

„Irgendwann müssen wir den Begriff [Evaluation] schützen.  
Und dann kann nicht mehr jeder  
Quacksalber als Evaluator hausieren gehen.“  
Lutz von Rosenstiel 2002

# Von der Messtechnik zur Mediation? Rollenerwartung und Selbstverständnis in der Medienevaluation

*Christine Schwarz*

*Universität Hannover/ University of Edinburgh*

## 1. Einleitung: Was ist Quacksalberei und was Evaluation?

Wer Stellenausschreibungen im Berufsfeld Evaluation überfliegt, gewinnt den Eindruck, was Evaluierende zu tun haben, sei ganz plausibel und selbsterklärend. Gewünscht ist eine formale Qualifikation: das Handwerkszeug des Messens und Bewertens. So wird mit dem Evaluieren häufig *irgendeine* Fähigkeit assoziiert, mehr oder weniger ‚wissenschaftlich‘ mit empirischen Daten und Datenverarbeitungsprogrammen umgehen zu können. Auch beansprucht das vorangestellte Motto von Lutz von Rosenstiel, dass Evaluieren eine Tätigkeit sei, dessen implizite Standards in fast ingenieurwissenschaftlicher Manier lediglich in explizite Formen transformiert werden müsse. Doch schenken wir Egon Gubas und Yvonna Lincolns These von der post-positivistischen Wende in der Evaluation Glauben (Guba/Lincoln 1989, 2001), so erschöpft sich das wissenschaftlich anerkannte Evaluieren nicht nur im Messen, Beschreiben und Urteilen nach einem Kanon kollektiv anerkannter Methoden, wie in den drei Generationen davor. Evaluation ist mehr als nur angewandte Sozialforschung, die Rolle der Evaluierenden verlagert sich immer stärker hin zur Vermittlung zwischen verschiedenen Akteurinnen: Von der Messtechnik zur Mediation. Wenn diese Rollenkomplexität tatsächlich besteht und gemäß veränderter Anforderungen auch neue Kompetenzen erwartet werden, was bedeutet das dann für die Aus- und Weiterbildung der Evaluierenden und für die betriebliche Personalpolitik von deren Auftraggeberinnen? Müssen Evaluierende Menschen mit ausgewiesener wissenschaftlicher Expertise sein? Oder sollen sie vor allem Projektmanagement-Erfahrung in dem jeweiligen Politikfeld haben? Wird jemand gesucht, der oder die „von unten“ kommt und beschlagen ist mit den praktischen Details, also über viel lokales Wissen über das zu evaluierende Projekt verfügt? Oder jemand, der oder die bereits „oben“ in der Karriereleiter ist, sich ganz selbstverständlich in den unvermeidlichen Institutionen des zu evaluierenden Politikfeldes bewegt, ein Renommee hat und mit globalem Wissen und Einschätzungsvermögen richtungsweisend sein kann? Müssen es verdiente Fachexpertinnen sein oder sog. Hybrid-Expertinnen, die an den Schnittstellen zwischen Themen und Disziplinen arbeiten?

Mittlerweile hat sich um die Deutsche Gesellschaft für Evaluation (DeGEval) eine Art Berufszunft unter Abnehmerinnen und Auftraggeberinnen von Evaluationen entwickelt, in der Anstrengungen übernommen werden, a) die aus dem Amerikanischen stammenden Qualitätsstandards für Evaluation nicht nur sprachlich, sondern auch kulturell zu übersetzen und anzupassen (DeGEval 2002: 13) sowie b) Qualifikationsprofile für Evaluierende zu erstellen, wie dies in der Diskussion um „Anforderungsprofile“ und „Qualitätsstandards“ im Arbeitskreis „Aus- und Weiterbildung in der Evaluation“ geschieht (Schmidt/Hennen 2001; Schmidt 2002; Hennen 2002). Auch wenn nach langen Forderungen in diesem Jahr in Deutschland und in der Schweiz die ersten Evaluations-Weiterbildungsstudiengänge starteten<sup>1</sup>, sind die Strukturierungs- und Verständigungsprozesse darüber, was Evaluierende können müssen, längst nicht abgeschlossen: „Der Weg zu einer vertieften Definition und einer Spezifikation der Aus- und Weiterbildung zum Evaluator (sic!) ist noch zu gehen, und es werden viele Verkrustungen in den bestehenden Anforderungsprofilen klassischer Berufsbilder aufzubrechen sein.“ (Hennen 2002: 196)

Der vorliegende Artikel zur Forschung über Evaluation in der Praxis hat sich ursprünglich aus dem DeGEval-Arbeitskreis Medienevaluation heraus entwickelt und möchte zur Verständigung über Ausbildungskomponenten und Qualitätsansprüche von Evaluation beitragen. Allerdings soll die bisher deduktiv von oben gestellte Frage „Was müssen Evaluierende können?“<sup>2</sup> umgekehrt werden, indem nach den bisherigen empirischen Erfahrungen mit der angeblich boomenden Evaluationstätigkeit gefragt wird: Mit welchen Rollen- und Kompetenzerwartungen werden Evaluierende in der konkreten Berufspraxis eigentlich konfrontiert? Und was ist ihr eigenes Selbstverständnis? Dass eine systematische Berufsfeldanalyse innerhalb der deutschsprachigen Evaluationsforschung noch fehlt, ist umso erstaunlicher, da eine solche Analyse eine notwendige Voraussetzung dafür ist, diese auf Praxisrelevanz und Auftragsgebundenheit verbrieft, neuere wissenschaftliche Aktivität als eigenes Berufsfeld abzustecken und ihre Leistungsfähigkeit erfolgreich nach außen zu kommunizieren – auch für Auftraggeberinnen. Im Folgenden soll deshalb anhand der exemplarischen Analyse des Evaluationsfeldes Medienevaluation, speziell der Implementierung von e-learning-Instrumenten in die Hochschullehre, nach dem Selbstverständnis von Evaluierenden und dem was sie tun gefragt werden sowie nach den an sie gestellten Rollenerwartungen und möglichen Rollenkonflikten. Dabei soll nicht von diesem spezifischen Evaluations-Politikfeld abstrahiert und auf andere geschlossen werden; aber möglicherweise ließen sich politikfeldbezogene Eigenheiten gerade im Vergleich genauer erkennen. Daher wird zunächst das methodische Vorgehen der Fallstudie beschrieben (1) sowie die rasante Entwicklung der Medienevaluation im Bereich öffentlich geförderter e-learning-Projekte skiz-

1 Das „Ergänzungsstudium Evaluation in Berufsbildung, Gesundheitswesen und Verwaltung“ (<http://www.cx.unibe.ch/kwb/wbprog/o8/p55.html>) sowie das „Fortbildungsprogramm in der Entwicklungszusammenarbeit“ (<http://www.feez.org/>).

2 Was Evaluierende eigentlich können müssen, war Thema der Abschlussdiskussion der DeGEval-Jahrestagung 2002. Von den geladenen ministeriellen Evaluations-Auftraggebern konnten aber lediglich allgemeinere, d.h. sämtliche Berufschancen verbessernden Qualifikationen genannt werden, wie Teamfähigkeit, Fremdsprachenkenntnisse und schriftliche Ausdrucksfähigkeit.

ziert (2). Nach der Beschreibung der komplexen Struktur von Virtuellen Universitätsprojekten (3) werden die Erfahrungen mit systematischer Nutzenbewertung aus der Sicht der Evaluierenden dargestellt (4) sowie deren Rollenverständnisse interpretiert (5). Dem Modell des Spannungsfeldes der Evaluation als eines der zentralen Ergebnisse der Studie wird die Fachliteratur zur Rollenkomplexität von Evaluation gegenübergestellt (6). Was vor diesem Hintergrund die These von Evaluation als Verhandlungsführerin für die Aus- und Weiterbildung von Evaluierenden bedeutet, ist Gegenstand der abschließenden Überlegungen (7): Als wie glaubwürdig erweist sich in der Evaluationspraxis das stets bei Evaluation mitschwingende Demokratisierungsversprechen durch wissensbasierte Politik? Und wo droht Evaluation möglicherweise zu einem modernen Ritual zu werden?

## 2. Methodisches Vorgehen der Fallstudie

In der hier vorzustellenden Fallstudie über die alltägliche Praxis des Evaluierens kam ein Mix von Erhebungs- und Analyseverfahren zur Anwendung: Dokumentenanalyse, Analyse von e-mail-Kommunikation, Leitfadeninterviews, Visualisierung<sup>3</sup> zur Arbeitssituation der Befragten sowie teilnehmende Beobachtung bei Projekttreffen. Das gewonnene Datenmaterial war so umfangreich, dass an dieser Stelle nur eine Teilauswertung vorgestellt werden kann, die sich schwerpunktmäßig bezieht auf die Analyse der leitfadengestützten Interviews mit Personen aus den sechs Akteurinnengruppen Kompetenzzentren (3), Ministerium (2), Beirat (4), Projektleitung bzw. -organisation (2), sozialwissenschaftliche Begleitung (10) sowie einem Hochschullehrer – im e-learning Jargon „Inhaltslieferanten“ genannt<sup>4</sup>.

Thema der Interviews, die im Zeitraum von Oktober 2001 bis Juni 2002 geführt wurden, war sowohl die subjektive Wahrnehmung des Modellprojekts als auch die Bedeutung, die Evaluation darin gespielt hat. Die Stichprobe der Befragung bezog sich auf alle der o.g. Akteursgruppen, wovon innerhalb des Projektes VUMS ca. 80% an der Befragung teilnahmen.<sup>5</sup> Die Einzel- und Gruppeninterviews wurden entlang eines identischen Interviewleitfadens geführt und waren zwischen 45 und 120 Minuten lang. Die Tonbandaufnahmen der Interviews wurden transkribiert; der Interviewtext wurde nach inhaltsanalytischem Verfahren (vgl. Mayring 1988) und unter Berücksichtigung der Visualisierungen ausgewertet.

Übergreifendes Ziel der Studie war es, die Rollen und Wirkungen von *Evaluation* zu untersuchen und weniger die konstatierten Wirkungen des Technikeinsatzes. Da virtuelle Hochschulprojekte in der ersten Pionierphase Länderprojekte waren, bestand anfangs erhebliche Konkurrenz zwischen den Projekten. Die Förderung der

3 Visualisierungen sind Bilder, die die Befragten im Anschluss an das Interview angefertigt haben zu der Frage „Wie sieht Ihr Bild von den Akteurskonstellationen im Virtuellen-Universitätsprojekt aus?“ Dieses Verfahren wurde im Laufe der Fallstudie entwickelt in Anlehnung an projektive Bilder in der Organisationsentwicklung und Dialogkonsens-Verfahren wie dem des Netzinterviews.

4 Zahl der befragten Personen in Klammern.

5 Eine Ausnahme bilden die sog. „Inhaltsanbieter“ von denen nur einer befragt wurde.

Studie von unabhängiger Seite<sup>6</sup> sowie die Anonymisierung waren deshalb eine wichtige, wenn nicht notwendige Voraussetzung für die Kooperation der Befragten.

### 3. Boomt die Medienevaluation oder expandiert nur das Label *Evaluation*?

Regelmäßig wurde auf die offenkundig schubweisen Evaluations-Booms im deutschsprachigen Raum hingewiesen (Rossi 1986; Beywl 1988, 1998; Stockmann 2000). Auch der „Quacksalber-Vorwurf“ wurde vor dem Hintergrund der Annahme steigender Nachfrage nach Evaluationen und einer damit verbundenen „Marktlücke“ an qualifizierten Evaluierenden formuliert. Aber so plausibel die Annahme des boomenden Berufsfelds Evaluation auch sein mag, es kann auch skeptisch gefragt werden, ob es sich hier nicht vielmehr um eine Tätigkeit handelt, die bislang nur anders und verschieden benannt wurde: Begleitforschung, angewandte Forschung, Begutachtung, didaktische oder technische Beratung. Möglicherweise wird jetzt, um die noch stärkere Anwendungs- und Nutzenorientierung deutlich zu machen, „Evaluation“ zu einem scheinbar alles sagenden Label und bläht sich zu einem Container-Begriff für sämtliche wissenschaftliche Tätigkeiten auf, die etwas mit Nutzenbewertung zu tun haben? Eben darin besteht die Gefahr: Dass der Terminus Evaluation zu einer Worthülse wird und damit selbst *wertlos*.

In der Tat setzte eine erhöhte Nachfrage an Evaluierenden im Medienbereich zum Jahrtausendwechsel ein. Damals wurden durch die Förderung von e-learning-Modellprojekten erstmals seit der Bildungsexpansion der 70er Jahre nennenswerte öffentliche Gelder für Hochschullehre zur Verfügung gestellt und der Aspekt der Nutzenbewertung extrem betont. So wurde meist von den Wissenschaftsministerien formuliert, dass diese Mittel *nur* in „Entwicklungs- und Umsetzungs- statt in Forschungsprojekte“ fließen sollten: „Wir sind hier nicht zum Forschen da. In der Virtuellen Universität geht es um harte Arbeit.“ oder „Ein schlecht erklärter Erfolg ist besser als ein gut erklärter Misserfolg.“ Die e-learning-Entwicklung an den deutschen Hochschulen war somit ab Mitte der 90er Jahre streng an Evaluation gebunden; für fast jeden Antragstellenden war die Entwicklung eines eigenen Evaluationskonzepts meist ebenso obligatorisch wie die Kooperation mit einer externen Begutachtung als Förderungsvoraussetzung.

Ein Höhepunkt der e-learning-Evaluation wurde in der Tat ca. Ende 2000 erreicht, als im Rahmen des Bundesförderprogramms „Neue Medien in Bildung“ (NMB) die Anzahl neu geschaffener Evaluationsstellen im Medienbereich noch einmal sprunghaft anstieg. Mit rund 130 Mio. Euro förderte allein dieses eine Programm – eines unter vielen anderen – rund 100 bundesweit verteilte Verbundprojekte zum Einsatz von Multimedia an den Hochschulen, an denen ca. 500 Universitäten beteiligt sind. Es entstanden schätzungsweise 250 Projektevaluations-Stellen in Teilzeit, die meistens zugleich mit didaktischer Beratung und teilweise auch mit Grundlagenforschung beauftragt waren; eine schlecht besoldete Mammut-Aufgabe,

6 Die Studie wurde gefördert von der Grünen Akademie der Heinrich Böll Stiftung.

zudem viele sich dafür nur zur Verfügung stellten, weil sie damit noch zusätzlich eine formale Weiterqualifikation verknüpfen wollten. Häufig fanden dort Berufsanfängerinnen aus den unterschiedlichsten wissenschaftlichen Disziplinen eine Erwerbsarbeit. Meistens waren dies wissenschaftliche Mitarbeiterinnen aus den einschlägigen evaluationsnahen Disziplinen wie Pädagogik, Psychologie und Soziologie. Aber es fanden auch Absolventinnen eine vorübergehende Anstellung aus den Fächern, die auf dem scheinbar immer verwertungsorientierter werdenden akademischen Arbeitsmarkt als problematisch gelten, wie z.B. Theologie, Philosophie oder Physik. So gab es bezüglich der Gestaltung von e-learning-Szenarien und erst recht bezogen auf deren Evaluation in den ersten Jahren der Medienevaluation zunächst ein babylonisches Sprachgewirr, was den interdisziplinärem Austausch betraf, und damit in etwa ein Szenario, das Michael Gibbons, Helga Nowotny und andere als Wissenschaft im „mode 2“ bezeichnen, also eine weniger hierarchische und vielmehr transdisziplinäre Wissenschaft (Gibbons u.a. 1994; Nowotny 1999). Mit diesen eher informellen Netzwerken zwischen den Disziplinen wurden die Probleme des bis heute mangelnden Qualifikationsangebots im Bereich Medienevaluation zunächst teilweise kompensiert. Jenseits von Fachverbänden und Zertifizierungsinstitutionen entwickelte sich im akademischen Mittelbau ein interessierter Erfahrungsaustausch zwischen unterschiedlichen wissenschaftlichen Zugangsweisen, was häufig auch die Möglichkeit bot, sich von den impliziten Werturteilen und unreflektierten Selbstverständlichkeiten der eigenen Fachkultur emanzipieren zu können.

Die These von der allgemeinen Unterversorgung mit qualifizierten Evaluierenden angesichts des Booms an Evaluationsnachfragen ist also durchaus anzuzweifeln und das Gegenteil kann behauptet werden: eine kritische Maße höchst qualifizierter arbeitssuchender Akademikerinnen findet in der Evaluationstätigkeit Unterschlupf und bildet sich autodidaktisch berufsbegleitend und auf ihre (Zeit-)Kosten fort. Angesichts der katastrophalen Stellenmisere an den deutschen Hochschulen strömt diese Lawine von qualifizierten Wissenschaftlerinnen in die vielen Evaluationsfelder, weil es lediglich hier noch Verwendung für traditionell (sozial-)wissenschaftlich Ausgebildete zu geben scheint. Für die Planung von Aus- und Weiterbildungsangeboten im Bereich Evaluation würde das bedeuten, dass sie sich auch mit den vorhandenen Qualifikationen der nicht Aus-, sondern Weiterzubildenden und deren Berufserfahrungen auseinander setzen müsste. Dabei kann es vorkommen, dass genau das Gegenteil von dem gelernt werden muss, was man in Studium oder Beruf bereits gelernt hat: vereinfachen statt differenzieren, verallgemeinern statt konkretisieren, Urteile fällen statt relativieren. Wissenschaft praktizieren im Sinne des verantwortlichen und raschen politischen Entscheidens ist etwas, das man in den Universitäten selten lernt. Auch anwendungsorientierte wissenschaftliche Disziplinen beschäftigen sich weniger mit den Voraussetzungen der technologischen Umsetzung ihrer Erkenntnisse in einem Politikfeld. Ein großer Gewinn für Evaluationscurricula wäre es deshalb, auch auf dieses Erfahrungswissen (Negt 1969) zwischen den Wissenschaftssphären zurückzugreifen – ob dieser Transformationsprozess nun im Sinne von Gibbons, Nowotny et al. begriffen wird oder auch anders. Wie etwas von dieser Gestaltveränderung von Wissenschaft auch sichtbar wird in der Aufgabenreformulierung im Zuge der Hochschul-Virtualisierung, soll im Folgenden erörtert werden.

#### 4. Virtualisierung und die Elastizität von Projektzielen

Das Ziel der ersten e-learning-Pionierprojekte in der Förderungsphase von ca. 1997 bis 2002 war es, mittels der IuK-Technologien die Universitätslehre in irgendeiner Weise zu verbessern. Doch wie häufig in der Geschichte soziotechnischer Reformen zu beobachten (Ortmann 1995), entwickelte sich dieses eher explorativ angelegte Unternehmen zu einem Magneten für übergreifende organisatorische und gesellschaftliche Zielsetzungen: Qualitätsverbesserung, Kundenfreundlichkeit bzw. Studierendenzentriertheit, Stärkung der Eigenverantwortung der Universitäten, Organisations- und Personalentwicklung, Praxisnähe des Studiums, Entmystifizierung des Professorenstatus, Effizienzsteigerung, Internationalisierung und nicht zuletzt Förderung der Hochschulkooperation (BLK 1998, 1999, 2000a, b, c; BMBF 1999, 2000a, b, BMBF 2001). Hinzu kommen eher inoffizielle Ziele der Beteiligten, die aber laut Angaben der Befragten in den Interviews mindestens ebenso handlungsleitend waren: die eigene Fachrichtung, Fachansicht oder Technik voranzubringen, wenigstens vorübergehend einige Stellen für Nachwuchswissenschaftlerinnen über Drittmittelprojekte schaffen, sich die Arbeit für die Lehre erleichtern oder ganz einfach den hohen Erwartungen an die Technik seitens der Politik mit empirischen Fakten zu begegnen.

Diese an Beliebigkeit grenzende Vielfältigkeit der Projektziele (Polytelie) ist zweifelsohne ein manifestes Problem der Evaluation, doch zeigt die Sozialgeschichte der Technik auch, dass es zeitweise sogar einer erheblichen Elastizität von Zielen bedarf, um verschiedenste Akteurinnen für die Mitarbeit an einem solchen Projekt zu mobilisieren (Callon 1986; Law 1994, 1999). Was also zeitweilig ein guter Kitt für das Projektmanagement sein mag, kann ein Sisyphos-Alptraum für diejenigen werden, die mit der Projektevaluation betraut sind. Mit der Schaffung sogenannter Virtueller Universitäten sollte die traditionelle Universität ins Netz kopiert und zugleich radikal verändert werden, da sie aus vielfältigen Gründen seit Jahren unter massivem Legitimationsdruck steht. Über die grundsätzlichen Barrieren der Medienevaluation (Baumgartner 1999; Schulmeister 1997; Schwarz 2001b) hinaus stellt diese komplexe Projektstruktur und Zieldefinition eine besondere Herausforderung für die Evaluation dar: Es sollte sowohl Altes evaluiert werden – die alte Struktur der Universitätslehre in all ihren sozialen Facetten – als auch Neues: eine mediengestützte standortübergreifende Lehrform vor einem ganz besonderen historischen Erwartungs-Hintergrund mit ganz vielfältigen und mit deren Fluktuation wechselnden Stakeholder-Interessen. Das zu evaluierende Projektziel ist somit ein *moving target*. Die Konstruktion der Virtuellen Universität war von Anfang an ein durchaus widersprüchliches Unterfangen, an dem sich sowohl reformfreundige als auch entschieden konservative Akteurinnen beteiligten. Aus heutiger Sicht sind diese enorm komplexen soziotechnischen Erwartungen besser nachzuvollziehen, wenn man den damaligen gesellschaftshistorischen Hintergrund mitbetrachtet: den atemberaubenden Boom der ‚new economy‘ und der dotcoms sowie den ungewöhnlichen Geldregen in den Staatskassen durch die Versteigerung der UMTS-Lizenzen, der einen Großteil der staatlichen e-learning-Förderung erst möglich machte. In ganz ähnlicher Weise wie die damaligen IT-start-ups standen auch diese Prestigeprojekte der Hochschulpolitik der Länder von Anfang an in harter Konkurrenz zueinander.

Viele wertvolle Erfahrungen wurden mit e-learning und Virtuellen Universitätsprojekten in den letzten ca. 10 Jahren gemacht, und rasant hat sich die Forschungsliteratur dazu entwickelt (Martens/Clement 1999; Pollock/Cornford 2003; Robin/Webster 2002; Ryan et al. 2000; Schwarz 2001a; Schulmeister 2001, 2003). Zweifellos ist aber eine der wichtigsten Erfahrungen, dass die überzogenen Erwartungen an die Technik einerseits den technischen Pioniergeist enorm befördert haben, andererseits aber im Bann der Technikutopien soziale und politische Barrieren unterschätzt wurden, was dazu führte, dass der soziale Reformgeist nach einer Anfangseuphorie auf der Strecke blieb: „Change is slow and not radical“ und „ICT in learning and teaching: widespread but a part of a blend“ schlussfolgert aktuell eine umfangreiche internationale Studie zum Einsatz von e-learning in der Hochschullehre (Collis/van der Wende 2003), eine alte Weisheit der Organisationssoziologie.

## 5. Die Fallstudie Virtual University Mixed Studies (VUMS)

Die Virtual University Mixed Studies<sup>7</sup> als eines der ersten staatlich geförderten Modellprojekte zur Etablierung von Virtuellen Universitäten<sup>8</sup> war eine der Fallstudien zum Berufsbild von Evaluierenden. VUMS wurde mit knapp 5 Millionen Euro vom Landesministerium gefördert und hatte anfangs zusätzlich ca. acht privatwirtschaftliche Partner aus der Telekommunikations- und Verlagsbranche. Das als Hochschul-Lehrverbund angelegte Projekt VUMS war über vier Universitätsstandorte verteilt und sollte ein Netzwerk bilden, in dem über Videokonferenz-Systeme klassische Vorlesungen als sogenannte Tele-Vorlesungen übertragen wurden. Darüber hinaus wurde angestrebt, mit Vorlesungs-CDs sowie sog. Studienmodulen ein Angebot der Fernlehre zu entwickeln. Das Ziel der gemeinsamen Virtuellen Universität sollte erreicht werden über die Strategie, dass über die vier Universitätsstandorte und insgesamt ca. 25 akademischen Disziplinen hinweg e-learning-Instrumente gemeinsam produziert, genutzt, evaluiert und optimiert werden sollten. Es wurden also für Hochschulverhältnisse extrem hohe Anforderung an Arbeitsteilung und Kooperation gestellt. Nicht nur verschiedene Standorte, sondern auch verschiedene Fach- und Hierarchieebenen sowie gesellschaftliche Institutionen sollten zusammenwirken wie ein ausgereiftes Datennetzwerk.

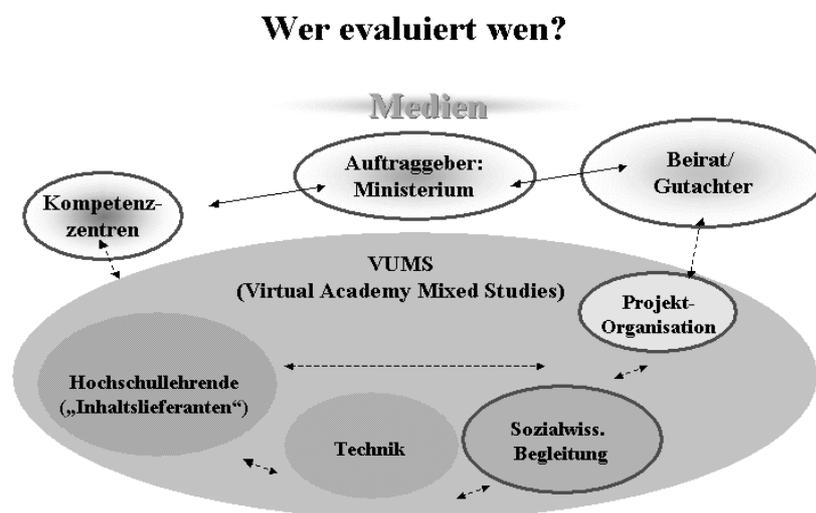
Angesichts der komplexen Projektstruktur und -ziele von VUMS sowie den ungewöhnlich hohen Landesaussgaben ist es nicht erstaunlich, dass von verschiedensten Seiten auch an deren Evaluation hohe Erwartungen geknüpft wurden. Wie bereits erwähnt, wurde die Bereitschaft zur systematischen Nutzenbewertung späte-

7 Der Name des Modellprojekts sowie Personen- und Ortsnamen wurde anonymisiert.

8 Unter der einstigen Vielzahl Virtueller Universitätsprojekte lassen sich verschiedene Konstruktionstypen wie folgt unterscheiden: Die virtualisierte klassische Fernuniversität (a) und die virtualisierte klassische Universität (b), die meist im "dual-mode"-Betrieb konzipiert ist, d.h. sowohl traditionelle Lehre als auch virtualisierte, d.h. netzgestützte oder mit elektronischen Medien). Auch gibt es neugegründete virtualisierte Universitäten (c) sowie „corporate universities“ (d), d.h. firmeneigene Weiterbildungsdienstleistungen (vgl. Schulmeister 2001). Die Fallstudie VUMS ist eher eine Mischung aus dem ersten Typ und der virtualisierten klassischen Universität, die aber im Projektverbund mehrerer Universitäten aufgebaut werden sollte.

stens Ende der 90er Jahre zur Bedingung öffentlich geförderter e-learning-Projekte. Gerade die Neuartigkeit des soziotechnischen Projekts Virtuelle Universität sowie die atemberaubenden Erwartungen, die daran geknüpft wurden (vgl. Schwarz 2001a), machen plausibel, von welchem hohem Interesse sowohl summative als auch formative Evaluation für alle Beteiligten war – auch seitens der Vertreter der Landespolitik, welche unter starkem Legitimationsdruck standen.<sup>9</sup> Im Gegensatz zu vielen anderen Virtuellen Universitätsprojekten wurde die interne Evaluation von VUMS finanziell mit einer knappen halben Million Euro gefördert und galt nicht wie so oft als eine Aktivität, die die Projektbeteiligten noch zusätzlich zu ihrer regulären Tätigkeit und der Projektarbeit tun müssen. Auch die ehrenamtlichen Gutachter hatten geringfügige, d.h. unkostendeckende Tagessätze.<sup>10</sup>

Abb. 1: Akteurskonstellation in der Virtuellen Universität



Die oben stehende Grafik verdeutlicht das Zusammenwirken der verschiedenen Akteursgruppen, v.a. derer die an Evaluation im weitesten Sinne aktiv beteiligt waren. Stakeholder des komplexen Projekts sind in erster Linie Inhaltslieferantinnen, Technikerinnen, Projektleiter, die sozialwissenschaftliche Begleitung sowie extern

9 „Qualitätssicherung und Evaluation müssen integraler Bestandteil der Projektkonzeption sein“ findet sich in der NMB-Ausschreibung als knappe Benennung der Notwendigkeit nicht aber des Ziels von Evaluation. Aufschlussreich ist, dass je knapper die Begründung des Auftraggebers ist, desto aufwendigere Konzepte zur Evaluation von den Auftragnehmenden erwartet wurden.

10 Neben der Reputation durch Begutachtungen wurde von den Gutachtern als Motivation für die aufwendige Tätigkeit, die intensive eigene Weiterbildung *durch* die Begutachtung benannt. Hier schließen sich selbstreferenziell Begründungszirkel zwischen Expertinnen und Laien, die mehr und mehr Gegenstand der Wissenschafts- und Techniksoziologie werden (so z.B. im Bereich Gesundheit Bal et al. 2002).

das Kompetenzzentrum, der Beirat und die Auftraggebenden. So gab es bei VUMS gleich mehrere Gruppen, die – z.B. laut Ausschreibung und Projektskizze – für „Evaluation“ zuständig waren: zum einen der vom Auftraggeber extern eingesetzte Kreis von vier Gutachtern. Er fertigte jährlich „Sachstandsberichte“ für den Auftraggeber an und entschied nach 3 Jahren Projektlaufzeit auf der Grundlage eines Projektbesuches des Projekts über dessen Weiterförderung. Zum anderen der Beirat, der im Laufe des Projekts auch den Namen „Zwischen-Evaluation“ bekam und die Gesamt-Entwicklung der Projektstruktur beurteilte, und zwar auf Grundlage der Angaben der Projektleiter. Zudem hatte er, auf eher abstrakterer Ebene, die Nachhaltigkeit des investierten Mitteleinsatzes für die Veränderung der Hochschulstrukturen zu begutachten und zu bewerten.

Auch ein angegliedertes Kompetenzzentrum sollte die Erfahrungen mit dem Virtuellen Hochschullehrverbund nach außen tragen, um das Wissen zu verbreitern und nicht involvierte Akteurinnen zu virtueller Lehre ermutigen. Es war ebenfalls für die redaktionelle Verarbeitung der Evaluationsergebnisse innerhalb und außerhalb des Projekts zuständig. Darüber hinaus gab es eine umfangreiche sozialwissenschaftliche Begleitung, bei der Evaluation nur eine Teilaufgabe war, zusätzlich zur didaktischen Beratung von Lehrenden und Studierenden sowie zur Grundlagenforschung. Ähnlich komplex wie der gesamte Projektarbeitsplan setzte sich auch die Begleitforschung aus vier verschiedenen Wissenschafts-Disziplinen zusammen und arbeitete standortverteilt in vier verschiedenen Instituten an drei verschiedenen Hochschulorten. Das im Vergleich zum Evaluationsdesign anderer Virtueller Hochschulen sehr anspruchsvolle Programm der Begleitforschung erwies sich anfangs als ein Konglomerat verschiedenster Zugangsweisen: die wirtschaftswissenschaftliche Evaluation konzentrierte sich auf Studierendenakzeptanz unter Marketinggesichtspunkten; die medienwissenschaftliche führte eine „Kontext-Evaluation“ durch, um Meinungsbilder von Studierenden und Lehrenden zu e-learning einzufangen; die didaktische Evaluation untersuchte, inwieweit traditionelle Lehre durch den Einsatz von IuK-Technologien verbessert werden konnte. Die ersteren beiden Evaluationszugänge wurden in der zweiten Projekthälfte nicht mehr weitergefördert, so dass nur die didaktische Begleitung weitergeführt wurde.

## 6. Methoden der Evaluation Virtueller Universitätsprojekte

Dieses facettenreiche Evaluationsdesign beinhaltete eine Fülle angewandeter Evaluationsmethoden, von denen hier nur die zentralsten genannt werden können.

Die didaktische Begleitforschung konzentrierte sich stark auf die konkreten Lehrveranstaltungen. Von den zum Teil stark fluktuierenden sieben wissenschaftlichen Mitarbeitenden und Hilfskräften wurden anfangs wöchentliche Intensivevaluationen verschiedenster Lehrveranstaltungen mit Technikeinsatz durchgeführt. Dazu kamen zwölf Fallstudien zu Teleseminaren und -vorlesungen sowie sogenannten Hybrid-Lehrveranstaltungen, bei denen die Integration von computer based trainings (CBT) optimiert werden sollte. Neben dem laufenden Lehrbetrieb, in den nach und nach internetgestützte Lehrformen integriert werden sollten, wurden elf quasi-experimentelle Untersuchungen zur CBT-Unterrichtsintegration angefertigt.

Über die Ergebnismeldung hinaus gab es Projekt screenings durch Abfragen mittels Fragebogen und Telefoninterview über den Stand der Teilprojekte. Waren die Datenanalysen von quantitativen Methoden (Fragebögen) dominiert, so wurden qualitative Methoden, wie die Szenario-Technik v.a. verwendet bei der Moderation von Sitzungen. Ein erheblicher Teil der Arbeit der internen Evaluation bestand zudem in dem Erstellen von Evaluationsberichten und Beiträgen für Konferenzen sowie der Kommunikation zwischen den verschiedenen Stakeholdern. In der zweiten Projekthälfte konzentrierte sich die Tätigkeit auf die intensive Zusammenarbeit mit einem Lehrstuhl.

Die Gutachter übernahmen nach Förderbeginn die Neu-Formulierung von Förderkriterien. Nachdem sich das ursprüngliche Ziel der „Schaffung der Virtuellen Universität“ implizit auf die Etablierung einiger Fernlehre-Elemente reduziert hatte. Damit aber wurden die Evaluationskriterien nach und nach abstrakter und entfernten sich mehr und mehr von der Lehre: „Nachhaltigkeit“ und „Übertragbarkeit“ sowie Auffinden von „Geschäftsmodellen“ und „Anerkennungsregularien“ verflüchtigten sich zu Worthülsen. Pro Jahr wurde vom Beirat eine sogenannte „Begehung“ des virtuellen Szenarios vorgenommen, bei der hauptsächlich mit der Projektleitung und den Teilprojektleitern verhandelt wurde. Nur in seltenen Fällen wurden vor der Begutachtung auch die öffentlichen Evaluationsberichte zur Kenntnis genommen. Den meisten Gutachtern war der Inhalt oder die Existenz der detaillierten Evaluationsberichte über virtuelle Lehre in VUMS nicht bekannt. Auch wurde mit keinem einzigen Studierenden von VUMS gesprochen.

Gegen die grundsätzliche Annahme des Evaluationsbooms spricht auch, dass im Zuge des 5 jährigen Projektverlaufs wurde die einst enorme Evaluationsaktivität erheblich gekürzt und umstrukturiert wurde: das Kompetenzzentrum wurde zur Projektmitte ersatzlos gestrichen, für alle Projektbeteiligten unverständlicherweise aber zum Projektende durch ein neu gegründetes Zentrum ersetzt. Die interne Evaluation wurde personell auf weniger als ca. ein Drittel gekürzt. Das Veto des Beirats war die Entscheidungsgrundlage für den Auftraggeber, wie VUMS in veränderter Form weitergeführt werden sollte. Auch als Reaktion darauf konzentrierte sich die Begleitforschung mehr und mehr auf ihre Grundlagenforschungsaspekte, die Diskussion ihrer Ergebnisse in den einschlägigen Forschungszirkeln sowie auf die stetige Verbesserung des Medieneinsatzes in den Lehrveranstaltungen eines Lehrstuhls.

Die Projektentwicklung von VUMS sowie die Entwicklung des Auftraggeber-Interesses führt plastisch vor Augen, dass es keine abgrenzbare Zuständigkeiten für Evaluation gab, sondern dass Evaluation eher als ein mikropolitisch Spiel<sup>11</sup> zu verstehen ist, bei dem das diffizile Wechselspiel von Wissen und Macht zum Tragen kommt: Welches Wissen ist relevant für welche Entscheidungen? Welche Entscheidungsträger, Hochschullehrende oder Bildungspolitikern brauchen wann welche Informationen? Wie detailliert dürfen die Antworten auf Evaluationsfragen sein, damit Entscheidungsträgerinnen sie wahrnehmen?

11 „Mikropolitik ist das Arsenal jener alltäglichen ‚kleinen‘ (Mikro-)Techniken, mit denen Macht aufgebaut und eingesetzt wird, um den eigenen Handlungsspielraum zu erweitern und sich fremder Kontrolle zu entziehen (vgl. Neuberger 1988).

Was wie eine regelrechte Über-Aktivität an Evaluationszuständigkeiten aussieht, zeigt, wie viele Aufgaben tatsächlich unter dem Label Evaluation zusammen gefasst werden: Begleitforschung und Beratungsservice (interne Evaluation), Ergebnisverarbeitung und -verbreitung (Kompetenzzentrum) und Begutachtung und Politik- bzw. Projektberatung (externe Evaluation).

## 7. Zur Rolle von Evaluierenden in virtuellen Hochschullehrverbänden

VUMS ist also ein sehr intensiv erforschtes und evaluiertes Projekt, aber es ist wenig bekannt über den Nutzen der Evaluation. Schwierigkeiten mit Evaluationsmethoden scheinen die geringsten Probleme der Evaluation ausgemacht zu haben, sondern vielmehr die Management- und Mikropolitik-Aspekte (Neuberger 1988; vgl. Schwarz 2003) einer sinnvollen Arbeitsteilung zwischen den verschiedenen Evaluationszugängen.

### 7.1 Die implizite Arbeitsteilung zwischen den Evaluationen: Evaluation als mikropolitisch Spiel

Im Folgenden soll detaillierter aus den Interviews berichtet werden, denn in den Gesprächen wurde rekonstruiert, wie sich Evaluations-, Beratungs- und Forschungsleistungen in VUMS überlappten. Dazu musste meist erst identifiziert werden, was mit „Evaluation“ eigentlich gemeint sei: die sozialwissenschaftliche Begleitung, der Beirat oder das Kompetenzzentrum oder das alltägliche Urteilen nach dem common sense? Die Lehrbücher dominierende begriffliche Trennung von externer und interner Begleitung hatte sich in der Praxis verflüssigt. Aus Sicht der Auftraggebenden, der Beiräte und des Kompetenzzentrums mag die sozialwissenschaftliche Begleitung zwar intern sein, von den Projektbeteiligten wurde sie aber meist als von außen in die eigene Lehre eingreifend betrachtet. Auch wurde das vom Ministerium als externe und unabhängige Beratungseinrichtung konzipierte Kompetenzzentrum als zu nah angesiedelt an den Projektbeteiligten wahrgenommen und deshalb schon nach der ersten Projektphase auf Empfehlung des Beirats aufgelöst.

So mag im Projektverlauf für viele Beteiligte offen geblieben sein, wer eigentlich evaluiert – die sozialwissenschaftliche Begleitung, das Kompetenzzentrum oder der Beirat – und was und wen sie evaluieren? Auch die Frage, wer eigentlich Adressat von Evaluation sei, und wer dazu verpflichtet wird, die vielen Evaluationsberichte auszuwerten, schien im Prozessverlauf immer unklarer zu werden: Dienen die Evaluationsergebnisse primär den Inhaltslieferanten zur Gestaltungsempfehlung ihrer Lehrveranstaltungen, der Wissensbündelung im Kompetenzzentrum, dem Beirat als Begutachtungsgrundlage oder dem Auftraggeber als Entscheidungsgrundlage für die Weiterförderung?

Auf halbem Wege des Projekts und damit zum Zeitpunkt der Weiterbewilligung wurde klar, dass es eine implizite Arbeitsteilung zwischen der Evaluation er-

ster und zweiter Ordnung gab. Zum einen gab es die flüchtige und eher die Hochschulstrukturen überschauende Begutachtung von oben durch den Beirat, deren Empfehlungen fast jedem Entscheidungsträger im Projekt bekannt waren. Das Gutachterurteil beruhte dabei mehr auf eigenen Erfahrungen mit e-learning als auf den detaillierten Forschungs- und Evaluationsergebnissen vom „Boden des Projekts“. Zum anderen gab es die aufwändige und detaillierte Evaluation von unten, deren umfangreiche Ergebnisberichte nur den wenigen bekannt waren, die sich dringlich für die Beantwortung der Frage interessierten: Wie ist traditionelle Universitätslehre durch Einsatz der IuK-Technologien konkret zu verbessern? Deren Ergebnisse wurden zwar von wenigern Hochschullehrenden genutzt, dies dann aber manchmal sehr intensiv und dankbar.

## 7.2 Die Rolle der Begleitforschenden: Zwischen Beratung und Forschung

Das Selbstverständnis der Begleitforschenden ist durchaus unterschiedlich, aber fast allen fünf wissenschaftlichen Mitarbeitenden und zwei Hilfskräften war eines gemeinsam: Sie sahen ihre Aufgabe vor allem in der Beratung durch Forschung. Unterschiedlich waren aber ihre Beratungsverständnisse, sei es, dass sie im Sinne des forschenden „Daten-Handwerks“ zur Entscheidungshilfe beitragen wollten, durch „Monitoring“ der Projektübersicht dienen oder durch „Reflexionen über Erfolge und Misserfolge“ virtueller Lehre begleiten wollten. Als Personen, die sich neben ihrer Tätigkeit als Evaluierende weiterhin wissenschaftlich durch Examens-, Doktor- oder Habilitationsarbeit qualifizierten, sahen sie ihre Aufgabe zwischen Forschung und Beratung positioniert. Lediglich ein wirtschaftswissenschaftlicher Mitarbeiter bildete eine Ausnahme: Für ihn diente Evaluation neben ihrer heuristischen Funktion im Kontext von e-learning als Element eines umfassenden „Controlling“-Ansatzes.

Anders als bei der Frage nach dem beruflichen Selbstverständnis fielen die Antworten auf die Frage danach, mit welchen Rollenerwartungen sich die Begleitforschenden im praktischen Berufsalltag konfrontiert sahen, sehr unterschiedlich aus.<sup>12</sup> Durchaus vergleichbar mit den Aussagen der meisten Evaluierenden drückte ein wissenschaftlicher Mitarbeiter der Psychologie aus, wie im Projekt auf die Evaluierenden reagiert wurde:

- 12 Dieses Machtverhältnis von Evaluierenden und die Unklarheit darüber, wer welche Rolle in der Virtuellen Universität spielt, spiegelte sich in den Interviews in den *Reaktionsweisen* auf meine Frage nach der Rolle und Funktion von „Evaluation“: Einige waren irritiert von der naiven Nachfrage und gingen davon aus, dass die Funktion und Rolle von Evaluation doch irgendwo festgeschrieben und schließlich aus den Antragspapieren ablesbar sei. Andere begriffen die Komplexität der Frage sofort und erzählten ausholend darüber, wie sich die Wahrnehmung dessen was als „Evaluation“ galt, im Laufe des Projekts verändert habe. Wiederkehrend wurden v.a. zwei hauptsächliche Gründe genannt, warum diese Frage schwer zu beantworten sei: „Für *wen* habe Evaluation welche Rolle?“. Denn darüber gäbe es durchaus verschiedene Sichtweisen. Und *wann* hatte Evaluation welche Rolle in den verschiedenen Projektetappen? Erst wenn man da angelangt sei, könne man zu der Frage vordringen: Welche Rolle *soll* denn Evaluation in solchen Projekten wie VUMS überhaupt haben?

„Das obere Drittel hat es gesehen als gewinnbringend-förderlich. (...) Das mittlere Drittel hat uns eher so als notwendiges Übel betrachtet. Und das letzte Drittel in der ersten Phase hat sich feindselig verhalten. So als Kontrollinstanz, auch mit Gerüchten schüren und Fehlinformationen geben, also kontraproduktiv.“

Eindrücklich schildert eine Gastwissenschaftlerin im Begleitforschungsteam zu welchen fast schon diskriminierenden Situationen das schlechte Image von Evaluation als Kontrollinstrument führen kann:

„Es bedarf so viel Einfühlungsvermögen auch, was die Leute grade von einem erwarten. Da gab es so eine Schlüsselsituation grade bei diesem ersten Projekt-Treffen, wo ich den Prof. Stolz kennen gelernt habe. Wir kamen so ins Gespräch und er meinte, woher ich käme und was ich machen würde. Und ich habe nur Begleitforschung gesagt und da meinte er gleich: ‚Oh! Ihr evaluiert uns ja!‘ Und dann hat er sich gleich so ein Bisschen weggestellt. (...) Ich wollte die Situation jetzt dazu nutzen, was er von den Evaluatoren hält. Ich weiß nicht, ob das alles so scherzhaft gemeint war oder ob das echt so ein Bisschen das Gefühl war: vor Euch muss man sich ja in Acht nehmen, denn im Prinzip habt Ihr das in der Hand. Ja soviel zu dieser Identifikation mit Identitäten, die einem da übergestülpt werden.“

Einige gingen angesichts der allgemein geringen Kooperationsbereitschaft mit der Begleitforschung so weit, zu sagen, dass Evaluation im Laufe des Projekts zum „Selbstzweck“ geworden sei:

„Vielleicht sehe ich das zu finster, aber die, die sich mal freiwillig gemeldet haben, weil sie evaluiert werden wollen, die haben das vielleicht auch so ein Bisschen so gemacht, weil Evaluation gleichzeitig auch so ein Stempel ist, so ein Gütezertifikat, das man aufbekommt, um zu sagen, ich habe mich dem ausgesetzt, also bin ich reflektiert. Da geht es aber auch mehr so um das Zertifikat und vielleicht weniger um die Inhalte dessen, was da eigentlich echt gelaufen ist.“

Fast alle im Begleitforschungsteam waren der Annahme, dass die erzielten Evaluationsergebnisse zwar extrem wichtig seien, wenn auch zum Teil für praktische Entscheidungen zu detailliert. So sei es ein wichtiges Ergebnis, dass gerade durch die Dominanz der Informatik die Gefahr bestünde, die tradierten kritikwürdigen Lehrveranstaltungen wie die Massenvorlesung durch die neuen Technologien festzuschreiben, anstatt sie abzulösen oder durch interaktivere Veranstaltungsformen zu ersetzen. Entscheidend für den Erfolg der Evaluation sei aber, ob die Evaluierten bereit sind, die Evaluationsergebnisse zu akzeptieren. Warum diese Beratungsangebote von den meisten Evaluierten verschmäht wurden, wurde auf unterschiedliche Gründe zurückgeführt, einer davon ist das Machtgefälle zwischen jungen Forscherinnen und alteingesessenen Hochschullehrenden:

„So ein 50 bis 60 jähriger Professor, der sein ganzes Leben so gelehrt hat, dem wird jetzt über die Evaluation gesagt: ‚Hör zu, so und so solltest Du es nicht machen. Mach es so!‘ Da weiß ich ja nicht, ob der das angenommen hat. Da ist Evaluation im Prinzip beratend, aber wohl auch nicht verpflichtend beratend gewesen.“

Etwas stärker noch klafften Selbstverständnis und Realität bei denen auseinander, die weniger mit der praktischen Umsetzung von Evaluation zu tun hatten, sondern lediglich konzeptionell für Evaluation verantwortlich waren, d.h. für das beantragte Untersuchungsdesign, die personelle Ausstattung und die repräsentativen Evaluationsaufgaben im Gesamt-Projekt. Die drei leitenden Professorinnen der Begleitforschung sahen ihre Aufgabe noch weniger in der Kontrollfunktion: Vielmehr sei Evaluation eine „Erkundung des gesellschaftlichen Kontextes“, eine „Moderation,

die dazu beiträgt aus Erfolgen und Fehlern zu lernen“ oder eine „Hilfestellung“ für die Verbesserung von Lehrveranstaltungen. Doch umso stärker stand das Selbstverständnis dieser federführenden Personen im Widerspruch zur Art und Weise, mit der das Beratungsangebot ihrer Mitarbeiterinnen im Projektverlauf angenommen wurde: alle drei dieser Professorinnen sahen sich mit *ihrem* Evaluationsangebot missverstanden. Entweder, weil sie abgewiesen wurden oder weil sie zur reinen Absegnung missbraucht wurden.

„Ich bin also häufig Gutachter und Experte für die Europäische Kommission in den Bereichen e-learning. Also ich habe, glaube ich, so einen Überblick über ganz Europa, mindestens 100 Projekte, die mindestens so groß sind wie VUMS, finanziell. Über mehrere Jahre! (...) Die haben ganz klare Erwartungen an das, was so ein Evaluator tun soll. Der kommt rein, wenn die Sache aus Sicht des Entwicklers oder des Entwicklungsteams fast perfekt ist. Weil sonst würden die sich gar nicht trauen, sich dem auszusetzen. Da ist so ein Wir-Ihr-Pool. (...) Die Evaluatoren lässt man in 99 % der Fälle nur herein, wenn das Produkt so gut wie fertig ist“.

Als Konsequenz aus dem verschmähten Beratungsangebot ergab sich für die Begleitforschenden zweierlei. Zum einen tat die Begleitforschung gut daran, wenn sie ihre Erwartungen an die Kooperationsbereitschaft mit Hochschullehrenden herunterschraubt.

„Also, ich glaube nicht, dass wir ein Feigenblatt sind. Nicht durchgängig, für manche ja. Aber für andere auch nicht. Und da man das nie letztlich weiß, orientiert man sich an denen, wo am meisten Ertrag für beide Seiten ist. Also wo der Inhaltsanbieter am meisten profitiert, profitieren wir am meisten von.“

Es gab so scheint es, in der Realität der Evaluation selten Kooperation mit allen zu Evaluierenden und im günstigsten Fall nur einen produktiven Austausch mit den wenigen, die Evaluation als zusätzliche Ressource und damit als Hilfestellung schätzten, auch wenn sie die „extrem problematische Doppelrolle von Beratung und Kontrolle“ sehen. Der Erfolg von Evaluationsangeboten scheint eben doch stark abzuhängen von der Form des persönlichen Aufeinanderzugehens. Und diese Erfahrung könnte selbst wieder als Ressource für die Verbesserung von Evaluationen genutzt werden.

Eine zweite Strategie, mit der mangelnden Kooperationsbereitschaft umzugehen, war der Rückzug auf die Forschung. Fragte man aber die Auftraggeber, Gutachter und das Kompetenzzentrum, so stellte sich heraus, dass sie genau diesen Rückzug auf Forschungsaspekte der Begleitung kritisch sahen oder sogar verübelten.

Die Projektleitung führte andere für sie bedeutende Funktionen der Begleitforschung an. Für sie habe sie vor allem den Zweck gehabt, einen Überblick über den Stand des Gesamtprojekts liefern zu können, also eine Monitoring-Funktion sowie eine gutes Marketing für die weitere Projektakquise: „Denn ohne Evaluation kriegen Sie heute kein Geld mehr rein“.

### 7.3 Die Rolle der Gutachter: Zwischen Vermittlung und Kontrolle

Erwiesen sich die Begleitforschenden im Projekt eher als Repräsentanten lokalen Wissens, so traten die viel reisenden Gutachter als Vertreter globalen Wissens auf. Dem eigenen Rollenselbstverständnis nach sahen sich alle vier Gutachter als Ver-

mittler zwischen Auftraggeber und Projekt und zwischen den verschiedenen Teilprojekten. Zwei dieser Gutachter sahen es zwar als problematisch an, in dieser Rolle neutral zu bleiben, betrachteten sich selbst aber als unparteiisch. Ein Ökonom verstand sich als „Transmissionsriemen“, ein Informatiker als „Übersetzungsorgan“. Ein weiterer Gutachter aus der Psychologie sah sich eher auf der Seite des „klassischen Politikberaters“, der sich dem Ziel des Auftraggebers am meisten verpflichtet fühle – dem entgegengesetzt verstand sich ein begutachtender Pädagoge mehr als Anwalt der Projekte. Die Probleme gutachterlicher Vermittlung wurden vor allem auf zwei Hemmnisse zurückgeführt. Das eine betraf die bereits erwähnten unklaren oder ungeteilten Projektziele:

„Man muss ja irgendwie zwischen den beiden vermitteln. Zumal wenn man sehr schnell merkt, dass der eine den anderen zwar fördert und der Geförderte auch sehr gerne die Förderung annimmt. Aber dass die eigentlich über ihre wirklichen Ziele nicht im Einverständnis sind.“

Ein weiteres Gutachterproblem sei die mangelnde Handlungsorientierung von Evaluationsergebnissen:

„Was noch mehr gilt bei der Evaluation als sonst bei empirischen Forschungen: Es ist sehr sehr schwer letztendlich, daraus Handlungsorientierungen und Handlungswissen zu ziehen. Und dann grade in einem Bereich, der sich so vehement verändert, wo wir heute noch nicht einmal wissen: Die Kisten, auf die wir das alles jetzt hin machen, ist das überhaupt irgendwo das Medium, mit dem man interagieren wird? Oder hat man vielleicht in drei Jahren ganz andere Dinger da?“

Hinzu komme der grundsätzliche Zweifel, wann überhaupt der rechte Zeitpunkt für die Evaluation eines Projektes sei:

„Evaluation könnte ja auch ganz anders sein. Wie eine Expo. Eine Weltausstellung virtueller Universitäten, wo man eingeladen wird, virtuelle Lehre zu erleben. Ich glaube, wir sollten mehr die Augen und Ohren öffnen. Solche innovativen Projekte brauchen Visionen. Es darf nicht sein, dass bei solchen Pionierprojekten die Schrauben immer gleich so fest angedreht werden! Man darf solchen innovativen Projekten mit Evaluation nicht die Luft zum Atmen nehmen. Damit ist niemandem geholfen!“

War das Selbstbild der Gutachter als Vermittler relativ homogen, so gingen die Bilder, die Gutachter bei den anderen Akteurinnen hinterlassen haben, stark auseinander: Die Auftraggebenden sagten, dass sie die Gutachter als sehr hilfreich empfunden haben und dass für sie die eigentlich relevante Erkenntnis war: „Die Evaluationsfunktion (sic!) vom Beirat wurde 100prozentig erfüllt“. Er habe für beide Seiten als Korrektiv gewirkt und Empfehlungen abgegeben. Besonders auch angesichts der geringfügigen finanziellen Vergütung des Ehrenamts der Gutachter waren die Ministerialbeamten dem Beirat voll und ganz dankbar und betonten, dass sie extrem auf deren Mithilfe angewiesen seien. Dieses Verständnis deckte sich mit dem Eindruck, den das Kompetenzzentrum vom Beirat hatte: „Der Beirat kam, sah und siegte.“

Die Projektleitung hatte ein eher ambivalentes Bild vom Beirat: Zwar wurde er von der Projektleitung von vornherein als Kontrollgremium verstanden. Aber er moderierte und vermittelte auch zwischen Ministerium, Hochschulleitung und Projekt und hielt den Projektleitern etwas „den Rücken frei“. Wichtig schien zu sein, dass die Projektleitung mit dem Beirat jemanden hatte, der zwischen wissenschaftlichen und politischen Handlungszwängen und Denkweisen vermitteln konnte.

Dennoch sprachen sich beide Projektleitenden dafür aus, dass der Beirat seine vermittelnde Funktion noch stärker hätte wahrnehmen können und die Expertise, die ja auch im Projekt entwickelt wurde, wenigstens hätte respektieren müssen.

Dass der Beirat die Evaluationsergebnisse kaum berücksichtigt hat und den meisten noch nicht einmal deren Existenz bekannt war, wurde von der Begleitforschung heftigst kritisiert. Er wurde als „lächerliche wissenschaftliche Prüfgruppe“ betitelt und als ein eingeflogenes „Gremium, das beeindruckt werden muss“, „willkürliche Absegner, die immer nur auf Strukturelles schauen und sich nie trauen, konkret zu werden“. Was den Gutachtern vor allem vorgeworfen wurde, war, dass sie extrem stark ihren eigenen Erfahrungshintergrund mit e-learning bzw. Projektorganisation einbrächten<sup>13</sup> und kaum auf die spezifischen Erfahrungen im Projekt schauten.

## 8. Das Spannungsfeld der Evaluation

In der Zusammenfassung der Interview-Aussagen aus der Fallstudie Virtueller Universitäten lässt sich feststellen, dass es sehr unterschiedliche Selbst- und Fremdverständnisse vom Evaluieren gibt, die lediglich analytisch, nicht aber in der Praxis zu trennen sind. Die daraus resultierenden ungeplanten und undurchsichtigen Formen der *Arbeitsteilung* bringen Missverständnisse und Enttäuschungen mit sich und schaffen eine paradoxe Situation, in der im Namen von Evaluation und wissenschaftsbasierter Politik aneinander vorbeigeredet werden kann.

Mehr oder weniger sehen sich alle Beteiligten selbst in erster Linie als Beratende. Doch nicht alle vertrauen den anderen oder erkennen die Bewertungskompetenz des anderen an, weil neben dieser Beraterrolle auch stets noch andere Rollen zu erfüllen sind, die untereinander konfliktieren. Evaluierende müssen nicht nur beraten, sie müssen auch zwischen verschiedenen Stakeholdern aus Politik und Universität vermitteln, müssen die Ziele und Maßstäbe der Forschung im Auge behalten, auch wenn diese mit dem Projekterfolg wenig zu tun haben, und müssen die Leistungen anderer kontrollieren und bewerten, was eine wenig förderliche Basis für Kooperation und ernst gemeinte Beratung darstellt. Offen bleibt die Frage, wie Evaluation sich glaubhaft positionieren kann zwischen dem ehrbaren Ziel der Beratung und der nicht zu leugnenden Pflicht zur Kontrolle in der praktischen Arbeit. Evaluierende sehen sich nicht selten in der Berufssituation, die in den Interviews als die „vier Jacketts-Situation“ beschrieben wurde: Für jedes der vier Evaluationsziele müsse man rasch in ein neues Jackett schlüpfen, um den andauernden Rollenwechsel deutlich zu machen. Vor lauter Jackettwechsel wird die Arbeit des Evaluierens zur Klamotte.

13 So zum Beispiel ein pädagogischer Gutachter, der einen Antrag auf Weiterförderung innerhalb eines Teil vom Hochschullehrverbund VUMS ablehnte mit einer Analogie zum Projektmanagement in der Biomedizin-Forschung.

Abb. 2: Das Spannungsfeld der Evaluation



Damit sind die vier Pole des Spannungsfeldes genannt, in dem Evaluation möglicherweise auch über Politikfelder hinweg stattfindet: *Beratung*, die gegenseitiges Vertrauen erfordert (wie bei der Lernpartnerschaft zwischen einem Inhaltsanbieter und einem Teil der Begleitforschung oder dem Ministerium und den hochkarätigen Gutachtern, nicht aber zwischen diesen beiden Hierarchien); *Kontrolle*, die die Anerkennung und Wirksamkeit von Macht voraussetzt; *Forschung*, die der Neutralität und anerkannten Expertise bedarf sowie *Vermittlung*, die eine extrem hohe soziale Intelligenz erfordert.

Ganz gleich, welchen methodischen oder theoretischen Evaluationszugang man als Einzelne wählt, Evaluierende befinden sich innerhalb dieses Spannungsfeldes stets in Doppel- und Mehrfachrollen. Nichtsdestotrotz müssen sich Evaluierende in diesem Feld durch eine Rollenklärung verorten und wissen, dass sie von anderen dort verortet werden.

## 9. Constructivist evaluation: Die Evaluierenden als Mediatorinnen?

Mit der Expansion der Chiffre Evaluation wächst ihre Scheinplausibilität und Beliebbarkeit. Trotz der Standardisierungsbemühungen der Evaluation (DeGEval 2002) wissen die Beteiligten in der Praxis immer weniger, was man von Evaluation zu erwarten hat, und ob und wozu sie genutzt wird; jede/r trägt seine Vorstellungen dazu auf den Evaluationsmarkt.

Gubas und Lincolns These von der der Evaluation als Verhandlungsführerin konnte zwar teilweise bestätigt werden. Allerdings lösen sich die Rollen Messen, Beschreiben, Urteilen und Vermitteln nicht ab – wie es ihr Modell der vier Evalua-

tionsgenerationen suggeriert – sondern diese Rollen addieren sich auf. Die Fallstudie VUMS zeigt nicht nur, dass Evaluierende *auch* als Vermittelnde nachgefragt werden könnten, sondern v.a. dass dieser Rollenwechsel höchst abhängig ist vom mikropolitischen Machtgefüge in einem Projekt und seines gesellschaftlichen Kontextes. Warum sollte Evaluation auch im machtfreien Raum stattfinden? Trotz der Utopie von wissensbasierten gesellschaftlichen Entscheidungen – dem Demokratisierungsversprechen von Evaluation – wird Evaluation stark bestimmt von den Drohgebärden und der Ignoranz inszenierter Machtsymbolik in und zwischen Wissenschaft und Politik. Entgleiten also sozialkonstruktivistische oder anti-positivistisch motivierte Evaluationsansätze zu einem politisch naiven Demokratisierungstraum, wo Evaluation sich in reiner Vermittlungstätigkeit auflöst? Dann droht Evaluation mittels dieser idealistischen Überbautheorie zum unglaublichen allseits politisch korrekten Selbstanstrich zu werden (vgl. Fischer 1995). Das macht wiederum auf der anderen Seite machtblindere methodenversessene Ansätze nicht besser. Auf beide ideologischen Verstrickungen von Evaluationstheorien hinzuweisen könnte ein dankbares Anwendungsfeld einer kritischen Theorie der so genannten Wissensgesellschaft sein: dass die als „mode 2“ bezeichnete Transformation von wissenschaftlichem Geschäft neben ihren demokratieförderlichen Kennzeichen – wie Interdisziplinarität und Ent-Hierarchisierung von Wissensansprüchen – dort problematisch wird, wo „Universitätsangehörige als Berater oder Produzenten in der kulturellen Imageproduktionsindustrie fungieren“ (Nowotny 1999: 102).

Die Mehrdeutigkeit und *Verhandelbarkeit von Evaluationsrollen* aber wurde in der us-amerikanischen Fachliteratur bereits mehrfach thematisiert (Weiss 1974; Rossi 1986). Nur selten werden hierzulande die Widersprüche zwischen Entwicklungs-, Forschungs- und Kontrollansätzen in der Evaluation angesprochen (Kromrey 2000, 2001), so dass es im deutschsprachigen Raum bisher keinen fokussierten wissenschaftlichen Diskurs darüber gibt. Das ist vor allem deshalb schade, weil es gerade in der Gestaltung des Aus- und Weiterbildungscurriculums für Evaluierende entscheidend wäre, für diese Rollenwidersprüche zu sensibilisieren und auf die mikropolitischen Anforderungen an Evaluierende vorzubereiten und zu begleiten.

Offensiver gehen dagegen andere Evaluationsforscherinnen das Problem der Mehrdeutigkeit von Evaluation an, indem sie das Rollenspektrum aufzeigen und mit dem überstrapazierten Gebrauch des Labels „Evaluation“ in Verbindung bringen. So zum Beispiel Stephen Morabito, der in Evaluierenden den „educator“, „consultant“, „facilitator“ und „counsellor“ entdeckt (Morabito 2002) oder Prudence Brown, die ein noch größeres Spektrum der Evaluationsrollen aufblättert:

“The use of the term ‘evaluation’ has undergone considerable expansion in its purpose, scope, and methods over the last twenty years. This expansion is reflected in the range of roles that evaluators currently play, or attempt to play, in carrying out their work: scientist, judge, educator, technical assistant, facilitator, documenter/historian and repository of institutional memory, coach, manager, planner, creative problem solver, co-learner, fund-raiser, and public relations representative.”

Die von Guba und Lincoln skizzierte konstruktivistische Trendwende aufnehmend, rückt Brown Evaluierende aus dem fälschlichen Bild einer technokratischen Randposition ins Licht der unvermeidlich aktiv-mitbeeinflussenden Tätigkeit. Aus positivistischer Sicht ein Alptraum, aus sozialkonstruktivistischer Sicht noch immer ein Problem, aber auch die Chance zur Gestaltung von Evaluationsrollen:

“While the field has moved toward a view of evaluation that is fundamentally normative rather than technical, political rather than neutral or value-free, many unresolved questions remain about how evaluators should operationalize such a view as they select and shape the role(s) they play in any particular evaluation enterprise.“ (Brown 1995: 1).

Genau dieses *shaping of evaluation roles* hat die Fallstudie VUMS als skizzenhafte und lediglich explorative Analyse von Berufsrollen vor Augen geführt: „Die“ Evaluation in einem Projekt gibt es nicht. Projekt-Evaluation ist veränderlich und muss und soll dies wohl auch sein, um den Projektveränderungen Rechnung zu tragen. Als zwei Folgewirkungen des Gebrauchs von Evaluation als Chiffre sind dabei besonders hervorzuheben: Der erhöhte Bedarf an erstens Arbeitsteilung und zweitens Rollenklärung. Evaluation, begriffen als ein mikropolitisch Spiel der Evaluationen vieler, macht sichtbar, dass relativ ist, was als intern und extern wahrgenommen wird. Für die Evaluationsforschung wie für die Aus- und Weiterbildung von Evaluierenden bedeutet dies die Not eines erhöhten Engagements für Fragen der Arbeitsteilung zwischen beratender, forschender, kontrollierender und vermittelnder Evaluation. Aber alle Funktionen in einer Person oder Institution sind höchst unglaubwürdig und können einem Projekt mehr schaden als nutzen.

Wünschenswert wäre, wenn neben der Festlegung auf Standards die Sicht auf die Vielfalt subjektiver Evaluations- und Beraterstile nicht verdeckt würde und so ein Diskurs zwischen „standards and diversity“ in konstruktive Vorschläge für die Aus- und Weiterbildung von Evaluierenden mündete. Das würde bedeuten über Methodenerörterungen hinweg nicht zu vergessen, die Anforderungen an die soziale Intelligenz von Evaluierenden deutlich zu machen und die Förderung sozialer Kompetenz nicht nur zur Randaufgabe zu machen. Uwe Schmidt und Manfred Hennen aus dem Arbeitskreis Aus- und Weiterbildung der DeGEval haben bereits „Klärungsprozesse“ gegenüber den Auftraggebererwartungen gefordert und damit die Notwendigkeit von „Situationsdefinitionen“ und die „Vermittlung divergierender Ansprüche“ (Schmidt, Hennen 2001: 33). Evaluation beinhaltet eben nur peripher „technische Verfahren“ und ist „in vielen Fällen nur durch die Auseinandersetzung in und mit der Praxis zu erlernen“ (Hennen et al 2002).

Wenn Evaluation – also das Bestreben, den Wert einer Sache auszumachen – überhaupt geschützt werden muss, dann davor, dass sie sich selbst entwertet. So lange die Chiffre Evaluation für alles und nichts erhalten muss, können Evaluationsprojekte mehr Arbeit an einer intensiven Arbeitsteilung und gemeinsam getragene Rollenklärung bedeuten als einen Nutzen für das Projekt stiften. Nichtsdestotrotz ist das „role shaping“ in der Evaluation eine Überlegung wert, denn sonst wäre Evaluation – wenn auch *methodisch* sauber – Quacksalberei.

## 10.Literatur

- Bal, Roland/Bijker, Wiebe E./Hendriks, Ruud (2003): Paradox van wetenschappelijk gezag. Over de maatschappelijke invloed van adviezen van de Gezondheidsraad. Den Haag: Gezondheidsraad.
- Baumgartner, Peter et al. (2002): Evaluation von Lernplattformen. E-Learning Praxishandbuch. Auswahl von Lernplattformen. Innsbruck: Studienverlag.
- Baumgartner, Peter (1999): 10 Todsünden in der Evaluation interaktiver Lehr- und Lernmedien. In: Lehmann, K. (Hg): Studieren 2000 – Alte Inhalte in neuen Medien? Münster: Waxmann.

- DeGEval (2002): Standards für Evaluation, <http://www.degeval.de/standards/Broschuere.pdf> [last checked: 7.3.03].
- Beywl, Wolfgang (1999): Zur Weiterentwicklung der Evaluationsmethodologie. Grundlegung, Konzeption und Anwendung eines Modells der responsiven Evaluation. Bern, New York, Paris: Lang (Reprint von 1988).
- BUND-LÄNDER-KOMMISSION (1998): Multimedia im Hochschulbereich. Heft 63. Bonn: <http://www.blk-bonn.de> [last checked: 14.06.01].
- BUND-LÄNDER-KOMMISSION (1999): Multimedia im Hochschulbereich. Heft 76. Bonn: <http://www.blk-bonn.de> [last checked: 14.6.01].
- BUND-LÄNDER-KOMMISSION (2000a): Digitale Revolution. Herausforderung + Chance für die Hochschulen. Bonn: <http://www.blk-bonn.de/presse00-12.htm> [last checked: 14.6.01].
- BUND-LÄNDER-KOMMISSION (2000b): Globalisierung des Bildungsmarktes durch Neue Medien. Auswirkungen auf die Hochschulen. BLK-Sondersitzung über Auswirkungen von Multimedia. Heft 81. Bonn: <http://www.blk-bonn.de> [last checked: 14.6.01].
- BUND-LÄNDER-KOMMISSION (2000c): Multimedia in der Hochschule. Heft 85. Bonn: <http://www.blk-bonn.de> [last checked: 14.6.01].
- BMBF (1999): Mut zur Veränderung. Deutschland braucht moderne Hochschulen. Vorschläge für eine Reform, April 1999. <http://www.bmbf.de> [last checked: 20.4.01].
- BMBF (2000a): Anschluß statt Ausschluß. IT in der Bildung. <http://www.bmbf.de> [last checked: 20.4.01].
- BMBF (2000b): Projektträger Neue Medien in der Bildung. Förderbereich Hochschulen. <http://www.gmd.de/NNM/Bereich%20Hochschulen/Hochschulen.htm> [last checked: 20.4.01].
- BMBF (2001): Bundesregierung plant neues Aktionsprogramm „Lebenslanges Lernen für alle“. <http://www.bmbf.de/presse01/299.html> [last checked: 20.4.01].
- Brown, Prudence (1995): The Role of the Evaluator in Comprehensive Community Initiatives. In: Connell, James P. et al. (Hg.): *New Approaches to Evaluating Community Initiatives. Concepts, Methods, and Contexts*. Washington D.C.: The Aspen Institute.
- Callon, Michel (1986): The Sociology of an Actor-Network. The Case of the Electric Vehicle. In: Callon/Law/Rip (Hg.): *Mapping the Dynamics of Science and Technology*. London: Macmillan.
- Collis, Betty/van der Wende, Marijk (Hg.) (2003): *Models of Technology and Change in Higher Education. An international comparative survey on the current and future use of ICT in Higher Education*. <http://www.utwente.nl/cheps/documenten/ictrapport.pdf>.
- Fischer, Frank (1995): *Evaluating Public Policy*. Wadsworth/ USA.
- Gibbons, Michael, et al. (1994): *The New Production of Knowledge. The Dynamics of Science and Research in Contemporary Societies*. London.
- Guba, Egon G./Lincoln, Yvonna S. (1989): *Fourth Generation Evaluation*. Newbury Park, London, New Delhi: Sage Publications.
- Guba, Egon G./Lincoln, Yvonna S. (2001): *Guidelines and Checklist for constructivist (a.k.a. Fourth Generation) Evaluation*. <http://www.wmich.edu/evalctr/checklists/constructivisteval.htm>.
- Hennen, Manfred (2002): Die Module Organisationswissen und Feldkenntnisse. In: *Zeitschrift für Evaluation*. Leverkusen: Leske + Budrich.
- Mayring, Philipp (1988): *Qualitative Inhaltsanalyse. Grundlagen und Techniken*. Weinheim.
- Martens, Bernd/Clemens, Ute (1999): Viel Lärm – um was? Zur Faszination von Multimedia in der Hochschullehre. In: *Leviathan*, S. 27.
- Morabito, Stephen (2002): Evaluator Roles and Strategies for Expanding Evaluation Process Influence. In: *American Journal of Evaluation* (No. 3, Summer 2002).
- Negt, Oskar (1969): *Soziologische Phantasie und exemplarisches Lernen: zur Theorie der Arbeiterbildung*, 3. Aufl., Frankfurt am Main: Europäische Verlagsanstalt.
- Neuberger, Oswald (1995): *Führen und geführt werden*. Stuttgart: Enke, Kapitel 10.
- Ortmann, Günther/Küpper, Willi (1988): *Mikropolitik, Macht und Spiele in Organisationen*. Opladen: Westdt. Verlag.
- Ortmann, G. (1995): *Formen der Produktion. Organisation und Rekursivität*. Opladen: Westdt. Verlag.
- Pollock, Neil/Cornford, James (2003): *Putting the university online: information, technology, and organisational change*. Buckingham: Open University.

- Nowotny, Helga (1999): Es ist so. Es könnte auch anders sein. Frankfurt/M: Suhrkamp.
- Kromrey, Helmut (2000): Qualität und Evaluation im System Hochschule. In: Stockmann, Reinhard (Hg.): s.u.
- Kromrey, H. (2001): Evaluation – ein vielschichtiges Konzept. In: Sozialwissenschaften und Berufspraxis Nr. 2/2001, S. 105-131.
- Law, John/Hassard, John (Hg.) (1999): Actor network theory and after. Oxford: Blackwell/The Sociological Review.
- Law, John (1994): Organizing modernity. Oxford: Blackwell.
- Rossi, Peter H. et al. (1988): Programm-Evaluation. Einführung in die Methoden angewandter Sozialforschung. Stuttgart: Enke.
- Robins, Kevin/Webster, Frank (2002): The Virtual University? Knowledge, Markets, and Management. Oxford, New York: Oxford University Press.
- Rosenstiel, Lutz von (2002): Die Quacksalber am Hausieren hindern. Evaluatoren plädieren für eigenen Studiengang. In: Mainzer Rheinzeitung vom 18.10.2002.
- Ryan, Steve et al. (2000): The Virtual University. The Internet and Resource-Based Learning. London, Sterling (US): Kogan Page.
- Schmidt, Uwe/Hennen Manfred (2001): Aus- und Weiterbildung in der Evaluation. In: DeGEval (Hg.): Evaluation. Reformmotor oder Reformbremse? Köln.
- Schmidt, Uwe (2002): Methodenkompetenz in der Evaluation. In: Zeitschrift für Evaluation. Leverkusen: Leske + Budrich.
- Schulmeister, Rolf (1997): Grundlagen hypermedialer Lernsysteme. Theorie – Didaktik – Design. München, Wien: Oldenbourg.
- Schulmeister, Rolf (2001): Virtuelle Universität – virtuelles Lernen. München: Oldenbourg.
- Schulmeister, Rolf (2003): Lernplattformen für das virtuelle Lernen: Evaluation und Didaktik. München: Oldenbourg.
- Schwarz, Christine (2001a): e-learning und Bildungspolitik. Von der Nachhaltigkeit hoher Erwartungen. In: Wagner, Erwin (Hg.): Virtueller Campus. Szenarien, Strategien, Studium, Tagungsband der Jahrestagung der Gesellschaft für Medien in der Wissenschaft (GMW). Band 14. Münster: Waxmann Verlag.
- Schwarz, Christine (2001b): Evaluation von e-learning in der Hochschullehre. Ein Experimentierfeld im Experimentierfeld. In: Deutsche Gesellschaft für Evaluation (Hg.): Evaluation – Reformmotor oder Reformbremse? Köln.
- Schwarz, Christine/Buchholz, Andrea (2002): Deferring Uncertainties in the Virtual University. About Trojan Horses and Hot Potatoes. European Association for the Study of Science and Technology (EASST) 2002. Conference Paper on CD Rom. <http://www.wa.uni-hannover.de/WIR/Schwarz/Schwarz-publik.html>.
- Schwarz, Christine (2003): Der Step-by-Step-Guide zur Mitarbeiterbefragung. In: Personal. Zeitschrift für Human Resource Management, 04/2003.
- Stockmann, Reinhard (2000): Evaluation in Deutschland. In: Stockmann, Reinhard (Hg.): Evaluationsforschung. Grundlagen und ausgewählte Forschungsfelder. Opladen: Leske+Budrich.
- Weiss, Carol H. (1974): Evaluierungsforschung. Methoden zur Einschätzung von sozialen Reformprogrammen. Opladen.