollen der Wissenschaft fruchtbar machen, so bietet sich v.a. die beschriebene Differenz zwischen „einfachem“ und „komplexem“ OL an. Sie korrespondiert idealtypisch der Differenz von Forschungsarbeit einerseits und Verwaltungs- sowie Dienstleistungsarbeit andererseits. Wenn letztere regelkonform geleistet werden, mag erstere „aus sich selbst heraus“, d.h. ohne gezielte externe Anreize, Lerneffekte im Sinne fortschreitender Erkenntnis produzieren. Deshalb müssten die entsprechenden Organisationsbereiche mit einem je besonderen Katalog von Leistungskriterien bedacht werden. Und es wäre durchaus diskutabel, Wissenschaftler im Sinne der Theorie des „komplexen“ OL zuvörderst als Initiator von Wandlungshandlungen – *change acts* – und folglich als *change agents* zu betrachten, statt v.a. auf quantitative Indikatoren ihrer Produktivität abzuheben.

Nun wäre es wohl etwas zu simpel, alle wissenschaftliche Tätigkeit mit dem Hervorbringen von Innovationen zu identifizieren, ist doch schon dann ein wertvoller Qualitätsgewinn des Systems Wissenschaft zu verzeichnen, wenn der jeweilige Forschungsstand allgemeine Kenntnis und Beachtung genießt. Evaluationen können dazu beitragen, diesem Zustand nahezukommen bzw. ihn – auf jeweils „höherem“ Niveau – zu stabilisieren. Die Konfrontation eingespielter Routinen des Denkens und der Orientierung mit den in einem größeren Horizont gewonnenen Wissensbeständen ist in diesem Sinne als Stimulus der Routinenrevision bzw. Entroutinisierung zu betrachten.

Evaluationen können, wenn ihre Instrumente zwischen kreativer und bürokratischer Praxis differenzieren, zum Abbruch unfruchtbarer Pfade wissenschaftlicher Arbeit beitragen und wichtige Anstöße für eine Neuausrichtung von Forschungsprogrammen – in sachlicher wie personeller Hinsicht – liefern. Vorausgesetzt ist, dass die Urteile und Empfehlungen auf anschlussfähige Einschätzungen und konforme Ambitionen auf Seiten der Evaluierten stoßen. Unter welchen Bedingungen, pfadabhängige Entwicklungen abbrechen bzw. ein Wechsel des Entwicklungspfades der Organisation wahrscheinlich wird, lässt sich vielleicht mit Blick auf eine Studie zum Konzept der Pfadabhängigkeit abschätzen (Beyer 2005).

Zu Pfadabbrüchen kommt es typischerweise, wenn eine von zwei Konstellationen vorliegt. Entweder haben sich die materiellen Bedingungen der Fortsetzung befestigter Handlungsweisen (d.h. Pfade) nachteilig verändert, z.B. infolge von drastisch gesunkenen Nutzenerwartungen, Funktionsverlusten gegenüber Dritten oder überschaubar gewordenen Kosten eines Pfadwechsels zu alternativen Programmen. Oder es haben sich Zweifel an den zuvor geteilten Überzeugungen und Orientierungen gebildet, z.B. aufgrund von Informationen über konkurrierende Forschungen, nach Legitimations- und Reputationseinbußen, aber auch aufgrund kontra-adaptiver Erwartungen bzw. einer Neigung zu nonkonformen Handeln. Es liegt auf der Hand, dass die Urteile und Empfehlungen externer Evaluatoren prinzipiell geeignete Anstöße zur Neubewertung von Nutzenerwägungen und Sinnzuschreibungen darstellen.

Ob Evaluation tatsächlich kognitive Innovationen auslöst, hängt allerdings weniger vom intrinsischen Wahrheitswert der Mitteilungen an die Evaluierten ab als von Reputationseinschätzungen und den in der Kommunikation evozierten Verständigungsbereitschaften. Es ist auch keineswegs so, dass sich neue und „bessere“ Informationen automatisch gegen "alten" Orientierungen durchsetzen. Vielmehr erlangen kognitive Innovationen regelmäßig erst dann einen privilegierten Status in organisierten Interaktionssystemen, wenn ihnen die Teilnehmer kognitiven *und* strategischen Wert beimessen. Wesentliche Momente sind zum einen Machtkommunikation und Metaentscheidungen, z.B. über Teilnahmerechte im Hinblick auf die Zusammensetzung von Entscheidungsgremien und "political coalitions" (i.S. von Cyert/March 1963), in denen auch über das Wirksamwerden neuen Wissens entschieden wird. Zum anderen kommt auch das Assoziationsbegehren von Individuen ins Spiel, die mit ihrem Wissen und Wollen nicht „allein dastehen“ mögen, sondern es vorziehen, Orientierungen mit anderen zu teilen – was die Bereitschaft zur Übernahme der Orientierungen relevanter Dritter einschließt sowie das damit verbundene Risiko, „victims of group think“ (Janis 1972) zu werden.

## **Schluss**

Evaluationsverfahren und insbesondere die Interaktion zwischen Evaluatoren und Evaluierten im Lichte der Forschungen zum Organisationslernen zu betrachten, scheint kein sinnloses Unternehmen. Evaluation als potentiellen Anlass und Stimulus des Organisationslernen zu analysieren, könnte vermutlich helfen, ungenutzte Kreativitätspotenziale ans Licht zu bringen. Das betrifft einerseits solche, die zwar regelmäßig als antizipative oder reaktive Resonanz auf Evaluationsverfahren auftreten, aber sich dank der kognitiven Autonomie und Kreativität individueller Wissenschaftler in beachtlicher Diversität manifestieren. Andererseits scheint es nicht aussichtslos, in einer stärker organisationssoziologisch ausgerichteten (und gruppendynamisch informierten) Perspektive die Bedingungen für einen gelingenden Sinn- und Orientierungstransfer zwischen Evaluatoren und Evaluierten weiter aufzuklären.



## Literatur

- Argyris, Chris, 1976: Single-Loop and Double-Loop Models in Research on Decision Making. *Administrative Science Quarterly* 21, 363-375.
- Beyer, Jürgen, 2005: Pfadabhängigkeit ist nicht gleich Pfadabhängigkeit! Wider den impliziten Konservatismus eines gängigen Konzepts. *Zeitschrift für Soziologie*, 34 (1), 5-21.
- Böhme, Gernot; van den Daele, Wolfgang; Krohn, Wolfgang, 1973: Die Finalisierung der Wissenschaft. *Zeitschrift für Soziologie*, 2 (2), 128-144.
- Cyert, Richard M./ March, James G., 1963: *A Behavioral Theory of the Firm*. Englewood Cliffs, N.J.: Prentice-Hall.
- Dierkes, Meinolf; Berthoin Antal, Ariane; Child, John; Nonaka, Ikujiro (eds.): *Handbook of Organizational Learning & Knowledge*. Oxford: Oxford University Press.
- Gherardi, Silvia; Nicolini, Davide, 2001: The Sociological Foundations of Organizational Learning. In: Dierkes, Meinolf; Berthoin Antal, Ariane; Child, John; Nonaka, Ikujiro (eds.): *Handbook of Organizational Learning & Knowledge*. Oxford: Oxford University Press, 35-60.
- Janis, Irving L., 1972: *Victims of Groupthink. A Psychological Study of Foreign Policy Decisions and Fiascoes*. Boston: Mifflin.
- Luhmann, Niklas, 1986: *Ökologische Kommunikation. Kann die moderne Gesellschaft sich auf ökologische Gefährdungen einstellen?* Opladen: Westdeutscher Verlag.
- Münch, Richard, 2006: *Drittmittel und Publikationen. Forschung zwischen Normalwissenschaft und Innovation*. Ms.
- Neidhardt, Friedhelm, 1988: *Selbststeuerung in der Forschungsförderung. Das Gutachterwesen der DFG*. Opladen: Westdeutscher Verlag.
- Stucke, Andreas, 1998: *Evaluation als Normalfall - die Bewertung von Einrichtungen der außeruniversitären Forschung durch den Wissenschaftsrat*. Impulsreferat auf der 1. Jahrestagung der Deutschen Gesellschaft für Evaluation, Köln 25.09.1998: [www.degeval.de/koeln1998/stucke.htm](http://www.degeval.de/koeln1998/stucke.htm).