

# PLATTFORM ■ TECHNOLOGIE ■ EVALUIERUNG

Nr. 11

Dezember 2000

## ■ THEMA

### **Projektevaluierung und Monitoring beim FFF**

Klaus Schnitzer, Reinhard Zeilinger

### **Ein Evaluationsmodell für Joanneum Research**

Dorothea Sturn

### **Bewertungen von Programmen der EU- Regionalpolitik**

Rudolf Schicker

## ■ BERICHTE

### **Konferenzberichte**

- The Impact on Academic  
Institutions of Research  
Evaluation Systems
- Strategic Evaluation: the  
governance of the research  
systems in the EU

Eine Initiative von  
BUNDESMINISTERIUM FÜR VERKEHR, INNOVATION UND  
TECHNOLOGIE  
JOANNEUM RESEARCH FORSCHUNGSGESELLSCHAFT M.B.H.  
ÖSTERREICHISCHES INSTITUT FÜR WIRTSCHAFTSFORSCHUNG  
ÖSTERREICHISCHES FORSCHUNGSZENTRUM SEIBERSDORF  
TECHNOLOGIE IMPULSE GESELLSCHAFT M.B.H.  
TECHNOPOLIS AUSTRIA  
FORSCHUNGSFÖRDERUNGSFONDF FÜR DIE GEWERBLICHE  
WIRTSCHAFT  
FONDS ZUR FÖRDERUNG DES WISSENSCHAFTLICHEN  
FORSCHUNG

## Vorwort

In unserer letzten, sehr gut besuchten Plattformveranstaltung ging es um das Thema Monitoring. Zwei Referenten, G. Fayl (DG Research, zum 5. EU-RP) und R. Schicker (Öst. Raumordnungskonferenz zu Strukturfondsförderungen) zeigten anhand sehr großer, facetten- und projektreicher Programme die Herausforderungen bei der Verfolgung solcher gewaltiger Förderinitiativen. Welche Ebene ist geeignet? Was kann man nur Ex ante absegnen, was kann man begleitend verfolgen? Welche Schlussfolgerungen kann man für das Management, für Folgeprogramme ziehen? Einen Überblick dazu gibt Schicker in dieser Ausgabe. Zwei Überblicks-Beiträge von Fayl sind bereits in Newsletter Nr.8 / Frühjahr 1999 erschienen.

Die Projektbewertung und das konkrete Monitoring von Projekten einer Förderstelle wurden von K. Schnitzer und R. Zeilinger anhand des FFF-Monitoringsystems verdeutlicht. Die Herausforderungen bestehen einerseits darin, zahlreiche Ex ante-Kriterien auf eine aussagekräftige Weise und in einer guten Gewichtung zueinander wirksam werden zu lassen, andererseits in der kontinuierlichen Generierung und Nutzung von Daten, die über den Zyklus hinweg entstehen. Damit kann ein integrierter, durchgängiger Prozess ermöglicht werden. Bereits vorhandene Daten dienen der Weiterverfolgung – und dies ermöglicht bei einer hohen Zahl von Projekten eine effiziente Förderabwicklung.

Zum Thema Institutionenevaluierung gibt D. Sturn einen Überblick auf die methodisch interessante und für Österreich teilweise pionierhafte Evaluierung des Joanneum Research. Den Abschluss machen zwei Konferenzberichte zu internationalen Evaluierungsfragen. Hoffentlich ist für Sie etwas Interessantes dabei!

## Preface

For our non German-speaking readers this issue of the Plattform newsletter is something of a challenge. If our level of contributions in English will remain on this all time low, we should send you a pocket dictionary instead. We sincerely apologise!

This newsletter is mainly about monitoring issues: K. Schnitzer and R. Zeilinger provide an insight into ex ante evaluations and monitoring procedures of Austria's Industrial Research Promotion Fund (FFF), our most important funding organisation for applied research projects. Due to the big number of applications, a standardised multi-criteria set is used for ex ante evaluations. In this context the ample set of data and information is used for the monitoring of the projects and for further evaluation steps. This system will in the near future comprise an even stronger link to ex post evaluations regarding effectiveness and efficiency of FFF funding.

R. Schicker gives us an insight into EU Structural Funds' Monitoring procedures, taking the case of the period 1995-99 in Austria. He highlights the big challenges, namely to close the policy cycle and to learn systematically from past experiences in such a big, multi-faceted programme structure.

Sturn describes the ongoing and methodologically interesting institutional evaluation of the Austrian contract research organisation Joanneum Research. Our issue closes with some remarks on international evaluation conferences: On University Evaluation in Lisbon in June and on Strategic Evaluation in Lyon last month.

If you are interested in more information, don't hesitate to contact the authors or the Newsletter team!

*Michael Stampfer für das Plattform-Team*

## Projektevaluierung und Monitoring beim FFF

### Analyse und neues Bewertungs- und Indikatorensystem

Klaus Schnitzer und Reinhard Zeilinger

Der FFF hat schon frühzeitig ein formalisiertes Projektbewertungssystem eingeführt und in den Jahren auch stets weiterentwickelt. Seit 1997 werden die Kriterien auch für ein laufendes Projektmonitoring benutzt, wobei zusätzlich noch erfolgsorientierte Indikatoren eingeführt wurden.

Dipl.-Ing. Zeilinger, der beim FFF Projekte aus dem Bereich Physik, Software, Messtechnik und Medizintechnik evaluiert, hat in seiner Diplomarbeit „Erfolgsabschätzung von Neuprodukten in der Frühphase“ (Fachhochschule für Management u. Unternehmensführung, Wien 2000) die Monitoringdaten im Hinblick auf die gesamte Evaluierung der Projekte untersucht. Die Ergebnisse sind u.a. in das neue Projektbewertungssystem des FFF eingearbeitet worden und sollen auch im Monitoringprozess verwendet werden. Derzeit läuft auch für eine kleine Anzahl von Projekten, die von der Oesterreichischen Nationalbank mitfinanziert wurden, der Versuch, inwieweit eine Ex-ante-Evaluierung, Monitoring und eine Ex-post-Evaluierung mit dem gleichen Set definierter Kriterien verbunden werden kann.

Im folgenden werden die Ergebnisse der Arbeit von Zeilinger, die Grundsätze des neuen Bewertungssets sowie das Bewertungsschema kurz vorgestellt.

Das FFF-Bewertungssystem umfasst 4 Bewertungsgruppen, für die jeweils maximal 50 Punkte vergeben werden können, wobei 50 die beste Bewertung wäre, 25 den Durchschnitt darstellt. Zwei der Gruppen setzen sich aus technisch orientierten Kriterien zusam-

men, nämlich die Projektbewertung an sich, weiters die Projektbewertung in bezug zur einreichenden Firma. Aus wirtschaftlicher Sicht werden die Performance der Firma (bei Neugründungen auch die zukünftige Performance) sowie die Marktchancen des Projektes bewertet. Sowohl bei den wirtschaftlichen als auch den technischen Kriterien werden auch Faktoren wie Ökologie, Forschungsdynamik, volkswirtschaftliche Kriterien sowie soziale Aspekte bereits mit bewertet.

Aus der schematischen Darstellung (*FFF-Monitoring System, im Anhang zu diesem Artikel*) ist zu sehen, dass in dem seit 1997 eingerichteten Monitoring-system eine jährliche Neubewertung dieser Kriterien erfolgt. Nach der Prüfung des Jahresberichtes sowie der technischen Kontrolle werden außerdem noch die Qualität der Forschungstätigkeit sowie die Ergebnisse bewertet; weiters auch die Möglichkeit der Verwertung der Ergebnisse. In der Regel wird jedes Projektjahr aus wirtschaftlicher Sicht und auf Kosten vor Ort geprüft. Dabei erfolgt eine Neubewertung der wirtschaftlichen Kriterien. Es werden auch das Projektmanagement, das Rechnungswesen, die Ausstattung sowie das Personal indikatormäßig erfasst. Bei mehrjährigen Projekten erfolgt vor Neugenehmigung der Fortsetzung wiederum eine Bewertung, ebenfalls im Jahresschritt das vorhin beschriebene Monitoring. Für das Einzelprojekt ergibt sich damit eine gute Beobachtbarkeit über den gesamten Förderungszeitraum. Weiter ist bei nicht ordnungsgemäß durchgeführten Projekten eine flexible Kontrolle und auch ein Abbruch der Förderung möglich. Interessant ist natürlich die Frage, inwieweit man diese Daten weiter nutzen kann. Zeilinger hat hier vor allem folgende Fragen untersucht:

- Kann mittels statistischer Verfahren eine komprimierte Darstellung der untersuchten Datenbasis erreicht werden?

- Kann die große Anzahl an Entscheidungskriterien auf wenige Faktoren reduziert werden?
  - Welche Zusammenhänge sind in den Daten zu entdecken?
  - Können Merkmale gemäß ihrer korrelativen Beziehungen in voneinander unabhängige Gruppen klassifiziert werden. Wenn ja, welche Bedeutung kann das für zukünftige Förderentscheidungen haben?
  - Es soll ermittelt werden, ob dieses Bewertungssystem Vorhersagekraft hat, oder ob die für einen externen Begutachter nicht fassbaren Einflussfaktoren jede Vorhersage zunichte machen.
  - Technologieentwicklung wird in der Literatur als ein strategisches Element für eine Firma betrachtet, um sich vom Wettbewerb abzuheben. Ist der technologische Vorsprung im Sinne von Neuheit primär ausschlaggebend für den Erfolg am Markt?
  - Bisher liegt das Hauptgewicht der Förderentscheidung bei den Merkmalen Neuheit, Risiko und Technologiesprung auf technischer und Marktaussicht und Verwertung auf wirtschaftlicher Seite. Haben andere der untersuchten abhängigen Variablen für die Unterscheidung der verglichenen Stichproben mehr Bedeutung?
  - Anhand der Projektdaten sollen die Erfolgs- und Misserfolgskriterien herausgearbeitet werden. Reichen die immer wieder in die Diskussion gebrachten technischen Faktoren (ausreichende Projektausstattung, Projektabwicklung, ...) und wirtschaftlichen Faktoren (z.B. Markterfahrung, Verwertungskompetenz, finanzielle Leistungsfähigkeit,...) zur Erklärung von Verwertung, Marktaussichten für das Produkt und den Erfolg bzw. Misserfolg aus?
  - Reicht die Trennschärfe aus oder müssen weitere Parameter herangezogen werden, um die Qualität der Entscheidung zu verbessern?
  - Welche Weiterentwicklungen des Bewertungsschemas hinsichtlich höherer Trennschärfe, Früherkennung von Misserfolgen u.s.w. erscheinen erfolgversprechend?
- Bei der Untersuchung wurden verschiedenste statistische Verfahren, wie z.B. Faktoranalyse, Clusteranalyse, Diskriminanz- und Regressionsanalyse verwendet. Die hauptsächlichen Ergebnisse lassen sich kurz, wie folgt, zusammenfassen:
- Die Faktoranalyse zeigte, dass man insbesondere die sechs technischen Variablen nur schlecht verdichten kann. Das bedeutet, dass diese Variablen unabhängige Größen sind und wenig redundante Informationen enthalten. Die wirtschaftlichen und die firmenbezogenen Variablen sind im Gegensatz dazu jedoch auf jeweils eine Komponente zusammenfassbar.
- Mit Hilfe der Diskriminanzanalyse konnte gezeigt werden, dass Erfolgsabschätzungen aufgrund der Ex-ante-Evaluierung möglich sind. Vorhersagen der Verwertung und des technischen Erfolges sind in rd. 60 % der Fälle möglich. Das mit den Daten für 1999 erstellte Modell bringt gute Vorhersagewerte, auch für die Probemenge der Daten des Jahres 1998. Die Datenbank des FFF kann daher als Managementinformationssystem für ganze Jahrgänge und einzelne Projekte genutzt werden.
- Die korrekte Klassifizierung von 60 % der Projekte zeigt, dass ein wesentlicher Anteil an der wirtschaftlichen Verwertung und dem technischen Erfolg durch die Variablen des FFF erfasst wird. Die schwer fassbaren Einflussfaktoren wie Unternehmerpersönlichkeit und Marktzugang können als zusätzliche Erklärungsvariablen nicht ausgeschlossen werden.

Für den technischen Erfolg hat sich das Kriterium „Zweckmäßigkeit“ des Projektes als wesentlicher Erfolgsfaktor herausgestellt. Interessant ist, dass es keine bemerkenswerte Korrelation zwischen Neuheit, Risiko sowie Technologiesprung und der daraus resultierenden Verwertung gibt. Aufschlüsse ergaben sich auch bezüglich des Evaluierungsverhaltens. Insgesamt bietet das Monitoringsystem noch eine Fülle zusätzlicher auswertbarer Möglichkeiten, die auch für die Bildung von Performanceindikatoren herangezogen werden können.

### **Neues Projektbewertungs- und Monitoringsystem des FFF**

Die bisher vom FFF angewendeten Bewertungskriterien werden seit mehr als 10 Jahren publiziert. Seit 1996 werden den Antragstellern auch die Konzepte, die hinter den einzelnen Kriterien stehen, erläutert. Intern gab es beim FFF für alle Kriterien Skalierungshilfen. Durch den relativ kleinen Stab an Begutachtern war eine gleichbleibende Qualität gewährleistet. Im Rahmen eines EU-Projektes wurde ein Technology-Rating-Tool entwickelt. Alle Evaluatoren des FFF arbeiteten in der Experimentalphase mit diesem Tool. Die dabei verwendeten Konzepte zeigten sich als praktikabel und wirkungsvoll, so dass beschlossen wurde, diese in vergleichbarer Form auch bei FFF-Projektevaluierungen anzuwenden. Wertvolle Erkenntnisse für Verbesserungen ergaben sich auch aus der Analyse des Monitorings.

Das neue Bewertungssystem kann nunmehr folgend charakterisiert werden: Die Hauptkriterien blieben im wesentlichen unverändert. Innerhalb ihrer vier Gruppen haben sie unterschiedliche Gewichtungen, entsprechend der ihnen zugemessenen Bedeutung. Diese Gewichtungen sind noch nicht endgültig fixiert. Man hat damit jedoch die Möglichkeit, Prioritäten, entsprechend sich ändernden Anforderungen, akzentuiert neu zu setzen. Jeder Faktor hat auch die

Möglichkeit einer Nullbewertung. Diese wirkt als K.o.-Faktor für das gesamte Projekt. Auch schwächer gewichtete Faktoren können durch die Ablehnungsmöglichkeit stark werden.

Neu ist nun, dass die Hauptkriterien durch Subkriterien ergänzt werden. Diese Subfaktoren wurden im Rahmen einer achtmonatigen Erprobungsphase ausgefiltert. Sie werden jeweils durch eine Frage, die eine Bewertung herausfordert, definiert. Für jedes Subkriterium wurde eine Skalierung festgelegt (siehe dazu das folgende Beispiel).

### **Schützbarkeit der Projektidee**

Wie gut ist die Projektidee schützbar (z.B. Patent, Gebrauchsmuster)? Bestehen wirksame Barrieren und Abwehrstrategien für Nachahmer auch ohne gesetzliche Schutzmaßnahmen? Verletzt die Entwicklung bestehende Schutzrechte?

Die Subkriterien haben jedoch keine Gewichtung, weil ihre Bedeutung von Projekt zu Projekt wechseln kann. Auch Subkriterien können als K.O.-Faktoren wirken. Wenn es nicht möglich ist, für ein Subkriterium eine Bewertung abzugeben, wird es auch nicht bewertet. Hauptfaktoren müssen jedenfalls bewertet werden. Sollte hier eine Bewertung nicht sinnvoll möglich sein (z. B. Umweltrelevanz für Softwareprojekte), ist eine durchschnittliche Bewertung vorzunehmen.

Bei allen Kriterien mit höheren Gewichtungen gibt es keine Möglichkeit einer „durchschnittlichen“ Bewertung. Der Evaluator muss entscheiden, ob das Projekt eher unterdurchschnittlich oder eher überdurchschnittlich zu bewerten ist. Alle Bewertungen werden in der FFF-Datenbank gesammelt und in der gleichen Tiefe auch beim Monitoringprozess weiter verfolgt.

KO	Bestehende Schutzrechte werden eindeutig verletzt.
----	--

--	Es besteht die Gefahr einer Schutzrechtsverletzung. Der Antragsteller setzt keine Schutzmaßnahmen, führt keine Recherche durch, obwohl dies sehr wichtig wäre. Eine Nachahmung ist sehr wahrscheinlich.
-	Der Antragsteller setzt keine Schutzmaßnahmen, führt keine Recherche durch, obwohl dies sinnvoll wäre. Es besteht Nachahmungsgefahr.
+	Der Antragsteller hat eine Schutzstrategie. Schutzrechtsansprüche anderer erscheinen nicht verletzt. Eine Nachahmung ist auch ohne Schutz schwierig.
++	Konkrete sinnvolle Schutzmaßnahmen sind eingeleitet.

was installed. All parameters of the yearly monitoring are part of a data base. In his master thesis, Reinhard Zeilinger checked the possibilities to predict the chances of success of projects using only data from the FFF data base in a statistical model. The results were satisfying to some extent. During the last 8 months FFF has tested an improved tool for project assessment and monitoring. This tool is connected with the old scoring system but goes more into detail. The main indicators are defined by supporting factors.

Bei der geplanten Verknüpfung von Ex-ante-Bewertung, Monitoring und schließlich Ex-post-Evaluierung ergeben sich damit interessante und aufschlussreiche Möglichkeiten.

Das Bewertungssystem wird mit den dahinterstehenden Konzepten ausreichend publiziert werden. Damit soll die Transparenz der Entscheidungsprozesse beim FFF erhalten, aber auch die Qualität der Entscheidungsfindung weiterhin gesichert werden.

Ein weiterer Vorteil ist auch, dass Teile der Bewertung problemlos in das Technology Rating übergeleitet werden können.

**ABSTRACT**

FFF, the Austrian Industrial Research Promotion Fund uses as common practice a scoring system for project assessment. Based on this system, in 1997 a project monitoring with an enlarged set of indicators

**Autoren-Hinweis**

Klaus Schnitzer, Reinhard Zeilinger, beide: FFF-Forschungsförderungsfonds für die Gewerbliche Wirtschaft, Kärntnerstraße 21 –23, A-1015 Wien  
 Tel.: ++43 1/512 45 84-34  
 Fax: ++43 1/512 45 84-41  
 E-mail: klaus.schnitzer@fff.co.at;  
 reinhard.zeilinger@fff.co.at

## Anhang

ANTRAG		50 % PROJEKT-KOSTEN		NACH 1 JAHR	SPÄT. NACH 6 MONATEN		FORTSETZUNGS-ANTRAG	
BEWERTUNG		ZWISCHEN-BERICHT & KONTROLLE		ENDBERICHT TECHN. KONTROLLE	RECHNUNGSPRÜFUNG ENTLASTUNG		BEWERTUNG DER FORTSETZUNG	
1. (P)	Neuheit Risiko Zweckmäßigkeit Umweltrelevanz	.. .. .. ..	Stop der Projektförderung möglich	1. P	.. .. .. ..	Stop der Projektförderung möglich	1. P	
2. (T)	Technologiesprung Dynamik Durchführbarkeit	.. .. ..		2. T	.. .. .. ..		2. T	
3. (F)	Leistungsfähigkeit Dynamik	.. ..		Ü	.. .. .. ..		3. F	
4. (W)	Markterfahrung Marktaussichten Verwertung Volkswirtschaft Soziale Aspekte	.. .. .. .. ..		3. F 4. W	.. .. .. .. .. ..		4. W	
		Besuch neuer Firmen und schwieriger Projekte		Qualität und Ergebnisse Verwertung	◆ ◆		Projekt-abwicklung Rechnungswesen Ausstattung Personal	◆ ◆ ◆ ◆

.. ) Teil der Datenbank

		<b>Wirtschaftliche Projektbewertung *)</b>					
Module	Wirtschaftliche Situation		KO	--	-	+	++
1	<b>Finanzielle Leistungsfähigkeit</b>						
	1.1. Umsatz						
	1.2. Innenfinanzierungspotenzial/Cash Flow						
	1.3. Eigenkapitalausstattung						
	1.4. Weitere Kapitalzuführung						
	1.5. Finanzierungsmöglichkeit des FFF-Projektes						
	1.6. Umsetzung der gesamten Innovation						
2	<b>Management</b>						
	2.1. Qualifikation und Praxiserfahrung						
	2.2. Controllinginstrumente/Business Plan						
	2.3. Strategie und Organisation						
	2.4. Innovationsorientierung,- verhalten						
	2.5. Stellenwert von F&E (Organisationsform)						
	2.6. Kooperationsbereitschaft						
3	<b>Markterfahrung</b>						
	3.1. Zielgruppe						
	3.2. Mitbewerber						
	3.3. Umsätze im Projektbereich						
	3.4. Synergien						
4	<b>Marktaussichten</b>						
	4.1. Marktpotential						
	4.2. Neues Marktsegment						
	4.3. Nutzen für den Kunden						



	4.4.	Konkurrenzfähiger Preis						
	4.5.	Mitbewerbersituation						
5		Verwertung						
	5.1.	Strategie für die Produkteinführung/Marketing						
	5.2.	Markteintrittsbarrieren						
	5.3.	Verhandlungsposition						
	5.4.	Marktgerechte Produktionsmöglichkeit/Kapazitäten						
	5.5.	Service						
6		Externe Effekte						
	6.1.	Arbeitsplätze erhalten bzw. neu schaffen						
	6.2.	Positive Wirkung auf die Leistungsbilanz						
	6.3.	Knowledge-flow						
	6.4.	Kooperationsansatz						
	6.5.	Regionalpolitische Aspekte						
7		Soziale Aspekte						
	7.1.	Wertehaltung der Gesellschaft						
	7.2.	Lebensqualität der Benutzer						
	7.3.	Qualifikationsniveau						
	7.4.	Einbindung der Mitarbeiter						
	7.5.	Arbeits- u. sozialrechtlichen Normen						
	7.6.	Verbesserung der Arbeitsbedingungen						

**\*)Gewichtungen noch in Diskussion**

Technische Projektbewertung *)							
Module	Technische Bewertung		KO	--	-	+	++
1	Technologische Neuheit						
	1.1.	Neuheit gegenüber Stand der Technik					
	1.2.	Schützbarkeit der Projektidee					
	1.3.	Wettbewerbsvorteil					
	1.4.	Zukunftspotenzial					
	1.5.	Vorbildcharakter für die Branche					
2	Schwierigkeit der Entwicklung						
	2.1.	Technische Probleme					
	2.2.	Einflussfaktoren, die die technische Zielsetzung gefährden					
	2.3.	Komplexität der Aufgabenstellung					
	2.4.	Versuchsaufwand					
3	Nutzen						
	3.1.	Qualität der Problemlösung					
	3.2.	Nutzen für den Anwender					
	3.3.	Anwendungsbereich der Technologie					
	3.4.	Gesetzliche Auswirkungen					
4	Umwelt						
	4.1.	Umweltbeeinflussung					
	4.2.	Ressourcenverbrauch					
	4.3.	Energieverbrauch					
5	Know-How-Zuwachs						
	5.1.	Wissenszuwachs durch neue Technologien					
	5.2.	Wissenszuwachs durch neue Anwendungsgebiete					

	5.3.	High-Tech Equipment						
	5.4.	Wissenstransfer durch Kooperation						
6		F & E - Dynamik						
	6.1.	Bisherige F&E-Aktivitäten						
	6.2.	F&E - Intensivierung durch Projekt						
	6.3.	Firmenstrategie und F&E						
7		Durchführbarkeit						
	7.1.	Qualifikation und Motivation der Mitarbeiter						
	7.2.	Personalressourcen für F&E						
	7.3.	Projektmanagement und Entwicklungsplanung						
	7.4.	Technische Ausstattung						
	7.5.	Produktions- und Umsetzungskompetenz						

\*)Gewichtungen noch in Diskussion

## Ein Evaluationsmodell für JOANNEUM RESEARCH

**Dorothea Sturn**

JOANNEUM RESEARCH, eine der größten außeruniversitären Forschungseinrichtungen Österreichs, entwickelt derzeit ein umfassendes Evaluationsmodell. Gerade angesichts dessen, dass die Evaluation von Programmen und Projekten zwar gängige Praxis ist, die von Institutionen jedoch noch in den Kinderschuhen steckt, ist diese Initiative als ein innovatives und ambitioniertes Vorhaben zu bewerten. Ausgangspunkt war die im Unternehmenskonzept vorgesehene Strategie, dass der Prozess der indirekten Evaluierung durch die Auftrags- und Förderungsgeber ergänzt werden muss durch ein an die Bedürfnisse des Unternehmens angepasstes Instrumentarium der Evaluierung, das neben Elementen der Selbstbeurteilung auch eine Fremdbeurteilung enthält. Im Zuge einer längeren Diskussion im Wissenschaftlichen Beirat der JOANNEUM RESEARCH wurde mit der schrittweisen Entwicklung eines Evaluationskonzepts begonnen.

Erste konzeptionelle Überlegungen erfolgten zunächst intern und wurden in einem zweiten Schritt der kritischen Analyse eines externen Evaluationsexperten unterzogen. Eines der größten Institute der JOANNEUM RESEARCH erklärte sich für eine Pilot-evaluation auf Grundlage des ausgearbeiteten Konzepts bereit. Dieses Vorhaben wird in Kürze abgeschlossen sein, die Erfahrungen sollen dann in ein Modell zur Evaluation des gesamten Unternehmens einfließen. Dies hat den Vorteil, dass mögliche Schwächen des Konzepts noch korrigiert werden können. Für die Pilotevaluation wurde die folgende grob dargestellte Vorgangsweise gewählt:

### 1. Grundsätze der Evaluation außeruniversitärer, wirtschaftsbezogener Forschung

Verfolgt Forschung bestimmte wirtschaftliche und gesellschaftliche Interessen und ist in diesem Sinne zielgerichtet, so sollte eine Evaluation dies in ihrem Design berücksichtigen. Im Gegensatz zu den meisten Forschungsaktivitäten an den Universitäten entstehen bei der außeruniversitären, wirtschaftsbezogenen Forschung Probleme und Fragen nicht nur innerhalb eines akademischen Diskussionszusammenhanges sondern vor allem im Kontext der Anwendung. Erfolg und Ergebnis hängt davon ab, in welcher Weise es den ForscherInnen gelingt, ihre Arbeit mit zentralen wirtschaftlichen und gesellschaftlichen Problemstellungen zu verbinden und wie ihr Umfeld die Ergebnisse akzeptiert und konsolidiert.

Entsprechend kann eine Evaluation nicht nur nach den internen Kriterien der „scientific community“ stattfinden, sie muss auch Elemente einer sozialen und ökonomischen Bewertung enthalten.

### 2. Evaluationsmethode

Insgesamt sollte die Evaluation nicht nur Vergangenes möglichst neutral bewerten, sondern vor allem auch strategische Hilfestellungen für die Zukunft geben (eher formative, weniger summative Evaluation). Um dieser Erwartung gerecht zu werden, gleichzeitig die genannten Grundsätze zu berücksichtigen und zudem den Aufwand des Vorgehens in Grenzen zu halten wurde wie folgt vorgegangen:

#### Modul 1: Indikatoren-gestützte Bewertung

Hier werden die Leistungen des Instituts mit - teilweise vorhandenen, teilweise zu diesem Zweck zu erhebenden - quantitativen Indikatoren möglichst "objektiv" bewertet.

#### Modul 2: Peer Review Bewertung

Eine Gruppe von Fachexperten erstellen schriftliche Bewertungen des Instituts auf Grundlage von eigens zu diesem Zweck zusammengestellten Informationen

und Materialien sowie eines halbstrukturierten Bewertungsbogens. Neben der schriftlichen Bewertung erfolgt auch eine Begehung des Instituts („visiting committee“), die ein fachunspezifischer externer Evaluationsexperte moderiert. Das Ergebnis der Begehung ist eine gemeinsame Stellungnahme der Expertengruppe.

### **Modul 3: Analyse der Kunden**

Hier wird eine Intensivbefragung ausgewählter, strategischer Kunden durchgeführt. 10 Kunden (6 öffentliche, 4 private) werden auf der Grundlage eines halbstrukturierten Leitfadens interviewt.

Der Vorteil dieses Methodenmix liegt darin, dass bekannte und vielfach dokumentierte Schwächen der einzelnen Methoden durch ergänzende Informationen der jeweils anderen Module relativiert werden können. So gibt es Evidenz dafür, dass Expertenbewertungen aufgrund ihres subjektiven Ansatzes bestimmte Forschungsrichtungen und Personen diskriminieren, andererseits der Einsatz quantitativer Indikatoren radikale und interdisziplinäre Innovationen benachteiligt, zu oberflächlichen Bewertungen und zu rigiden Handlungsschemata führt. Durch Einbezug der Kunden soll der Tatsache Rechnung getragen werden, dass die Forschungsaktivitäten der JOANNEUM RESEARCH im Kontext der Anwendung stehen und daher die Gruppe jener, die an einer Bewertung und Qualitätsdefinition zu beteiligen sind, legitimerweise um die Kunden zu erweitern ist.

Nach Abschluss der Arbeiten an den drei Modulen werden die Ergebnisse intern ausgewertet, entsprechende Maßnahmen formuliert und gemeinsam mit dem Institut ein Plan für die Umsetzung der Evaluationsergebnisse erstellt. Gleichzeitig erfolgt die Formulierung des Evaluationskonzepts für das gesamte Unternehmen, wobei sowohl Zielsystem als auch Methodik und Inhalte der Evaluation aufgrund der Erfahrungen in der Pilotphase mögliche Überarbeitungen erfahren.

### **Autoren-Hinweis**

Dorothea Sturn, Joanneum Research, Institut für Technologie- und Regionalpolitik, Wiedner Hauptstrasse 76, A-1040 Wien  
Tel ++43 1/581 75 20-2814  
E-mail: dorothea.sturn@joanneum.ac.at

## Bewertungen von Programmen der EU-Regionalpolitik – Verwertbarkeit bei Um- und Neuprogrammierungen in Österreich

**Rudolf Schicker**

Die EU-Regionalpolitik als Teil der Strukturpolitik der Europäischen Union unterliegt seit 1994 der Verpflichtung, Programme Ex ante, begleitend und Ex post zu bewerten (Art. 6 der VO(EWG) 2081/93). Prüfungsmaßstab ist die Erfüllung der sozialen und regionalen Kohäsion.

Österreich hat 1995 zwölf Regionalprogramme (Ziel 1, 2, 5b) und 23 Programme aus Gemeinschaftsinitiativen für die Laufzeit bis 1999 eingereicht, die alle einer von der EK beauftragten Ex ante-Evaluierung unterzogen worden sind. Die Auswirkungen auf die letztlich von der Europäische Kommission genehmigten Programme waren in folgenden Programmteilen sichtbar:

- Reduzierung der Programmschwerpunkte auf eine überschaubare Anzahl
- Festlegung der Bewertungsindikatoren und deren Aufgliederung in Ergebnis-, Wirkungs- und Outputindikatoren und Quantifizierung von Zielwerten
- Technische Details der Programmabwicklung

Mit den Programmergebnissen zu Jahresende 1996 waren in allen Mitgliedstaaten die Programme zu evaluieren. Die Zwischenevaluierung oblag den Mitgliedstaaten im Einvernehmen mit der Europäische Kommission. In Österreich bestand die Problematik der kurzen faktischen Programmlaufzeit. Die Beauftragung erfolgte bei den meisten Programmen im Wege der ÖROK. Ziel war es, nicht

nur den Stand der Umsetzung zu dokumentieren sondern v.a. die Schwachstellen der Umsetzungsstrukturen im Vergleich mit anderen Programmen herauszuarbeiten. Dieser Vergleich der österreichischen Umsetzungsstrukturen ist bei den Ziel 1, 2, Retex, Rechar, Resider und INTERREG-Programmen auch gelungen und hat zu Adaptierungen in den Programmen geführt. Die GD Landwirtschaft hat auf Einzelevaluierungen der 5b-Programme bestanden, was die Vergleichbarkeit deutlich erschwerte.

Die **Ergebnisse der Zwischenevaluierungen** ergaben zusammenfassend, dass

- die Bewertungsindikatoren vielfach nicht aussagekräftig waren
- quantitative Indikatoren viel zu vorsichtig angesetzt waren (z.B. Beschäftigungswirkungen in Ziel 2-Gebieten)
- die Vielzahl der eingesetzten Förderungsinstrumente die Programmumsetzung erschwerten
- die oftmaligen Abänderungen von Führungsrichtlinien großen organisatorischen Aufwand verursachten
- die Auswirkungen von konjunkturellen Sondersituationen (z.B. Einbrüche in der Tourismuswirtschaft) zu wenig antizipiert worden waren
- innovative Projekte offenbar eine sehr lange Anlaufphase haben.

Daraus konnte folgende Verbesserungen in den Programmen vorgenommen werden:

- Anpassung der Zielindikatoren an realistische Werte

- Ergänzung der Indikatoren um darstellbare Bewertungsgrößen
- Reduzierung der Programmänderungen auf max. 2 pro Jahr und Programm
- Umschichtungen von Finanzmitteln

Problematisch dabei war, dass die Programmierfordernisse (z.B. keine Veränderung der Förderungsrichtlinien) oft aus politischen Gründen missachtet wurden und die Umschichtung der Finanzmittel zu besser nachgefragten Programmteilen entweder aus Ressort- oder föderalen Egoismen nicht oder sehr spät erfolgte oder aus Schutz für innovative Vorhaben nicht vorgenommen wurde.

Die ex post-Bewertung der Programme aus der Periode 1994/5- 1999 ist dem Streit über die Finanzierung zwischen den Mitgliedstaaten und der EK Großteils zum Opfer gefallen. Die GD Regio führt lediglich thematische oder Stichproben-Bewertungen durch. Die GD Agri hat auf einer Ex post-Bewertung bestanden, die für die 5b-Programme gerade erst angelaufen ist. Die Ergebnisse werden aber nahezu keine Bedeutung haben, da sie für die neuen Programme keine Relevanz haben.

In der aktuellen Strukturfonds-Periode 2000 – 2006 ist die ex ante-Evaluierung Bestandteil der Programmierung und fließt in den Genehmigungsprozess der Programme durch die EK ein. Die Verantwortung liegt beim Mitgliedsstaat. Zur Programmhälftezeit ist eine Zwischenevaluierung zu erstellen, welche dann Grundlage der Nachjustierung der Programme sein muss. Dafür sind auch für jedes Programm Kernindikatoren zu verwenden, die einen europaweiten Vergleich der Programme ermöglichen soll. Die Daten dafür werden aus den offiziellen Statistiken und aus den Monitoring-Daten gezogen. Das Monitoring ist entsprechend aufgebaut.

Auf die neuen Programme 2000 – 2006 haben die Ergebnisse der Zwischenevaluierung der alten Periode insofern Einfluss gehabt, als

- die Indikatoren tatsächlich darstellbar sein werden
- die Zielwerte realistischer gewählt wurden
- die involvierten Förderungsstellen reduziert wurden
- eigene schwerpunktverantwortliche Förderstellen installiert wurden
- innovative Projekte auch den nötigen begleitenden Support aus der Technischen Hilfe erhalten können.

#### **Autoren-Hinweis**

Rudolf Schicker, Österreichische  
Raumordnungskonferenz (ÖROK),  
Hohenstaufengasse 3 , A-1010 Wien,  
Tel ++43 1/5353444  
E-mail: oerok@oerok.gv.at

## Konferenzberichte

### Konferenz „The Impact on Academic Institutions of Research Evaluation Systems“

**Lissabon 23. / 24. 6. 2000**

Ein Tagungsband wird erscheinen; anwesend waren etwa 130 Teilnehmer, vorwiegend Universitätsvertreter. Als Veranstalter traten die Konföderation der Rektorenkonferenzen der EU-Staaten die EU-Kommission und mehrere portugiesische Institutionen auf. Aus Österreich waren neben dem Berichterstatter v.a. mehrere hochrangige Universitätsvertreter und ein Mitarbeiter der Rektorenkonferenz anwesend.

Kernthema der Konferenz bildete die Frage der Wechselwirkungen zwischen den sich wandelnden Universitätssystemen der europäischen Staaten und Evaluierungsmechanismen, die eben diese akademische Forschung und ihre Rahmenbedingungen bewerten sollen. Dabei ist klarerweise von unterschiedlich ausgestalteten rechtlichen, organisatorischen und kulturellen Eckpunkten in den verschiedenen Ländern auszugehen.

Die Rektorenkonferenzen der einzelnen Länder, ihre Dachorganisation(en) und natürlich einzelne initiative Universitäten bemühen sich aktiv, der Herausforderung gerecht zu werden, die in einem miteinander verknüpften stärkeren Reformdruck und erhöhten Evaluierungsanforderungen liegen. Im Rahmen einer mehrjährigen analytischen und best practice-Studie durch das CHER wurden mehrere Fallstudien näher angesehen und kritische Fragenkomplexe beleuchtet. In der Konferenz erfolgte dies durch drei Elemente:

- Tomlin stellte die bisherigen Studienergebnisse vor. Ziel der Arbeit ist es, den Status der universitären Forschungsevaluierung in Europa zu erheben und unterschiedliche Ansätze zu vergleichen. Ein besonders wichtiger Punkt dabei ist die Wirkung (das passendere englische Wort "im-

pact" bezeichnet übrigens auch den Einschlag einer Bombe) von Evaluierungen auf das universitäre Getriebe und die jeweilige akademische Kultur in einem Land. In einem weiteren Schritt geht es darum, best practices zu analysieren und Wege für ihre Verbreitung darzustellen. Dabei geht es vor allem um den Einfluss auf akademische Karrieren, um Erkennen und Anerkennen von Forschungsexzellenz, um damit zusammenhängende Mittelallokationen, um das gewollte/ungewollte Auslösen von Wettbewerben, um Schaffung / Legitimierung von kräftigen universitären Entscheidungsstrukturen und damit verbunden um das Entstehen neuer universitärer Managementfunktionen. Die Universität und ihre Träger / stakeholders müssen sich eine Strategie zurechtlegen und aktiv und unter Zuhilfenahme von geeigneten Evaluierungsmechanismen planend vorgehen. Dass dies in einem jedenfalls für unsere Begriffe neoliberalen angelsächsischen Umfeld zu Auswüchsen führen kann, versteht sich offenbar ebenso von selbst wie das Zögern mitteleuropäischer Humboldt-Hüter.

- In vier parallelen Arbeitsgruppen wurden Fragen angesprochen, die im Zusammenhang mit der Evaluierung universitärer Forschung von Bedeutung sind. Erstens können bei mangelnder Beachtung Fragen der Autonomie der Universität (aber auch der verbrieften Freiheiten der an der Universität Forschenden) durch Evaluierungsprozesse in Frage gestellt werden. Vielfach handelt es sich freilich um eine Schutzbehauptung zur Abwehr von Überprüfungen von Leistung; mögliche Gefahren lassen sich minimieren, wenn die Universität selbst eine aktive Rolle spielt und das Evaluieren nicht passiv dem Staat überlässt. Zweitens ist die Wahl der Prozesse und Methoden entscheidend: Eine Reihe von Garantien (wie Unabhängigkeit der Evaluatoren oder



Transparenz der Prozesse) und Methoden (Peer Review, Bibliometrie) stehen erprobt zur Verfügung. Drittens wurde in einer zusammengefassten Gruppe die Frage der institutionellen Perspektiven und der universitären Strategien behandelt: Aktive Schwerpunktsetzung auf universitärer Ebene ja oder nein, Verhinderung des Auseinanderdriftens von Forschung und Lehre, Rolle externer Faktoren, Rolle der Financiers, Schwierigkeiten für "Voll"-Universitäten bildeten Kernpunkte der Diskussion. Viertens schließlich geht es darum, wie die Universität mit den Evaluierungsergebnissen umgeht, was wiederum Fragen der Autonomie, der Methoden und der Strategie einschließt usw.

- In einer zweiten Arbeitsgruppenrunde wurden fünf Fallstudien vorgestellt; neben den Universitäten von Helsinki, Aveiro, Pavia und Antwerpen wurde auch die TU Graz untersucht. Schreiterer und Kahlert präsentierten dazu als Hintergrund die zunehmende Autonomie des universitären Sektors und die etwas traurige Vorgeschichte versandeter heimischer Disziplinenevaluierungen. Hier drehte sich die Diskussion wieder stark darum, wer die governance über Evaluierungen haben soll und ob sie als regelmäßiges oder als anlassbezogenes Instrument zum Einsatz kommen sollen. Über alle Fallstudien fallen die großen Unterschiede in den nationalen Rahmenbedingungen auf: manche Regierungen geben strikte Evaluierungsbedingungen vor, andere geben den Universitäten die Verpflichtung, selbst

für Evaluierungen zu sorgen, wieder andere üben sich im laissez-faire.<sup>1</sup>

Rund um diese zentrale Debatte wurden einerseits der Ansatz zum Europäischen Forschungsraum inkl. europäischer Arbeitsteilung und Vergleichbarkeit in der Forschung (Routti) diskutiert sowie Evaluierungsmethoden und Wahl der Indikatoren durch einen Mix quantitativer und qualitativer Ansätze vorgestellt (Magalhaes). Methodische Ratschläge kommen auch deshalb stark "von außen", da die Universitäten Spieler in einer übernationalen Arena sind (siehe Europäischer Forschungsraum, Routti) und weil andere Einrichtungen und Gestaltungsebenen mitunter in der Entwicklung geeigneter Methoden schon mehr Erfahrungen sammeln konnten als dies dem Universitätssektor bislang möglich war (*Barabaschi*). Dazu wurden zwei Bereiche besonders vorgestellt:

Zum einen ging es bei "Meta evaluation of impact assessment approaches" (Georghiou) darum, dass vor allem für Programmevaluierungen doch einige robuste - letztlich meist qualitative, ja geradezu "soziale" - Wirkungsindikatoren zur Verfügung stehen. Hier können gerade Lehren von Kooperationsprogrammen Wissenschaft-Wirtschaft auf die dahinterstehenden akademischen Institutionen zurückgespiegelt werden. Zum anderen können Evaluierungen aus dem Bereich der regionalen Struktur- und Innovationspolitik (Tsipouri) reiche Erfahrungen über die Wirkung von Infrastrukturen auf die Entwicklung des jeweiligen Umfeldes geschöpft werden.

---

<sup>1</sup> Was bedeutet das angesichts der Initiative zum Europäischen Forschungsraum und zur Ausgestaltung europäischer Forschungsinfrastrukturen? Wohl einerseits eine nicht zu übersehende Erschwernis für gesamteuropäische Initiativen; andererseits kann mit der Forschungsraum-Initiative auch ein sich doch abzeichnender Trend zu mehr Universitätsautonomie und einheitlicheren, verbindlicheren Evaluierungsmechanismen verstärkt werden.

Ein weiterer Diskussionspunkt betraf die Frage der demokratischen Rechenschaftslegung von Institutionen: Hier ist eine weitere Legitimation für Evaluierungen von (meist teuren) öffentlichen Forschungsinfrastrukturen gegeben.

Die gesamte Konferenz wurde durchzogen vom Bewusstsein der unterschiedlichen Evaluierungskulturen vor dem Hintergrund unterschiedlicher nationaler Regulierungssysteme. Als ein gemeinsamer Trend konnte freilich festgestellt werden, dass mehr Eigenverantwortung und etwas erhöhter Außendruck auf die Universitäten in ganz Europa zukommt (oder schon zugekommen ist). Damit gewinnen Evaluierungen an Bedeutung, ebenso wie die Frage, wessen Instrument die Evaluierung ist: Wenig verwunderlich überwog ein starkes Plädoyer dafür, dass die Universitäten selbst und aktiv das akademisch-institutionelle Evaluierungsgeschehen in die Hand nehmen (Tomlin), bevor es der möglicherweise nicht immer wohlmeinende Staat monopolisierend aufgreift und die bisher häufige direkte Lenkung durch ein subtileres Instrument ersetzt.

### **Konferenz „Strategic Evaluation: the governance of the research systems in the EU“**

**Lyon 9. / 10.10. 2000**

Der Berichterstatter war der einzige österreichische Teilnehmer dieser Veranstaltung. Unter den ca. 80 Teilnehmern waren einige sehr gut ausgewiesene EvaluierungsforscherInnen und viele höherrangige Policy Maker. Als Organisatoren traten das frz. Forschungsministerium, die Ecole Normale Supérieure Lyon, das Observatoire des Sciences et des Techniques (OST) und die Europäische Kommission auf. Das Thema der Konferenz umfasste Fragen, die über Projekt-, Programm- und Institutionenevaluierungen vielfach hinausgingen. Vielmehr stand quasi ein

Stockwerk darüber die Rolle von Evaluierungen als politisches Steuerungsinstrument im Vordergrund. Das Programm enthielt dazu drei Schwerpunkte: Societal Relevance (=Science for Policy), Performance von FTE-Systemen und Governance derselben. Der Bericht gliedert sich dementsprechend; einige allgemeine Aussagen sind vorangestellt.

#### *Allgemeines*

Ausgangspunkt ist die Erhöhung der Legitimation von Forschung. Doch wie macht man eine solche Legitimation sichtbar, wenn die Forschung nicht mehr automatisch und überzeugend (nur) als per se etwas Gutes dargestellt werden kann?<sup>2</sup> Wenn Forschung um einer Politik, um eines wirtschaftlichen und / oder sachpolitischen Ziels, bzw. zur Lösung eines Problems betrieben werden soll, dann genügen Kriterien wie Exzellenz alleine nicht mehr. Das bedeutet, dass Evaluierungen von der Reichweite der Fragestellungen her und vom Adressatenkreis breiter anzulegen sind. "Die Gesellschaft", wie auch immer sie sich im konkreten Fall zusammensetzt, will Antworten darauf, ob Forschungsprogramme, -institutionen etc. in der Lage sind, das definierte Problem auch zu lösen. Wo es etwa um einen neuen Impfstoff geht, ist eine solche Frage noch halbwegs leicht zu beantworten. Wo Forschung Teil eines Maßnahmenbündels zur Bearbeitung komplexer sozioökonomischer Zusammenhänge mit vielen Akteuren ist, dort wird die Feststellung des Beitrages und der Wirksamkeit dieser Forschung schwierig. Noch einmal komplexer wird es, wenn nicht nur Forschungsarbeiten, sondern ganze Politikfelder zu evaluieren sind.

---

<sup>2</sup> Damit ist auch gemeint: als "endless frontier", die auszudehnen nicht hinterfragt wird und werden soll; und das kennt man aus vielen Systemen: Werden die Ziele einer öffentlichen Aufgabe nicht auch öffentlich diskutiert, so tendieren die Insider auch dazu, die Instrumente, Wege und Nebenwirkungen dem öffentlichen Raum zu entziehen und sie mit Tabus zu belegen.

Da Evaluierungen Spiegelbilder ihrer Gegenstände sind, hat gerade die zunehmende Beladung der FTE-Politiken mit sozialen, wirtschaftlichen, sonstig politischen Zielen und Aufgabenstellungen eine zunehmende Anreicherung und Erhöhung der Anforderungen der dazugehörigen Evaluierungen zur Folge. Konkret heißt das: Wenn im Rahmenprogramm der Europäische Mehrwert zu beachten ist, muss es auch Evaluierungskriterien dafür geben - so unfassbar dies auch manchmal zu sein scheint. Wenn das Rahmenprogramm ein Instrument auch der Verkehrs-, Gesundheits- oder sonst einer Politik ist, dann muss diese zusätzliche Politikebene sich auch in den Evaluierungen wiederfinden. Auf der Projekt- und Programmebene sind es Kriterien, die die konkrete Zielerfüllung etc. zu messen versuchen. Auf der *Ebene der Politiken* sind Fragen der grundsätzlichen Eignung des Herangehens, der Aufstellung der Akteure und schließlich auch der Fähigkeit der Einlösung genuin politischer Ziele durch ein Zusammenwirken von Instrumenten und Akteuren zu beantworten. Wichtig in diesem Zusammenhang ist die Feststellung, dass Evaluierung von Politikfeldern und -mechanismen auch ein anderes Instrumentarium erfordern als die Evaluierung von Projekten und Programmen: Einbeziehung breiterer Kreise wie Nutzer (→ *Societal Relevance*), Bewertungsmechanismen für Leistung und Zusammenspiel von Institutionen (→ *Research Performance*) und solche für die Steuern und ihre Mechanismen (→ *Governance*).

Ein Anwendungsgebiet, das sich auf europäischer Ebene expliziter herausbildet, ist der *Europäische Forschungsraum*, in dem es nicht nur um bessere Koordination von Projekten und Programmen, sondern auch von Politiken geht: Koordinierung von 15+1 FTE-Politiken; zugleich inhaltlich Koordinierung der FTE-Politik(en) mit anderen Politikfeldern (eben Gesundheit, Verkehr, Umwelt etc.) Letzteres geschieht auch unabhängig vom Europäischen Forschungs-

raum auf nationaler Ebene, auch im 5. Rahmenprogramm etc. Will man hinterher wissen, ob diese so anspruchsvollen kombinierten Operationen auch den erwarteten Terraingewinn gebracht haben, muss man sich schon vorher darüber klar sein, welche Parameter, Messgrößen, Kriterien man anwenden will und wer auf welcher Ebene was betrachtet.<sup>3</sup>

Für die nationalen Mitgestalter des Europäischen Forschungsraumes - und ebendieser ist nicht nur (nicht so sehr) eine Unionsangelegenheit, sondern stark auch eine (Zueinander-) Entwicklung der nationalen FTE-Politiken - bedeutet dies, Orientierung auch bei Evaluierungsfragen zu schaffen. Man muss wissen, wie man Aktivitäten im Forschungsraum mittels Evaluierungen bewertet und dieser Wissensbildungsprozess sollte nicht in 15+1 verschiedene Richtungen laufen.

#### *Societal Relevance of Research*

Der erste der drei Schwerpunkte der Konferenz ging vom "Science for Policy" - Ansatz aus: Zunehmend wird (nicht nur) auf europäischer Ebene versucht, Forschung bewusst zur Lösung von Problemen in anderen gesellschaftlich wichtigen Politikfeldern einzusetzen. Das heißt: noch mehr Programmcharakter, Interdisziplinarität, stärkere Frage nach Nützlichkeit (eben der societal relevance). Evaluierungen müssen darauf reagieren und entsprechende Fragen stellen.

Eine wichtige Voraussetzung ist der Versuch, mit Evaluierungen und verwandten Instrumenten nicht nur die unmittelbaren technischen Fragen zu erfassen. Ex ante heißt das etwa, sozio-ökonomische Bewertungskriterien einzuführen und auch durchaus

---

<sup>3</sup> Eine Illustration sind die jetzt schon sehr komplizierten integrierten Evaluierungs- und Monitoringabläufe im 5. Rahmenprogramm auf Projekt-, Programm- und Politikebene; letztere Ebene wird etwa durch das Fünfjahresassessment repräsentiert.

bei größeren Entscheidungen nicht nur die drei spezialisiertesten Spezialisten zu fragen. Ex post kann man darunter Versuche verstehen, die Wirkung nicht nur hinsichtlich einer konkreten Produktverbesserung darzustellen, sondern auf ein Segment der Gesellschaft, auf ein umfassenderes Problem einzugehen. Hier wird es schnell beliebig und bei der Evaluierung kleinerer Vorhaben unbezahlbar. Das Erzählen plastischer Geschichten durch Evaluatoren ist ein Schritt zu größerer Einbeziehung von Öffentlichkeiten, aber eben auch zweischneidig (*Rip*). Breitere qualitative Ansätze werden empfohlen und auch lernende Netzwerke und nicht zuletzt eine enge Verbindung von FTE-Management und Evaluierung (*Molas-Gallart*). Mit klassischen Fragebögen (*Manglis*) und vielen quantitativen Ansätzen kann man nur schwer solche Impacts messen - damit wird die Sache auch wieder teuer. Erfolgversprechend scheint eine Verbreiterung der Fragestellungen um gesellschaftliche Abschätzungen vor allem bei der Ex ante Evaluierung größerer Vorhaben (*Luukkonen*). Messungen von Öffentlichkeitswirksamkeit können durchaus in einfachen Dingen liegen: So analysiert der Wellcome Trust,<sup>4</sup> der größte britische private FTE-Förderer seinen societal impact auch durch systematische Durchsicht von Tageszeitungen: Alle Artikel über medizinische FTE-Ergebnisse werden danach ausgewertet, ob sie von Wellcome-geförderten WissenschaftlerInnen stammen (*Lewison*). Eine solche Stiftung bekommt den societal impact übrigens mitunter frei Haus geliefert - etwa durch Aktivistengruppen, wenn Projekte mit Tierversuchen gefördert werden.

Insgesamt geht es hin zu mehr *gesellschaftlicher Relevanz*, Systemanalysen, Ex ante Bewertungen und Prozessevaluierungen (*Smits*); Forderungen

nach mehr Anbindung an die Politikplanung und zu ähnlicheren Methoden und Ansätzen quer über Europa. Schließlich wird nicht nur von Smits darauf hingewiesen, wie stark sich die FTE-Szene in Richtung Projekte, Programme und Institutionen bewegt, die von einem "Multi-Akteurs - Multi-Maßnahmen - Ansatz" geprägt sind: Wieder eine neue Evaluierungsherausforderung.<sup>5</sup>

#### *Evaluierung der FTE-Performance*

Performance Evaluation ist natürlich auf allen Ebenen (Projekte - Programme - Institutionen - Politiken) zuhause, hier geht es wieder um die beiden letzterwähnten.

Ein Ausgangspunkt hier ist - aus europäischer Perspektive - das Vorhandensein völlig unterschiedlicher *Evaluierungskulturen* in den einzelnen Himmelsrichtungen und sogar in den einzelnen Mitgliedsstaaten. Das garantiert zusammen mit den unterschiedlichen Institutionensystemen bereits eine sehr spannende vielstimmige Diskussion, etwa um Instrumente und Indikatoren für den schon erwähnten Europäischen Forschungsraum. Zugleich findet mit der stärkeren Globalisierung der FTE-Akteure und -Felder auch eine solche der Evaluatoren und damit auch eine breitere Instrumentendiskussion statt. Dabei steht die Performance nicht nur von Projekten und Programmen, sondern von Infrastrukturen und Netzwerken sowie von ganzen Systemen als Aufgabe im Vordergrund und eine Antwort von der Evaluierungsseite ist erst im Werden (*Georghiou*). Da dies nicht nur anspruchsvoll ist, sondern dem Staat - wenn er selbst als Evaluator auftritt - auch gerne eine hidden agenda wie etwa Sparen vorgeworfen wird, er somit ein Glaubwürdigkeitsproblem aufweisen kann, stehen "Selbstverwaltungsmodelle" wie die holländische

<sup>4</sup> "With an asset base of £13 billion and an estimated expenditure in 1999/2000 of some £600 million, the Wellcome Trust is the world's largest medical research charity". So ist das nämlich. Siehe [www.wellcome.ac.uk](http://www.wellcome.ac.uk)

<sup>5</sup> Dazu wird es demnächst Publikationen auch zur österreichischen Landschaft geben.

Universitäts-evaluierung durch ihre gemeinsame Evaluierungsorganisation VSNU in einem hohen Ansehen (Smits).

Der Bogen spannt sich dabei von Accountability-Ansätzen (wie viel Output um so und so viel öffentliches Geld; v.a. im angelsächsischen Raum) über Evaluierungen als bewusstes Politikinstrument beim Institutionenumbau in den mittel- und osteuropäischen Beitrittsländern (*Kroo*), den umfangreichen periodischen Hinterfragungen aller US-amerikanischen Agencies durch den Government Performance and Results Act (GPRA, mit Evaluierung strategischer, managementbezogener und performanceorientierter Ziele; siehe auch mein Lissabon Bericht vom Sommer d.J.) bis hin zur Verknüpfung der Evaluierungen mit strategischer Planung (auch GPRA, VSNU und andere). Instrumente wie das laufende Science - Industry Benchmarking der OECD (*Barré für J. Guinet*) sollen helfen, ein Bild und bessere Indikatoren von Performance nicht nur in einem Teilsystem wie "Universitäten", sondern insgesamt in einem Innovationssystem zu schaffen.

Für alle, die all die erwähnten Anforderungen nervös machen: "Man stelle sich eine Welt ohne Evaluierungen vor!" (*Kuhlmann*). Zugleich darf man auch nicht vergessen, wie sehr gerade in der Frage der Performance das Instrument der Evaluierungen und seine Kriterien die Erwartungen und Handlungen der zu Evaluierenden beeinflussen. Schließlich geht es gerade im Zusammenhang mit dem Europäischen Forschungsraum darum, Instrumente weiterzuentwickeln, mit denen man die Performance von Politiken und Institutionen zumindest begleiten kann.

#### *Governance*

Das Hauptreferat hielt *J. Majo*, der Chairman des Panels zum Five Year Assessment der EU-FTE-Politik. Er schilderte den Weg der Panelmitglieder ins

Freie, nämlich von einer ursprünglich gedachten Bewertung der Programme hin zu einer Politikevaluierung. Die einzelnen Empfehlungen sollten ohnehin von allen im O-Ton gelesen und hier nicht im Detail ausgewälzt werden: Der Punkt ist, dass mit dieser Übung nicht Vorschläge innerhalb eines bestehenden Regimes / Systems erfolgen, sondern Ansätze zu Änderungen des Systems selbst gemacht werden. Die "Governance", eben die Legitimation für und die Ausführung von steuernden Eingriffen auf der Systemebene, liege nicht nur bei den Forschungsministern, sondern auch bei den Staats- und Regierungschefs, die die selbstgestellten Ziele (→ Lissabon Beschlüsse März 2000, Europa als Innovationsweltmeister) eben auch breit umsetzen müssen. Wesentliche Empfehlungen gehen auch in die Richtung, die Strategiefähigkeit wichtiger Akteure, v.a. der Kommission zu erhöhen, indem man Abwicklungsaufgaben bewältigbarer gestaltet oder sie teilweise auslagert. Ein Kritikpunkt zum ersten Themenkreis "Science for Policy" ist der, dass lt. Majo auf europäischer Ebene solche Sektorpolitiken keinen Plan haben, welche langfristige strategische Forschung sie brauchen.

Auf nationaler Ebene gibt es den Ansatz, Agenturen zu schaffen, die Informationen für den Governance-Prozess liefern können (für Spanien *Hermenegildo* über die National Agency for Evaluation and Prospective), auf europäischer Ebene soll eine Europäische Akademie für Advise and Consent sorgen (*Fenger*, s.a. unten *Ourisson*). Auf Unionsebene sind es vor allem die Monitoring-Abläufe, die derzeit für Evaluierungs-Input in die Entscheidungsstrukturen sorgen (*Fayl*).

Es gibt den Anspruch, strategische Entscheidungen an Evaluierungen zu koppeln und somit Evaluierungen und verwandten Instrumenten wie Foresight Studies einen expliziten Platz im Politikgestaltungs-

prozess einzuräumen (*Barré*). Zugleich und täglich sichtbar gibt es auch in den Akteuren und Institutionen selbst mehrere Schichten von "Beliefs" und Steuerungsinstrumentarien, die aus unterschiedlichen Zeitschichten und Philosophien herrühren: Big Science, reines bottom up, Programme, Systeme ... (*Georgiou*), was Governance-Ansätze wieder schwieriger macht.

### **Resumée**

Beide Konferenzen zeigten die Bedeutung von Evaluierungen als Politikinstrument, die Wichtigkeit der Frage, wer mit welchen Instrumenten diese vornimmt und ganz besonders die Abhängigkeit von Kontexten; sei es institutionell, sei es im Rahmen von unterschiedlichen Politiken.

### **Autoren-Hinweis**

Michael Stampfer, Technologie Impulse Gesellschaft m.b.H., Grillparzerstr. 7/8, A-1010 Wien, Tel. ++43 1/5132627-20  
E-mail: michael.stampfer@tig.or.at

**Für den Inhalt dieser Ausgabe verantwortlich:**

BUNDESMINISTERIUM FÜR VERKEHR, INNOVATION  
UND TECHNOLOGIE; Renngasse 5, A-1010 Wien  
Dr. Rupert Pichler (rupert.pichler@bmv.gv.at)

TECHNOLOGIE IMPULSE GESELLSCHAFT M.B.H.,  
Grillparzerstr. 7/8, A-1010 Wien  
Dr. Michael Stampfer (michael.stampfer@tig.or.at)

JOANNEUM RESEARCH, Institut für Technologie- und  
Regionalpolitik, Wiedner Hauptstraße 76, A-1040 Wien  
Dr. Oliver Fritz (oliver.fritz@joanneum.ac.at)

**Weitere Plattformmitglieder:**

ÖSTERREICHISCHES INSTITUT FÜR  
WIRTSCHAFTSFORSCHUNG, A-1103 Wien, PF 91  
Mag. Gernot Hutschenreiter (ger-  
not.hutschenreiter@wifo.ac.at)

ÖSTERREICHISCHES FORSCHUNGSZENTRUM  
SEIBERSDORF, Systemforschung Technik-Wirtschaft-  
Umwelt, A-2444 Seibersdorf  
DI Anton Geyer (anton.geyer@arcs.ac.at)

FFF-Forschungsförderungsfonds für die Gewerbliche  
Wirtschaft, Kärntnerstraße 21-23, A-1015 Wien  
Mag. Klaus Schitzer (klaus.schnitzer@fff.co.at)

TECHNOPOLIS AUSTRIA, Prinz Eugen Straße 80/12, A-  
1040 Wien  
DI Fritz Ohler (fritz.ohler@technopolis-group.com)

FWF-Fonds zur Förderung der wissenschaftlichen  
Forschung, Weyringergasse 35, A-1040 Wien  
Dr. Rudolf Novak (novak@mails.ffw.univie.ac.at)

**Medienhinweis:**

PLATTFORM•TECHNOLOGIE•EVALUIERUNG ist ein unregel-  
mäßig erscheinendes offenes Forum zur Diskussion metho-  
discher und inhaltlicher Evaluierungsfragen in der Techno-  
logiepolitik. © Wien 2000

Herausgabe und Versand: Joanneum Research, Institut für  
Technologie- und Regionalpolitik, A-1040 Wien, Wiedner  
Hauptstraße 76.

**PLATTFORM■TECHNOLOGIE■EVALUIERUNG**

Die PLATTFORM•TECHNOLOGIE•EVALUIERUNG ist  
eine Initiative des österreichischen  
BUNDESMINISTERIUMS VERKEHR, INNOVATION  
UND TECHNOLOGIE. In der Zusammenarbeit von  
drei österreichischen Forschungseinrichtungen,  
JOANNEUM RESEARCH, WIFO und  
FORSCHUNGSZENTRUM SEIBERSDORF, der  
TECHNOLOGIE IMPULSE GESELLSCHAFT,  
TECHNOPOLIS AUSTRIA sowie den beiden Fonds  
FFF und FWF werden Themenstellungen zur Tech-  
nologieevaluierung erarbeitet und - z.T. mit Ein-  
beziehung namhafter ExpertInnen - in einem Fach-  
kreis diskutiert.

Der Newsletter beinhaltet Fachbeiträge zu Fragen der  
technologienpolitischen Evaluierung. Die Herausgabe  
erfolgt in zeitlicher als auch inhaltlicher Abstimmung  
mit Plattform-Veranstaltungen, um die Synergie-  
wirkungen eines breiten Austauschforums zu nutzen.

**Internet-Adresse der Plattform:**

<http://www.bmwf.gv.at/4fte/wirtech/techeva/>