Institut für Betriebswirtschaft (BW)
Institut für Ländliche Räume (LR)
Institut für Marktanalyse und Agrarhandelspolitik (MA)
Bundesallee 50, 38116 Braunschweig

Arbeitsbericht zum Workshop

Zwischenbewertung der Programme zur Entwicklung des ländlichen Raumes nach der Verordnung (EG) Nr. 1257/1999 –

Erfahrungsaustausch und Verbesserungsansätze

durchgeführt am 27./28. Januar 2004 in Braunschweig

Braunschweig, im Februar 2004

Impressum

Herausgeber und Redaktion:
Bernhard Forstner (Institut für Betriebswirtschaft, FAL, bernhard.forstner@fal.de)
Josef Efken (Institut für Marktanalyse und Agrarhandelspolitik, FAL, josef.efken@fal.de),
Regina Grajewski (Institut für Ländliche Räume, FAL, regina.grajewski@fal.de)

unter Mitwirkung von:
Sabine Bresemann (Bundesanstalt für Forst- und Holzwirtschaft, Hamburg)
Detlev Clemens (Europäische Kommission, Brüssel)
Edwin Clerckx (ADE, Brüssel)
Sebastian Elbe (Sprint consult, Darmstadt)
Helena Hefti (Bundesamt für Landwirtschaft, Bern)
Hans Hochberg (Thüringer Landesanstalt für Landwirtschaft, Wandersleben)
Helmut Karl (Universität Bochum)
Birgit Koch (Institut für Ländliche Räume, FAL)
Rainer Lüdigk (Institut für Strukturpolitik und Wirtschaftsforschung, Halle)
Kathrin Peter (Büro für Evaluationen, Projektmanagement und Consulting, Bern)
Reiner Plankl (Institut für Ländliche Räume, FAL)
Andrea Pufahl (Institut für Ländliche Räume, FAL)
Karin Reiter (Institut für Ländliche Räume, FAL)
Wolfgang Roggendorf (Institut für Ländliche Räume, FAL)
Maria Schinke (Bundesministerium für Verbraucherschutz, Ernährung und Landwirtschaft, Bonn)
Heinz Wendt (Institut für Marktanalyse und Agrarhandelspolitik, FAL)

Inhaltsverzeichnis Seite

1 Einleitung (Forstner) 1

2 Evaluation aus Sicht der Europäischen Kommission (Clemens) 5

3 The use of common evaluation questions for evaluating rural development programmes: 1994 to 2006 (Clerckx) 9

4 Darstellung der Arbeitsgruppenergebnisse 17

4.1 Ergebnisse der AG_1 „Förderung von Investitionen in landwirtschaftlichen Betrieben und in Unternehmen der Verarbeitung und Vermarktung landwirtschaftlicher Produkte“ (Efken, Forstner) 17

4.2 Ergebnisse der AG_2 „Ausgleichszulage“ (Klare, Plankl) 22

4.3 Ergebnisse der AG_3 „Agrarumweltmaßnahmen“ (Reiter, Roggendorf) 27

4.4 Ergebnisse der AG_4 „Dorf- und ländliche Regionalentwicklung“ (Koch, Grajewski) 34

4.5 Ergebnisse der AG_5 „Forstwirtschaft“ (Bresemann) 39

4.6 Zusammenfassung der Ergebnisse der Arbeitsgruppen (Grajewski) 44

5 Programmbewertung und Perspektiven der Evaluation - Darstellung der Impulsreferate der Plenumsitzung des zweiten Tages 51

5.1 Besonderheiten der Ziel 1-Bewertung: Erfahrungen der Halbzeitbewertung des EAGFL/A im Rahmen der Evaluation des Gemeinschaftlichen Förderkonzeptes 2000-2006 für den Einsatz der Strukturfonds in den neuen Bundesländern und im Ostteil Berlins (Karl) 51

5.2 Zentrale/dezentrale Bewertung 53
5.2.1 Impulsreferat I (Hochberg) 53
5.2.2 Impulsreferat II (Schinke) 56

5.3 Summative/formative Bewertung (Grajewski) 58

5.4 Idealer Bewertungszyklus 60
5.4.1 Impulsreferat I „Idealer Bewertungszyklus? - Erfahrungen aus dem FuE Vorhaben BNPro“ (Elbe) 60
5.4.2 Impulsreferat II zum Thema „Idealer Bewertungszyklus“ (Lüdigk) 63

5.5 Programmansatz und kapitelübergreifende Bewertung (Grajewski) 64

5.6 Erfahrungsaustausch (Reiter) 66

5.7 Evaluation der Ökomaßnahmen 1994 bis 2005 in der Schweiz (Peter, Hefti) 68

6 Schlussbemerkungen (Pufahl, Forstner) 73

Anhang

Anhang I: Workshopprogramm 77

Anhang II: Teilnehmerliste 79

Anhang III: Ergebnisfolien der Arbeitsgruppen 85

1 Einleitung (Forstner)

Die Zwischenbewertungen der Programme zur Entwicklung des ländlichen Raumes und der Operationellen Programme für den Förderzeitraum 2000 bis 2006 wurden zum Ende des Jahres 2003 abgeschlossen.[1] Die Weiterleitung der Bewertungsberichte durch die Auftraggeber an die Europäische Kommission ist inzwischen erfolgt. Sowohl zur Reflexion der abgeschlossenen Evaluation als auch zur Vorbereitung auf zukünftige Bewertungen bestand ein vielfach geäußerter Wunsch nach einem Erfahrungsaustausch zwischen den beteiligten Evaluatoren. Dabei sind sowohl inhaltlich-methodische als auch organisatorische Fragen von Relevanz. Bislang existiert keine institutionalisierte Plattform, die den Austausch der Evaluatoren untereinander oder der Evaluatoren mit den Auftraggebern ermöglicht.

Bereits während der Bewertungsarbeit wurden zwar von jedem Bewerter zahlreiche Kontakte mit anderen Bewertungseinrichtungen und Evaluatoren geknüpft, die jedoch aus Zeitgründen meistens nicht systematisch weitergeführt werden konnten. Angesichts der sich bereits in Sichtweite befindenden Aktualisierung der Zwischenbewertung war es aus Sicht der an der FAL tätigen Evaluatoren wichtig, sehr frühzeitig, d.h. vor der Inangriffnahme der nächsten Bewertungsaufgaben, einen strukturierten Erfahrungsaustausch mit Evaluatoren der verschiedenen Bundesländer und benachbarter Mitgliedstaaten durchzuführen. Daher beschlossen die drei ökonomischen Instituten der FAL[2] die Vorbereitung und Durchführung eines derartigen Workshops für den 27./28. Januar 2004. Um den organisatorischen Aufwand relativ gering zu halten, wurde der Workshop trotz des großen Interesses an einem Erfahrungsaustausch auch in anderen EU-Mitgliedstaaten auf deutschsprachige Interessenten begrenzt. Der Einladung zum Gedankenaustausch sind schließlich insgesamt 60 Teilnehmer gefolgt, wobei es sich ganz überwiegend um Evaluatoren und teilweise um für Evaluation zuständige Personen aus dem Bereich der Verwaltung handelte. Aus dem benachbarten Ausland (Österreich, Belgien, Schweiz) haben 13 Personen teilgenommen.

In fünf unterschiedlichen Arbeitsgruppen wurde diskutiert:

AG_1: Investitionen in landwirtschaftlichen Betrieben (Kap. I) und in den Bereichen Verarbeitung und Vermarktung landwirtschaftlicher Erzeugnisse (Kap. VII)

AG_2: Ausgleichszulage in benachteiligten Gebieten (Kap. V)

AG_3: Agrarumweltmaßnahmen(Kap. VI)

---

[1] Der zugrunde liegende Rechtsrahmen resultiert aus den Verordnungen (EG) Nr. 1257/1999 und Nr. 1260/1999.

[2] Institut für Betriebswirtschaft (BW), Institut für ländliche Räume (LR) und Institut für Marktanalyse und Agrarhandelspolitik (MA)

AG_4: Anpassung und Entwicklung von ländlichen Gebieten (Kap. IX)

AG_5: Forstwirtschaft (Kap. VIII)

Eine spezielle Arbeitsgruppe zu kapitelübergreifenden Fragen und zur Programmbewertung konnte nicht durchgeführt werden, da die Programmbewerter jeweils zugleich mit der Bewertung einzelner Kapitel befasst waren und daher eine zeitgleiche Arbeitsgruppe zu kapitelübergreifenden Inhalten nicht zustande gekommen wäre. Der Aspekt der Programmbewertung wurde daher nur am zweiten Tag im Plenum angerissen.

Für die Gruppenarbeit am ersten Tag war ein Zeitrahmen von 14:00 bis 18:00 Uhr vorgesehen. Die inhaltliche Diskussion orientierte sich an gemeinsamen Diskussionspunkten, die für alle Arbeitsgruppen einheitlich festgelegt wurden. Dabei handelte es sich um folgende Punkte:

1. Inhalte der kapitelspezifischen Fragen, Kriterien und Indikatoren: Relevanz, Vollständigkeit und notwendige Ergänzungen;
2. Beantwortung der kapitelspezifischen Fragen, Kriterien und Indikatoren: Datenverfügbarkeit, Repräsentativität, Methodik, Referenzsituation;
3. Bewertungsergebnisse und deren Belastbarkeit sowie Handlungsempfehlungen;
4. Schaffung einer Basis für nachfolgende Bewertungen: Begleit- und Monitoringsysteme, Untersuchungsmethodik;
5. Eignung der Programmplanungsdokumente für die Kapitelbewertung: Inventionslogik, Zielbeschreibung und –quantifizierung, Ex-ante-Bewertung;
6. Evaluationskultur in den Ländern: Verhältnis Auftraggeber – Evaluator, formativer Prozess oder ‚Hofberichterstattung‘.

Die jeweiligen Punkte wurden jeweils durch mindestens ein Impulsreferat (max. fünf Minuten) eingeleitet und anschließend zielorientiert in der Arbeitsgruppe diskutiert. Durch die Impulsreferate war jeder Teilnehmer aktiv in den Workshop eingebunden.

Die relativ starke Vorstrukturierung der Diskussionen diente dem Zweck, die Ergebnisse am nächsten Tag im Plenum in einer vergleichbaren Form präsentieren zu können. Aufgrund des knappen Zeitrahmens erschien den Organisatoren dieser formale Rahmen erforderlich.

Am zweiten Tag wurden von 9:00 bis 12:00 Uhr die Diskussionsergebnisse der Arbeitsgruppen im Plenum vorgestellt und gegebenenfalls einer kurzen Erläuterung oder Ergänzung unterzogen. Die Ergebnisse finden sich in Kurzform in Anhang III.

Am Nachmittag (13:30 bis 15:30 Uhr) wurden im Plenum Aspekte der Programmbewertung und Perspektiven der Evaluation erörtert. Folgende Punkte standen dabei zur Diskussion:

1. Besonderheiten der Bewertung im Ziel 1-Gebiet
2. Zentrale/dezentrale Bewertung
3. Summative/formative Bewertung
4. Idealer Bewertungszyklus
5. Programmansatz und kapitelübergreifende Bewertung
6. Erfahrungsaustausch der Evaluatoren sowie der Evaluatoren und Auftraggeber

Die Diskussion der einzelnen Punkte wurde jeweils wieder durch ein oder mehrere Impulsreferate vorbereitet. Einen Schwerpunkt bildeten Fragen zur Kombination der Programmbewertung aus zentraler und dezentraler Bewertung, wie sie in Deutschland bei der Zwischenbewertung erfolgte und auch bei einer Aktualisierung der Zwischenbewertung und der Ex-post-Bewertung vorgesehen ist.

Der Workshop wurde ergänzt durch zwei Vorträge. Den Auftakt zur Veranstaltung am ersten Tag machte Herr Clemens als Vertreter der Europäischen Kommission und zuständig für Evaluation der Politik zur Förderung des ländlichen Raums. Herr Clemens gab einen interessanten und sehr aufschlussreichen Überblick über den gegenwärtigen Stand der Zwischenbewertung, die weitere Verwendung der nun vorliegenden Bewertungsberichte und die Vorstellungen der Kommission hinsichtlich der künftigen Bewertungsschritte (‚update‘ und Ex-post-Bewertung). Die Ausführungen von Herrn Clemens wurden im Anschluss an seinen Vortrag ausführlich diskutiert.

Am zweiten Tag präsentierte Herr Clerckx (ADE, Brüssel) in einem etwa 30-minütigen Vortrag zur Evaluationsmethodik, in dem er einen Vergleich der Ziel-5b-Bewertung (1994 bis 1999) mit der aktuellen Zwischenbewertung durchführte.

Der vorliegende Tagungsband enthält – weitgehend entsprechend der zeitlichen Abfolge des Workshops – zunächst die überarbeiteten Beiträge von Herrn Clemens (Kapitel 2) und Herrn Clerckx (Kapitel 3). Anschließend werden die Beiträge und Diskussionsergebnisse der einzelnen Arbeitsgruppen, ergänzt durch eine Zusammenfassung über die Arbeitsgruppen hinweg, wiedergegeben (Kapitel 4). Der folgende dritte Abschnitt enthält die Einzelbeiträge und Diskussionsergebnisse zu Fragen der Programmbewertung und Bewertungsorganisation (Kapitel 5). Den Schluss (Kapitel 6) bildet ein kurzes Resümee zum Workshop, dem sich einige Vorschläge für die Organisation eines weiterführenden Gedankenaustausches anschließen. Im Anhang befinden sich das Workshopprogramm, die Teilnehmerliste sowie die in den Arbeitsgruppen erarbeiteten ausführlichen Ergebnisfolien, die im Hauptteil nur kurz zusammengefasst werden.

2 Evaluation aus Sicht der Europäischen Kommission (Clemens)

Herr Clemens nahm als Vertreter Europäischen Kommission am Workshop teil. Er arbeitet seit dem Jahr 2002 in der Generaldirektion Landwirtschaft (DG Agri) im Referat Evaluierung und ist zuständig für die Evaluation im Bereich der ländlichen Entwicklung.

Zu Beginn äußerte sich Herr Clemens sehr positiv über den stattfindenden Workshop. Dieses Beispiel sollte möglichst von anderen Mitgliedstaaten aufgegriffen werden, da die EU Kommission auf diese Weise unmittelbar und effizient die Situation, Probleme und Anpassungswünsche aus den Mitgliedsstaaten erfahren könne. Das bestehende Bewertungsraster mit den darin enthaltenen Kernfragen, Kriterien und Indikatoren werde seitens der EU Kommission als flexibel interpretiert und die Evaluationskonzeption werde in einem prozessualen Kontext gesehen. Insgesamt seien die ihm bekannten Rückmeldungen aus den Mitgliedstaaten zum Bewertungsraster überwiegend positiv gewesen, was wünschenswerte Änderungen in Einzelfragen – etwa hinsichtlich bestimmter Indikatoren – keineswegs ausschließe.

Als besonders erfreulich wertete Herr Clemens, dass die Ergebnisse der Zwischenbewertung erstmals dezidiert im Prozess der Entscheidungsfindung für die Weiterentwicklung der Politik für den ländlichen Raum genutzt würden. Entsprechende Politikvorschläge würden bis zum Sommer 2004 vorgelegt. Dies stelle einen erheblichen Fortschritt gegenüber den vorangegangenen Perioden dar. Jedoch seien inhaltlich von der Zwischenbewertung im Hinblick auf die Beantwortung der wirkungsbezogenen Fragen noch keine substanziellen Antworten zu erwarten, da der Abstand zwischen der Durchführung der Interventionen und der Bewertung vielfach noch zu gering sei.

Die Synthese der Zwischenbewertungsberichte könne nach Ansicht von Herrn Clemens in der von der Kommission bislang vorgesehenen Form nicht durchgeführt werden, da ihre Ergebnisse zu spät für nützlichen Input in die politische Entscheidungsfindung vorliegen würden. Stattdessen werde zunächst eine ‚light'-Version der Synthese erstellt, die auf circa 30 ausgewählten Bewertungsberichten aufbaue, von denen rund 20 Berichte dem Bereich Garantie und rund 10 dem Bereich Ausrichtung zuzuordnen seien. Inhaltlich würden nicht alle, sondern nur ein Ausschnitt der Evaluationsfragen aufgegriffen. Die Auswahl orientiere sich an den Kernthemen der Konferenz von Salzburg, die im November 2003 stattfand. Diese sind:

- Wettbewerbsfähigkeit von Betrieben,
- Vielfalt im ländlichen Raum (Diversifizierung),
- Lebendiger ländlicher Raum (Soziale, gemeindliche, kulturelle Infrastrukturen),
- Mobilisierung lokaler Akteure (Mainstreaming, LEADER),

– Vereinfachung der Durchführung (Verwaltung, Kontrolle, Finanzen, Programmierung).

Die geplante umfangreiche Synthese werde zu einem späteren Zeitpunkt erstellt (Plan: 6/2004 bis 8/2005).

Herr Clemens berichtete, dass gegenwärtig im Auftrag der Kommission ein sog. Quality Assessment aller Evaluierungsberichte durchgeführt werde, das auf den Kriterien der MEANS-Handbücher aufbauend die formale Qualität der Bewertungsberichte prüfe. In der Folge werde den Mitgliedstaaten die Entscheidung über die Annahme der Berichte schriftlich mitgeteilt. Dieser Prozess solle bis Ende Februar 2004 abgeschlossen sein. Spätestens nach Annahme der Berichte werde von den zuständigen Ebenen (Mitgliedstaat, Land) erwartet, die Bewertungsberichte zu veröffentlichen.

Die EU-Kommission werde, so Herr Clemens, auf der Basis der Bewertungsberichte die Sinnhaftigkeit der Fragen, Kriterien und Indikatoren des Bewertungsleitfadens überprüfen und gegebenenfalls anpassen. Aufgrund des prozesshaften Ansatzes, in welchem die Zwischenbewertung 2003 und deren Aktualisierung 2005 in einem engen Zusammenhang gesehen werden, sei jedoch nicht davon auszugehen, dass es zu wesentlichen Änderungen bei den Bewertungsfragen komme. Mithin gehe er auch davon aus, dass der bestehende Leitfaden auch für die gemäß der VO (EG) Nr. 445/2002 „nötigenfalls durchzuführende“ Aktualisierung (‚update‘) der Zwischenbewertung nicht verändert werde.

Herr Clemens verwies darauf, dass das ‚update‘ für den Bereich Garantie nicht verpflichtend sei, ihre Durchführung aber von der Kommission sehr empfohlen werde. Zwar käme der Zeitpunkt des ‚update’ für ein direktes Einfließen der Ergebnisse in politische Entscheidungen auf EU-Ebene zu spät, und die Kommission sehe daher auch keine EU-weite Synthese der Berichte vor. Auf nationaler/regionaler Ebene sei das ‚update’ jedoch am besten geeignet, Resultate und Wirkungen der gegenwärtigen Maßnahmen und Programme abzubilden und den Programmierungsprozess für die Programmperiode 2007 bis 2013 zu beeinflussen. Zudem dürfte die Durchführung eines ‚update’ die Erstellung der verpflichtenden Ex-post-Bewertung wesentlich erleichtern.

Die Ex-post-Bewertung wird nach Ansicht von Herrn Clemens vielfach kritisiert. Sie sei aber aus Rechenschaftsgründen unabdingbar. Der verbindliche Finanzanteil von insgesamt 40 % für die Ex-post-Bewertung sei nicht mehr in der Durchführungsverordnung verankert. Daher können die Mitgliedstaaten die zur Verfügung stehenden Mittel auf das ‚update‘ und die Ex-post-Bewertung nach ihren Vorstellungen aufteilen. Die verpflichtend vorgeschriebene Synthese der Ex-post-Bewertung auf EU-Ebene sei für 2009 geplant.

Im Hinblick auf die neue Programmplanungsperiode ab 2007 sollten in den neuen Verordnungen zur Entwicklung des ländlichen Raumes die Widersprüche bezüglich der Zuständigkeiten und der Anforderungen an die Evaluierungen, die derzeit zwischen der VO (EG) Nr. 1260/1999 und der VO (EG) Nr. 1257/1999 existieren, beseitigt werden. Ferner sei es Ziel der EU-Kommission, die Vorgaben für das Evaluierungssystem rechtzeitig vor der Programmierungsphase an die Mitgliedstaaten herauszugeben, das heißt auch, die Leitfäden für Monitoring und Evaluierung den Mitgliedstaaten rechtzeitig zur Verfügung zu stellen. Nachdem die Bewertungsberichte hinsichtlich der Kommentare zum Bewertungssystem gesichtet wurden und ‚im Feld' ausreichend Feed-back gesammelt wurde, soll den Mitgliedstaaten ein neuer Leitfaden für die künftige Programmierung noch im Jahr 2005 an die Hand gegeben werden. An der bestehenden klassischen Binnengliederung des Evaluierungsprozesses mit den Elementen Ex-ante-, Zwischen- und Ex-post-Bewertung soll nach Ansicht von Herrn Clemens auch in der nächsten Förderperiode 2007 bis 2013 festgehalten werden, doch können auch Elemente einer ‚kontinuierlichen' Evaluierung diskutiert werden.

Da das ‚update' erst bis zum Ende 2005 und damit zu spät für die Kommission im Hinblick auf die kommende Programmplanungsperiode komme, konnte sich Herr Clemens vorstellen, noch während der Bewertungsphase ein Seminar im Brüssel durchzuführen, an dem die Evaluatoren der Mitgliedstaaten teilnehmen und ihre Erfahrungen und Empfehlungen an die Europäische Kommission weitergeben könnten.

3 The use of common evaluation questions for evaluating rural development programmes: 1994 to 2006 (Clerckx)

Summary of the presentation made by Edwin Clerckx at the workshop on the mid-term evaluations of rural development programmes 2000-2006, held at the FAL in Braunschweig on 27 and 28 January 2004. The presentation was based on a paper written by Jean-Marie Wathelet (jean-marie.wathelet@ade.be) and Edwin Clerckx (edwin.clerckx@ade.be) of ADE s.a. Consulting services. The full paper is available on http://www.ade.be.

Introduction

The presentation concerned the evaluation practices – and more specifically the use of common evaluation questions (CEQs) – in the field of rural development and their evolution from the 1994-1999 to the 2000-2006 programming period.

The analysis was based on a comparison of the Ex-post evaluation at EU level[3] (or 'synthesis' evaluation) of the Objective 5b programmes 1994-1999[4] and the *mid-term* evaluations of two rural development programmes (RDPs) 2000-2006.[5] EU rural development programmes are indeed implemented at country, region or 'zone' level. This means that across Europe, a large number of individual programmes are supported. These programmes are submitted to *ex ante*, *mid-term* and *ex-post* evaluations, undertaken at the level at which the programmes were implemented, but also at EU level, at least for the *mid-term* and *ex-post* evaluations. It is precisely with a view to facilitating such EU-level or 'synthesis evaluations' that the CEQs[6] were developed.

---

[3] This *ex-post evaluation of Objective 5b programmes 1994-1999* was carried out by a consortium led by ADE (Belgium). The other members of this consortium were: Agriconsulting S.p.A. (Italy), EFLR Universität Hohenheim (Germany), Edater (France), IM Gruppen (Sweden), University of Oulu (Finland), Regional Consulting (Austria) and the Universidad Autonoma de Barcelona, Fundacio Empresa I Cienca (Spain). The report is available at: http://europa.eu.int/comm/agriculture/eval/reports/objective5b/index_fr.htm.

[4] The Objective 5b programmes 1994-1999 were part of the Structural Policy of the EC and participated to what can be called a 'territorial' approach. They provided financial aid to rural areas outside Objective 1 regions confronted with specific difficulties (such as a low GDP, a high share of agricultural emloyment, a low level of agricultural income, a low population density or a depopulation trend.

[5] Unlike the Objective 5b programmes, the RDPs 2000-2006 do not participate to a regional (territorial) approach, or at least not to the same extent (an exception is the measure aiming at "promoting the adaptation and development of rural areas" that is indeed part of a regional approach). This is one of the important differences between both programmes.

[6] A set of common questions, identical for all implemented programmes.

The purpose of this comparison was – without questioning the approach through CEQs itself – to draw a number of preliminary conclusions concerning the methodology applied, more specifically in terms of its implications for the future *mid-term* and *ex-post* evaluations at EU level (but also for the evaluations at country, region or zone level), and to propose a number of directions to explore in order to optimise these future evaluations.

The present summary starts with a presentation of the extent to which the imposed CEQs were addressed in the *ex-post* evaluations of the Objective 5b programmes 1994-1999 and the reaction of the authors of these evaluations on this methodology. It then analyses the methodological changes introduced in the 2000-2006 programming period and examines how these changes were applied in two of the recent *mid-term* evaluations of the RDPs 2000-2006. To conclude, it proposes a number of 'directions to explore' with a view to optimising future evaluations, with a specific focus on the evaluations at EU-level.

The use of common evaluation questions in the ex-post *evaluations of the Objective 5b programmes 1994 to 1999*

In order to ensure uniformity of structure of the various evaluation reports and to facilitate the later evaluation at EU-level, the evaluations of the Objective 5b programmes 1994-1999 had to **answer to a set of 17 CEQ**[7]. The *ex-post* evaluation at EU level of these programmes (the synthesis evaluation) carried out in 2003 showed nevertheless that, despite the instructions given, **the CEQs method had not led to the expected results**. From the analysis of the 53 reports (covering 63 Objective 5b zones) it appeared that:

- For the majority of zones (32 on 63), the CEQs had not been addressed and for the remaining ones only some questions had been addressed;
- Even when the questions had been used, an analysis of the answers showed that:
- There were many diverging interpretations of the meaning of the questions;
- (Impact) indicators had rarely been used and, when used, differed from report to report and were not often accompanied by baseline and (realistic) target values;
- The link between the evaluation question (expressing an expected effect) and the actions or measures funded by the programme was often not clear.
- There was a strong dichotomy between the answers to the CEQs and the evaluation in terms of the classical evaluation criteria, leading to a sort of 'double' evaluation.

[7] Two other central requirements were: to reconstruct and examine the validity of the programme's intervention logic and to carry out an evaluation in terms of the 'classical' evaluation criteria (relevance, efficiency, effectiveness, etc.).

Interviews with the evaluators showed that this was **due to both the lack of data and shortcomings in terms of understanding how to use the CEQs**. With respect to the latter, the evaluators underlined that:

- Questions were not always clear and gave rise to different interpretations;
- They did not know how to structure the responses and what (impact) indicators to use;
- They had difficulties in determining the link between the measures, sub-measures or actions and the evaluation questions;
- The relation between the CEQs and the evaluation criteria was not clear to them.

Nevertheless, the evaluators **considered the approach through the CEQs as most valuable**, as a way of structuring the exercise by expressing the expected effects of the programmes.

The methodology for the evaluation of the RDPs 2000-2006 and its implementation in two* mid-term *evaluations

The programming period 2000-2006 introduced some substantial changes in terms of evaluation methodology, notably with respect to the CEQs. The new methodology is explained in a detailed manual that lists 53 common evaluation questions and provides for each question:

- The link between the nature of the supported activity and the expected effect (the intervention logic in which the question participates);
- The judgment criteria and (mostly quantitative) indicators to use to answer the questions (while stressing the need to define baseline and target levels);
- Comments on the question and the criteria (how to understand them), on the indicators (the meaning of the indicator, how to use it, etc.), contextual information (the role of possible exogenous factors, etc.) and information on the sources to use.

In other words, **the evaluation strategy for the 2000 to 2006 programming period manages to remedy - at least *ex ante* - most of the shortcomings observed for the 1994 to 1999 Objective 5b programmes**.

A first experience with two *mid-term* evaluations carried out for the RDPs of Wallonia (Belgium) and the Grand-Duchy of Luxembourg[8] reveals nevertheless a **number of remaining difficulties**.

The first series of difficulties concerned the **data collection**:

- The indicators imposed by the EC manual needed to be broken down further in primary data requested to calculate their value;
- As many of these primary data were not available as such, specific data collection systems needed to be developed to obtain them. Given the volume of data requested (174 indicators, often broken down in several categories) and the scarcity of resources available to collect them (an additional task for the scheme managers), delays were observed and many data could not be obtained;
- The costs for the collection of some data were prohibitive compared to the support received.

The **use of indicators** to address the CEQs also raised a number of problems:

- A limited number of indicators appeared to be **inadequate** for responding to the questions;
- The **interpretation** of the values reached by the indicators was hampered by difficulties in making comparisons in the context of the RDPs:
  - Many beneficiaries received similar aid since several decades so that there was not really a situation 'before and after the intervention', as requested for *temporal comparisons*. Baseline values were also missing;
  - It was often difficult to obtain information about non-beneficiaries at a reasonable cost in order to use *counterfactual situations* (policy on/off); in some cases these situations simply did not exist;
  - The use of a *norm* was not possible as best practices had not been identified and networks to realise benchmarking in the field did not exist at EU-level;
  - *Target values* that can be used as a means for comparison had often not been fixed at the *mid-term* stage.
- Establishing **causality** was an important challenge not only given the well-known difficulty of isolating exogenous factors, but also because:

[8] The first evaluation was carried out by ADE (Belgium) in collaboration with the University of Gembloux (Belgium), the second by the ADE (Belgium) – ECAU (Grand-Duchy of Luxembourg) consortium.

  - ÷ Beneficiaries from the RDPs receive support through different channels (different measures within the same programme, national aid programmes, etc.), so that it becomes difficult, if not impossible, to isolate the effects of a specific intervention;
  - ÷ Expectations are sometimes too high with respect to the effects that can be attributed to a particular intervention, or in terms of the visibility of the causal link.
- In general **information provided was more qualitative** than could have been expected given the predominance of quantitative indicators:
  - ÷ Qualitative information was used to fill the gaps in quantitative data;
  - ÷ To understand the outputs, results and effects identified through quantitative indicators, it is indispensable to rely on qualitative information.

In terms of the **nature of the CEQs**[9] and the required **structure of the answers**, the **dichotomy between the CEQs and the classical evaluation criteria remains** (the two evaluations under discussion have treated them separately). As the CEQs concern nearly exclusively the effectiveness criterion and much less matters of efficiency, coherence, etc.:

- Privilege is given to the evaluation of measures, and not of the programme, which is surprising since the RDPs remain part of an integrated approach (although not territorial), supposing an intervention based on a specific mix of measures, i.e. a plan or programme;
- While in the same time information on the measures (in terms of relevance, efficiency) is restricted, which is detrimental to their understanding (even in terms of effectiveness).

Finally, as the questions focus on effects, they may induce answers of a **high level of abstraction** in the sense of focussing on *measuring* the effects targeted by the questions and less on the actions implemented, nevertheless indispensable to *understand* properly these effects.

Directions to explore to optimise future (synthesis) evaluations

Bearing in mind the experience of the synthesis evaluation of the Objective 5b programmes and taking into account the preliminary findings drawn from the experience with the *mid-term* evaluations of two RDPs, a number of directions to explore to optimise future synthesis evaluations (at *mid-term* and *ex-post* stage) should be taken into consideration.

---

[9] This issue is only addressed here in methodological terms and does not try to verify, for instance, whether the questions covered all the themes present in these specific RDPs.

As a preliminary remark, it should be underlined that the **new methodology and the richness of tools developed for its application constitute a clear improvement compared to the previous programming period**: the exercise is far better structured and *a priori* more suited for generating useful information, notably for synthesis evaluations.

A number of difficulties have been revealed by the *mid-term* evaluations, some of which will be amplified in the synthesis exercises. The actions suggested to deal with these difficulties are listed hereafter.

With respect to the **data**:

- In order to avoid that the primary data resulting from the breakdown of indicators differ from zone to zone and thus hamper the aggregation and comparison of indicators at EU-level it might be envisaged to develop a ***common framework for the calculation of indicators*** or to make sure that this exercise is standardised;
- Strategies to ***improve the planning of resources to manage the data collection systems***, eventually foreseeing a separate amount in the programme budget for the data collection, might be an appropriate means to deal with the problems encountered in terms of data collection.

In order to deal with the difficulties related to the **use of the indicators**, it seems advisable:

- As some indicators are not adequate to answer the CEQs, to ***optimise the set of indicators***;
- To develop the tools for interpreting the indicators: fix adequate baseline and ***target values***, promote ***benchmarking*** between regions (as an option for fixing the targets, but also to share experience, take over best practices) and set up and monitor ***networks*** to do so, promote ***benchmarking between programmes***;
- To ***adjust the expectations*** with respect to the information provided by (quantitative) indicators and stress the importance of ***relying on qualitative information***

With respect to the **nature of the CEQs**:

- As the dichotomy between the CEQs and the 'classical' evaluation criteria is persistent (questions focus on effectiveness) and in order to avoid that evaluators continue to treat separately the analysis in terms of the evaluation criteria and the answers to the CEQs, it seems indispensable to ***clarify the relation between the CEQs and the criteria***. This is even more important given the fact that the RDPs still participate to a programme or 'integrated' approach;
- The level of abstraction of the CEQS (i.e. the fact that reference to the measures or actions may be lacking in the description of effects) will probably be amplified in the

synthesis analysis. This may be counterbalanced through ***complementary thematic studies***, which may also serve comparisons across different geographical areas and thus provide a valuable feedback to the Member States.

Overall, the performances of the evaluations would certainly benefit from the establishment of a **network of evaluators at EU level** in order to ensure the constant coordination and monitoring of the exercise. Such a network could be responsible for enhancing a common understanding of the exercise, solving specific difficulties met by the evaluators, sharing best practices among evaluators, etc.

In terms of time, it should be possible to carry out most of these tasks soon enough to benefit from their results at the stage of the *ex-post* evaluation (the synthesis of the *mid-term* evaluation could already be the occasion to implement some of the tasks, for instance the exchange of best practices between evaluators). The results of some other tasks, such as the reflection on the relation between CEQs and classical evaluation criteria, are difficult to integrate in the exercise of the current period, but they should nevertheless be available for the next programming period.

4 Darstellung der Arbeitsgruppenergebnisse

4.1 Ergebnisse der AG_1 „Förderung von Investitionen in landwirtschaftlichen Betrieben und in Unternehmen der Verarbeitung und Vermarktung landwirtschaftlicher Produkte“ (Efken, Forstner)

TeilnehmerInnen:

- Herr Efken, Herr Forstner, Herr Wendt (jeweils FAL, Braunschweig)
- Herr Jacobs, Frau Enderling (jeweils Landgesellschaft Sachsen-Anhalt)
- Herr Janetschek, Herr Pistrich (Bundesanstalt für Agrarwirtschaft, Wien)
- Frau Beck (Universität Gembloux)
- Herr Pölking (agroplan Wolfenbüttel)
- Herr Mark (Landesanstalt für Entwicklung der Landwirtschaft und der Ländlichen Räume, Schwäbisch-Gmünd)
- Frau Schinke (Bundesministerium für Verbraucherschutz, Ernährung und Landwirtschaft, Bonn)

Die AG_1 repräsentierte die Evaluatoren der investiven Fördermaßnahmen der Programme. Zum einen handelt es sich um die einzelbetriebliche Investitionsförderung in landwirtschaftlichen Unternehmen (EIF) und zum anderen um die Förderung von Investitionen zur Verbesserung der Verarbeitung und Vermarktung landwirtschaftlicher Erzeugnisse (VV), die hauptsächlich in der Landwirtschaft nachgelagerten Unternehmen und Erzeugerkooperationen durchgeführt werden.

Konkret sind es bezüglich der landwirtschaftlichen Betriebe vornehmlich Investitionen in Gebäude und damit verbundenen technischen Anlagen (Stallbau, Melkanlage, Gülletechnik etc.). Investitionen im Bereich Verarbeitung und Vermarktung decken insbesondere die Bereiche Lagerung, innerbetriebliche Logistik sowie Verarbeitungsanlagen für die landwirtschaftlichen Rohstoffe (Milch, Fleisch, Obst, Gemüse etc.) ab.

Spannt man einen Bogen von den Inhalten der kapitelspezifischen Fragen über deren Beantwortungseignung bis hin zu den konkreten Bewertungsergebnissen, so nehmen nach Ansicht der AG die Probleme, Unzulänglichkeiten und Lücken zu. Dies ist auch nicht weiter überraschend, denn in den Inhalten der Bewertungsfragen und den damit verknüpften Kriterien sowie Indikatorvorschlägen spiegeln sich eher die Ansprüche an die Bewertung wider; d.h. was *soll* ermittelt und beurteilt werden.

Hinsichtlich der Beantwortung der Fragen tun sich deutlichere Lücken und Probleme auf, da hier die Suche nach den Möglichkeiten zur Beantwortung schnell erkennen lässt, dass

das verfügbare Datenmaterial und auch die im Rahmen der Möglichkeiten zusätzlich erhebbaren Informationen nur begrenzte Analysen erlauben.

So können vor diesem Hintergrund und unter Berücksichtigung des kurzen Beobachtungszeitraumes auch kaum Bewertungsergebnisse präsentiert werden, zumal der ‚Impact' von Investitionen überwiegend erst nach 2 bis 5 Jahren belastbar ermittelt werden kann.

1. Inhalte der kapitelspezifischen Fragen

Die Inhalte der kapitelspezifischen Fragen decken im Bereich der Investitionen in landwirtschaftlichen Betrieben die Aspekte Einkommen, Rationalisierung, Qualität, Neuausrichtung/Diversifizierung, Umwelt, Tierschutz und Arbeitsbedingungen ab. Im Bereich Verarbeitung und Vermarktung zielen die Fragen auf die Wettbewerbsfähigkeit der landwirtschaftlichen Erzeugnisse durch (a) verbesserte rationelle Verarbeitungs- und Vermarktungsbedingungen und (b) verbesserte Produktqualität ab. Weiterhin wird nach den positiven Effekten für die Landwirtschaft (v.a. Wertschöpfung) sowie die Umwelt, die Gesundheit und den Tierschutz gefragt.

Vor dem Hintergrund der klassischen Evaluationskriterien Relevanz, Effizienz und Effektivität werden nach Ansicht der AG-Mitglieder die relevanten Punkte abgedeckt. Die Fokussierung auf den Begriff Qualität stellt im Bereich Verarbeitung und Vermarktung eine unnötige Einengung dar; Innovation und Marktausrichtung sind ebenso wettbewerbsrelevante Aspekte. Diese Einengung wird im Rahmen der vorgeschlagenen Kriterien und Indikatoren zudem nicht eingehalten. Überdies taucht der Aspekt Qualität bei den vorgeschlagenen Kriterien und Indikatoren kaum auf, wohingegen Innovationen hier abgefragt werden.

Folglich wurde der Analysekette Fragen – Kriterien – Indikatoren eine Unausgewogenheit attestiert und zugleich ein Mangel an definitorischer Schärfe der verwendeten Begriffe festgestellt. Insbesondere hinsichtlich der Indikatoren wird die Unausgewogenheit auch dahingehend deutlich, dass teilweise sehr detaillierte und spezifische Indikatoren vorgeschlagen werden (z.B. Ernährungswert bei VV), die kaum zu ermitteln sind, und teilweise andere Bereiche wie etwa Qualität im Indikatorenset (bei VV) nur unzulänglich berücksichtigt sind.

Als Dilemma stellt sich der Konflikt dar, einerseits dem Anspruch an präzisen, das Problemfeld abdeckenden Indikatoren gerecht zu werden, andererseits die Praktikabilität, Durchführbarkeit und nicht zuletzt Akzeptanz bei den Bewilligungsstellen und Unternehmen noch sicherzustellen.

2. Beantwortung der kapitelspezifischen Fragen

In der Diskussion um die Beantwortung der kapitelspezifischen Fragen wurde auf das völlig unzureichende System bei der Datenerfassung und –bereitstellung hingewiesen. Amtliche Statistiken und andere Daten sind häufig nur sehr begrenzt für die speziellen Bedürfnisse der Evaluation problemadäquat zu nutzen, weil ihre ursprüngliche Zielsetzung nicht mit dem aktuellen Bewertungszweck übereinstimmt. Selbst das programminterne Monitoringsystem wird den Ansprüchen der Evaluation nur in geringem Umfang gerecht, da der Indikatorenumfang viel zu gering ist. Dies gilt überwiegend auch für die neuen Länder Deutschlands, wo jeweils ein zusätzliches Begleitsystem installiert wurde, das jedoch hinsichtlich des für die Bewertung erforderlichen Datenumfangs nicht ausreicht.

Die Datenprobleme stellen aus Sicht der Evaluatoren ein entscheidendes Problem bei der Generierung von repräsentativen (validen) Ergebnissen dar. Ein weiteres methodisches Kernproblem ist der geforderte und derzeit nur marginal erfüllbare Vergleich zwischen geförderten und nicht geförderten Unternehmen bzw. das Fehlen von Referenzszenarien. Mithin ist der geforderte Ausweis von Nettoeffekten in der Regel kaum zu ermitteln.

Zur Beantwortung der Bewertungsfragen werden häufig Fallstudien und zusätzliche, speziell für die Evaluation konzipierte, Datenerhebungen durchgeführt. Der damit verbundene Aufwand stößt auf wenig Gegenliebe bei der Administration und den begünstigten Unternehmen mit entsprechend begrenzten Erfolgen bei der Datenerhebung.

Die Bewertung des Fördererfolges der einzelbetrieblichen Investitionsförderung baut in *Deutschland* in wesentlichen Teilen auf dem Investitionskonzept und auf Jahresabschlüssen der Auflagenbuchführung auf. Dabei ist zu berücksichtigen, dass die Jahresabschlussdaten erst mit erheblichem zeitlichem Verzug vorliegen, wenn der geforderte Abstand von mindestens zwei Jahren zum Abschluss der geförderten Investition eingehalten werden soll. Auch zum Zeitpunkt des ‚update' werden daher nur in unzureichendem Umfang Daten vorliegen. In *Belgien* werden dagegen Investitionskonzepte (Betriebsverbesserungspläne) nicht für die Bewertung verwendet, da diese Daten aus Sicht der Bewerter zielorientiert gestaltet und daher kaum belastbar sind. Für die Beantwortung einkommens- und produktivitätsbezogener Fragen wird daher auf Buchführungsdaten der geförderten Betriebe und des Informationsnetz landwirtschaftlicher Buchführungen (INBL) zurückgegriffen. Bei dieser Vorgehensweise treten jedoch erhebliche Schwierigkeiten hinsichtlich der Repräsentativität und der Bildung von Vergleichsgruppen auf.

Mit den speziell erhobenen Förderdaten ist es jedoch zumindest in Deutschland im Bereich Verarbeitung und Vermarktung gelungen, einen Vorher–Nachher–Vergleich aufzustellen. Methodisch wurde diesbezüglich jedoch auf die grundsätzliche Schwierigkeit des angemessenen Zeithorizontes bei Investitionen verwiesen, die nur mit erheblicher Zeit-

verzögerung eine Wirkungsmessung zulassen. Darüber hinaus besteht oftmals das Problem der wiederholten Förderung von Unternehmen durch dasselbe Programm (dieselbe Maßnahme) oder durch andere Programme (Maßnahmen), sodass eine isolierte Wirkung vielfach gar nicht ermittelbar ist.

3. Bewertungsergebnisse

Insgesamt führte die Zwischenbewertung im Bereich der einzelbetrieblichen Investitionsförderung von landwirtschaftlichen Unternehmen wie auch in den ihnen vor- und nachgelagerten Unternehmen nur zu sehr begrenzten Ergebnissen, die für eine Wirkungsbeurteilung eindeutig nicht ausreichen. In der AG-Diskussion wurde darauf hingewiesen, dass die Ermittlung von Wirkungen, wie sie von der EU-Kommission vorgesehen ist, in manchen Bereichen Dissertationscharakter habe und damit dem gängigen Auftragsdesign der Evaluation überhaupt nicht entspreche. Die Ermittlung und Beurteilung von Wirkungen ist dementsprechend gering. Sie beschränkt sich vornehmlich auf einzelne Indikatoren, zu denen ausreichend Daten gefunden bzw. ermittelt werden konnten.

Die Struktur der Programme wie auch der Halbzeitbewertung in deutschen Ziel 1-Gebieten ist nochmals komplexer und drängt die Bewertungsbereiche jeweils in einen ‚Nischenstatus'. Detaillierte Ergebnisanalysen sprengen unter diesen Umständen den gesetzten Rahmen. Verknüpft mit der unbefriedigenden Datenbasis werden Ergebnisse mit stark eingeschränkter Validität erzeugt, die eher Tendenzaussagen erlauben. Dies gilt für Einkommens- und Produktivitätsindikatoren ebenso wie für die Schaffung oder Erhaltung von Arbeitsplätzen. Ein derzeit ebenfalls nicht zu lösendes Problem stellt die halbwegs präzise Ermittlung von Mitnahmeeffekten dar. Hier fehlt es an eindeutig anwendbaren Analyseinstrumenten und Beurteilungsrahmen.

Die Erfahrungen in der AG führten zu den Empfehlungen, die Förderung zu entflechten, die Komplexität der Programme zu verringern sowie wesentlich präzisere (operationalisierbare) Ziele zu formulieren. So könnte die Bewertungsfähigkeit verbessert oder überhaupt erst hergestellt werden.

Daneben wird eine EU-weite Diskussion zwischen den Evaluatoren empfohlen. Es sollte ein Prozess des gegenseitigen Austausches in Gang gesetzt bzw. intensiviert werden. Auf diesem Wege sollten insbesondere praktisch anwendbare Grundlagen geschaffen werden, wie die klare Definition von Indikatoren oder die Erstellung von Indikatorensets etc. Daneben ermöglicht ein Diskussionsprozess die Ermittlung von ‚best practice-Verfahren' sowie allgemein einen Prozess der Qualitätssicherung und –verbesserung der Evaluation. Nicht zuletzt könnten so die Vereinheitlichung oder zumindest Vergleichbarkeit der Evaluationen verbessert werden und eventuell gemeinsame Verfahren der Datenzusammensetzung und –analyse initiiert werden. In diesem Zusammenhang wurde eine Kategorisierung der Kriterien und v.a. Indikatoren vorgeschlagen: Je nach Betrachtungsebene

und Beurteilungsbedürfnissen sollte ein (gegenüber dem jetzigen Stand verringerter) Kern von Indikatoren mit EU-weiter Gültigkeit aufgestellt werden, der je nach Bedarf um Indikatoren mit nationaler Gültigkeit und mit regionaler Gültigkeit ergänzt werden könnte.

4. Grundlagen für weitere Evaluationen

Die Einschätzung der Evaluatoren, inwiefern durch die Zwischenbewertung eine gute Grundlage für weitere Bewertungen geschaffen werden konnten, war nicht einheitlich. Die gravierenden Defizite des bestehenden Monitoringsystems hinsichtlich seiner Verwendbarkeit für Evaluierungszwecke und der Mangel an anderen nutzbaren Datenquellen führten zu einem zusätzlichen Datenerhebungsaufwand. Während einzelne Evaluatoren hierauf mit der Entwicklung differenzierter Datenerfassungssysteme reagierten, die erst künftig voll nutzbar sein werden, verweisen andere Evaluatoren lediglich auf ad hoc-Ansätze, die nur von geringem Wert für Folgebewertungen sein dürften.

Generell enthalten alle Zwischenbewertungen eine Diskussion der Fragen, Kriterien und Indikatoren, die teilweise in konkrete Handlungsempfehlungen münden. Insofern wurden auch inhaltlich und methodisch Grundlagen für nachfolgende Evaluationen geschaffen.

5. Qualität der Programmplanungsdokumente

In der Diskussion über die Eignung der Programmplanungsdokumente für eine fundierte Kapitelbewertung wurden die Begrenzungen des im Zeitraum 2000 bis 2006 etablierten Programmplanungsansatzes offensichtlich. Das teilweise zwanghaft anmutende Verknüpfen und in Beziehung setzen der Maßnahmeninhalte erfolgte auf Kosten der Genauigkeit und Nachvollziehbarkeit. In den Programmen sind die Interventionslogik, die SWOT-Analyse und die Ziele nur auf einer vergleichsweise hoch aggregierten Ebene formuliert. Sie sind daher zwar bei oberflächlicher Betrachtung nachvollziehbar, entziehen sich jedoch einer präzisen Analyse der damit verbundenen Konsequenzen für die einzelnen Maßnahmen.

Bedeutsam war auch der enorme Zeitdruck bei der Erstellung der Programme, der häufig zu Lasten einer fundierten Ex-ante-Bewertung ging. Eine belastbare Ex-ante-Bewertung im formativen Sinn war unter diesen Umständen nur sehr begrenzt möglich. Die Evaluatoren fordern daher bei künftigen Ex-ante-Bewertungen eine realistischere und den Aufgaben angemessenere Zeitplanung.

Die Integration der Ex-ante-Bewertung in den Programmantrag ist nach Auffassung einzelner AG-Mitglieder ungünstig, da vermutlich ohnehin nur positiv mit den Programminhalten korrespondierende Ergebnisse in den Programmantrag übernommen werden und daher eine ergebnisoffene Herangehensweise erschwert ist. Eine vorausschauende Bewertung sollte daher unabhängig vom Programmantrag von externer Stelle durchgeführt werden.

Auch vor diesem Hintergrund empfahlen die Evaluatoren, die Komplexität der Programme zu reduzieren, d.h. den integrativen Ansatz der Programme zurückzufahren sowie die Zielformulierung deutlich zu verbessern in Richtung Genauigkeit und Messbarkeit. Eventuell sollte eine Veränderbarkeit von Zieldefinitionen im Zeitablauf möglich sein, um neuen Entwicklungen Rechnung tragen zu können.

6. Evaluationskultur

Unabhängig von den Einstellungen der Beteiligten ließ der späte Vertragsabschluss zur Evaluation mit dem damit verbundenen Zeitdruck kaum Platz für eine eher formative Evaluation.

Vielfach bestand bei den Evaluatoren der Eindruck, die Evaluation werde von Seiten der Auftraggeber als ein weiteres Kontroll- und Berichtsverfahren angesehen. Nun gelte es, diese von der EU-Kommission auferlegte Verpflichtung möglichst unkompliziert und kostengünstig zu erfüllen. Die Evaluation wurde kaum als formatives Verfahren zur möglichen Verbesserung von Sachverhalten und Prozessen interpretiert. Dies zeigt sich unter anderem auch darin, dass nach der Abgabe der Evaluationsberichte vielfach keinerlei Interesse an einer Ergebnisdiskussion bzw. -nachbereitung von Seiten der Auftraggeber mehr bestand.

In Deutschland bestanden zwischen den Bundesländern große Unterschiede bezüglich der Unterstützung der Evaluation und der Offenheit gegenüber einer eher formativen Evaluation. Von einzelnen Mitgliedern der AG wird auf eine deutliche Einflussnahme der Auftraggeber auf die Berichterstellung im inhaltlichen Bereich hingewiesen.

4.2 Ergebnisse der AG_2 „Ausgleichszulage“ (Klare, Plankl)

TeilnehmerInnen:

- Herr Klare, Herr Plankl, Frau Rudow (jeweils FAL, Braunschweig)
- Herr Hovorka (Bundesanstalt für Bergbauernfragen, Wien)
- Herr Clerckx (ADE, Brüssel)
- Herr Hochberg (Thüringer Landesanstalt für Landwirtschaft, Wandersleben)
- Frau Deimer (Landgesellschaft Sachsen-Anhalt, Halle)
- Herr Clemens (Europäische Kommission, Brüssel)

Die Ergebnisse der Diskussion im Arbeitskreis Ausgleichszulage sind von der teilweise erheblichen Bedeutung der Ausgleichszulage geprägt. Sowohl in Deutschland, Österreich als auch Luxemburg und Belgien hat die Ausgleichszulage aus historischer, räumlicher

und finanzieller Sicht eine hohe Bedeutung im Transferzahlungssystem an den landwirtschaftlichen Sektor. Durch die kritische Auseinandersetzung des EU-Rechnungshofs mit dem Instrument Ausgleichszulage stehen diese Transfers verstärkt auf dem Prüfstand. Kritisiert wird der Mangel an aussagekräftigen und allgemein gültigen Kriterien zur Erfassung der Benachteiligung, und es wird eine verstärkte Effektivitätsprüfung verlangt.

1. Inhalte der kapitelspezifischen Fragen/Kriterien/Indikatoren

Die vier Bewertungsfragen aus Kapitel V – benachteiligte Gebiete - wurden von allen Teilnehmern des Arbeitskreises als relevant angesehen. Teilweise wurden diese Fragen in den genannten Ländern durch Ergänzungen erweitert, wenn man dadurch den regionalen Bedingungen besser gerecht werden konnte. Auch wurden, wenn es das Zielsystem des Landes erforderte, zusätzliche Fragen in die Bewertung einbezogen. Hierbei handelt es sich vor allem um Fragen zur Erhaltung der Kulturlandschaft im Kontext touristischer Anforderungen, zur Aufrechterhaltung von Ackerbau und Viehzucht als eine Einheit und zur Stabilisierung strukturschwacher Räume. Außerdem wurden die Bewertungsfragen, hier vor allem die Bewertungsfrage nach der Einkommenswirkung, von den verschiedenen Evaluatoren für relevante Betriebsgruppen, Rechtsformen und Produktionsrichtungen erweitert und aufgesplittet. Von allen Teilnehmern wurde die Beantwortung der Bewertungsfrage nach der Wirkung der Ausgleichszulage auf eine lebensfähige Gesellschaftsstruktur im ländlichen Raum (V.3) aufgrund ihrer Komplexität als sehr schwierig angesehen. Hier müssten weitere Überlegungen zur Operationalisierung dieser Frage angestellt werden. Außerdem wären auch einige grundsätzliche Untersuchungen außerhalb der Evaluation zu den Faktoren, die eine lebensfähige Gesellschaftsstruktur ausmachen, wünschenswert. Einschränkend wurde zugleich der Nutzenbeitrag der kontingenten Bewertungsansätze zur Erfassung derartiger Leistungen in Frage gestellt.

Hinsichtlich der Kriterien wurde im Arbeitskreis überwiegend die mangelhafte Quantifizierung der angestrebten Ziele kritisiert. Dies betrifft vor allem die Bewertungsfrage zur Kompensation der Einkommensdefizite (V.1). In diesem Zusammenhang ist bezüglich der von der EU vorgegebenen Kriterien anzumerken, dass diese nur einen Ausschnitt aus der realen Lage der Betriebe abbilden. So werden nach diesen Kriterien Betriebe mit einem Kompensationsgrad von mindestens 100 % nicht erfasst. Der Anteil dieser Betriebe ist jedoch nicht unerheblich. Außerdem ist für eine bessere Vergleichbarkeit der Bewertungsergebnisse eine stärkere Diskussion um eine geeignete Referenzgruppe notwendig. Je nach Datenlage und Vorgehensweise des Bewerters können sich die Ergebnisse in qualitativer und quantitativer Hinsicht erheblich unterscheiden. Für die nächste Förderperiode sind Ziele verstärkt quantitativ festzulegen.

Bezüglich der Indikatoren war festzustellen, dass vor allem die Indikatoren für die Bewertungsfrage nach dem Schutz der Umwelt (V.4) aufgrund nicht vorliegender Daten kaum abzubilden und auch schlecht zu substituieren waren. In einigen Fällen mussten sie

deshalb durch Hilfsindikatoren unterlegt werden. Insbesondere trifft dies für Umweltindikatoren zu, da Umweltwirkungen der Ausgleichszulage nur indirekt zugeschrieben werden können, sodass der Wirkungsbeitrag im Vergleich zu direkt wirkenden Maßnahmen generell schwer zu definieren und zu messen ist.

2. Beantwortung der kapitelspezifischen Fragen/Kriterien/Indikatoren

Alle Evaluatoren empfanden die verfügbaren Daten des EU-Monitorings als unzureichend. Trotz Schließung der Datenlücken durch national verfügbare einzelbetriebliche Daten sowie verschiedene Sekundärdaten war die Repräsentativität der Daten nicht immer gegeben. Speziell der Bedeutung der Nebenerwerbsbetriebe konnte bei der Beantwortung der verschiedenen Bewertungsfragen nicht voll entsprochen werden. Chancen für eine substanzielle Verbesserung der Antworten werden in einer verstärkten Unterlegung der Ergebnisse durch Fallstudien, Vernetzung geeigneter Infopools auf regionaler Ebene sowie in der Option einer besseren Verschneidung von einzelbetrieblicher Sekundärdaten gesehen.

Neben Mit-Ohne Vergleichen sowie Vorher-Nachher-Vergleichen als Untersuchungsmethoden werden Vergleiche mit Zielgrößen durchgeführt. Während im letzteren Fall die Schwierigkeiten in einer nicht immer hinreichenden Quantifizierung der Zielgröße liegen, liegen die Schwierigkeiten für erstere Methoden bei der Wahl der richtigen Vergleichsregion und betrieblichen Vergleichsgruppe sowie in der Erfassung einer Null- bzw. Ausgangssituation bei laufenden Änderungen der Förderbedingungen. Unterschiedliche Datenqualitäten und Untersuchungsmethoden, sowie Abweichungen bei der Definition und den Modifizierungen einiger Indikatoren und Bezugsgrößen erschweren Aussagen auf interregionaler Vergleichsebene erheblich. Dies wird besonders deutlich beim Vergleich des Einkommens landwirtschaftlicher Betriebe mit dem in verwandten Sektoren. Beim Vergleich mit Arbeitnehmereinkommen stellt sich das Problem einem damit korrespondierenden Einkommensmaßstab bei landwirtschaftlichen Unternehmen. Beim Vergleich mit anderen Gruppen von Unternehmern stellt sich das Problem der Auswahl einer geeigneten Vergleichsgruppe.

3. Bewertungsergebnisse

Zum Zeitpunkt der Zwischenbewertung lagen noch nicht zu allen Bewertungsfragen belastbare Ergebnisse vor. Für die Bewertungsfragen V.1 (Einkommensdefizit) und V.4 (Schutz der Umwelt) waren basierend auf Mit-Ohne-Vergleichen erste Ergebnisse für die Beschreibung der Ausgangslage verfügbar. Die Ergebnisse wiesen in Abhängigkeit vom räumlichen und betrieblichen Differenzierungsgrad eine unterschiedliche Sensibilität hinsichtlich ihrer Belastbarkeit auf. Zwischen der Bedeutung der Ausgleichszulage als Einkommensbestandteil und der Ausgestaltung der Förderung an Kriterien zum Ausgleich der unterschiedlichen Bewirtschaftungserschwernisse konnte speziell in Österreich eine positive Korrelation nachgewiesen werden. Generell lässt sich jedoch kein signifikanter

Nachweis der Wirksamkeit der Ausgleichszulage in Abhängigkeit von ihrer unterschiedlichen Förderausgestaltung erbringen. Hierfür ist die Datengrundlage in einigen Untersuchungsregionen nicht ausreichend, sodass keine allgemein gültigen Aussagen getroffen werden können.

Als mögliche Handlungsempfehlungen zur Verbesserung der Belastbarkeit der Ergebnisse wurden drei Wege gesehen.

- Modellrechnungen können für entsprechende Betriebsgruppen gute Erkenntnisse liefern; bei einer breiten Streuung der Merkmalsausprägung sind jedoch Repräsentativität und Allgemeingültigkeit eingeschränkt.
- Verbesserung des Testbetriebsnetzes und der Versuch einer Verschneidung von InVeKoS- und Testbetriebsdaten.
- Regionale Fallstudien sind ebenfalls eine mögliche Alternative, um die in einigen Regionen unbefriedigende Datensituation zu verbessern.

Für die Beantwortung der Bewertungsfragen V.2 (dauerhafte Flächennutzung) und V.3 (lebensfähige Gesellschaftsstruktur) liegen wegen fehlender Zeitreihendaten zum Zeitpunkt der Zwischenbewertung noch keine hinreichenden Ergebnisse vor. Bislang wurden verschiedene Datenoptionen und mögliche Kontextindikatoren speziell für die schwer zu operationalisierenden Bewertungsfragen überprüft. Es wurde angeregt, auch Sonderauswertungen der amtlichen Agrarberichterstattung sowie verbesserte Darstellungen in den Agrarberichten hinsichtlich ihrer Eignung als Monitoringdatenbasis zu prüfen und bewährte regionale Monitoring-Infopoolnetze fortzuführen. Strukturbrüche sollten so weit wie möglich vermieden werden.

4. Wurden Grundlagen für die Ex-post-Evaluation geschaffen?

Nachdem von allen Bewertern das bestehende EU-Monitoringsystem als unzureichend beurteilt wurde und sich die EUROSTAT-Daten in der Zwischenbewertung partiell als zu alt erwiesen, wurde von den Evaluatoren zumeist ein eigenes System mit einzelbetrieblichen und Sekundärdaten der amtlichen Agrar- und Regionalstatistik zur Auswertung erstellt. Durch einen Datenmix wurde die Eignung der Sekundärdaten weitgehend ausgeschöpft. Vereinzelte regionale Datenlücken müssen noch geschlossen werden. Regional hat sich die Erhebung von Ergänzungsdaten aus Fallstudien bewährt. Der nachrangige Einsatz von Fallstudien als Ergänzung zu den sekundärstatistischen Auswertungen wird befürwortet. Der Fallstudienumfang hat sich jeweils am Aufwand und Ertrag sowie an einer möglichst repräsentativen Auswahl von Untersuchungsregionen zu orientieren.

Mit der Überprüfung verschiedener Datengrundlagen auf Ihre Eignung zur Beantwortung der Bewertungsfragen und der Darlegung der methodischen Aufbereitungs- und Vorgehensweise wurde eine hinreichende Basis für die Ex-post-Evaluation geschaffen. Länder

mit Lücken im Datenmaterial wurden gezielt auf die Notwendigkeit und die Möglichkeiten hingewiesen, diese Datenlücken zu schließen. Für länderübergreifend relevante Fragestellungen ist zu prüfen, ob ergänzende Grundlagenstudien durchzuführen sind.

Eine bereits in der Anfangsphase der Bewertung getroffene Entscheidung zugunsten einer vorgezogenen Ex-post-Bewertung (up-date) wäre in Bezug auf Aussagen über die problemadäquate methodische Vorgehensweise vermutlich sehr hilfreich gewesen.

5. Programmplanungsdokumente als Grundlage für die Kapitelbewertung

Die Programmplanungsdokumente wurden analytisch durchgehend als gut bezeichnet. Auch die Situationsbeschreibungen wurden dem Instrument der Ausgleichszulage gerecht. Teilweise wäre eine nach benachteiligten Gebietskategorien differenziertere Stärken-Schwächen-Analyse (SWOT) wünschenswert gewesen. Die Förderstrategie war jeweils nachvollziehbar.

Die Zielbeschreibung hing in starkem Maße vom jeweiligen Programmplanungsdokument ab. Wenngleich vereinzelt gute Konkretisierungen vorhanden waren, war die Zielquantifizierung regelmäßig nicht ausreichend. Ebenso fehlten Zielspezifizierungen, die mit operationellen Indikatoren auf der Grundlage der vorhandenen Daten überprüft werden können. Darüber hinaus wurde auf eine Wichtung der Ziele in den Programmplanungsdokumenten verzichtet. Erst durch eine nachträgliche Zielanalyse konnte diese Lücke geschlossen werden. Hier hat die Zwischenbewertung eine geeignete Grundlage für die Ex-post-Bewertung geschaffen. Generell ist zu fordern, dass die Behörden die Ziele vorgeben und soweit wie möglich quantifizieren. Strikt abzulehnen ist eine Zielquantifizierung durch den Evaluator. Daraus folgt mit Blick in die Zukunft, dass Programmerstellung und Ex-ante-Bewertung nicht in Personalunion erfolgen dürfen.

6. Zusammenarbeit Auftraggeber - Evaluator

Die Zusammenarbeit zwischen Auftraggeber und Evaluator war weitgehend konstruktiv. Vereinzelte anfängliche Auffassungsunterschiede hinsichtlich der Rolle und Aufgabe der Bewerter bei der Ziehung von Schlussfolgerungen konnten zufrieden stellend beigelegt werden. Durch einen offenen Dialog konnten Einflussnahmen auf den methodischen Ansatz umgangen werden. Durch den Ansatz der zentralen Evaluation in Deutschland mit mehreren Vertragspartnern war der Konsensfindungsprozess teilweise zeitintensiver. Generell war eine enge Korrelation zwischen den Erwartungen an die Ergebnisse der Evaluation und der regionalen Bedeutung der Ausgleichszulage zu beobachten. Dies kam verstärkt in Regionen zum Tragen, in denen Parallelbewertungen der Ausgleichszulage durch den Landesprogrammevaluator und den zentralen Evaluator implementiert wurden.

In Ländern mit schlechter Datenbasis ist zur Verbesserung der Bewertungsgrundlagen eine intensivere Zusammenarbeit notwendig. Initiativen für einen kontinuierlichen Dialog

zwischen Auftraggeber und Evaluator, der in der „heißen“ Analyse- und Bewertungsphase vielfach zu kurz kam, sollten durch die Schaffung der personellen und finanziellen Voraussetzung implementiert und verstetigt werden.

Die Einsicht, Evaluation als Prozess zu verstehen und den notwendigen Rahmen hierfür zu schaffen, ist bei den Auftraggebern noch nicht stark ausgeprägt. Das Verständnis, Evaluation nicht als Pflichtaufgabe, sondern als Lernprozess für eine effiziente Förderpolitik zu verstehen, ist deshalb zu schärfen. Dies würde auch zu einer Verbesserung der Evaluationskultur beitragen.

Die Vorteile einer Etablierung der Evaluation als Prozess, die u.a. in der Verhinderung von know-how-Verlust durch das Auseinanderfallen von Bewertungsteams und der Verhinderung von Strukturbrüchen in der Methodik liegen, sind gegen Nachteile wie beispielsweise „Betriebsblindheit“ aufzurechnen. Ob eine kontinuierliche oder diskrete Evaluation die geeignetere Form darstellt, ist sicherlich auch von der jeweiligen Maßnahme abhängig. Die diskrete Vorgehensweise kann u.U. die Basis für die bereits oben angesprochenen Begleituntersuchungen schaffen.

4.3 Ergebnisse der AG_3 „Agrarumweltmaßnahmen“ (Reiter, Roggendorf)

TeilnehmerInnen

- Frau Essmann, Herr Preising, Frau Pufahl, Frau Reiter, Herr Roggendorf (jeweils FAL, Braunschweig)
- Herr Geberth (Landesanstalt für Entwicklung der Landwirtschaft und der Ländlichen Räume, Schwäbisch-Gmünd)
- Frau Matzdorf (ZALF, Müncheberg)
- Herr Horlitz, Herr Sander (jeweils ARUM, Hannover)
- Herr Heyer (Universität Halle)
- Herr Wagner (Bundesanstalt für Agrarwirtschaft, Wien)
- Frau Tripolt (Bundesministerium für Land- und Forstwirtschaft, Wien)
- Herr Schramek (Institut für Ländliche Sozialforschung, Frankfurt)
- Frau Eckstein (Technische Universität Weihenstephan, Freising)
- Frau Hefti (Bundesamt für Landwirtschaft, Bern)
- Frau Peter (Büro für Evaluationen, Projektmanagement und Consulting, Bern)

Die Agrarumweltmaßnahmen (AUM) der deutschen Bundesländer zeichnen sich durch eine große Bandbreite aus. Gemeinsam ist allen Programmen[10], dass sie sich aus zwei Maßnahmengruppen zusammensetzen.

– Extensivierungsmaßnahmen, die durch eine Reduzierung der Produktionsmittel oder eine Begrenzung der Viehbesätze je Flächeneinheit gekennzeichnet sind,
– Vertragsnaturschutzmaßnahmen.

1. Inhalte der kapitelspezifischen Fragen/Kriterien/Indikatoren

Die von der EU-Kommission vorgegebenen, kapitelspezifischen Fragen zu den Agrarumweltmaßnahmen konzentrieren sich alleinig auf deren Ressourcenschutzwirkung. Analysiert werden soll der Beitrag der Agrarumweltmaßnahmen zum Schutz der abiotischen Ressourcen Boden und Wasser sowie der biologischen Vielfalt. Zusätzlich sind Landschaftsschutzaspekte und der kulturhistorische Beitrag der AUM zu bearbeiten. Aspekte des Luft- und Klimaschutzes sowie sozioökonomische Aspekte werden im kapitelübergreifenden Fragenbereich behandelt. Gegenüber der vorangegangenen Förderperiode (VO (EWG) Nr. 2078/1992) mit den Zielsetzungen Umweltschutz, Marktentlastung und Einkommenssicherung (für die Landwirtschaft) erfolgt mit dem Bewertungsraster zur VO (EG) Nr. 1257/1999 für die "neuen" Maßnahmen eine eindeutige Prioritätensetzung auf ihre Ressourcenschutzwirkung.

Die Bewertungsfragen unterteilen sich in zwei Kategorien. Erstere, die deutlich überwiegt, zielt auf die Ausweisung von Flächensummen ab. Darzustellen sind die Flächen, die zum Schutz der jeweiligen Ressourcen (Vermeidung von Belastungen und/oder Ressourcenentlastung) gefördert werden. In der zweiten Kategorie von Bewertungsfragen wird nach dem Ausmaß der Ressourcenschutzwirkung gefragt, jedoch nur für ausgewählte Kriterien wie beispielsweise die Ausweisung von Stickstoffsalden bei spezifischen Wasserschutzmaßnahmen.

Das Bewertungsraster wurde von den AG-Teilnehmern als wichtige und bedeutende Arbeitsgrundlage zur Evaluierung der Agrarumweltmaßnahmen eingeschätzt. Alle beteiligten Evaluierer haben das Bewertungsraster angewendet und nur in Teilen, auf der Ebene von Kriterien und Indikatoren, modifiziert. Im Ideal schafft es Transparenz und lässt sowohl einen Vergleich der Mitteleffizienz zwischen den einzelnen Förderbereichen als auch einen Länder- und Staatenvergleich zu. In der Praxis ist aber zur Operationalisierung des Bewertungsrasters durch die verschiedenen Evaluierer ein Bündel ganz unterschiedlicher Methoden eingesetzt worden, wodurch eine Vergleichbarkeit der Bewertungsansätze

[10] Die einzige Ausnahme stellt das Agrarumweltprogramm Berlins dar. Zum Zeitpunkt der Programmgenehmigung waren Vertragsnaturschutzmaßnahmen nicht vertreten.

und –ergebnisse sehr in Frage zu stellen ist. Auch eine bundes- oder gar europaweite Synthese von Bewertungsergebnissen dürfte aus diesem Grund wissenschaftlich kaum haltbar sein. Auf jeden Fall wird das Aggregationsniveau in der Synthese durch die „schwächste“ Auswertung bestimmt werden.

Die mit den kapitelspezifischen Fragen verbundene Konzentration auf den Ressourcenschutz wird von der AG zwar als sinnvoll, in Hinblick auf die Programmbewertung aber als nicht ausreichend erachtet. Neben der Ressourcenschutz-Sicht ist die maßnahmenbezogene Sicht der Programmplanung zu beachten. Die in einer bundesweiten Arbeitsgruppe organisierten Evaluierer der Agrarumweltmaßnahmen haben bereits in früheren Treffen darauf verwiesen, dass in der Bewertung ein nachvollziehbarer Zusammenhang zwischen Maßnahmen und Ressourcenschutzwirkung hergestellt werden muss, damit die Bewertungsergebnisse auf die Konzeption und Steuerung der Maßnahmen Einfluss ausüben können.

Probleme wirft das Bewertungsraster im Detail auf, was die Verwendbarkeit in Teilbereichen erheblich in Frage stellt. Trotz der vorhandenen Evaluierungsleitfäden entstanden bei der Anwendung des Bewertungsrasters Verständnis- und Abgrenzungsprobleme zu einer Reihe von Kriterien und Indikatoren. Zwar hat sich die bereits erwähnte bundesweite Arbeitsgruppe in Deutschland in mehreren gemeinsamen Sitzungen über die Interpretation und Beantwortung der Bewertungsfragen/-kriterien/-indikatoren verständigt. Aber dennoch sind in der Folge Abgrenzungs- und Definitionsprobleme besonders bei den Bewertungsfragen zur Biodiversität und zum Schutz besonderer Arten aufgetreten. Die AG regt an, bei einer Überarbeitung der Leitfäden die Definitionen zu den Fragen/Kriterien/Indikatoren in Zukunft klarer zu fassen.

Anmerkungen zu einzelnen Fragen/Kriterien/Indikatoren:

- Zur Frage des Bodenschutzes wird eine Aufnahme von Kriterien/Indikatoren zu den Themen Bodenstruktur/Humuswirtschaft sowie bodenkundliche Sonderstandorte gewünscht.
- Die KOM-Frage zum Schutz der Grundwasserquantität wurde bereits in der o.a. bundesweiten AG als irrelevant für Deutschland identifiziert und mithin in allen Gutachten nicht beantwortet.
- Meist nicht beantwortet werden konnte der auf der Analyse von Nachbarschaftsbeziehungen beruhende Indikator zum Schutz wertvoller Habitate vor Stoffeinträgen. Hier werden in Zukunft mit GIS-gestützten Förderdaten (InVeKos-GIS) Ergebnisse abzuleiten sein.
- Verschiedene Evaluierer plädieren für die Berücksichtigung der Themen Klima/Luft sowie Tierschutz in den kapitelspezifischen Fragen.

– Das in den Kommissionsfragen enthaltene Begriffsverständnis über das Schutzgut Landschaft (reduziert auf Landschaftsbild) wird von Teilen der AG hinterfragt, ein mehr ökosystemarer sowie soziokultureller Ansatz wird empfohlen.

Außerdem halten die Evaluierer die Untersuchung sozio-ökonomischer Wirkungen innerhalb der kapitelspezifischen Fragen weiterhin für sehr wichtig, weil die Höhe der Prämien eine bedeutende Lenkungsfunktion in Hinblick auf die Teilnahme einnimmt. Die Indikatorenfrage muss dabei in sehr enger Abstimmung mit verwandten Förderbereichen geklärt werden, um programmübergreifenden Effekte besser einschätzen zu können.

In Hinblick auf die zukünftige Weiterentwicklung des Bewertungsrasters gab es innerhalb der AG verschiedene Standpunkte: Neben den bereits beschriebenen Wünschen in Hinblick auf eine Ergänzung plädieren einige Teilnehmer für eine ‚Verschlankung' des Rasters durch Konzentration auf wenige Fragen/Kriterien/Indikatoren von ‚zentraler' Relevanz. Von der Kommission werden in diesem Zusammenhang klare Vorgaben gewünscht, welche Fragen/Kriterien/Indikatoren später in einer Synthese der Länderbewertungen Verwendung finden sollen. Andere Positionen fordern eine insgesamt größere Flexibilität in der Handhabung des Bewertungsrasters, um stärker als bisher auf regionale Besonderheiten eingehen zu können.

2. Beantwortung der kapitelspezifischen Fragen/Kriterien/Indikatoren

Daten/Repräsentativität

Auch hinsichtlich der für die Evaluierung verwendbaren Datensätze bestand in der AG weitgehend Einigkeit. Das von der EU-KOM vorgesehene Monitoring zur Begleitung der Fördermaßnahmen wird von allen Evaluierern wegen des hohen Aggregationsgrades als ungeeignet eingestuft. Zudem ist nicht der Zustand der Bewilligung abgebildet, weil Antragsdaten ausgewertet werden. Bislang haben die Daten daher kaum Verwendung gefunden. Die AG sieht hier erheblichen Überarbeitungsbedarf.

Die Palette der verwendeten Sekundärdaten ist groß: Programm-/Förderdaten, Buchführungsergebnisse aus den Testbetriebsnetzen (auch für physische Input/Outputfaktoren), Landes- und Bundesstatistiken, sonstige Begleitstudien, Literatur. Die Verfügbarkeit und Qualität/Repräsentativität dieser Daten in den einzelnen Bundesländern ist sehr unterschiedlich. Auch Förderdaten sind in einigen Fällen nur eingeschränkt verfügbar oder weisen eine für Evaluierungszwecke zu geringe Detaillierung auf. Die angestrebte Einbeziehung von Umweltdaten zur Verbesserung der Wirkungskontrolle scheitert noch in vielen Bereichen/Ländern an Datenlücken.

Als Primärdatenquelle von hoher Bedeutung wurden Befragungen (Begünstigte/Experten/Multiplikatoren) benannt. Allerdings bringt die Befragung von Begünstigten häufig wenig eindeutige Ergebnisse, insbesondere im Bereich quantitativer Wirkindikato-

ren. Geeigneter scheint sie für die Evaluierung der administrativen Umsetzung im Verbund mit weiteren Befragungen nach dem Prinzip der Triangulation zu sein.

Methodik

Über den Bewertungsansatz der EU-KOM wird überwiegend die Wirkung zum Ressourcenschutz durch die Aufrechnung von potenziell wirksamen Förderflächen ermittelt (... ha Fläche, die zum Schutz von ... beiträgt). Die AG stimmt überein, dass diese quantitative Herangehensweise zu stark vereinfacht. Es fehlt eine nach dem tatsächlich Ausmaß der Wirkung differenzierte Betrachtung. Außerdem bleiben ausschließlich qualitativ beschreibbare Wirkungen oft völlig unberücksichtigt. Einige Teilnehmer monieren auch die mit dem Ansatz verbundene Doppelzählung von Flächen für den Schutz unterschiedlicher Ressourcen. Zwar haben einige Evaluierer versucht, durch die Unterscheidung von Haupt- und Nebenwirkungen eine Differenzierung vorzunehmen, in der Flächenaufrechnung blieb dieser Aspekt dann aber meist wieder unberücksichtigt.

Um die vornannten Beschränkungen zu überwinden, haben die Evaluierer den Versuche unternommen, das Bewertungsraster der EU-KOM weiter zu entwickeln. Als Ziel stand dabei im Vordergrund, die Frage der ökologischen Effizienz von Maßnahmen in der Bewertung zu berücksichtigen. Auch hier divergieren die methodischen Ansätze deutlich (Treffsicherheit/räumliche Äquivalenz/Bewertungsstufen). Es erscheint angeraten, u.a. in Hinblick auf die angestrebte Vergleichbarkeit der Bewertungsansätze den Informationsaustausch – auch mit der EU-KOM - zu diesen Aspekten weiter voran zu treiben (Frage der Messbarkeit/Indikatoren).

Eine Reihe von weiteren methodische Schwierigkeiten, vor allem aus dem Bereich der Wirkungsanalyse, wurde aufgezählt:

- Methodische und praktische Probleme werden von vielen AG-Teilnehmern zunächst im Zusammenhang mit der Wahl des geeigneten Referenzsystems erwähnt. In den MEANS-Handbüchern empfiehlt die EU-Kommission als Analyseinstrument sowohl den Mit-Ohne- als auch den Vorher-Nachher-Vergleich. Beide sind aber in der Evaluierungspraxis oft kaum anwendbar. In einigen Bundesländern kann aufgrund der hohen Teilnahme an den AUM keine repräsentative Referenzgruppe gebildet werden. Weil viele der Teilnehmer an den AUM zudem schon in der vorangegangenen Förderperiode an vergleichbaren Maßnahmen teilgenommen haben, stehen auch für den Vorher-Nachher-Vergleich zu wenig repräsentative Betriebe zur Verfügung.
- Aufgrund des Time-lags treten Schwierigkeiten bei der Nachweisbarkeit von Wirkungen auf, weil selten entsprechend lange Zeitreihen in den Daten zur Wirkungsanalyse vorliegen.
- Vielfach aber, vor allem in Hinblick auf die abiotischen Ressourcen, sind Erhebungen zur Abschätzung der Wirkungen aus Zeit- und Kostengründen gar nicht durchgeführt

worden, sodass auf die Ergebnisse von Fremdstudien zurückgegriffen werden musste. Deren Übertragbarkeit ist aber oft in Frage zu stellen.

– Schließlich wird darauf verwiesen, dass die Ermittlung von Nettoeffekten meist nicht zu leisten ist, weil die Wirkungen exogener Faktoren kaum abgrenzbar sind.

3. Bewertungsergebnisse und Handlungsempfehlungen

Die Evaluierer in der AG haben trotz der aufgezeigten Probleme mit Bewertungsraster und -methoden zu den meisten Fragen, Kriterien und Indikatoren Ergebnisse erarbeiten können. Die Belastbarkeit der Ergebnisse ist unterschiedlich. Die meisten quantitativen Angaben zur Anrechnung von Flächen für die Ressourcenschutzziele sind unter den genannten methodischen Einschränkungen durchweg als belastbar einzustufen. Allerdings ist die Belastbarkeit meist nur innerhalb des jeweils gewählten methodischen Ansatzes gegeben. Zur Gewährleistung der Transparenz über die Ableitung der Ergebnisse sind die gewählten Methoden zu dokumentieren. Zu einigen Indikatoren, vor allem im Bereich der biotischen Ressourcen, mussten Hilfsindikatoren zur Beantwortung herangezogen werden, die ebenso wie die häufig schlechte Datenlage (vor allem Umweltdaten) die Belastbarkeit der Ergebnisse stark einschränken. Auch die Angaben zum Umfang eingesparter Produktionsmittel beruhen in vielen Fällen auf groben Schätzungen.

Alle Evaluierer haben Handlungsempfehlungen aus ihren Ergebnissen abgeleitet und diese in den Ländern mit den zuständigen Fachreferenten diskutiert.

4. Grundlagen für die Ex-post-Evaluierung

In vielen Bundesländern wurde begleitende Systeme zur Wirkungsevaluierung der AUM aufgebaut, um vorher vorhandene Lücken zu schließen und belastbare Einschätzung der Wirkungen zu gewährleisten. Ein Schwerpunkt der Begleituntersuchungen liegt eindeutig im biotischen Bereich, wo i.d.R. die zuständigen Landesämter oder –anstalten oder auch die jeweiligen Programmevaluierer entsprechende Monitoringkonzepte erarbeitet und umgesetzt haben. Große Defizite bestehen im Bereich abiotischer Wirkungen. Hervorzuheben ist die Konzeption der Länder Sachsen und Sachsen-Anhalt, die zusammen mit der Universität Halle-Wittenberg ein Netz repräsentativer Untersuchungsbetriebe zur biotischen und abiotischen Wirkungskontrolle aufgebaut haben.

Die AG sieht für zukünftige Evaluierungsphasen einen hohen Bedarf, zusammen mit den Ländern über eine Nutzen-Kosten-Abwägung einen möglichst breiten Konsens sowohl hinsichtlich der Begleituntersuchungen als auch zu Fragen der Bewertungsmethoden herzustellen. Erfahrungen haben gezeigt, dass eine Beteiligung von Wirtschafts- und Sozialpartnern zu diesen Fragen eingefordert wird.

Alle Teilnehmer der AG sind der Auffassung, dass die entwickelten Methoden ein gute Grundlage für weitere Evaluierungsschritte bilden. Über die Fortführung der Arbeiten im Rahmen eines Up-dates können noch vorhandene Lücken geschlossen und neuentwickelte Konzeptionen breit erprobt werden.

5. Programmplanungsdokumente als Grundlage für die Kapitelbewertung

Interventionslogik und Ex-ante-Bewertung

Die Interventionslogik in der Reihenfolge Bedarf - Ziele - Maßnahmen wurde bei der Programmerstellung häufig nicht eingehalten. So musste in Einzelfällen die SWOT-Analyse für die Agrarumweltmaßnahmen zur Programmgenehmigung nachgearbeitet werden.

In der Regel lag in den Programmplanungsdokumenten eine Ex-ante-Bewertung vor, jedoch von sehr unterschiedlicher Qualität. Für die Wirkungen der Maßnahmen hat die Ex-ante-Bewertung kaum Hinweise geliefert.

Zielbeschreibung und Quantifizierung

Zielsystem und operationelle Ziele der Entwicklungspläne sind für die Evaluation nur bedingt geeignet. Die Neuausrichtung auf Umweltziele hat nur z.T. stattgefunden, Ziele und Maßnahmen wurden nicht unbedingt bedarfsorientiert entwickelt. Daraus ergibt sich einerseits ein Problem der Nachvollziehbarkeit der Interventionslogik und der Evaluation anhand der Zielsetzungen. Demzufolge musste das Ziel-Maßnahmen-System zusammen mit den Ländern für die Evaluierung häufig überarbeitet werden. Andererseits führten diese Schwächen der Programmplanung dazu, dass die Maßnahmen häufig gar nicht erst die eigentlichen Problemgebiete erreichen.

Darüber hinaus besteht eine einfache Möglichkeit der Beeinflussung des Maßnahmenerfolgs, wenn der Zielerreichungsgrad an den aufgestellten Zielen gemessen wird: Kleine Ziele = hoher Zielerreichungsgrad, unabhängig von der Problemlage. Beklagt wurde auch, dass z.T. quantitative Ziele nicht vorliegen bzw. nicht zu den Indikatoren passen.

Empfehlungen für die zukünftige Programmplanung

Die AG sieht die Notwendigkeit, dass die Programmplanung in Zukunft die Umweltzielsetzungen ernster nehmen muss. Umweltziele sind nach Möglichkeit zu quantifizieren. Die Zielsetzungen sollten stärker differenziert und konkretisiert werden; ggf. muss auch formuliert werden, was nicht erreicht werden soll (positive und negative Zielformulierungen). Zudem sollten die beabsichtigten Wirkungen der Maßnahmen genauer beschrieben werden. Nur anhand solcher Vorgaben lässt sich auch eine sinnvolle Ex-ante-Evaluierung durchführen.

Die idealtypische Interventionslogik bedarf zudem eines besseren Timings der einzelnen Arbeitsschritte. Außerdem wurde in der AG gefordert, dass die Interventionslogik der Entwicklungsplänen von der EU-Kommission in den Genehmigungspraxis einer gründlicheren Prüfung unterzogen werden sollte.

6. Evaluationskultur in den Ländern

Zu dieser Themenstellung hat die FAL ausführlich ihre Erfahrungen aus der Evaluierung der sechs nordwestdeutschen Bundesländer berichtet. Ähnlich wie auch in anderen Bundesländern konnte ein sehr konstruktives Verhältnis zwischen Programmbewerter und Auftraggeber hergestellt werden. Beide Seiten haben die Evaluierung schon frühzeitig als fortlaufenden Prozess aufgefasst, der durch einen ständigen gegenseitigen Austausch belebt wurde. Der zu diesem Zweck gegründete Lenkungskreis konnte überdies den Programmkoordinatoren als Plattform zum Erfahrungsaustausch dienen. Ein wichtiger Erfolgsgarant zur Stabilisierung des gegenseitigen Verhältnisses war ein Verhaltenskodex, der zu Beginn des Prozesses im beiderseitigen Einvernehmen definiert wurde und dessen Einhaltung als verpflichtend galt. Darin war festgelegt, dass als Grundlage der Evaluierung die sachbezogene Bewertung dient, die Evaluierung den Ansatz zur ständigen Verbesserung verfolgt, damit sowohl Erfolge als auch Kritikpunkte aufzeigt und der Auftraggeber keinen Einfluss auf die Ergebnisse ausübt.

Die AG teilt insgesamt das Plädoyer der FAL für Evaluierung als Lernprozess. Da in Deutschland die Evaluierungskultur noch in den Kinderschuhen steckt, sind die Erfahrungen aus den einzelnen Bundesländern durchaus als Erfolge zu werten, da durchweg die Chance für eine Fortentwicklung der Programmplanung genutzt wurde. Meist haben sich die Auftraggeber sehr konkrete Empfehlungen mit Umsetzungshinweisen gewünscht. Es hat sich allerdings auch gezeigt, dass Bedenken auf der Umsetzungsebene schnell ‚zur Schere im Kopf' werden können.

4.4 Ergebnisse der AG_4 „Dorf- und ländliche Regionalentwicklung" (Koch, Grajewski)

TeilnehmerInnen:

- Herr Becker, Herr Eberhardt Frau Grajewski, Frau Koch, Frau Raue, Herr Schrader, Herr Tietz, (jeweils FAL)
- Herr Großkopf (Uni Hohenheim),
- Herr Elbe (SPRINT GbR),
- Frau Trefflich (Landgesellschaft Sachsen-Anhalt),
- Herr Thissen (BMVEL),

– Herr Ortner (Bundesanstalt für Agrarwirtschaft, Wien),
– Herr Thomas (FH Weihenstephan),

In der AG_4 waren Evaluatoren von Maßnahmen aus dem Förderkapitel 9 „Förderung der Anpassung und Entwicklung von ländlichen Gebieten“ versammelt. Dieses Förderkapitel beinhaltet eine große Bandbreite an Maßnahmen, beginnend bei der Förderung von Flurbereinigung, Dorferneuerung, Diversifizierung landwirtschaftlicher Betriebe und Vermarktung landwirtschaftlicher Produkte bis hin zum Küstenschutz. Entsprechend breit gefächert waren auch die Arbeitsschwerpunkte und methodischen Ansätze der teilnehmenden Evaluatoren.

1. Inhalte der kapitelspezifischen Fragen/Kriterien/Indikatoren

Die Inhalte der fünf kapitelspezifischen Fragen (Einkommen, Lebensqualität, Beschäftigung, Strukturmerkmale, Umwelt) wurden von allen Evaluatoren als relevant erachtet. Die Fragen decken die thematische Bandbreite möglicher Wirkungen dieses Förderkapitels voll ab, sodass in keinem Fall eine Ergänzung von Fragen für notwendig gehalten wurde.

Auch die Kriterien sind überwiegend relevant, wenngleich einzelne Kriterien (wie z.B. „Verringerung der Abgelegenheit“) auf die in den verschiedenen Ländern angebotenen Maßnahmen nicht zutrafen. Vereinzelt wurden Kriterien ergänzt, wie z.B. die Verbesserung der Standortfaktoren bei der Frage 4 „Strukturmerkmale der ländlichen Wirtschaft“.

Der wichtigste Diskussionspunkt bei diesem Thema waren die kapitelspezifischen Indikatoren. Sie sind nicht für alle Maßnahmen relevant und zeichnen sich insgesamt dadurch aus, dass sie schwierig zu erheben sind und einen hohen Detaillierungsgrad aufweisen. Die einzelnen Evaluatoren haben die Bearbeitungsschwierigkeiten mit den Indikatoren sehr unterschiedlich gehandhabt. So wurden z.B. in Sachsen-Anhalt für die Maßnahme Flurbereinigung keine Indikatoren ergänzt, während der Indikatorenkatalog der EU-Kommission in Österreich als Vorschlagsliste begriffen, die nach eigenem Ermessen teils nicht beantwortet, teils durch zahlreiche eigene Indikatoren ergänzt wurde. Allen Evaluatoren erschien die feine Aufgliederung der Indikatoren, wie z.B. der Beschäftigungsmöglichkeiten (nach Geschlecht, Alter, Sektoren) nicht sachgerecht, da schon die Anzahl der geschaffenen Beschäftigungsmöglichkeiten insgesamt nur unter großen methodischen Schwierigkeiten darstellbar ist. Viele Indikatoren wurden daher nicht beantwortet, andere wurden im Sinne einer Beschreibung von Wirkungen anstelle einer Quantifizierung abgeändert. Vielfach wurden Realisierungsindikatoren (z.B. über die Art und Anzahl durchgeführter Projekte) hinzugefügt, um die vorgegebenen Wirkungsindikatoren indirekt beantworten zu können. Zudem wurde der Indikatorenkatalog von fast allen Evaluatoren um zusätzliche, meistens beschreibende Indikatoren ergänzt, um die Band-

breite der mit den Fragen bzw. Kriterien angesprochenen Wirkungen abdecken zu können.

Insgesamt stand für die Evaluatoren die prinzipielle Beantwortung der Bewertungsfragen im Vordergrund, während die Kriterien und Indikatoren diesem Zweck untergeordnet und gegebenenfalls verändert wurden.

2. Beantwortung der kapitelspezifischen Fragen/Kriterien/Indikatoren

Bei der Beantwortung der kapitelspezifischen Fragen, Kriterien und Indikatoren sind verschiedenste Probleme aufgetreten. Die Evaluatoren haben teils unterschiedliche, teils ähnliche Lösungswege gewählt, um mit diesen Problemen umzugehen.

Für alle Evaluatoren lag ein Problem in der unzureichenden Datenlage. Nach einhelliger Meinung sind die Monitoringdaten der EU aufgrund ihrer hohen Aggregationsstufe für eine Evaluation gänzlich ungeeignet. In allen Ländern wurden daher in einem ersten Schritt einzelprojektbezogene Daten, die teilweise von verschiedensten Stellen gehalten werden, angefordert und zusammengestellt. Bei dem teilweise erst sehr geringen Umsetzungsstand einiger Maßnahmen stellen Bewilligungsdaten die einzige Quelle für Aussagen über die Projekte dar. Daher hat der größte Teil der Evaluatoren für die Zwischenbewertung Bewilligungsdaten genutzt. Dies birgt die Gefahr, dass nach Beendigung der Projekte zwischen Bewilligungsdaten und tatsächlicher Realisierung Unterschiede vorliegen. Eine Evaluatorengruppe[11] konnte dagegen weitestgehend mit Daten abgeschlossener Projekte arbeiten, da in den bearbeiteten Ländern bereits eine umfangreichere Anzahl abgeschlossener Projekte vorlag.

Methodische Probleme wurden in der AG_4 unter anderem bei der geforderten Darstellung von Netto- und Mitnahmeeffekten identifiziert. Hierzu hat kein Evaluator eine Methodik gefunden, die geeignet ist, diese Effekte abschließend und umfassend darzustellen. Allerdings gab es Übereinstimmung darin, diese Effekte so weit möglich darzustellen und zu beschreiben. Hierbei müssen die getroffenen Annahmen transparent gemacht werden.

Einen weiteren methodischen Problempunkt stellt die Erfassung und Darstellung von indirekten Wirkungen dar. Im Verlauf der Diskussion wurde deutlich, dass alle Evaluatorengruppen hier den Weg der Beschreibung gewählt haben. Einzig der baden-württembergische Bewerter hat über die Beschreibung hinausgehend versucht, den entstandenen Nutzen auch wertmäßig auszudrücken.

---

[11] Dies betrifft die Evaluatoren der FAL, die die Länderprogramme aus Bremen, Hamburg, Hessen, Niedersachsen, Nordrhein-Westfalen und Schleswig-Holstein bewerteten (Sechs-Länder-Evaluation).

Die Problematik der in den EU-Vorgaben geforderten Referenzsituation hat keiner der anwesenden Evaluatoren lösen können. Für die im Förderkapitel 9 enthaltenen Maßnahmen ist es in der Regel nicht möglich, eine Vergleichsgruppe zu finden. Gerade für die bedeutsamen Maßnahmen Dorferneuerung und Flurbereinigung ist dies leicht nachvollziehbar. Aktuell lassen sich keine Dörfer oder Gemarkungen finden, die mit den geförderten vergleichbar sind, in denen aber noch nie Förderung stattgefunden hat.

Einerseits ist man als Evaluator bestrebt, zuverlässige Aussagen zur Beratung von Politik und Verwaltung zu treffen. Andererseits empfanden viele Evaluatoren die Beantwortung der Fragen als „Gratwanderung“ zwischen wissenschaftlicher Genauigkeit und Belegbarkeit und einer hinreichenden Aussagekraft.

Im Hinblick auf die kapitelspezifischen Fragen, Kriterien und Indikatoren wurde vielfach die Haltung der EU-Kommission kritisiert, die zunächst einen „verbindlichen“ Katalog aufgestellt und später, wenn auch nicht offiziell, mehr Flexibilität in der Beantwortung zugestanden hat. Einige Evaluatoren hätten sich viel Arbeit in der Auseinandersetzung mit den Indikatoren erspart, wenn der Grad ihrer Verbindlichkeit von Anfang an klarer gewesen wäre.

3. Bewertungsergebnisse (qualitativ/quantitativ) und Handlungsempfehlungen

Die zur Zwischenbewertung vorliegenden Bewertungsergebnisse variieren je nach Umsetzungsstand der einzelnen Maßnahmen. Für Maßnahmen, bei denen bisher nur wenige Projekte bzw. Projekte mit längerer Laufzeit begonnen oder umgesetzt wurden, liegen nur erste vorläufige Ergebnisse vor. Bei anderen Maßnahmen mit einer größeren Anzahl abgeschlossener Projekte existieren dagegen bereits belastbarere Ergebnisse. So konnte z.B. der österreichische Evaluator für die Maßnahme Vermarktung landwirtschaftlicher Qualitätsprodukte bereits gut belastbare Ergebnisse in fast allen Fragebereichen mit Ausnahme der Fragen zur Einkommensentwicklung erzielen.

Handlungsempfehlungen bezogen auf die Ausgestaltung von Maßnahmen wurden vor dem Hintergrund der wenigen und vorläufigen Ergebnisse der Förderung nur sehr vereinzelt ausgesprochen. Bezüglich einzelner Maßnahmen kommt die Schwierigkeit hinzu, dass die Länder neben den Förderprogrammen aus der VO (EG) Nr. 1257/1999 auch eigene Maßnahmen zur ländlichen Entwicklung (Art.-52-Maßnahmen) durchführen. Die Evaluation erfasst somit nur einen Teil des Wirkungsspektrums, sodass mögliche Handlungsempfehlungen „ins Leere laufen“.

Alle deutschen Evaluatorenteams haben einen Schwerpunkt ihrer Arbeit auf die Untersuchung der administrativen Umsetzung der Maßnahmen gelegt. Hierzu wurden vielfach Handlungsempfehlungen gegeben, ebenso zur Haltung und Aufbereitung von Förderdaten für weitere Evaluationen. In der Sechs-Länder-Evaluation konnten aufgrund der Ver-

gleichsmöglichkeiten häufig Empfehlungen in Form von „best-practice-Beispielen“ abgeleitet werden.

Eine Handlungsempfehlung, die in der Diskussion zwischen den Evaluatoren entstand, ist nach innen gerichtet und betrifft das System der Bewertungsfragen, das auch im Hinblick auf den Austausch zwischen den Evaluatoren weiter diskutiert und angepasst werden sollte.

4. Welche Grundlagen wurden für die Ex-post-Evaluierung geschaffen?

Dieses Thema wurde in der AG_4 nicht diskutiert.

5. Programmplanungsdokumente als Grundlage für die Kapitelbewertung

Dieses Thema wurde in der AG_4 nicht diskutiert.

6. Evaluationskultur in den Länder

In der Diskussion über das Verhältnis zwischen Auftraggebern und Evaluatoren wurde deutlich, dass die Evaluatoren insgesamt sehr zufrieden mit diesem Verhältnis sind. Der Kontakt ist durch eine kooperative Haltung auf beiden Seiten und Kooperationsstrukturen mit viel Austausch geprägt. Im Zusammenhang mit der Zwischenbewertung stellt die Präsentation und Diskussion der Ergebnisse bei den meisten Evaluatoren den Schlusspunkt der Bearbeitung dar. Informationen über den konkreten Umgang mit den erarbeiteten Ergebnissen und Empfehlungen im Land liegen bei den meisten Evaluatoren nicht vor. Daher wurde diskutiert, ob die EU-Kommission die Umsetzung von Empfehlungen stärker überwachen sollte. Im Unterschied zu vergangenen Förderperioden nimmt sie diese Aufgabe jedoch nicht mehr wahr.

Über diese kapitelspezifischen Erfahrungen hinaus wurden in der Arbeitsgruppe auch einige grundsätzlichere Aspekte hinsichtlich der Unabhängigkeit der Evaluatoren angesprochen. Hier wurde unter anderem diskutiert, dass das Land die Ergebnisse der Bewertung z.B. durch die Auswahl des Evaluators, durch die Weitergabe von Informationen und die Vergabe von Folgeaufträgen beeinflussen kann. Die Frage ist allerdings, welche Alternativen z.B. zur Wahl des Evaluators durch das Land vorliegen. In diesem Zusammenhang wurde eine stärkere Mitsprache der EU-Kommission angesprochen. Allerdings ist wurde als Gegenargument vorgebracht, dass einem von Außen bestimmten Evaluator weniger Vertrauen im Land entgegengebracht wird. Ein abschließendes Ergebnis wurde in der Diskussion nicht erzielt.

7. Sonstige diskutierte Themen

Außerhalb der vorgegebenen Tagesordnung wurde das kapitelspezifische Problem des großen Maßnahmenspektrums innerhalb des Kapitels 9 diskutiert. Mit der Anforderung,

das gesamte Förderkapitel zu bewerten, sind die Evaluatoren sehr unterschiedlich umgegangen. Überwiegend wurde der Schwerpunkt auf die Evaluation einzelner Maßnahmen (Flurbereinigung, Dorferneuerung etc.) gelegt, wobei jeweils eine kurze kapitelbezogene Zusammenfassung am Schluss vorgenommen wurde. Lediglich in der Sechs-Länder-Evaluation wurde der Versuch einer Synthese aller Maßnahmen im Bericht unternommen. Dies zog jedoch die Notwendigkeit nach sich, einen umfangreichen Materialband anzuhängen, in dem die einzelnen Maßnahmen dargestellt sind.

Die Bundesländer als Auftraggeber sind in erster Linie an Ergebnissen der Maßnahmen interessiert, weniger an einer Gesamtbewertung des Kapitels insgesamt. Dem steht die Zielsetzung der EU-Kommission entgegen. Der Bericht kann dadurch, dass das Kapitel als Ganzes bewertet wird, eine größere Wirkung erzielen und die Zusammenhänge zwischen den Maßnahmen besser darstellen. Die große Heterogenität der Maßnahmen im Kapitel 9 hat allerdings bisher eine solche gemeinsame Bewertung erschwert. Daher hat sich in der Diskussion die Frage gestellt, wie das Maßnahmenspektrum zukünftig zusammengesetzt werden sollte. Eine Möglichkeit wäre hier, sich im Kapitel 9 auf Maßnahmen zu konzentrieren, die stärker auf das Ziel der ländlichen Regionalentwicklung ausgerichtet sind.

4.5 Ergebnisse der AG_5 „Forstwirtschaft“ (Bresemann)

TeilnehmerInnen:

- Frau Bresemann, Herr Englert (jeweils BFH)
- Herr Dr. Otto (Landgesellschaft Sachsen-Anhalt mbH, Halle)
- Herr Dr. Lüdigk (ISW, Institut für Strukturpolitik und Wirtschaftsforschung, Halle)
- Herr Dr. Jirikowski (BFW Forstliche Ausbildungsstätte, Gmunden)
- Herr Majer (BMLFUW, Wien)
- Herr Bieling (Universität Freiburg)

Das Angebot forstlicher Fördermaßnahmen gliedert sich in zwei für die Evaluation getrennte Bereiche der Erstaufforstung und der sonstigen forstwirtschaftlichen Maßnahmen. Dabei handelt es sich einerseits um Maßnahmen, die im Rahmen der Gemeinschaftsaufgabe zur Verbesserung der Agrarstruktur und des Küstenschutzes (GAK) gefördert werden und andererseits um reine Landesmaßnahmen, die sich in der Ausgestaltung stärker unterscheiden. Der Schwerpunkt der Umsetzung liegt in der Regel bei den im Rahmen der GAK geförderten Maßnahmen und reicht von waldbaulichen Maßnahmen wie den Umbau von Mono- in Mischkulturen oder Jungbestandspflege über Maßnahmen aufgrund neuartiger Waldschäden, bis hin zur der Förderung von forstlichem Wegebau oder der Förde-

rung forstwirtschaftlicher Zusammenschlüsse. Speziellere Maßnahmen, wie Vertragsnaturschutz, die Nutzung von Holz als energetische Ressource oder aber Maßnahmen, die in der GAK-Ausgestaltung eher unattraktiv sind, werden alternativ als EU-kofinanzierte Landesfördermaßnahme angeboten. Abgesehen davon gibt es abhängig von der naturalen Ausstattung des jeweiligen Landes bzw. Bundeslandes inhaltliche Förderschwerpunkte. So spielt in Gebirgslagen der forstliche Wegebau (in Österreich der Forststraßenbau) eine größere Rolle, als im gut erschlossenen Flachland.

Zusammenfassend lässt sich sagen, dass die forstliche Förderung überwiegend den Charakter einer investiven Förderung hat, deren Ergebnisse oder Effektivität oft erst langfristig konkret beurteilt werden können. Viele Wirkungen, die in den kapitelspezifischen Fragen und im Rahmen der Entwicklung des ländlichen Raums eine größere Rolle spielen, wie z.B. sozioökonomische Effekte, stellen sich mit der Umsetzung der Maßnahmen kaum messbar und lediglich indirekt ein. Das machte eine Bewertung nach den klassischen Evaluationskriterien oftmals schwierig.

Da die Evaluation zur Begleitung forstlicher Förderprogramme ein relativ neues Instrument darstellt, mussten die Grundlagen für eine Bewertung (sowohl die Datenbereitstellung als auch methodische Grundlagen) erst geschaffen werden. Dabei erwies sich die Datenerhebung als sehr zeitintensiv.

1. Inhalte der kapitelspezifischen Fragen

Die kapitelspezifischen Fragen gliedern sich für den Bereich der forstlichen Förderung in drei Komplexe. Es werden erstens die Erhaltung und Entwicklung forstlicher Ressourcen (z.B. Holzvorräte, Kohlenstoffspeicherung), zweitens sozioökonomische Effekte (z.B. Beschäftigung, Einkommen, Erholungswirkungen) und drittens die Erhaltung und Entwicklung der vielfältigen Schutzfunktionen der Wälder (Schutz gegen biotische und abiotische Gefahren sowie Schutzleistungen wie z.B. Biotop- und Artenschutz) abgefragt. Grundsätzlich sind alle Fragen und Kriterien im Sinne des EU-Zielsystems für die forstliche Förderung relevant, jedoch spiegeln die Fragestellungen nicht immer die Zielsysteme der Länder wieder. Vielfach schließen die in den Programmen genannten forstlichen Ziele die Fragestellungen lediglich indirekt ein. Dies betrifft beispielsweise sowohl die Erhöhung der Holzvorräte, die Fragen zur Kohlendioxidbindung, wie auch die meisten sozioökonomischen Fragestellungen. Die Relevanz ergibt sich in diesen Fällen aber durchaus im größeren Rahmen, also im Zusammenhang mit der Entwicklung des ländlichen Raums.

Im Gegensatz zu den Fragen und Kriterien erwiesen sich die Indikatoren zum größten Teil als problematisch. Je nach dem Angebot der Maßnahmen in den entsprechenden Bundesländern waren einzelne Indikatoren nicht relevant (z.B. Indikatoren zu dem Kriterium 2.C.2 – Schutz von Flächen, die keine Holzflächen sind), bzw. konnten so nicht beant-

wortet werden. Vielfach litt die Relevanz unter der teilweise sehr speziellen Formulierung oder den geforderten indikatorbezogenen Einheiten (z.B. Kostensenkung/m^3 statt Kostensenkung/ha). Andere Formulierungen waren demgegenüber nicht hinreichend präzisiert (Indikator 3.A-2.1 zu den „kritischen Standorten“), oder Fragestellungen wiesen Überschneidungen auf (insbesondere zu Schutzfunktionen). Inhaltliche Ergänzungen oder Änderungen zu den kapitelspezifischen Fragen wurden seitens der Evaluatoren in die Berichte eingearbeitet. So wurde beispielsweise für den Bereich der Erstaufforstung eine regionale Verteilung nach waldarmen und waldreichen Gebieten vorgeschlagen.

Für alle drei Ebenen des kapitelspezifischen Fragenkatalogs und für den Großteil der Fragestellungen selbst war der abgelaufene Förderzeitraum zu kurz, um die Effizienz und Effektivität konkret zu bewerten. Dies betrifft alle drei genannten Komplexe zu den forstlichen Fragestellungen, insbesondere aber die Fragen zu der Erhaltung und Entwicklung forstlicher Ressourcen.

2. Beantwortung der kapitelspezifischen Fragen

Die Beantwortung der kapitelspezifischen Fragen gestaltete sich aufgrund der schlechten Dokumentation der Ausgangslage für alle drei forstlichen Themenkomplexe schwierig. Einige der notwendigen Daten lagen in nicht auswertbarer Form (z.B. in Förderanträgen bei den Forstämtern) vor. Die zur Verfügung stehenden Sekundärdaten aus dem begleitenden Monitoring erwiesen sich oft als unzureichend. Dabei handelte es sich in der Regel um hochaggregierte Outputdaten, die in Bezug auf die Fragestellungen wenig Aussagekraft hatten. Vielfach wurde eine nachträgliche Erhebung von (Output-) Daten notwendig, die mit einem erheblichen Aufwand für alle Beteiligten verbunden war.

Die Erhebung weiterer Primärdaten, beispielsweise im Zuge von Befragungen, konnte teilweise nicht die gewünschte inhaltliche und formale Datenqualität erbringen. So stießen Befragungen zu den sozioökonomischen Themenkomplexen an ihre Grenzen. Ebenso ergab die Diskussion in der Arbeitsgruppe, dass aufgrund der ungenügenden Datengrundlage und der zeitlichen Restriktionen eine repräsentative Befragung nicht möglich war.

Aus den genannten Gründen lagen wesentliche Informationen zur Beantwortung der kapitelspezifischen Fragen in der geforderten Tiefe nicht vor. Anhand der vorhandenen Daten konnten lediglich Hochrechnungen, oftmals mit Durchschnittswerten, vorgenommen werden. Schriftliche und mündliche Befragungen stellten trotz der genannten Nachteile eine wichtige Informationsquelle dar. In einigen Bundesländern und in Österreich wurden Fallbeispiele durchgeführt, die wiederum mit einem hohen Aufwand für die Evaluatoren verbunden waren.

Eine Referenzsituation lag für die Beantwortung der forstlichen Fragestellungen nicht vor. Es konnten daher weder vorher-/nachher- noch mit-/ohne-Vergleiche durchgeführt werden.

Erste Lösungsansätze, wie die standardisierte Auswertung von Antragsdaten, sind für nachfolgende Evaluationen hilfreich, können jedoch die vorhandenen Defizite nicht vollständig beheben. Aufgrund des hohen Aufwandes zur Erfassung zusätzlicher Antragsdaten, der folgenden Auswertung dieser Daten und der mangelnden Repräsentativität von Befragungen wurde in der Arbeitsgruppe diskutiert, ob die Nutzung von Fallbeispielen in Zukunft nicht einen höheren Stellenwert einnehmen sollte. Die Unterstützung durch andere Institutionen während der Evaluation in vielen Fällen weiterhin wichtig bleiben. In Österreich wird ein Monitoringsystem installiert, welches Vorher-/Nachher- und Mit-/Ohne-Vergleiche ermöglichen soll.

3. Bewertungsergebnisse

Die Teilnehmer der Arbeitsgruppe bestätigten, dass zu fast allen Fragestellungen quantitative Angaben möglich waren. Das betrifft vor allem die Indikatoren, deren Beantwortung Outputdaten, wie Flächenangaben, zur Grundlage hatten. In anderen Fällen konnten Hochrechnungen oder forstlich anerkannte Durchschnittswerte herangezogen werden. Verbale Beschreibung von Wirkungsketten finden sich vermehrt bei den sozioökonomischen Fragestellungen, so z.B. bei Fragen zu den Vermarktungsmöglichkeiten von bestimmten Holzsortimenten.

Die Belastbarkeit der Ergebnisse nach Hochrechnungen oder Durchschnittswerten ist befriedigend, auch wenn z.B. die Holzvorräte nur sehr grob berechnet werden konnten, weil für die sonstigen forstwirtschaftlichen Maßnahmen keine Angaben zu den Baumartenanteilen vorlagen. Immerhin handelt es sich bei den meisten Vorgehensweisen und Grundlagen um allgemein anerkannte forstliche Standards, so dass eine Vergleichbarkeit hergestellt werden kann.

Weniger belastbar sind die Ergebnisse zu den sozioökonomischen Fragestellungen und der Berechnung der Kohlendioxidbindung, da die Vorgehensweise hier nicht immer einheitlich ist und nur wenige auf die Anforderungen zugeschnittenen Vorgaben vorlagen. Ein besonderes Problem betraf die Fragen zu den Maßnahmen in Natura 2000- oder anderen Schutzgebieten, da die durchgeführten Maßnahmen größtenteils keinem Schutzgebiet zugeordnet werden konnten. Diese Fragestellungen nehmen im kapitelspezifischen Fragenkatalog jedoch eine nicht unbedeutende Stellung ein. Dementsprechend lautete eine Handlungsempfehlung, die Angaben zur Lage der Förderfläche in die Begleit-/Monitoringsysteme aufzunehmen. Für Österreich wurde die Empfehlung gegeben, das Controlling im Schutzwald umzusetzen.

Kontrovers wurde die Einrichtung von maßnahmenbezogenen Gebietskulissen in der Arbeitsgruppe diskutiert. Während zum einen die Meinung geäußert wurde, Gebietskulissen würden die Möglichkeiten der forstlichen Förderung nur unnötig einschränken, sahen andere die Chance zur gezielten Steuerung der Maßnahmen im Sinne klarer Zielsetzungen und damit auch zu einer verbesserten Prüfung von Effektivität und Effizienz.

4. Grundlagen für weitere Evaluationen

Wie schon unter 2. erwähnt, wurden teilweise erweiterte oder neue Datenerfassungssysteme angeschoben. In einigen Fällen werden brauchbare Antragsdaten neuerdings EDV-technisch erfasst und aufgearbeitet. Dafür war es den Teilnehmern der Arbeitsgruppe wichtig, dass die Vorgaben für die Evaluation möglichst erhalten bleiben, um diesbezüglich Kontinuität für alle Beteiligten zu gewährleisten. In Österreich wurden die beiden Zahlstellen zusammengelegt, um die Abwicklung zu erleichtern.

Die Defizite bezüglich der Datengrundlage für die Evaluation können diese Lösungsansätze nicht vollständig ausgleichen. Ein Großteil der Informationen, wie sie in den kapitelspezifischen Fragen gefordert werden, werden auch für weitere Evaluationen nicht zur Verfügung stehen. Dafür wäre ein spezielles Begleitsystem erforderlich, welches in keinem Bundesland entwickelt wurde.

Grundlagen für weitere Evaluationen wurden jedoch vor allem methodisch geschaffen. Das betrifft standardisierte Hochrechnungen, Musterfragebögen und langfristige Fallbeispiele genauso wie den Methodikaustausch zwischen den Evaluatoren.

5. Qualität der Programmplanungsdokumente

Obwohl die Interventionslogik der Länder mit der Programmgenehmigung seitens der EU akzeptiert wurde, empfanden die Teilnehmer der Arbeitsgruppe häufig die jeweilige Interventionslogik weder als stringent noch als den neuen strategischen Anforderungen angepasst.

In allen Programmplanungsdokumenten wurden Ziele beschrieben und in den meisten Fällen wurde ein entsprechender Output quantifiziert. Wie sich bei der Evaluation herausstellte, waren diese Quantifizierungsansätze sehr oft unrealistisch. Es wurden zwar SWOT-Analysen durchgeführt, diese Ergebnisse mündeten jedoch nicht in eine steuerbare Strategie, die beispielsweise eine Darstellung von Zielregionen (Gebietskulissen) oder Zielgruppen beinhalten würde. Die Ex-ante-Evaluierung (falls vorhanden) war in den wenigsten Fällen hilfreich. Dort fanden sich nur vereinzelt konkrete Vorgaben.

Empfehlungen für die zukünftige Programmplanung beziehen sich zum einen auf eine stärkere Anpassung der Maßnahmen/Strategie an den neuen Rahmen der forstlichen Förderung, nämlich an die Entwicklung des ländlichen Raums und damit an eine verstärkt

sozioökonomische Ausrichtung. Maßnahmen mit einer anderen als einer forstlichen Intention sollten ausgelagert und in anderen Bereichen gefördert werden (z.B. Hochwasserschutz). Konkret auf die Maßnahme bezogen sollte die Erstaufforstung attraktiver gestaltet und in Regionen gefördert werden, in denen die Waldvermehrung vorrangiges Ziel ist. Eine Vereinfachung der Abwicklung der forstlichen Förderung wurde von allen Teilnehmern der Arbeitsgruppe begrüßt.

6. Evaluationskultur

Für den forstlichen Bereich stellt die Evaluation ein relativ neues Instrument zur Begleitung der Förderprogramme dar. Folglich ergaben sich für alle Beteiligten neue Anforderungen und sowohl Auftragnehmer als auch Auftraggeber fassten dieses als einen intensiven Lernprozess auf. Die Zusammenarbeit gestaltete sich nach Meinung aller Teilnehmer der Arbeitsgruppe dementsprechend gut, auch weil größtenteils ein ausgeprägtes Eigeninteresse an den Ergebnissen bei den Auftraggebern (Bundesländer) über die reine EU-Berichterstattung hinaus vorhanden war.

4.6 Zusammenfassung der Ergebnisse der Arbeitsgruppen (Grajewski)

In diesem Kapitel erfolgt die Zusammenfassung der Ergebnisse der Diskussionen in den einzelnen Arbeitsgruppen. Dabei zeigt sich, dass einige Themen über alle Arbeitsgruppen hinweg zu tendenziell ähnlichen Ergebnissen geführt haben, andere dagegen grundsätzlich unterschiedlich gehandhabt wurden.

1. Inhalte der kapitelspezifischen Fragen/Kriterien/Indikatoren

Die kapitelspezifischen Fragen wurden von allen Arbeitsgruppen grundsätzlich als relevant eingeschätzt. Ergänzungen erfolgten nur im Bereich der AG_2 „Ausgleichszulage“ im Hinblick auf die Erfordernisse der Zielsysteme einzelner Länder. Die AG_3 „Agrarumweltmaßnahmen“ regt zukünftig die Ergänzung einer Frage zu den sozioökonomischen Wirkungen ihrer Maßnahmen an.

Anders ist die Einschätzung bei den Kriterien. Sie waren nicht in jedem Fall relevant, wurden teilweise modifiziert und durch neue Kriterien ergänzt. Zusätzlich sind mit der Bearbeitung der Kriterien in den einzelnen Arbeitsgruppen unterschiedliche Probleme verbunden gewesen. Vor allem die AG_2 „Ausgleichszulage“ hat hier angemerkt, dass bestimmte Kriterien ihres Kapitels wichtige Betriebsgruppen nicht erfassen und die Kriterien in den Ländern nur mangelhaft mit quantifizierten Zielen hinterlegt sind.

Die größten Schwierigkeiten in der Bearbeitung haben in allen AGs die Indikatoren bereitet. Sie wurden durchweg von allen Gruppen als oftmals zu kompliziert und zu speziell empfunden. Die Datengrundlagen, um die Indikatoren wie gefordert darzustellen, fehlten

bei den Arbeitsgruppen für viele Indikatoren weitgehend. In diesen Fällen wurden zumeist Hilfsindikatoren herangezogen. Zusätzlich erfolgte die Ergänzung von Indikatoren, wenn die vorgegebenen nicht ausreichten, um die Wirkungen der Maßnahmen abzubilden.

Zudem berichten einige Arbeitsgruppen von Verständnis- und Abgrenzungsproblemen bei Kriterien und Indikatoren. Die in den Leitfäden enthaltenen Erläuterungen reichen nicht aus, um Kriterien und Indikatoren eindeutig zu definieren; die verwendeten Begriffe sind zudem oftmals auch unklar.

2. Beantwortung der kapitelspezifischen Fragen/Kriterien/Indikatoren

Ein grundsätzliches Problem bei der Beantwortung der Fragen, Kriterien und Indikatoren wurde in allen Arbeitsgruppen thematisiert: Die Verfügbarkeit und Qualität von Daten. Die Daten des EU-Monitoringsystems wurden dabei von allen Arbeitsgruppen als unzureichend eingestuft. Die in diesem System vorhandenen Daten wiesen für die Halbzeitbewertung ein zu hohes Aggregationsniveau auf. Alle Evaluatoren haben daher andere Sekundärdatenquellen genutzt, um die Bewertungsfragen zu beantworten. Zudem wurden in den meisten Fällen eigene Primärerhebungen durchgeführt. Beide Datenquellen führen jedoch zu neuen Problemen, z.B. bezüglich der Repräsentativität von Daten. Die Sichtung und Aufarbeitung zusätzlicher Sekundärdaten und die Erhebung eigener Daten ist außerdem mit einem großen Zeit- und Kostenproblem behaftet.

Ein weiteres grundsätzliches methodisches Problem, dass alle Arbeitsgruppen beschäftigt hat, ist die Herleitung von Nettowirkungen. Die EU-Kommission sieht hierfür methodisch insbesondere Mit-Ohne-Vergleiche oder Vorher-Nachher-Vergleiche vor. In den meisten Arbeitsgruppen herrschte Einigkeit zwischen den Evaluatoren, dass es nicht möglich ist, geförderte und nicht-geförderte Betriebe abzubilden oder geeignete Referenzszenarien zu entwickeln. Dies ist teilweise darauf zurückzuführen, dass vielfach alle potenziell förderfähigen Betriebe oder Flächen auch gefördert werden oder dass für Vergleichsgruppen keine oder keine hinreichend differenzierten Daten vorliegen. Daher schied bei der Bearbeitung der meisten Förderkapitel die Durchführung eines Mit-Ohne-Vergleichs aus. Ähnliches gilt für den Vorher-Nachher-Vergleich. In vielen Förderkapiteln besteht schon eine lange Fördertradition, sodass die Begünstigten der jetzigen Förderperiode häufig auch schon vorher Zuwendungen bekamen und somit eine geeignete Ausgangssituation fehlt.

3. Bewertungsergebnisse (qualitativ/quantitativ) und Handlungsempfehlungen

Die Darstellung und Belastbarkeit der bisherigen Ergebnisse ist zwischen den Arbeitsgruppen sehr unterschiedlich. Bei den von AG_1 und AG_2 bearbeiteten Förderkapiteln waren zur Halbzeitbewertung nur sehr begrenzt Ergebnisse darstellbar. Dies sind dann im Wesentlichen Aussagen zu einzelnen Indikatoren, Tendenzaussagen oder auch Ergebnisse zu einzelnen Fragen. Die Evaluatoren der AG_4 und AG_5 berichten dagegen von großen

Variationen bei der Beantwortung der kapitelspezifischen Bewertungsfragen. Teilweise sind hier für einzelne Fragen bzw. Maßnahmen bereits belastbare quantifizierte Ergebnisse geliefert worden, in anderen Bereichen sind nur verbale Beschreibungen von Wirkungsketten mit einer entsprechend geringeren Belastbarkeit möglich gewesen. Bei den von der AG_3 bearbeiteten Agrarumweltmaßnahmen wurden Ergebnisse zu allen Bewertungsfragen in den Halbzeitbewertungen dargestellt. Ihre Belastbarkeit ist innerhalb des jeweils gewählten methodischen Ansatzes als gegeben einzustufen.

Im Hinblick auf Handlungsempfehlungen wurden unterschiedliche Aspekte und Ebenen in den Arbeitsgruppen diskutiert. AG_1 richtet sich hier vor allem an die EU-Kommission, Bund und Länder mit den Empfehlungen, die Förderung zu entflechten, die Komplexität der Programme zu verringern und präzisere Ziele zu formulieren. AG_2 hat Empfehlungen zur Verbesserung der Belastbarkeit ihrer Ergebnisse formuliert. Bei den AGs 3 bis 5 standen konkrete Empfehlungen zu Maßnahmen und der Umgang mit ihnen in den Ländern im Vordergrund der Diskussion.

Handlungsempfehlungen, die sich aus mehreren AGs ergeben haben, betreffen zum enen die Diskussion zwischen den Evaluatoren. Eine solche Diskussion über die Handhabung der Indikatoren, methodische Probleme usw. sollte zukünftig auf nationaler wie auf internationaler Ebene verstärkt werden. Zum anderen resultierte die Empfehlung an die EU-Kommission, die Verbindlichkeit der Fragen, Kriterien und vor allem Indikatoren eindeutiger zu definieren und gegebenenfalls danach zu kategorisieren, ob sie in jedem Fall oder fakultativ zu bearbeiten sind. Zudem wird eine Konzentration auf wenige Fragen, Kriterien und Indikatoren von zentraler Relevanz empfohlen.

4. Welche Grundlagen wurden für die Ex-post-Bewertung geschaffen?

Die Halbzeitbewertung sollte u.a. auch dem Zweck dienen, Grundlagen für die Ex-post-Bewertung zu schaffen. Diese Grundlagen beziehen sich auf zwei Bereiche: Einerseits die Einrichtung von Bewertungssystemen bzw. Überprüfung von bestehenden Monitoringsystemen auf deren Eignung und andererseits die Schaffung einer methodischen Grundlage zur Bewertung.

Das einzige Monitoringsystem, das über alle Förderkapitel existiert, ist das EU-Monitoring, dem mehr oder weniger von allen Arbeitsgruppen eine erhebliche Unzulänglichkeit attestiert wurde. Neben dem EU-Monitoring existieren weitere Datenhaltungssysteme in den Mitgliedstaaten (Projektdatenbanken, Investitionskonzepte, naturschutzfachliches Monitoring usw.). Diese Systeme wurden gesichtet und hinsichtlich ihrer Eignung für Bewertungszwecke analysiert. Bestehende Datenunzulänglichkeiten wurden im Wesentlichen mit drei unterschiedlichen Vorgehensweisen zu beheben versucht:

– Von den Evaluatoren selbst wurden differenzierte Datenerfassungssysteme entwikkelt;

– die bestehenden Datenlücken wurden mit ad-hoc-Ansätzen zu schließen versucht;
– im Agrarumweltbereich wurden von den Fachbehörden in Absprache mit den Evaluatoren Begleituntersuchungen entwickelt, deren Schwerpunkt allerdings im biotischen Bereich liegt.

Die Sichtung und Prüfung der vorhandenen Daten hat, und dies wurde in allen Arbeitsgruppen bestätigt, einen großen Teil der Arbeitskapazitäten der Halbzeitbewertung beansprucht.

Die Fragen, Kriterien und Indikatoren wurden in allen Arbeitsgruppen sowohl inhaltlich als auch methodisch intensiv diskutiert. In den meisten Arbeitsgruppen richtete sich der Blick nicht nur auf die Generierung von Ergebnissen zur Halbzeitbewertung, sondern auch auf die umfassendere Beantwortung von Fragen zur Ex-post-Bewertung. In diesem Sinne wurden konzeptionelle Grundlagen für die Ex-post-Bewertung geschaffen. Im Rahmen einer Aktualisierung der Halbzeitbewertung könnten noch vorhandene Lücken geschlossen und neu entwickelte Konzeptionen breit erprobt bzw. einer vertieften Auswertung zugeführt werden.

5. Programmplanungsdokumente als Grundlage für die Kapitelbewertung

Die Genehmigung der ländlichen Entwicklungsprogramme zu Beginn des Förderzeitraums 2000 bis 2006 erfolgte auf der Grundlage umfassender Programmplanungsdokumente, deren einzelne Bestandteile durch die Durchführungsverordnung (EG) Nr. 1750/1999 vorgegeben waren. Ob diese Programmplanungsdokumente eine Grundlage für die Kapitelbewertungen darstellen, wurde anhand folgender Aspekte diskutiert:

– Interventionslogik des Landes;
– Zielbeschreibung und -quantifizierung;
– Ex-ante-Bewertung.

Aus den Ergebnissen wurden Empfehlungen für die zukünftige Programmplanung gezogen.

Die **Interventionslogik** der einzelnen Länderprogramme wurde von einzelnen Arbeitsgruppen als zu oberflächlich bezeichnet. Gerade die Darstellung der Stärken und Schwächen erfolgt auf einer teilweise sehr hoch aggregierten Ebene, wie sie für ein Programm, nicht aber für einzelne Maßnahmen adäquat ist. Hier wären (zusätzliche) spezifischere und auch stärker regionalisierte Aussagen sachgerechter gewesen. Die bestehenden Mängel einer klaren Beziehung zwischen Bedarf - Strategie - Zielen - Maßnahmen sind auch auf einen starken politischen Einfluss in der Programmerstellung zurückzuführen.

Die **Zielsystematik** der einzelnen Förderkapitel weist nach Auffassung aller Arbeitsgruppen teilweise erhebliche Mängel auf. So sind die Ziele nicht hinreichend genau beschrieben oder es wurde keine Gewichtung zwischen multiplen Zielen vorgenommen. Die Zielsysteme mussten in vielen Fällen nachgearbeitet werden. Wenn Ziele quantifiziert sind, dann zumeist auf der Outputebene. Beabsichtigte Wirkungen wurden in der Regel nicht quantifiziert. Eine Quantifizierung von Zielen macht allerdings nur bei einem kontinuierlichen Controlling und einer möglichen Veränderbarkeit von Zielen im Zeitablauf einen Sinn. Außerdem müssen Ziele genau formuliert werden und auch Auskunft geben können über den Erfolg von Maßnahmen.

Das Fehlen einer fundierten **Ex-ante-Bewertung** wurde in allen Arbeitsgruppen als Problem benannt. Die Mängel bei der Ex-ante-Bewertung sind auch auf die Rahmenbedingungen bei der Programmerstellung, z.B. den enormen Zeitdruck, zurückzuführen. Außerdem hat es sich nicht als sinnvoll erwiesen, die Programmerstellung mit der Ex-ante-Bewertung so stark zu verknüpfen.

Für **zukünftige Programmerstellungen** bedeutet dies, stärker als bisher den tatsächlichen Bedarf, die Ziele und die Maßnahmen zu beschreiben, um damit eine bessere Grundlage für die Bewertung zu schaffen. Die Ex-ante-Bewertung sollte von externer Seite unabhängig vom Programmantrag durchgeführt werden, wobei der stark formative Ansatz hinsichtlich einer Verbesserung des Programms beibehalten werden sollte.

6. Evaluationskultur in den Ländern

Unter diesem Punkt wurden das Verhältnis zwischen Auftraggeber und Auftragnehmer sowie der Zweck der Evaluation – als Teil eines Lernprozesses oder reine Pflichtübung gegenüber der EU-Kommission – diskutiert.

Insgesamt wurde in allen Arbeitsgruppen die Meinung vertreten, dass sich die Entwicklung einer Evaluierungskultur noch in den Anfängen befindet. Teilweise wurde die Evaluation ausschließlich auf die Erfüllung einer Berichtspflicht reduziert, was sich auch auf das Verhältnis zwischen Auftraggeber und -nehmer auswirkt. Geht es v.a. um das formale Abarbeiten der Vorgaben der EU-Kommission, erwartet der Auftraggeber einen möglichst unkomplizierten und kostengünstigen Bericht. Das Interesse an einer intensiven Ergebnisdiskussion ist demnach auch gering. Diese Einschätzung wurde v.a. von den zentralen Bewertern in Deutschland in AG I und II getroffen.

In den anderen Arbeitsgruppen, in denen Programmbewerter zahlenmäßig stärker vertreten waren, wurde ein eher konstruktives und auf Verbesserungen der Maßnahmen ausgerichtetes Verhältnis hervorgehoben. Will man die Ergebnisse auch zur Fortentwicklung der eigenen Maßnahmen bzw. des Programms nutzen, ist das Interesse an einem formativen, prozessorientierten Evaluierungsansatz größer.

Die Frage der Unabhängigkeit der Evaluatoren gegenüber den Auftraggebern wurde in allen Arbeitsgruppen problematisiert. In einigen Bereichen wurde auf eine deutliche Einflussnahme der Auftraggeberseite auf die Berichterstattung im inhaltlichen Bereich hingewiesen. Im Sinne der geltenden Evaluierungsstandards sind hier zukünftig klare Absprachen zwischen Auftraggebern und Auftragnehmern einzuhalten.

5 Programmbewertung und Perspektiven der Evaluation - Darstellung der Impulsreferate der Plenumsitzung des zweiten Tages

Nachfolgend werden die Impulsreferate zu den einzelnen Programmpunkten des Nachmittags des zweiten Tages in der von den Autoren erstellten Fassung wiedergegeben. Da zu den einzelnen Referaten aus Zeitgründen überwiegend keine ausführliche Diskussion stattfinden konnte, wird auf eine zusätzliche ergebnisbezogene Berichterstattung zur Diskussion verzichtet.

5.1 Besonderheiten der Ziel 1-Bewertung: Erfahrungen der Halbzeitbewertung des EAGFL/A im Rahmen der Evaluation des Gemeinschaftlichen Förderkonzeptes 2000-2006 für den Einsatz der Strukturfonds in den neuen Bundesländern und im Ostteil Berlins (Karl)

Neben der Evaluation der Operationellen Programme der Bundes bestand die Halbzeitbewertung in einem ersten Schritt in der fondsspezifischen Bewertung in den einzelnen neuen Bundesländern, der Zusammenfassung der fondsspezifischen Ergebnisse und der Bewertung der Operationellen Programme der Länder. In einer zweiten Stufe wurde auf der Grundlage der Länderevaluation eine zusammenfassende Halbzeitbewertung für die Gesamtheit der Interventionen der Europäischen Strukturfonds und in diesem Rahmen auch für den EAGFL/A erstellt.**12** Fasst man die Erfahrungen mit der Halbzeitbewertung des EAGFL/A zusammen, können folgende Eindrücke und Ergebnisse hervorgehoben werden:

– Ausgangspunkt der Evaluierung war, dass auf der Grundlage der Operationellen Programme der Bundesländer ein Set von Indikatoren zusammengestellt wurde, das in allen Länderevaluierungen zum Zuge kommen sollte. Dieses Set von Indikatoren ist eine sehr kleine Teilmenge der Indikatoren, wie sie von Seiten der Europäischen Kommission in den „Gemeinsamen Bewertungsfragen“[13] sowie im „Leitfaden für die Halbzeitbewertung der Programme zur Entwicklung des ländlichen Raums 2000-2006

[12] GEFRA et al.: Erfahrungen der Halbzeitbewertung des EAGFL/A im Rahmen der Evaluation des Gemeinschaftlichen Förderkonzeptes 2000-2006 für den Einsatz der Strukturfonds in den neuen Bundesländern und im Ostteil Berlins, Endbericht, Münster 2004.

[13] GD Landwirtschaft (1999): Bewertung von Programmen zur Entwicklung des ländlichen Raums im Zeitraum 2000-2006 mit Unterstützung des Europäischen Ausrichtungs- und Garantiefonds für die Landwirtschaft, Brüssel,), VI8865/99.

mit Fördermitteln des EAGFL/A“[14] vorgeschlagen werden. Im Interesse von Vergleichbarkeit wurde auf der Ebene der GFK-Evaluierung versucht, vorab gemeinsame einheitliche Indikatoren festzulegen. Allerdings muss konstatiert werden, dass dies lediglich für eine begrenzte Zahl einheitlich und flächendeckend gelungen ist; dies gilt insbesondere für Finanzindikatoren. Zum Teil konnten auch die in den Operationellen Programmen aufgeführten Indikatoren nicht immer abgedeckt werden. Die von Seiten der Europäischen Kommission beispielsweise in den „Gemeinsamen Bewertungsfragen“[15] vorgeschlagenen Indikatoren wurden, wenn überhaupt, nur bei der Prüfung der Begleit- und Monitoringsysteme berücksichtigt.

– Da sowohl die Vereinheitlichung auf Basis der Indikatoren nur bedingt möglich war als auch die Problemstellung auf Seiten des EAGFL/A in den Ländern sehr unterschiedlich ausfiel, waren die Länderevaluationen aus Sicht des GFK im Wesentlichen gut ausgebaute und fundierte Fallstudien. Diese haben neben den bekannten Vorteilen von Fallstudien aber auch deren typische Nachteile, weil die Aussagen in nur begrenztem Umfang verallgemeinerbar sind. Vor diesem Hintergrund überwiegt auf der Ebene der GFK-Evaluation insbesondere im Hinblick auf die Wirkung der Maßnahmen die dezentrale Bewertung, die auch Grundlage für eine GFK-Bewertung war. Zentrale, länderübergreifende Momente in der Bewertung kamen vor allem dann zum Zuge, wenn neben maßnahmenspezifischen Wirkungen die Rückwirkungen der EAGFL/A-Intervention in Hinblick auf das regionalwirtschaftliche Konvergenzziel der Strukturfonds beurteilt wurde.

– Vor dem Hintergrund des Evaluationsprozesses ging es darum, im Sinne der Fondsverordnung erste Wirkungen und Ergebnisse zu dokumentieren. Darüber hinaus sollte auf möglicherweise erkennbare Anlaufschwierigkeiten und Umsetzungsprobleme hingewiesen werden. Insofern lagen summative Ergebnisse auf der Ebene der finanziellen Indikatoren sowie auf der Ebene des Maßnahmevollzugs vor. Eine formative Bewertung konnte auf der Ebene der initiierten Lernprozesse, des Begleit- und Monitoringsystems sowie der strategischen Orientierung angestoßen werden. Hier wurde deutlich, dass insbesondere die aus dem Ökoaudit bekannte Methodik der Systemprüfung bei der formativen Bewertung hilfreich ist.

– Für eine formative Bewertung sind der Bewertungszyklus nach dem Muster einer Halbzeitbewertung sowie deren Aktualisierung sinnvoll. Für zahlreiche Indikatoren der „Gemeinsamen Bewertungsfragen“ sind der Bewertungszyklus der Halbzeitbewertung bzw. selbst ein 6-jähriger Bewertungszyklus problematisch, da zahlreiche

---

[14] GD Landwirtschaft (2002): Gemeinsame Indikatoren zur Begleitung der Programmplanung für die Entwicklung des ländlichen Raumes, Arbeitspapier der Kommission VI/43512/02 endg. v. 26.2.2002.

[15] GD Landwirtschaft (2000): Gemeinsame Bewertungsfragen mit Kriterien und Indikatoren, Arbeitspapier der Kommission VI/12004/00 ENDG.

langfristige ökonomische Wirkungen in der Strukturentwicklung der Landwirtschaft und ländlicher Räume in diesen Zeiträumen nicht erfasst werden können. Darüber hinaus ist zu berücksichtigen, dass methodisch bei fortlaufenden Agrarinterventionen eine Separation von Wirkungen in Hinblick auf einzelne Maßnahmenbündel der verschiedenen Programmperioden nicht mehr möglich ist. Insofern ist eine kontinuierliche Ausrichtung der summativen Bewertung notwendig, wobei hier Planungszeiträume von bis zu 10 Jahren ins Auge gefasst werden müssen.

– Was den Erfahrungsaustausch der Evaluatoren auf der Ebene des GFK angeht, so ist dieser insgesamt positiv zu bewerten. Alle Treffen beim Bundesfinanzministerium haben zur Verbesserung der GFK-Evaluation beigetragen und auch die Länderevaluation stimuliert. In Zukunft wird es darum gehen, sich über die Verständigung auf ein Indikatorenset hinaus stärker methodisch aufeinander abzustimmen, um zum einen die Fallstudien auf Länderebene zu vereinheitlichen und zum anderen die Konsistenz der synthetisierenden Halbzeitbewertung zu verbessern.

Fasst man zusammen, steht die EAGFL/A-Evaluation im Rahmen des GFK sicherlich noch ganz am Anfang. Erst ein kleines Set von Indikatoren wird tatsächlich abgefragt und erhoben und die Länderevaluationsberichte besitzen aus der Perspektive des GFK den Charakter spezifischer Länderfallstudien, die insbesondere im Hinblick auf deren dezentrale Bewertung als auch im Hinblick auf die formative Bewertung aussagekräftig sind. Sieht man die Halbzeitbewertung als Teil eines Evaluationsprozesses, konnten im Sinne der Verordnung 1257/99 erste Ergebnisse, die Relevanz und Kohärenz der Strategie und die Verwirklichung der angestrebten Ziele bewertet werden; zudem wurde die Wirtschaftlichkeit der Haushaltsführung sowie die Qualität der Begleitung und Durchführung untersucht. In Zukunft wird es vor allem darum gehen, die Länderevaluation stärker methodisch zu harmonisieren und bei den Bewertungen noch deutlicher als bisher die spezifischen Strukturfondsziele (insbesondere das regionalwirtschaftliche Konvergenzanliegen neben den spezifischen Zielen des EAGFL/A) in den Mittelpunkt der Aufmerksamkeit zu rücken.

5.2 Zentrale/dezentrale Bewertung

5.2.1 Impulsreferat I (Hochberg)

a) Ausgangssituation

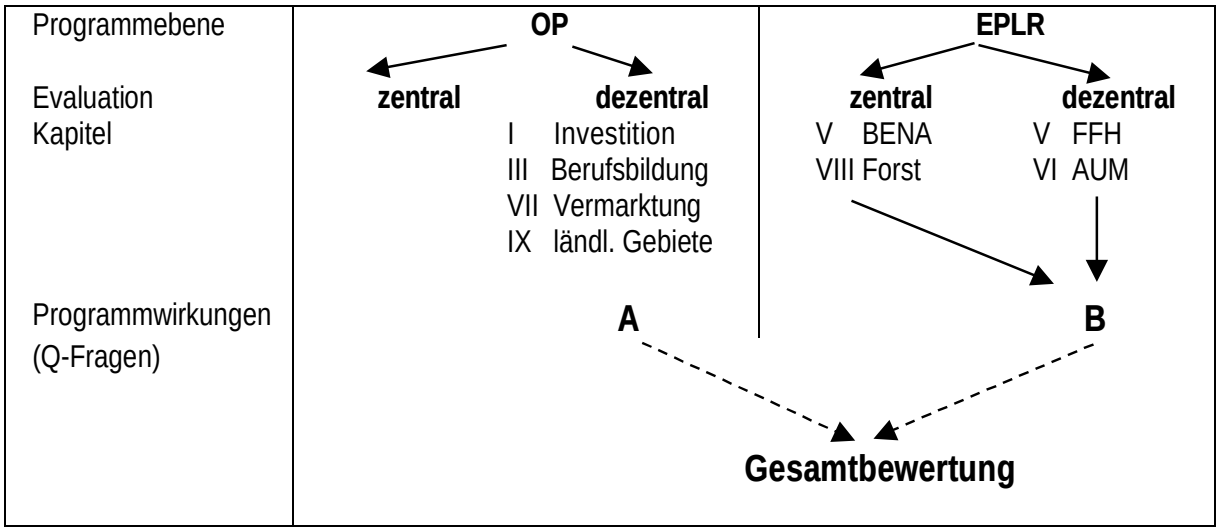
⇒ Programmwirkungen auf Landesebene resultieren aus unterschiedlicher Anzahl von Maßnahmen mit regionalspezifischen Wirkungen

⇒ Überlagerungs-, Katalysatoreffekte kapitelspezifischer Maßnahmen untereinander nicht quantifizierbar

⇒ Ziel 1-Gebiete mit OP und EPLR ausgestattet

⇒ Zentrale Evaluation kapitelspezifischer Maßnahmen und dezentrale Bewertung weiterer kapitelspezifischer Maßnahmen sowie Programmwirkungen

Beispiel Thüringen:



b) Botschaften aus Mid-term Review

⇒ Einsatz mehrerer Evaluatoren ist verbunden mit

unterschiedlichen Ansprüchen an die

+ Datenbereitstellung vom Land,
+ Belastbarkeit der Evaluierungsergebnisse,
+ Verwertbarkeit für Programmbewertung

grundsätzlich

+ höherem Aufwand für den Koordinator;
+ sehr schwieriger Gesamtbewertung (OP + EPLR)

⇒ Abstimmung der Evaluatoren untereinander (Kapitelbewerter) mit Defiziten behaftet, weil:

+ kein Mindestmaß an länderübergreifend vereinbarten Bewertungsmaßstäben erreicht wurde und
+ Bearbeitungszeitraum angesichts der Aufgabe zu kurz war

⇒ Kombination von zentraler und dezentraler Evaluation hat Qualität des Mid-term Review maßgeblich beeinflusst (vom Methodenbericht bis zur Ergebnisbewertung)

c) Ausblick - Thesen

⇒ Kontinuität der Evaluation auf Landesebene (Mid-term ➔ up date ➔ Ex-post),

vorausgesetzt, der Evaluator hat

+ zweckdienliche Datenerfassungssysteme geschaffen (Monitoring, ...)
+ zumindest die Ausgangssituation repräsentativ abgebildet und
+ die Ergebnisse einer Bewertung (Zielerfüllungsgrad) unterzogen.

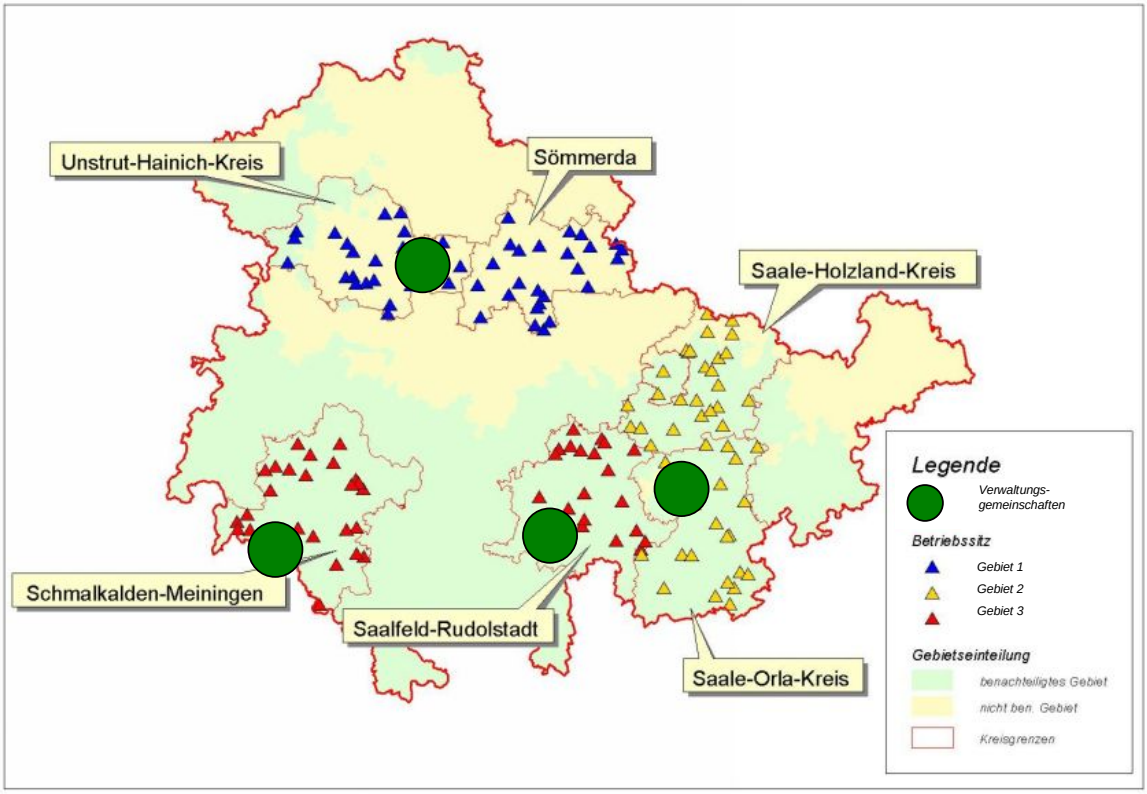
⇒ Nicht zentral versus dezentral sondern ein Evaluator für eine Region;

⇒ Mindestmaß an regional übergreifend vereinbarten Bewertungsmaßstäben

Dezentrale Evaluation der Regionalprogramme mit weitestgehend einheitlicher Bewertungsmethodik

Bestimmungsgründe:

⇒ Bewertung der sozioökonomischen Wirkungen (Q1 und Q2) erfordert eine für alle Kapitelmaßnahmen repräsentative Datenbasis (eigenständige Repräsentativanalysen)



Gebietsauswahl

Gebiet	Landkreis[1)]	LF (Tha)	BENA (% LF)[1)]	FFH (% LF)	KULAP (% LF)	Relevanz des EPLR
1	UH/WE	130,1	3,5	2,1	11,0	**gering**
2	SOK/SHK	91,3	75,5	2,8	25,6	**mittel bis hoch**
3	SM/SLF	78,9	90,0	4,5	61,7	**hoch**

1) Landkreise Thüringens mit einer Bevölkerungsdichte unter 150 Einwohner/km²; 2) bewilligte Fläche

⇒ Bewertung der Einkommenswirkung (Q3) setzt regionalspezifisch Modifizierung der Indikatoren bzw. Unterindikatoren voraus.

⇒ Der Programmbeitrag zum Umweltschutz (Q5) war nur unter Einbeziehung ergänzender Studien zu bewerten (Wasserschutz, Emission von Treibhausgasen, Landschaftsschutz)

5.2.2 Impulsreferat II (Schinke)

Wenn im Rahmen dieses Workshops von „zentraler Evaluation“ gesprochen wird, ist damit zuweilen die übergreifende Bewertung z.B. der Operationellen Programme in den Ziel-1-Gebieten gemeint gewesen. Davon soll hier nicht die Rede sein.

Es geht vielmehr darum, dass in Deutschland zwei agrarstrukturpolitische Instrumente miteinander verwoben sind: die ländlichen Entwicklungsprogramme nach VO (EG) Nr. 1257/1999 (EPLR) und die Gemeinschaftsaufgabe „Verbesserung der Agrarstruktur und des Küstenschutzes“ (GAK). Dementsprechend ist auch die Evaluation dieser Instrumente verzahnt:

– die Evaluation der GAK ist erforderlich (und vom Bundesrechnungshof gefordert), um Rechenschaft über die Wirksamkeit der GAK und ihrer Maßnahmen abzulegen

– die Evaluation der EPLR wird vom EU-Recht gefordert.

Eine parallele Evaluation beider Instrumente in organisatorisch völlig getrennten Verfahren wäre unzweckmäßig, zumal mit den Leitfäden und Evaluationsfragen, -kritierien und –indikatoren der KOM ein inhaltlicher Rahmen vorgegeben ist.

Eine von Bund und allen Ländern gemeinsam in Auftrag gegebene und aus der GAK finanzierte Evaluation bestimmter GAK-Maßnahmen hat es bereits in der Vergangenheit gegeben:

– für den Förderzeitraum 1994 bis 1999 wurden die Maßnahmen nach VO (EG) 950/97 und 951/97 – also AFP, Junglandwirteförderung, Ausgleichszulage und Marktstrukturverbesserung - zentral durch die FAL evaluiert

- seit 2000 wurden die Halbzeitbewertungen für AFP (einschließlich Junglandwirteförderung), Ausgleichszulage, Marktstrukturverbesserung und Erstaufforstung zentral erarbeitet.
  - ÷ Dazu wurde für jedes Bundesland ein Bericht erstellt, der dann als Baustein in die Evaluation des Landesprogramms einfließen konnte.
  - ÷ Außerdem wurde ein länderübergreifender Bericht erarbeitet, der die Ergebnisse für die GAK zusammenfasst.

Bei der Halbzeitbewertung der GAK-Maßnahmen haben sich folgende Probleme gezeigt:

- es mussten alle Maßnahmen einbezogen werden, die aus der GAK finanziert werden, unabhängig davon ob und wie sie vom EAGFL kofinanziert werden –Landesmaßnahmen außerhalb der GAK (z.B. Investitionsförderungsrichtlinien in Brandenburg, ergänzende Landesprogramme zur Ausgleichzulage in Thüringen und Sachsen) wurden hingegen nicht einbezogen, obwohl sie teilweise mit EU-Mitteln kofinanziert werden. Das führte dazu, dass der zentrale Bewerter und der Programmevaluator unnötigerweise parallel arbeiten mussten.
- Die Bausteine für die Landesprogramme konnten aus unterschiedlichen Gründen nicht immer fristgerecht geliefert werden.
- Es war nicht allen Landesevaluatoren problemlos möglich, die Bausteine der zentralen Evaluation in die Evaluation des Landesprogramms zu übernehmen.
- Teilweise wurden die Evaluatoren der Landesprogramme so spät benannt, dass eine Zusammenarbeit mit dem zentralen Evaluator aus Zeitgründen schwierig war.

Bei den Beratungen zum Rahmenplan 2004 der GAK haben alle Bundesländer eine Fortsetzung der zentralen Evaluation bestimmter GAK-Maßnahmen bis zur Ex-post-Bewertung der Landesprogramme befürwortet.

Es beginnt also jetzt ein neues Verfahren:

- die <u>inhaltlichen</u> Vorgaben orientieren sich wie bisher an den Leitfäden der KOM
- die <u>organisatorischen</u> Vorgaben werden in den nächsten Wochen erarbeitet und entschieden; dabei werden Erkenntnisse aus der Halbzeitbewertung berücksichtigt:
  - ÷ die Zusammenarbeit zwischen zentralem Bewerter und Landesevaluatoren muss über das bisher bestehende hinaus intensiviert und institutionalisiert werden. Die zentrale Evaluation bietet hier die Möglichkeit, eine Plattform zu schaffen, die bei den übrigen (dezentral und nur durch die Landesevaluatoren bewerteten) Maßnahmen in dieser Form nicht existiert.

- Die Aufnahme inhaltsgleicher Landesrichtlinien, die außerhalb der GAK entweder als reine Landesmaßnahme (Staatsbeihilfe) oder mit EAGFL-Kofinanzierung durchgeführt werden, in die zentrale Evaluation ist anzustreben
- Aufwendungen für gesonderte Datenauswertungen müssen berücksichtigt werden.

Bundesländer, zentrale Bewerter und Landesevaluatoren sind aufgefordert, ihre Anregungen und Wünsche zeitnah in das laufende Verfahren einzubringen.

5.3 Summative/formative Bewertung (Grajewski)

Bei der Durchführung von Evaluierungen ist ihr jeweils vorrangiger Zweck festzulegen. Daran orientieren sich die Evaluierungsfragen und die voraussichtlichen Nutzer. Nagarajan und Vanheukelen (1997) treffen folgende Unterscheidung:

Bei **formativen Evaluierungen** steht die Untersuchung von Möglichkeiten zur Verbesserung und Aufwertung des Managements und der Durchführung von Programmen im Vordergrund. Hier geht es um Verbesserungsmöglichkeiten innerhalb der Organisation. Die Evaluierung wird als Lernmedium begriffen.

Summative Evaluierungen sollen die Wirkungen und Effizienz von Programmen abbilden. Im Vordergrund steht die Leistungsmessung und die Rechenschaft nach außen. Der Nutzen liegt weniger beim Programmmanagement sondern bei externen Akteuren (d.h. Gruppen, die nicht unmittelbar am Programmmanagement beteiligt sind).

Obgleich die Unterscheidung klar definiert ist, ist in der praktischen Durchführung von Evaluierungen eine Trennung zwischen formativen und summativen Funktionen oftmals schwierig. Grundsätzlich ist die Frage zu stellen, ob eine klare Abgrenzung sinnvoll ist oder ob eine Kombination von beiden Funktionen nicht zur Verbesserung von Evaluierungen beitragen kann.

Stellt man die summative Funktion einer Evaluierung in den Vordergrund, so bedeutet dies für die Halbzeitbewertung der Programme zur Entwicklung des ländlichen Raums, dass vorrangig das Bewertungsraster der Europäischen Kommission abzuarbeiten ist. Dabei werden erste Ergebnisse und Wirkungen aufgezeigt. Des Weiteren sind die methodischen und datentechnischen Grundlagen für die Ex-post-Bewertung zu schaffen. Der Analyse des Programmmanagements und der verwaltungsmäßigen Abwicklung kommt nur ein untergeordneter Stellenwert zu. Die Umsetzung der gewonnenen Ergebnisse steht nicht im Mittelpunkt der Evaluierung. Wenn Evaluierung als reiner Rechenschaftsbericht (an die EU-Kommission) verstanden wird, ist das Interesse einer breiten Einbindung aller Beteiligten und Betroffenen gering. In diesem Fall wird die Auftragsvergabe von der Pro-

grammkoordination vorgenommen mit der Maßgabe, einen EU-kompatiblen Bericht zu erstellen. Die verschiedenen Beteiligten und Betroffenen haben demzufolge nur ein geringes Interesse, sich in den Evaluierungsprozess einzubringen, da sie selbst die Evaluierungsfragen nicht mitbestimmen konnten. Daraus resultiert für den Evaluator eine geringe Akzeptanz bei den Evaluierten, die den Zugang zu Informationen und auch informellen Netzwerken erschwert.

Eine Evaluierung, die formativ angelegt ist, weist eine starke Umsetzungsorientierung auf. Da Verbesserungen der Umsetzung und der Ausgestaltung der Förderung im Vordergrund stehen, ist eine kontinuierliche Reflektion der im Evaluierungsprozess gewonnenen Ergebnisse mit den Auftraggebern und den für die Fördermaßnahmen zuständigen Personen erforderlich. Diese Reflektion kann in Form von begleitenden Workshops durchgeführt werden. In diesen Workshops kommen dem Evaluator zwei Rollen zu: Zum einen muss er einen Input als Diskussionsanstoß geben, zum anderen fungiert er als Moderator. Um eine Diskussion anzuregen, ist die Ableitung von Schlussfolgerungen und Empfehlungen aus dem im Rahmen der Evaluierung gewonnenen Ergebnisse erforderlich. Ein lernprozessorientierter Ansatz kann auch durch eine Moderation von Arbeitsgruppen überstützt werden, in denen beispielsweise Fachreferenten, nachgeordnete Behörden oder Zuwendungsempfänger aus verschiedenen Bundesländern oder anderen EU-Staaten ihre Erfahrungen austauschen.

Eine formative Evaluierung setzt voraus, dass in der Leistungsbeschreibung der Zweck der Evaluierung genau beschrieben ist. In den Leistungsbeschreibungen der Halbzeitbewertungen wurde diesen formativen Aspekten nicht ausreichend Rechnung getragen. Das Aufgabenspektrum war auf die Abarbeitung des Bewertungsrasters der Europäischen Kommission ausgerichtet, sodass der zeitliche und inhaltliche Aufwand für ein lernprozessorientiertes Vorgehen nicht ausreichend personell und finanziell berücksichtigt werden konnte. Eine formative Evaluierung setzt zudem die Offenheit für Lernprozesse voraus. Dies gilt für beide Seiten, Auftraggeber und Auftragnehmer. Nur wenn diese Offenheit existiert, ist ein Zugang zu Informationen und auch eine partielle Einbindung in bestehende (informelle) Netzwerke gegeben. Eine zu starke Betonung des formativen Charakters einer Evaluierung kann aber zu Betriebsblindheit führen und zu einer Verlagerung des Blicks von übergreifenden Aspekten auf Detailprobleme von Fachreferaten oder Behörden.

Eine summative Evaluierung als reiner Rechenschaftsbericht verfehlt wegen der fehlenden Umsetzungsorientierung seine Wirkung. Die Einbeziehung formativer Elemente kann dieses Defizit ausgleichen. Evaluierung sollte demzufolge als Prozess (siehe auch Eser, 2001) verstanden werden. Dieser Prozess beinhaltet, dass nicht nur adhoc, sondern stärker begleitend evaluiert wird und eine ständige Rückkoppelung mit den Auftraggebern und Fachreferaten stattfindet. In diesem Dialog ist es wichtig, die unabhängige Position des

Evaluators zu wahren und auch bei Kenntnis der Innensicht nicht zu einem Teil der Organisation zu werden.

Eser, T. W. (2001), Evaluation und Qualitätsmanagement - Anforderungen und Konsequenzen für die EU-Strukturpolitik. Informationen zur Raumentwicklung, Heft 6/7. 2001, S. 327-339.

Nagarajan, N. und Vanheukelen, M. (1997): Evaluierung der ausgabenwirksamen Programme der EU. Ein Leitfaden zur Halbzeit- und Ex-post-Bewertung. Generaldirektion XIX - Haushalt.

5.4 Idealer Bewertungszyklus

5.4.1 Impulsreferat I „Idealer Bewertungszyklus? - Erfahrungen aus dem FuE Vorhaben BNPro“[16] (Elbe)

Ausgangslage – oder: wieso der Bewertungszyklus derzeit nicht wirklich funktioniert

Der Bewertungszyklus der laufenden Förderperiode war und ist durch die Verordnungen per Stichtag terminlich fixiert. So hat die Halbzeitbewertung zum 31.12.2003 und die Aktualisierung bis zum 31.12.2005 zu erfolgen. Dabei wird nicht berücksichtigt, ob die Programme tatsächlich auch zum 01.01.2000 gestartet sind oder nicht.

In der laufenden Förderperiode starteten die Programme gemäß VO (EG) Nr. 1257/1999 nicht wie in den Verordnungen festgelegt, sondern erst sehr viel später. Hinzu kommt, dass die Umsetzung der Förderprogramme, insbesondere bei veränderter Fördersystematik wie im Fall der VO (EG) Nr. 1257/1999, mit Anlaufschwierigkeiten verbunden ist. Dies führt zu Evaluierungen, die nicht wie vorgesehen 3 und 5 Jahren Programmlaufzeit um-

[16] Das FuE Vorhaben **BN$_{Pro}$** „Berücksichtigung von Naturschutzzielen in der Programmbewertung und -entwicklung“ verfolgt das Ziel, Handlungsempfehlungen für eine stärkere Berücksichtigung von Naturschutzaspekten in der Programmevaluierung zu entwickeln, um so letztendlich eine bessere Verankerung dieser Belange bei der Gestaltung und Umsetzung von Förderprogrammen zu erreichen. Das vom Bundesamt für Naturschutz geförderte Projekt wird von SPRINT (GbR) in Zusammenarbeit mit dem Zentrum für Naturschutz, Forschungsstelle Naturschutzpolitik (Uni Göttingen) und dem nova-Institut (Bonn) durchgeführt (Laufzeit 08/2002 bis 04/2004). Hierfür werden Programme zur Entwicklung des ländlichen Raums nach EU VO 1257/99, die EU-Gemeinschaftsinitiative LEADER+, die Ziel 1 Förderung (soweit vorhanden) und die GAK in den Bundesländern Bayern, Sachsen und Niedersachsen vertiefend untersucht.

fassen, sondern einen wesentlich verkürzten Zeitraum[17]. Kurz gesagt: ⇒ ***Der Zeitpunkt bestimmt die Inhalte!***

Daraus resultieren Folgen für alle Bewertungszeitpunkte der „Evaluations-Trias".

– Die **Ex-Ante-Bewertung** hatte keine feststehende Grundlage, da die Programme gemäß der VO (EG) Nr. 1257/1999 der Länder und die Verordnungen der EU teilweise parallel entwickelt wurden. Die Ex-Ante kam demnach zu früh und wurde deshalb im Evaluations-Leitfaden der EU-Kommission inhaltlich abgespeckt.

– Die **Halbzeitbewertung** kommt eigentlich ebenfalls zu früh und kann die Messung von Wirkungen nur eingeschränkt (wenn überhaupt) leisten.

– Der Schwerpunkt – so wird immer wieder betont – soll nun auf der **Aktualisierung** zum 31.12.2005 liegen.

  Aber: Welche Programmänderungen sollen auf der Grundlage der Aktualisierung noch erfolgen? Und kann man zu diesem Zeitpunkt noch realistisch davon ausgehen, dass die Ergebnisse für die neue Förderperiode verwendet werden (können)?

– Die **Ex-Post-Bewertung** kommt für einen Lernprozess systematisch zu spät und wurde/wird inhaltlich „abgespeckt"[18].

Lösungsansätze – oder: wie der Bewertungszyklus verbessert werden könnte

Bei aller gebotenen Vorsicht kann man davon ausgehen, dass es auch in Zukunft keine verbindlichen Zeitpläne für Programmgenehmigungen geben wird und es immer Anlaufschwierigkeiten bei (neuen) Programmen geben wird. Trotzdem sollte gelten:
⇒ ***Die Inhalte bestimmen den Zeitpunkt!***

Damit dies möglich ist, müsste aus meiner Sicht zunächst geklärt werden, welche Funktion/welche Ziele mit der Evaluierung, insbesondere mit der Halbzeitbewertung, erreicht werden sollen. Derzeit soll sie eine Doppelfunktion übernehmen: Zum einen die Steuerung des aktuellen Programms und zum anderen als Grundlage für die neue Förderperiode. Neben den zeitlichen und methodischen Problemen bei der Wirkungsmessung (s.o.) führt diese Doppelfunktion zu einer Überforderung der Bewertung.

---

[17] In extremen Fällen wie bspw. im EAGFL/A, bei dem das Wirtschaftsjahr der Landwirtschaft auch noch versetzt ist, führte dies in der Halbzeitbewertung dazu, dass nur Daten für 1 Wirtschaftsjahr zur Verfügung standen. Es handelt sich also eher um eine „Jahresbewertung" denn um eine Halbzeitbewertung.

[18] Teilweise waren die Ex-Post der vorausgegangenen Förderperiode bei der Halbzeitbewertung noch nicht vorhanden.

Folgende Fragen können hieraus abgeleitet werden:

- Zielrichtung der Evaluierung: Aktuelles Programm oder neue Förderperiode?
- Möglichkeiten der Evaluierung: Müssen je nach Zielrichtung unterschiedliche Fragen gestellt werden und was ist methodisch überhaupt machbar (z.B. Nettoeffekte)?
- Kosten/Nutzen der Evaluierung: Steht der Aufwand im Verhältnis zu den produzierten Ergebnissen bzw. deren Einfluss auf das laufende Programm oder die neue Förderperiode?
- Umgang mit den Ergebnissen der Evaluierung: Ist der Bericht „für die EU“ und somit reine Pflichtübung oder soll/wird er im jeweiligen Land als Steuerungsinstrument anerkannt.

Würden diese Fragen Top Down durch die EU-Kommission in die eine oder andere Richtung beantwortet werden, so wäre zu befürchten, dass die Akzeptanz von Evaluierungen in den Programmplanungsbehörden der Ländern nicht erhöht würde. Die Länder sind aber in diesem System der Dreh- und Angelpunkt – der „Flaschenhals“. Sie müssen eingebunden und vom Nutzen der Evaluierungen überzeugt werden. In Teilen ist dies bereits gelungen.

Was könnte dies für die Bewertungszeitpunkte bedeuten? Ein Vorschlag:

- Die **Ex-Ante-Bewertung** wird von Beginn an formativ angelegt und hat als eine Hauptaufgabe auch die Operationalisierung der Ziele des Programms.
- Die **Halbzeitbewertung** bleibt summativ, aber nicht stichtagsbezogen. Sie wird nach tatsächlicher Programmlaufzeit und Zielrichtung der Halbzeitbewertung (s.o.) durchgeführt.
- Die **Aktualisierung** könnte aufgrund der o.a. zeitlichen Rahmenbedingungen entfallen.
- Die **Ex-Post-Bewertung** wird zu einem mehr technischen „Rechenschaftsbericht“.

Fazit: Synchronisierung ist notwendig

Zusammenfassend kann festgehalten werden, das die derzeitige Evaluations-Trias weder synchron mit dem aktuellen Programm- noch mit dem Programmierungszyklus der neuen Förderperiode ist. Die Abstimmung dieser Zyklen ist die Grundlage für einen erfolgreichen Einsatz von Evaluierungen bei der Programmsteuerung und Programmkonzeptionierung. Dies kann auch bedeuten, dass es, wenn es um die neue Förderperiode geht, zeitlich versetzt eigene Fragen geben kann. Es ist davon auszugehen, dass jeder Evaluator diese Chance und Herausforderung der Mitgestaltung und Verbesserung von Förderprogrammen gerne annehmen wird.

5.4.2 Impulsreferat II zum Thema „Idealer Bewertungszyklus“ (Lüdigk)

1. Die Überlegungen setzen am Begriff der *Halbzeit*bewertung an:

Die Sacharbeit war im Juni 2003, d.h. zur Mitte der Förderperiode faktisch abgeschlossen.

Aber: Da die Operationellen Programme erst Ende 2000 verabschiedet (bzw. von der Kommission bestätigt) wurden, konnte das Jahr 2000 zumindest in einigen Bereichen nicht als „vollständiges“ Bewertungsjahr genutzt werden.

2. Die Landwirtschaft unterscheidet sich von allen Bereichen in der Wirtschaft dadurch, dass sie neben dem Kalenderjahr ein *Wirtschaftsjahr* kennt. Damit nimmt sie eine Sonderstellung ein. Dieser zeitliche Auseinanderfall muss hingenommen werden. Das Wirtschaftsjahr endet am 30.6.

Damit standen für Teile der Auswertung nur Daten eines Wirtschaftsjahres zur Verfügung: im Wirtschaftsjahr 2000/2001 war das OP noch nicht in Kraft; 2002/2003 lagen teilweise noch keine Daten vor.

3. Schon die Notwendigkeit einer Nachbewertung (‚*update*‘) belegt, dass der vorgefundene Bewertungszyklus *nicht* ideal ist.

4. Wie kann ein idealer Bewertungszyklus aussehen?

Diese Frage setzt die Beantwortung der Frage voraus, was mit der Evaluation eigentlich erreicht werden soll. Bei der ex-ante-Bewertung und bei der Ex-post-Bewertung sind die Evaluationsziele eindeutig.

Wie sich gezeigt hat, soll demgegenüber die Halbzeitbewertung offensichtlich *zwei Zielen* genügen:

a) Eine Einschätzung des Anlaufs und der Annahme der Maßnahmen, ferner soll sie Hinweise auf drohende Fehlsteuerungen und Vorschläge zur Strategieänderung geben.

Aber: gerade Fehlsteuerungen können zu unterschiedlichsten Zeitpunkten erkannt werden und Strategieänderungen erforderlich machen.

b) Vorbereitung der nächsten Programmperiode

Die „quasi-punktförmige“ Evaluierung kann auftretende zeitliche Auseinanderfälle nur partiell widerspiegeln; d.h. *mit der Erfüllung beider Zielstellungen ist die Halbzeitbewertung überfordert.*

5. Daraus folgt:

a) Für die Erfüllung der erstgenannten Zielstellung ist ein *ständiger Prozess der Begleitung und Bewertung* am besten. Dazu gehört die Aktualisierung der Datenstände einerseits und eine kurze Einschätzung zu Problemfällen andererseits.

Eine derartige Verfahrensweise würde möglicherweise auch der Halbzeitbewertung zugute kommen. Denn für ein echtes Gegensteuern kommt die Halbzeitbewertung möglicherweise schon zu spät.

b) Der punktförmige Bericht könnte sich dann in erster Linie auf die *Bewertung der Gesamtstrategie* konzentrieren (das entspräche etwa der Beantwortung der Querschnittsfragen). Um der zweiten Zielstellung zu genügen, könnte dieser Bericht um etwa ein halbes bis ein Jahr nach hinten verlagert werden (und damit ein ‚update‘ verzichtbar machen).

6. Bei diesem Vorschlag ist freilich zu hinterfragen, in welchem Umfang für die Halbzeitbewertung eine zentral durchgeführte Evaluierung erforderlich ist (noch dazu in einer faktisch gegebenen Parallelität).

5.5 Programmansatz und kapitelübergreifende Bewertung (Grajewski)

Mit der Programmbewertung sollen nach den Vorstellungen der EU-Kommission Qualität und Relevanz der Programme verbessert und Programmänderungen ermittelt werden, „die möglicherweise erforderlich sind, um die ursprünglich gesteckten Ziele tatsächlich erreichen zu können“ (Dok. VI/3300/02). Neben programmspezifischen Fragen dienen dazu vor allem die kapitelübergreifenden Bewertungsfragen, die die EU-Kommission in ihrem Leitfaden zur Bewertung mit entsprechenden Kriterien und Indikatoren hinterlegt hat. Diese gliedern sich in zwei Bereiche:

Ein Bereich beschäftigt sich im Sinne einer summativen Evaluation mit der thematischen Verdichtung der Ergebnisse der Förderkapitelbewertungen und der wesentlichen, positiven wie auch negativen zusätzlichen Wirkungen, die im Rahmen der Maßnahmenkonzeption nicht beabsichtigt sind. Dazu zählen bspw. auch die Beschäftigungswirkungen von Agrarumweltmaßnahmen, die nicht durch kapitelspezifische Fragen abgedeckt sind. Au-

ßerdem ist zu untersuchen, ob der Ansatz der Programmplanung tatsächlich zu einem Mehrwert gegenüber maßnahmenbezogenen Ansätzen geführt hat.

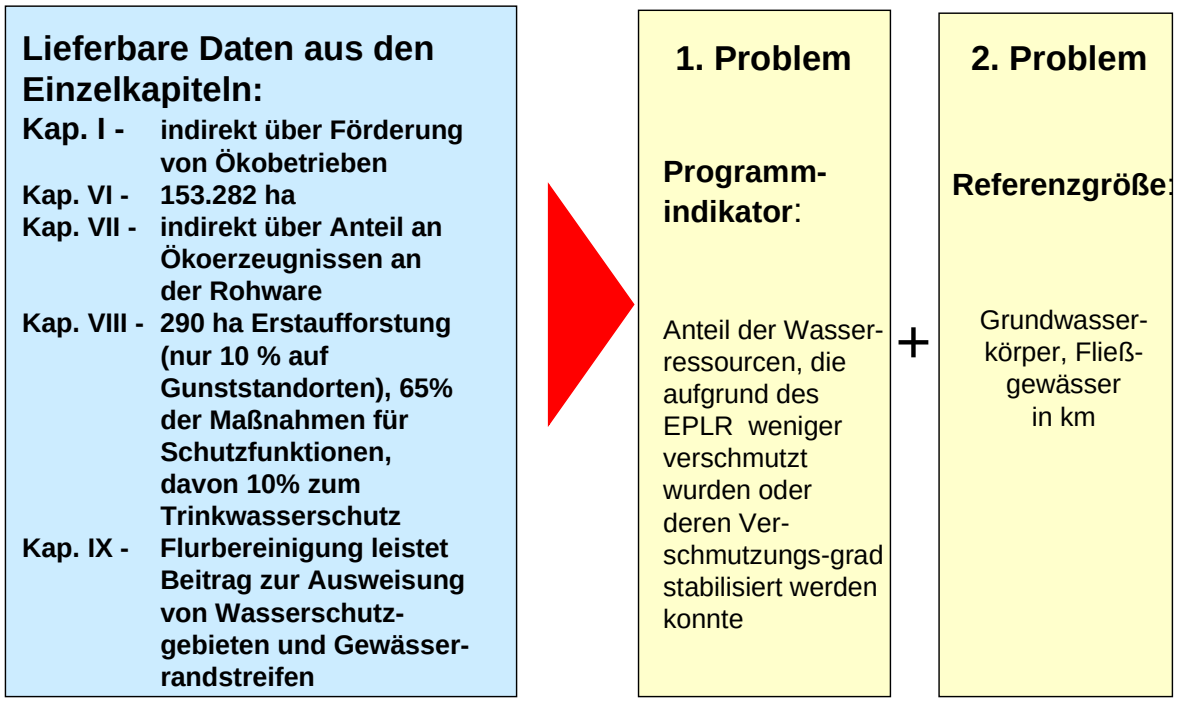
Ein weiterer Bereich behandelt die verwaltungsmäßige Umsetzung des Programms. Hier geht es im Rahmen einer formativen Evaluation v.a. um die Identifikation von Umsetzungshemmnissen und Anpassungsmöglichkeiten, die noch im laufenden Programmplanungszeitraum vorgenommen werden können.

Die kapitelübergreifende Bewertung ist mit verschiedenen Problemen verbunden. Die kapitelüberübergreifenden Fragen decken sich nur zum Teil mit den Zielkategorien auf der Kapitelebene. Mit jeder neuen Zielkategorie müssen neue Wirkungsbeziehungen geprüft werden. Durch die zu erwartenden Interdependenzen – beispielsweise positive Wirkungen bestimmter Investitionsvorhaben auf die wirtschaftliche Entwicklung versus Inanspruchnahme der Umwelt – wird der Analyseaufwand überproportional erhöht. Die Gewichtung der Ergebnisse aus den verschiedenen Zieldimensionen wird komplizierter: Wie viel Umweltbelastung sollte toleriert werden, um eine Maßnahme im Bereich der Produktionsausweitung als erfolgreich zu werten? Solche Fragestellungen wären von den Evaluatoren nur beantwortbar, wenn diese Querschnittszielsetzungen konsequent in das Zielsystem der Interventionen auf allen Ebenen integriert wären und ihre politische Gewichtung von den Trägern der Förderung klargestellt würde. Auf Programmebene fehlt diese Messlatte jedoch weitgehend.

Wenn man sich in der Programmbewertung darauf beschränkt, die Ergebnisse der Förderkapitel und zusätzlich gelieferte Informationen ausschließlich zusammenzufassen, ohne eine Aussage zur Effektivität zu treffen, steht man vor einem weiteren Problem der Aggregierbarkeit von Daten. Abbildung 1 zeigt am Beispiel der Umweltfrage, wie heterogen die Ergebnisse der einzelnen Förderkapitel bezüglich des Schutzes einer ausgewählten Ressource (hier: Wasserqualität) sind. Auf dieser Grundlage lässt sich der auf Programmebene vorgesehene Indikator nicht quantifizieren.

Die EU-Kommission sieht in ihrem Bewertungsleitfaden analog zu den kapitelspezifischen Bewertungsfragen eine vorrangig auf quantitative Indikatoren gestützte Beantwortung der kapitelübergreifenden Fragen als Aggregation von der Mikro- zur Makroebene vor. Diese Zusammenschau wird durch den unterschiedlichen Vollzug in den einzelnen Förderkapiteln bzw. eine unterschiedliche Datenlage und Bearbeitungsstände erheblich erschwert und verzerrt. Ungenauigkeiten und Annahmen auf dieser Ebene werden hierdurch aufsummiert. Eine zusätzliche Schwierigkeit ergibt sich durch den teilweise subjektiven Charakter von qualitativen Einschätzungen, denen zumeist eine rein ordinale Skalierung ohne äquidistante Intervalle zugrunde liegt.

Abbildung 1: Problemlage bei der Verdichtung von Untersuchungsergebnissen der Kapitel auf einen Querschnittsindikator aus dem Bereich Umwelt



Aufgrund der genannten Schwierigkeiten wurde für die Programmbewertung im Rahmen der 6-Länder-Halbzeitbewertung ein vorrangig qualitativer Ansatz gewählt, mit dem Wirkungsrichtungen auf Kriterienebene dargestellt werden. Da die oben skizzierten Schwierigkeiten einer Programmbewertung auch weiterhin Bestand haben, ist ein solcher qualitativer Ansatz zielführender als ein Versuch, quantitative (Teil-) Ergebnisse unterschiedlich messbarer Wirkungen mit vorläufigem Charakter zusammenzuführen. Des Weiteren wurde ein deutliches Gewicht auf die Analyse der Programmgenese und der Umsetzungsbedingungen auf Programmebene (Finanzmanagement und Verwaltung, Kontrolle, Sanktionen) gelegt. Dagegen konnten die traditionellen Aufgaben der Programmbewertung wie Prüfung von Relevanz, interner und externer Kohärenz sowie Effektivität und Effizienz u.a. wegen des Fehlens in sich schlüssiger Programmstrategien nur partiell untersucht werden.

5.6 Erfahrungsaustausch (Reiter)

Perspektiven der Zusammenarbeit der Evaluatoren in Deutschland

Das Personalmanagement der Evaluierungsgruppen unterliegt engen Finanzrestriktionen. Für die Mehrheit der Programmevaluierungen und für alle zentralen Evaluierungen gilt, dass die Evaluierungsaufträge nicht den gesamten Evaluierungszyklus umfassen, sondern z.T. nur Einzelevaluierungen wie die jetzt vorliegenden Zwischenbewertungen. Folgeauf-

träge für ein up-date werden nicht im unmittelbaren Anschluss an die Zwischenbewertung vergeben, sodass „Auftragslücken“ entstehen. Hieraus resultiert, dass einige Arbeitsverträge auslaufen und Humankapital abwandert.

Der beschriebene Prozess ist zwar aus Sicht der Auftragnehmer misslich, spiegelt jedoch den Alltag von Projektgruppen am Wissenschaftsmarkt wider. Positiv ist sicherlich, dass Konkurrenz das Geschäft belebt. Negativ erscheint, dass eine Permanenz der Evaluierung kaum gegeben ist, insbesondere was übergeordnete Evaluierungsaspekte wie bspw. die Ableitung von Vorschlägen zur Fortentwicklung des Bewertungsrasters aus deutscher Sicht betrifft. Erschwerend wirkt weiterhin, dass ein institutionalisierter Austausch zwischen den Evaluierungsgruppen in Deutschland nicht besteht. Dieser findet zwar in Teilen selbst organisiert statt, bleibt jedoch eher zufällig. Kein Evaluatorenteam verfügt über die notwendigen Finanzen und personellen Ressourcen, um entsprechende Veranstaltungen mit der notwendigen Kontinuität inhaltlich und organisatorisch vorzubereiten und auszurichten.

Der Mangel eines institutionalisierten Austauschs spiegelt sich auf den folgenden Ebenen wider:

- Die Kommunikation mit der EU-Kommission hinsichtlich einzelner Evaluierungsfragen wird seitens der Evaluierer als unzureichend eingestuft,
- Reibungsverluste zwischen zentralen Evaluatoren und Programmevaluatoren entstehen,
- eine Methodendiskussion und deren Fortentwicklung findet nicht statt.

Aus den benannten Defiziten leitet sich die Forderung nach einer institutionalisierten Plattform zum Zweck des Austauschs und der teilweisen Koordination der Evaluatorengruppen ab. Die Aufgabe der „Koordinierungsstelle Evaluierung“ liegt darin

a) den Kontakt zwischen EU-Kommission und Evaluierung herzustellen und zu bündeln;

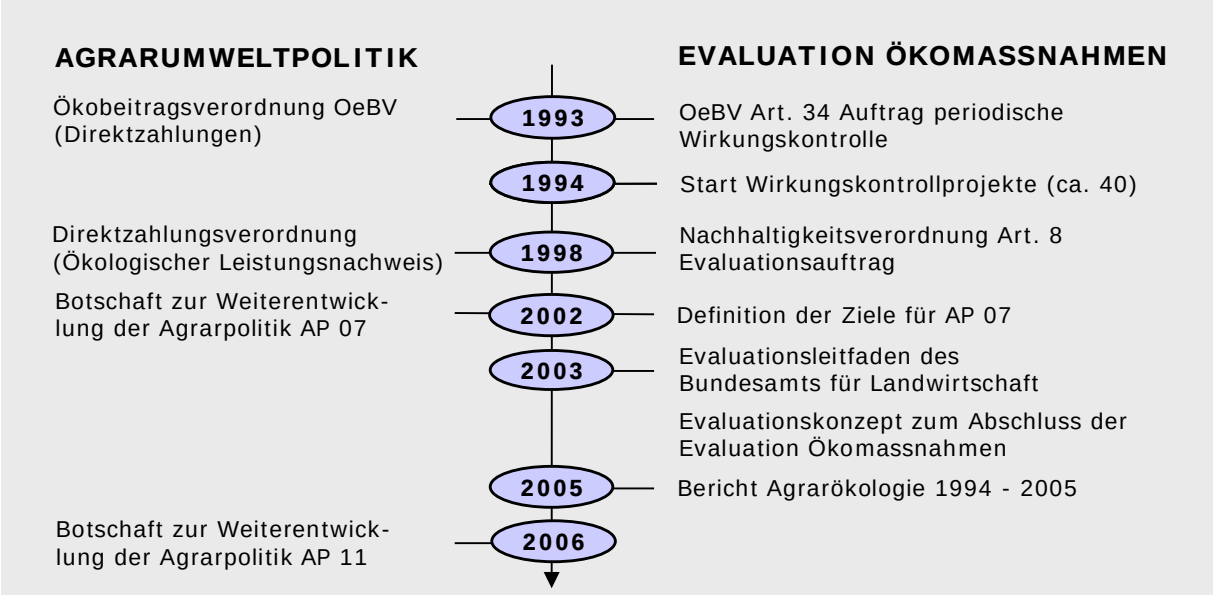
b) den Austausch zwischen den Evaluatoren auf den unterschiedlichen Bewertungsebenen zu organisieren und zu moderieren. Die inhaltliche Gestaltung des Austausch obliegt den Mitgliedern der jeweiligen Arbeitsgruppen, also den Evaluatoren.

Um Missverständnissen vorzubeugen: Ziel einer nationalen „Koordinierungsstelle Evaluierung“ soll es nicht sein, über die Hintertür die methodisch-inhaltlichen Freiheiten der Programm- bzw. der zentralen Evaluatoren zu beschneiden. Stattdessen soll die Möglichkeit eines kontinuierlichen Austauschs geschaffen werden, um langfristig die Qualität der Evaluierung zu erhalten und möglichst zu verbessern.

Selbstverständlich ist auch, dass eine solche Koordinierungsfunktion finanziert werden muss. Anzuraten ist eine anteilige Finanzierung durch die Länder und den Bund als Verantwortliche für die Evaluierung.

5.7 Evaluation der Ökomaßnahmen 1994 bis 2005 in der Schweiz (Peter, Hefti)

Anlass



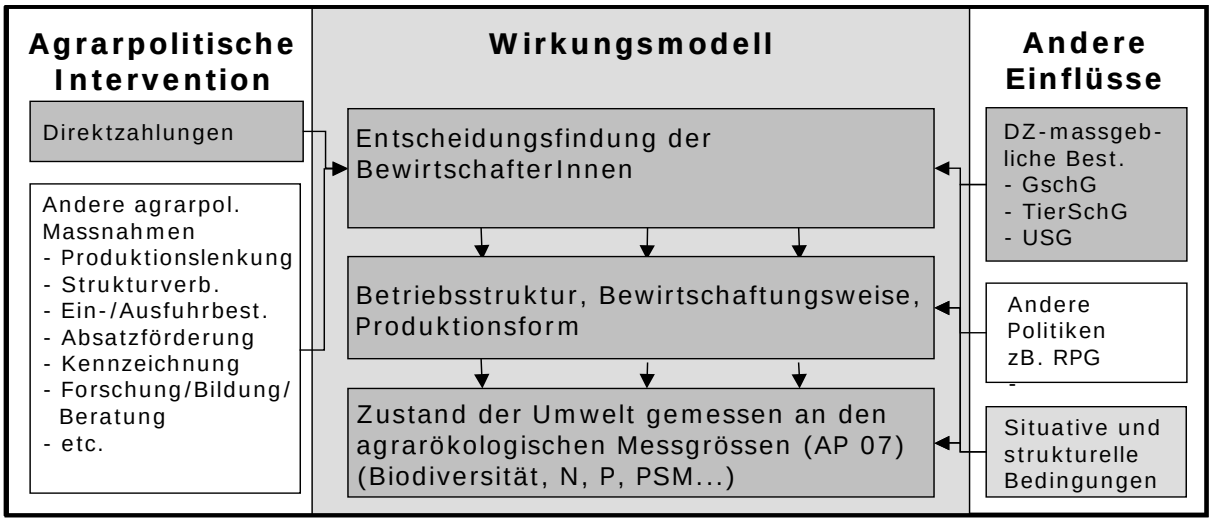
Ziele

Die Resultate der Evaluation werden im Bericht Agrarökologie 1994-2005 zusammengefasst. Die umfassende Bewertung und Analyse der agrarökologischen Ziele (Botschaft AP 07) stützt sich auf Erfahrungen und Ergebnisse aus den Wirkungskontrollprojekten der Ökomaßnahmen und Tierhaltungsprogramme 1994 – 2005 ab.

Die Evaluation dient der Berichterstattung und soll ebenfalls Antworten auf spezifische und zukünftige Fragestellungen der Landwirtschaftspolitik liefern. Die Evaluation verhilft zu verbesserten Maßnahmen und Programmen in der schweizerischen Agrarpolitik.

Abgrenzung des Evaluationsgegenstands und Definition des Wirkungsmodells

Die Direktzahlungen für den ökologischen Leistungsnachweis der Betriebe und deren Auswirkungen auf den Zustand der Umwelt sind Gegenstand der Evaluation.



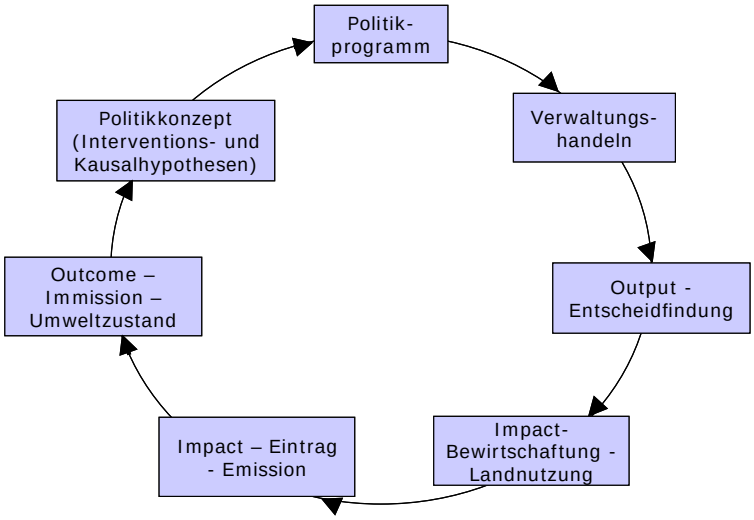
Das Wirkungsmodell zeigt die Intervention, ihre Wirkung auf die Entscheidfindung und damit auf die Ökologisierung von Bewirtschaftung und Produktion. Mit dem Wirkungsmodell wird anderen Einflüssen auf den Erfolg der Politik auf allen Ebenen Rechnung getragen.

Die Methode der Politikevaluation

Der Politikzyklus

Die Evaluationsmethode nach Bussmann et al.[19] orientiert sich am Politikzyklus. Evaluiert wird der ganze Umsetzungsprozess. Die Evaluation beginnt mit der Prüfung des Politikkonzeptes, analysiert die Verwaltungstätigkeiten für den Vollzug und bewertet die Wirkungsebenen der Politik: Output, Impact und Outcome. Die Evaluation schließt den Zyklus, indem sie Ergebnisse und Empfehlungen zur Verbesserung der Maßnahmen in das Politikkonzept einbringt.

[19] Bussmann W., Klöti U., Knoepfel P. (Hrsg.), 1997: Einführung in die Politikevaluation; Verlag Helbling & Lichtenhahn Basel



Analyse des Politikprogramms und Formulierung von Detailfragen

Das Politikprogramm bezeichnet die Vorgaben für den Vollzug und gibt damit die Messlatten vor. Das Programm wird in seine Elemente zerlegt, um die Detailfragen zu formulieren.

Einteilung des Politikprogramms nach:	Programmvorgaben (Beispiele zu Nitrat im Wasser):	Detailfragen (Beispiele):
Zielbeschrieb	Die Nitratgehalte im Grund- und Quellwasser dürfen 40mg $N0_3$/l nicht mehr überschreiten.	Wie hoch ist die $N0_3$-Belastung des Grund- und Quellwassers?
Evaluatorische Elemente, Methoden, Messungen	Messungen werden in den Trinkwasserfassungen durchgeführt.	Stammen die Messungen aus repräsentativen Trinkwasserfassungen?
Emissionen - Verhalten - Landnutzung (Impacts)	Gezielte Bewirtschaftung d.h. geeignete Kombinationen im Bereich der Kulturwahl, der Fruchtfolge und der Anbautechnik.	Wie ist der Zusammenhang zwischen Bewirtschaftungsweise und dem Eintrag von N in Gewässer?
Instrumente (Outputs)	Abgeltung gezielter Bewirtschaftung (Ökobeiträge).	Sind die Beiträge ausreichend um eine angestrebte Beteiligung zu erhalten?
Kompetenzen, Ressourcen	Der Bund bezahlt 80%, die Kantone 20% der Beiträge.	Verfügen die Kantone über genügende Ressourcen um die Mitfinanzierung zu sichern?
Verfahrensregeln, Programme	Wird die Zielvorgabe mit freiwilligen Vereinbarungen nicht erreicht wird die Bewirtschaftungsänderung verfügt.	Kann die Zielerreichung durch ungenügende Beteiligung erklärt werden?

Die Evaluationsfragen

Aussagen zu Relevanz, Effektivität und Effizienz der Maßnahmen ermöglichen die Weiterentwicklung der Agrarpolitik. Die Mehrheit der Evaluationsfragen nimmt diese Kriterien gemäss dem Evaluationsleitfaden des Bundesamtes für Landwirtschaft auf.

Relevanz: Trägt die Massnahme wirklich zur Erreichung der übergeordneten politischen Ziele bei?

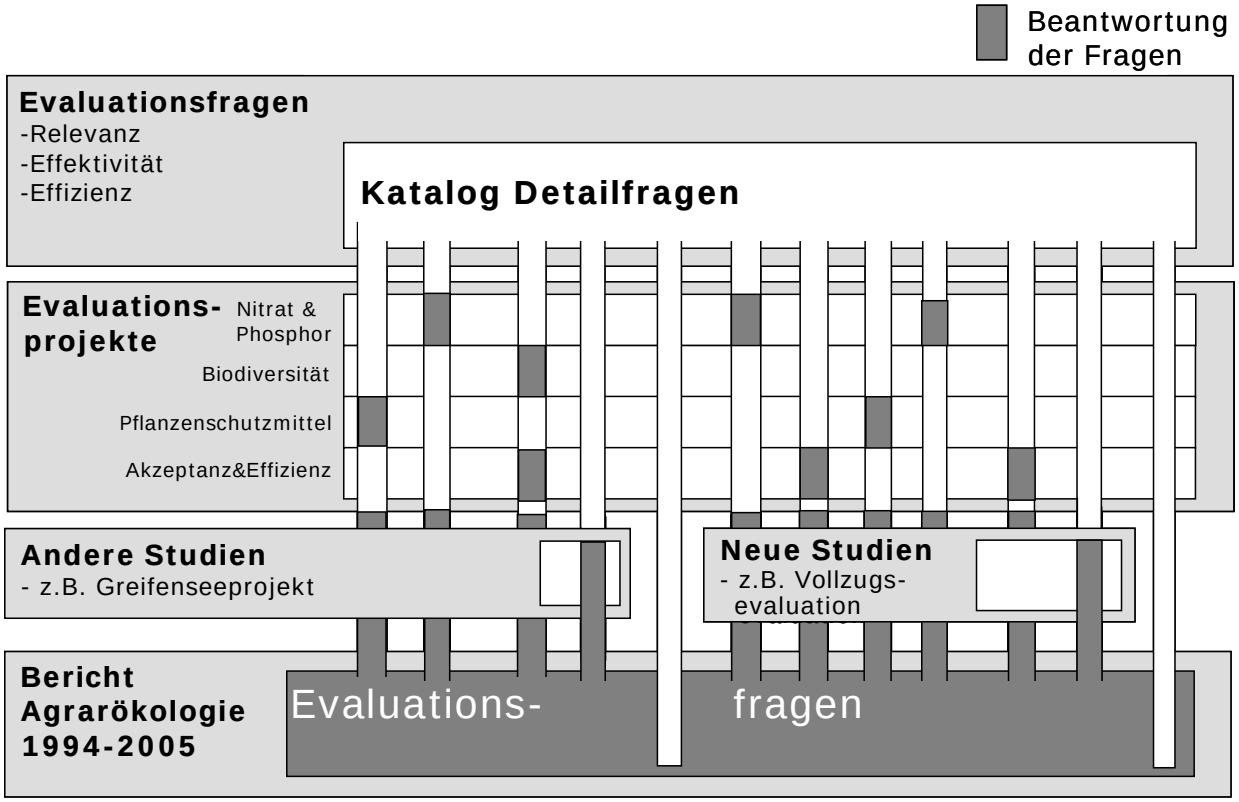
Effektivität: Können mit der Massnahme die definierten Sachziele erreicht werden (=Wirksamkeit)?

Effizienz: Werden die Mittel optimal eingesetzt? Ist der Vollzug optimal gestaltet?

Zusätzliche Evaluationsfragen wurden anlässlich eines internen Workshops im Bundesamt für Landwirtschaft gesammelt. Es handelt sich dabei um aktuelle und potenzielle zukünftige Fragen zu Ökomaßnahmen.

Von den Evaluationsfragen zu den Detailfragen

Die Detailfragen dienen als Grundlage für die Beantwortung der Evaluationsfragen. In der Regel erfordern die Antworten Informationen aus mehreren Bereichen und damit die Behandlung mehrerer Detailfragen. Die Evaluation soll so weit wie möglich das Wissen aus den seit 1994 laufenden Wirkungskontrollprojekten verwerten.



Der Katalog mit Detailfragen und eine Anleitung für die ProjektverfasserInnen stellen sicher, dass die Erkenntnisse für die nach der Methode der Politikevaluation konzipierten Auswertung verwendbar sind. Anleitung und Katalog werden in einem interaktiven Pro-

zess auf Verständlichkeit geprüft und bereinigt. Der Zusammenhang zwischen Evaluations- und Detailfragen wird von EvaluationsspezialistInnen hergestellt. Fragen, zu denen keine Daten vorliegen oder welche nur teilweise beantwortet werden können, geben Anhaltspunkte über zusätzlichen Forschungs- und Evaluationsbedarf.

6 Schlussbemerkungen (Pufahl, Forstner)

Am Ende des Workshops stellt sich die Frage, inwieweit die mit dessen Durchführung verbundenen Erwartungen erfüllt werden konnten und welche Ergebnisse und Empfehlungen im Hinblick auf die künftige Evaluationsarbeit festzuhalten sind. Wesentliche Ziele des Workshops waren

(1) ein strukturierter Erfahrungsaustausch der Evaluatoren über Inhalte und Ergebnisse der abgeschlossenen Halbzeitbewertung,

(2) eine Diskussion über organisatorische Fragen der Bewertung,

(3) das Knüpfen persönlicher Kontakte mit dem Ziel einer Verbesserung des Erfahrungsaustauschs.

Wenngleich nicht alle Ziele vollständig erreicht werden konnten, ziehen die Veranstalter in Übereinstimmung mit den Teilnehmern doch ein überwiegend positives Resümee. Besonders erfreulich ist das rege Interesse am Workshop, das sich in der großen Teilnehmerzahl zeigte. Sie ist als eindeutiges Indiz für den Bedarf eines fachlichen Austausches zwischen den Evaluatoren – und dies nicht nur auf nationaler Ebene – zu werten. Die offene und stets konstruktive Diskussionsatmosphäre trug wesentlich zum Erfolg des Workshops bei. Sicherlich war es ein besonderer Pluspunkt der Veranstaltung, dass die Europäische Kommission bei vielen der beteiligten Evaluatoren erstmals ein konkretes Gesicht erhielt und durch Herrn Clemens sehr offen über Beweggründe bestimmter Evaluationsvorgaben und über künftige Schritte im Zusammenhang mit den anstehenden Bewertungsaufgaben Auskunft gab.

Inhalt des Erfahrungsaustausches

Der Erfahrungsaustausch machte deutlich, dass die von der Europäischen Kommission erarbeiteten Bewertungsfragen von den Evaluatoren begrüßt und als ein nützliches Instrument zur Standardisierung der Bewertung eingeschätzt werden. Die Bearbeitung der Fragen war jedoch bei allen Evaluatoren von ähnlichen Problemen begleitet. Schwierigkeiten ergaben sich insbesondere durch das weit gehende Fehlen von repräsentativen Daten sowie die meist nur schwer konstruierbare Referenzsituation. Als häufig gewählte Lösungsansätze zur Bewältigung der Datenmängel wurden Befragungen, Workshops, Interviews und Fallstudien gewählt. Während die Evaluatoren ganz überwiegend davon ausgingen, dass das von der Europäischen Kommission vorgegebene Bewertungsraster mit Fragen, Kriterien und Indikatoren einen weitgehend verbindlichen Charakter hat, wies Herr Clemens darauf hin, dass die vorgegebenen Indikatoren ein erhebliches Maß an Flexibilität beinhalteten. Diesbezüglich bestand offensichtlich ein erhebliches Informationsvermittlungsproblem.

Hinsichtlich der von der Europäischen Kommission vorgeschlagenen Kriterien und Indikatoren wurde kritisiert, sie seien nicht ausgewogen, unvollständig und teilweise auch viel zu anspruchsvoll. Die von Herrn Clemens formulierte Aufforderung an die Evaluatoren, Verbesserungsvorschläge in diesem Bereich zu erarbeiten und an die Europäische Kommission weiterzuleiten, sollte daher seitens der Evaluatoren unbedingt aufgegriffen werden. Der derzeitige Stand des Bewertungsrasters dürfte ansonsten dazu führen, dass jeder Evaluator in die Lösung des Problems lückenhafter oder ungeeigneter Fragen-Kriterien-Indikatoren-Ketten erhebliche Mühen und Ressourcen investiert, um zu individuellen Ergebnissen zu kommen, die in vielen Fällen dennoch unbefriedigend sind. Dies ist ein eher ineffizientes Verfahren, dass nur einem anfänglichen Lernprozess geschuldet sein sollte. Die Nutzung dieses erheblichen Spielraumes zur Beantwortung der Fragen durch sehr individuelle Ansätze erschwert eine Vergleichbarkeit zwischen Mitgliedstaaten oder macht sie unmöglich. Die Evaluatoren favorisierten die Festlegung einiger weniger Kernfragen, die jedoch für alle Mitgliedstaaten nach einheitlichem Muster (Daten, Methodik) beantwortet werden sollten. Darüber hinaus sollte möglichst große Flexibilität für die Evaluatoren herrschen.

Die thematisch gegliederten Diskussionen und deren Ergebnisse hatten aufgrund der zeitlichen Begrenzung vielfach nur einen allgemeinen Charakter. Nur so war ein Überblick über die Stärken, Schwächen und Kritikpunkte der gegenwärtigen Praxis der Programm- und Kapitelbewertung möglich. Auf Grund der Unterschiedlichkeit der in den Kapiteln verwendeten Daten, Methoden sowie der Ergebnisse hat dies eine konkretere Diskussion über inhaltlich-methodische Fragen der Kapitelbewertung nicht in ausreichendem Maß ermöglicht. Neue Lösungsansätze oder „Best-Practise“ Beispiele für zukünftige Evaluierungen konnten nicht im erhofften Umfang entwickelt werden. Der inhaltlich-methodische Erfahrungsaustausch sollte daher zukünftig auf Ebene der einzelnen Kapitelbewerter bzw. der kapitelübergreifenden Programmbewerter organisiert werden.

Organisation des Erfahrungsaustausches

Der Workshop wurde durch die Bildung von Arbeitsgruppen und die Vorgabe von zahlreichen Diskussionspunkten in starkem, vielleicht zu starkem Maß vorstrukturiert. Das insgesamt sehr ehrgeizige Arbeitsprogramm hat rückschauend sicherlich an manchen Stellen den Raum für eine tiefergehende inhaltliche Diskussion beschnitten. Folglich ist bei ähnlichen Veranstaltungen in der Zukunft darauf zu achten, das Programm nicht zu überfrachten und einer breiten inhaltlichen Diskussion bei wichtigen Sachfragen den Vorzug vor einer scheinbaren und standardisierten Vollständigkeit zu geben.

Die Vergabe von Impulsreferaten zur inhaltlichen Vorbereitung der Teilnehmer auf den Workshop hat sich für den Zweck des Erfahrungsaustausches bewährt. Die kapitelübergreifende Zusammenfassung der Diskussionsergebnisse am Vormittag des zweiten Tages konnte nur verallgemeinernden Charakter haben und ging zu Lasten der kapitelspezifi-

schen Fragestellungen. Diskussionen haben zu den einzelnen Bereichen meist nicht mehr stattgefunden, sodass der Mehrwert dieses Programmteils begrenzt war.

Die vorgegebenen thematischen Diskussionspunkte waren durchweg relevant. Einige Themen, z.B. über das Verhältnis zwischen zentraler und dezentraler Bewertung, waren der föderalen Struktur Deutschlands geschuldet. Nicht zuletzt aus diesem Grund kam die überwiegende Anzahl der Teilnehmer aus Deutschland, insbesondere von der FAL. Dies führte teilweise zur Dominanz der „spezifisch deutschen Sichtweise“. Teilnehmende Ländervertreter aus Österreich, Belgien und der Schweiz haben schon auf Grund der strukturellen Unterschiede in der Organisation der Evaluierung – Österreich besitzt beispielsweise nur ein Entwicklungsprogramm – einen anders gelagerten Diskussionsbedarf zu organisatorischen Fragen der Evaluierung.

Die Zusammenstellung der Arbeitsgruppen hatte vor allem pragmatische Gründe, weil für die Einzelmaßnahmen nicht genügend Diskussionsteilnehmer vorhanden waren. Eine fruchtbringende Diskussion müsste aus fachlicher Sicht jedoch vielfach auf der Ebene der Einzelfördermaßnahmen stattfinden.

Aus all den Erfahrungen dieses Workshops ziehen die Veranstalter der FAL den Schluss, dass der Austausch zwischen Evaluatoren zukünftig in einem kleineren, dezentralen Rahmen erfolgen sollte. Hier bietet sich die Ebene der einzelnen Kapitel bzw. der kapitelübergreifenden Fragen an. Dies hätte den Vorteil, dass

- sich der organisatorische Aufwand für eine überschaubare Anzahl von Kapitelbewertern in Grenzen hält,
- gemeinsame inhaltliche Arbeitsschwerpunkte gesetzt werden können und
- eine „Doppelzuständigkeit“ einzelner Evaluierer für Einzelkapitel und kapitelübergreifende Fragen nicht automatisch zum Ausschluss bei anderen Arbeitsgruppen führt.

Daneben sollten diese Kapitelworkshops von Zeit zu Zeit von einem übergreifenden Workshop begleitet werden, um inhaltlich und methodisch voneinander lernen zu können sowie um Synergien im Bereich der Datensammlung und –auswertung entwickeln zu können.

Generell sind bei dieser Vorgehensweise jedoch einer Diskussion auf nationaler Ebene enge Grenzen gesetzt, da häufig nur eine sehr begrenzte Anzahl an Bewertern existiert. Dies gilt vor allem für die Bereiche, in denen in Deutschland eine zentrale Bewertung durchgeführt wird. Es ist es daher unbedingt erforderlich, einschlägig erfahrene Bewerter aus verschiedenen Mitgliedstaaten zusammenzubringen. Die für diesen Workshop getroffene Einschränkung auf deutschsprachige Bewerter reicht dabei keineswegs aus.

Die Tatsache, dass ein Erfahrungsaustausch seitens der Evaluatoren, der Europäischen Kommission und des BMVEL als sinnvoll erachtet wird, zeigt, dass die Politikevaluation zunehmend als kontinuierlicher Lernprozess aufgefasst wird. Dennoch wurde auf dem Workshop zu Form und Trägerschaft eines solchen Prozesses keine weiterführenden Diskussion geführt. Ein BMVEL-Vertreter regte lediglich an, dass die Durchführung von Workshops zur Verbesserung des Informationsflusses zwischen zentralen und Programmbewertern künftig Bestandteil der Evaluationsverträge werden solle. Eine Verantwortlichkeit der Europäischen Kommission auf EU- und des Bundes auf nationaler Ebene für die Organisation eines künftigen Erfahrungsaustausches wurde dagegen leider nicht erkennbar.

Anhang I

Programm

27. Januar 2004

12:30 **Einchecken** (Kaffee, Getränke)

13:00 **Begrüßung**

13:05 **Überblick über Inhalt, Ziele und Ablauf des Workshops**

13:15 **Evaluation aus Sicht der EU-Kommission** (Clemens, EU-KOM)

14:00 Arbeitsgruppen: **Kapitelbewertung - Methodik und Ergebnisse**

Arbeitsgruppen:

AG_1: Investitionen in landwirtschaftlichen Betrieben; Verarbeitung & Vermarktung

AG_2: Ausgleichszulage

AG_3: Agrarumweltmaßnahmen

AG_4: Dorf- und ländliche Regionalentwicklung

AG_5: Forstwirtschaft

Diskussionspunkte für alle Arbeitsgruppen (jeweils mit fünfminütigem Impulsreferat):

1. Beantwortung kapitelspezifischer Fragen
2. Inhalte kapitelspezifischer Fragen
3. Bewertungsergebnisse
4. Umsetzungsanalyse
5. Programmqualität
6. Verhältnis Auftraggeber-Evaluator

Erläuterung im Begleitschreiben

15:30 **Kaffeepause**

16:00 **Forts. der Gruppenarbeit**

18:00 **Ende**

Abendprogramm

19:30 **Abendessen**

28. Januar 2004

9:00 Plenum: **Präsentation der Ergebnisse der Arbeitsgruppen** (Strukturierung anhand der Diskussionspunkte vom Vortag)

12:00 **Vortrag zur Evaluationsmethodik**: Vergleich der Ziel 5b-Bewertung mit den aktuellen Anforderungen der VO (EG) Nr. 1257/1999 (Clerckx, Belgien)

12:30 **Mittagspause**

13:30 Plenum: **Programmbewertung und Perspektiven der Evaluation** (jeweils mit fünfminütigem Impulsreferat):

1. Besonderheiten der Ziel 1-Bewertung
2. Zentrale/dezentrale Bewertung
3. Summative/formative Bewertung
4. Idealer Bewertungszyklus
5. Programmansatz und kapitelübergreifende Bewertung
6. Erfahrungsaustausch

Erläuterung im Begleitschreiben

15:30 **Ende des Workshops**

Ziel und Ablauf des Workshops

Ziel: Der Workshop wendet sich an Evaluatoren in Deutschland und benachbarten EU-Mitgliedsstaaten, die an der Zwischenbewertung der Programme zur Entwicklung des ländlichen Raumes oder der Operationellen Programme für den Förderzeitraum 2000 bis 2006 mitgewirkt haben. Ziel des Workshops ist ein Erfahrungsaustausch zwischen Evaluatoren zur Verbesserung der Programmbewertung (up-date, ex-post). Aus organisatorischen Gründen ist die **Konferenzsprache Deutsch.**

1. Tag: Der Workshop beginnt mit einer zentralen Auftaktsitzung und gliedert sich dann in die vorgeschlagenen Arbeitsgruppen. Die festgelegten Diskussionspunkte geben jeder Arbeitsgruppe einen einheitlichen Rahmen. Jeder Diskussionspunkt wird durch ein fünfminütiges Impulsreferat eingeleitet und anschließend diskutiert. Jeder Teilnehmer sollte mindestens ein Impulsreferat halten (Thesen, besonders relevante Aspekte, Schwierigkeiten). Die Zuordnung zu einer Arbeitsgruppe und Bereichswahl für das Impulsreferat erfolgt mit der Anmeldung.

2. Tag: Am zweiten Tag werden die Ergebnisse der Gruppenarbeit anhand der strukturierten Diskussionspunkte im Plenum diskutiert. Am Nachmittag stehen die Programmbewertung und die Perspektiven der Evaluation im Mittelpunkt. Jedes Thema wird wiederum durch ein Impulsreferat eingeleitet. Die Workshopergebnisse werden aufbereitet und veröffentlicht.

Anmeldung und Rückfragen

Eine Anmeldung zum Workshop ist erforderlich. Anmeldeschluss ist der 12. Dezember 2003 (siehe beiliegendes Anmeldeformular). Rückfragen und Anmeldung bei:

Regine Albert regine.albert@fal.de
Tel + 49 (0)5 31/5 96 5316, Fax -5399
Bernhard Forstner
bernhard.forstner@fal.de
Tel + 49 (0)5 31/5 96 5233, Fax -5199

Übernachtungsmöglichkeiten entnehmen Sie bitte dem beiliegenden Blatt „Hotelübersicht“.

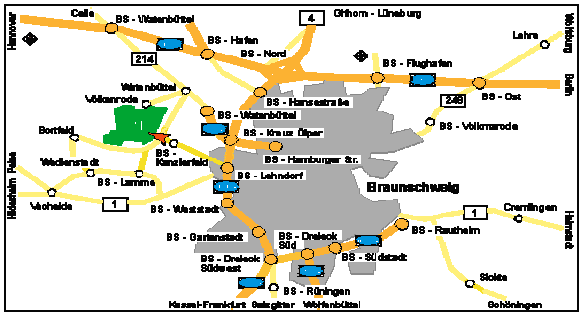
Veranstaltungsort

Institut für Betriebswirtschaft (BW), Institut für Ländliche Räume (LR) und Institut für Marktanalyse und Agrarhandelspolitik (MA) der FAL.

Organisation

Bernhard Forstner (BW), Regina Grajewski, Birgit Koch, Dr. Reiner Plankl, Andrea Pufahl (jeweils LR), Dr. Josef Efken, Dr. Heinz Wendt, Regine Albert (jeweils MA).

Anreise



***Bus**: Linien 411(Lamme),416 (PTB) bis Haltestelle FAL, dann 10 Min. Fußweg*

FAL
Bundesforschungsanstalt für Landwirtschaft

Workshop zur

Zwischenbewertung der Programme zur Entwicklung des ländlichen Raumes nach VO (EG) Nr. 1257/1999

Erfahrungsaustausch und Verbesserungsansätze

27./28. Januar 2004

Anhang II

Teilnehmerliste Evaluatoren Workshop
Zwischenevaluation VO 1257/1999
27./28. Januar 2004, FAL Braunschweig

Beck, Monika
Universität Gembloux
Unite d'Economie Developpement rural
Passage des Deportes, 2

B-5030 Belgien
Telefon: 0032-81612371
e-mail: econet@fsagx.ac.be

Dr. Becker, Heinrich
Bundesforschungsanstalt für Landwirtschaft (FAL)
Institut für Ländliche Räume
Bundesallee 50
38116 Braunschweig
Telefon: 0531-596-5165
e-mail: heinrich.becker@fal.de

Bieling, Alfons
Institut für Forst- und Umweltpolitik, Universität Freiburg
Tennenbacherstr. 4
79085 Freiburg
Telefon: 0761-203-8493
e-mail: alfons.bieling@ifp.uni-freiburg.de

Bresemann, Sabine
Bundesanstalt für Forst- und Holzwirtschaft (BFH)
Abt. Ökonomie
Leuschnerstr. 91
21031 Hamburg
Telefon: 040-73962-320
e-mail: s.bresemann@holz.uni-hamburg.d

Clemens, Detlev
EU-Kommission
rue de la Loi 102, 01/11

B-1040 Brüssel
Telefon: 0032-2-2993730
e-mail: detlev.clemens@cec.eu.int

Dr. Clerckx, Edwin
ADE s.a. (AIDE a la Decision Economique)
Rue de Clairvaux 40
B-1348 Belgien
Telefon: '0032-10454510
e-mail: edwin.clerckx@ade.be

Deimer, Cornelia
Landgesellschaft Sachsen-Anhalt mbH
Neustädter Passage 6
06122 Halle
Telefon: 0345-6911123
e-mail: deimer.c@lgsa.de

Eberhardt, Winfried
Bundesforschungsanstalt für Landwirtschaft (FAL)
Institut für Ländliche Räume
Bundesallee 50
38116 Braunschweig
Telefon: 0531-596-5161
e-mail: winfried.eberhardt@fal.de

Eckstein, Karin
Lehrstuhl Wirtschaftslehre des Landbaus,
Technische Universität Weihenstephan
Alte Akademie 14

85350 Freising
Telefon: 08161-713492
e-mail: eckstein@wzw.tum.de

Dr. Efken, Josef
Bundesforschungsanstalt für Landwirtschaft (FAL)
Institut für Marktanalyse und Welthandelspolitik
Bundesallee 50
38116 Braunschweig
Telefon: 0531-596-5307
e-mail: josef.efken@fal.de

Elbe, Sebastian
SPRINT GbR
Feldbergstraße 40

64293 Darmstadt
Telefon: 06151-6677801
e-mail: elbe@sprintconsult.de

Enderling, Antje
Landgesellschaft Sachsen-Anhalt bmH
Große-Diesdorfer-Str. 56/57

39109 Magdeburg
Telefon: 03909-47040
e-mail: enderling.a@t-online.de

Essmann, Sandra
Bundesforschungsanstalt für Landwirtschaft (FAL)
Institut für Betriebswirtschaft
Bundesallee 50
38116 Braunschweig
Telefon: 0531-596-5167
e-mail: sandra.essmann@fal.de

Forstner, Bernhard
Bundesforschungsanstalt für Landwirtschaft (FAL)
Institut für Betriebswirtschaft
Bundesallee 50
38116 Braunschweig
Telefon: 0531-596-5233
e-mail: bernhard.forstne@fal.de

Grajewski, Regina
Bundesforschungsanstalt für Landwirtschaft (FAL)
Institut für Ländliche Räume
Bundesallee 50
38116 Braunschweig
Telefon: 0531-596-5217
e-mail: regina.grajewski@fal.de

Hefti, Helena
Bundesamt für Landwirtschaft
Mattenhofstraße 5
CH-3003 Bern
Telefon: 0041-31-3225733
e-mail: helena.hefti@blw.admin.ch

Dr. habil. Hochberg, Hans
Thüringer Landesanstalt für Landwirtschaft
Referat 720
Bahnhofstraße 1a
99869 Wandersleben
Telefon: 036202-9780
e-mail: h.hochberg@wandersleben.tll.de

Englert, Hermann
Bundesanstalt für Forst- und Holzwirtschaft (BFH)
Abt. Ökonomie
Leuschnerstr. 91
21031 Hamburg
Telefon: 040-73962-322
e-mail: h.englert@holz.uni-hamburg.de

Fährmann, Barbara
Bundesforschungsanstalt für Landwirtschaft (FAL)
Institut für Betriebswirtschaft
Bundesallee 50
38116 Braunschweig
Telefon: 0531-596-5179
e-mail: barbara.faehrmann@fal.de

Geberth, Franz
Landesanstalt für Entwicklung der Landwirtschaft und der Ländlichen Räume (LEL)
Oberbettringstr. 162

73525 Schwäbisch-Gmünd
Telefon: 07171-917-221
e-mail: franz.geberth@lel.bwl.de

Prof. Dr. Großkopf, W.

70599 Stuttgart
Telefon: 0711-459-2631
e-mail: apo420@uni-hohenheim.de

Dr. Heyer, W.
Universität Halle
Ludwig-Wucherer-Str. 2
06108 Halle
Telefon: 0345-5522632
e-mail: abraham@landw.uni-halle.de

Dr. Horlitz, Thomas
ARUM / entera
Alte Herrenhäuser Straße 32

30451 Hannover
Telefon: 0511-16789-17
e-mail: horlitz@entera.de

Dr. Hovorka, Gerhard
Bundesanstalt für Bergbauernfragen
Marxergasse 2
A-1030 Wien
Telefon: 0043-1-5048869
e-mail: gerhard.hovorka@babf.bmlfuw.gv.at

Janetschek, Hubert
Bundesanstalt für Agrarwirtschaft
Marxergasse 2
A-1030 Wien
Telefon: 0043-1-8773651-7414
e-mail: hubert.janetschek@awi.bmlfuw.gv.a

Prof. Dr. Karl, H.

44780 Bochum
Telefon: 0234-3225332
e-mail: helmut.karl@ruhr-uni-bochum.de

Koch, Birgit
Bundesforschungsanstalt für Landwirtschaft
(FAL)
Institut für Ländliche Räume
Bundesallee 50
38116 Braunschweig
Telefon: 0531-596-5175
e-mail: birgit.koch@fal.de

Dr. Lüdigk, Rainer
ISW (Institut für Strukturpolitik und
Wirtschaftsforschung)
Heinrich-Heine-Straße 10
06114 Halle
Telefon: 0345-5213618
e-mail: r.luedigk@isw-institut.de

Mark, Steffen
Landesanstalt für Entwicklung der Landwirtschaft
und der Ländlichen Räume (LEL)
Oberbettringstr. 162
73525 Schwäbisch-Gmünd
Telefon: 07171-917-221
e-mail: steffen.mark@lel.bwl.de

Ortner, Karl
Bundesanstalt für Agrarwirtschaft
Marxergasse 2
A-1030 Wien
Telefon: +431-0877 3651 7482
e-mail: karl.ortner@awi.bmlfuw.gv.at

Jacobs, Ronald
Landgesellschaft Sachsen-Anhalt mbH
Am Eichengrund 3
38486 Klötze
Telefon: 03909-4704417
e-mail: jacobs.r@lgsa.de

Dr. Jirikowski, Wolfgang
BFW Forstliche Ausbildungsstätte
Johann Orth Allee 16
A-4810 Gmunden (A)
Telefon: 0043-7612-64419-0
e-mail: wolfgang.jirikowski@bfw.gv.at

Klare, Klaus
Bundesforschungsanstalt für Landwirtschaft
(FAL)
Institut für Ländliche Räume
Bundesallee 50
38116 Braunschweig
Telefon: 0531-596-5241
e-mail: klaus.klare@fal.de

Dr. Lins, Thomas
Bundesministerium für Verbraucherschutz,
Ernährung und Landwirtschaft (BMVEL)
Referat 421
Rochusstraße 1
53107 Bonn
Telefon: 0228-529-3986
e-mail: thomas.lins@bmvel.bund.de

Majer, Christoph
BMLFUW
Abteilung IV/3
Marxergasse 2
A-1030 Wien
Telefon: 0043-1-71100-7317
e-mail: christoph.majer@lebensministeriu

Matzdorf, Bettina
ZALF e.V. Müncheberg
Eberswalder Str. 84

15374 Müncheberg
Telefon: 033 432 82 150
e-mail: matzdorf@zalf.de

Dr. Otto, Rainer
Landgesellschaft Sachsen-Anhalt mbH
Neustädter Passage 6
06122 Halle
Telefon: 0345-6911151
e-mail:

Dr. Peter, Kathrin
Büro für Evaluationen, Projektmanagement und Consulting
Spitalackerstrasse 65
CH-3013 Bern
Telefon: 0041-31-3722025
e-mail: kathrin.peter@evaluationen.ch

Dr. Plankl, Reiner
Bundesforschungsanstalt für Landwirtschaft (FAL)
Institut für Ländliche Räume
Bundesallee 50
38116 Braunschweig
Telefon: 0531-596-5235
e-mail: reiner.plankl@fal.de

Preising, Andreas
Bundesforschungsanstalt für Landwirtschaft (FAL)
Institut für Betriebswirtschaft
Bundesallee 50
38116 Braunschweig
Telefon: 0531-596-5335
e-mail: andreas.preising@fal.de

Raue, Petra
Bundesforschungsanstalt für Landwirtschaft (FAL)
Institut für Betriebswirtschaft
Bundesallee 50
38116 Braunschweig
Telefon: 0531-596-5217
e-mail: petra.raue@fal.de

Roggendorf, Wolfgang
Bundesforschungsanstalt für Landwirtschaft (FAL)
Institut für Ländliche Räume
Bundesallee 50
38116 Braunschweig
Telefon: 0531-596-5171
e-mail: wolfgang.roggendorf@fal.de

Rudow, Katja
Bundesforschungsanstalt für Landwirtschaft (FAL)
Institut für Ländliche Räume
Bundesallee 50
38116 Braunschweig
Telefon: 0531-596-5235
e-mail: katja.rudow@fal.de

Pistrich, Karlheinz
Bundesanstalt für Agrarwirtschaft
Marxergasse 2

A-1030 Wien
Telefon: +431-0877 3651 7482
e-mail: karlheinz.pistrich@awi.bmlfuw.gv.at

Dr. Pölking, Andreas
agroplan
Am Exer 10

38302 Wolfenbüttel
Telefon: 05331-902771
e-mail: info@agroplan.de

Pufahl, Andrea
Bundesforschungsanstalt für Landwirtschaft (FAL)
Institut für Ländliche Räume
Bundesallee 50
38116 Braunschweig
Telefon: 0531-596-5177
e-mail: andrea.pufahl@fal.de

Reiter, Karin
Bundesforschungsanstalt für Landwirtschaft (FAL)
Institut für Ländliche Räume
Bundesallee 50
38116 Braunschweig
Telefon: 0531-596-5221
e-mail: karin.reiter@fal.de

Sander, Achim
ARUM / entera
Alte Herrenhäuser Straße 32

30451 Hannover
Telefon: 0511-16789-16
e-mail: sander@entera.de

Dr. Schinke, Maria
Bundesministerium für Verbraucherschutz,
Ernährung und Landwirtschaft (BMVEL)
Referat 521
Rochusstraße 1
53107 Bonn
Telefon: 0228-529-4328
e-mail: maria.schinke@bmvel.bund.de

Schramek, Jörg
Institut für Ländliche Strukturforschung (IfLS)
Zeppelinallee 31

60325 Frankfurt
Telefon: 069-97785779
e-mail: schramek@ifls.de

Thomas, Michael
FH Weihenstephan, Abt. Triesdorf
Forschungsgruppe Agrar- und
Regionalentwicklung
Steingruberstraße 2
91746 Weidenbach
Telefon: 0171-1415530
e-mail: michael.thomas@fh-weihenstephan

Dr. Trefflich, Annette
Landgesellschaft Sachsen-Anhalt mbH
Neustädter Passage 6

06122 Halle
Telefon: 0345-6911120
e-mail: trefflich.a@lgsa.de

Wagner, Klaus
Bundesanstalt für Agrarwirtschaft
Marxergasse 2

A-1030 Wien
Telefon: +431-87-736-517-428
e-mail: klaus.wagner@awi.bmlfuw.gv.at

Zucker, Judith
Bundesforschungsanstalt für Landwirtschaft
(FAL)
Institut für Betriebswirtschaft
Bundesallee 50
38116 Braunschweig
Telefon: 0531-596-5254
e-mail: judith.zucker@fal.de

Prof. Dr. Schrader, Helmut
Bundesforschungsanstalt für Landwirtschaft
(FAL)
Institut für Ländliche Räume
Bundesallee 50
38116 Braunschweig
Telefon: 0531-596-5237
e-mail: helmut.schrader@fal.de

Dr. Thissen, Gerhard
Bundesministerium für Verbraucherschutz,
Ernährung und Landwirtschaft (BMVEL)
Referat 522
Rochusstraße 1
53107 Bonn
Telefon: 0228-529-3330
e-mail: gerhard.thissen@bmvel.bund.de

Tietz, Andreas
Bundesforschungsanstalt für Landwirtschaft
(FAL)
Institut für Ländliche Räume
Bundesallee 50
38116 Braunschweig
Telefon: 0531-596-5169
e-mail: andreas.tietz@fal.de

Tripolt, Margarete
Bundesministerium für Land- und
Forstwirtschaft, Umwelt und Wasserwirtschaft
Stubenring 1
A-1012 Wien
Telefon: +43 1 71100 6775
e-mail: Margarethe.Tripolt@bmlfuw.gr.at

Dr. Wendt, Heinz
Bundesforschungsanstalt für Landwirtschaft
(FAL)
Institut für Marktanalyse und Welthandelspolitik
Bundesallee 50
38116 Braunschweig
Telefon: 0531-596-5312
e-mail: heinz.wendt@fal.de

Anhang III

AG_1

1. Inhalte der kapitelspezifischen Fragen/Kriterien/Indikatoren

	Fragen	Kriterien	Indikatoren
Relevanz	- gegeben	- gegeben	- teilweise gegeben - je nach Adressat unterschiedliche Ansprüche
Vollständigkeit	- gegeben	- überwiegend	- nein - unausgewogen (teilweise sehr differenziert/oberflächlich)
Ergänzungen		- Begriffe klarer definieren	- zahlreiche Ergänzungen notwendig, aber Flexibilität der Bewerter erhalten - Aggregationsproblem - Praxiserfahrung nutzen

AG_2

1. Inhalte der kapitelspezif. Fragen/Kriterien/Indikatoren

	Fragen	Kriterien	Indikatoren
Relevanz	- hohe Relevanz - nicht gleichmäßig gut zu beantworten	- nicht immer hinreichend (Gesellschaftsstruktur)	- grundsätzlich geeignet, jedoch stärker spezifizieren (teils nur Hilfsindikatoren)
Vollständigkeit	- weitgehend	- kausale Zusammenhänge bleiben offen	- nur bedingt mit verfügbaren Daten abzubilden
Ergänzungen	- stärkere Betonung des Zielerreichungsgrades - Ergänzungsfragenblock mit landespolitischer Bedeutung	- stärkere Ausdifferenzierung der Benachteiligung	- Indikatoren und Kontextindikatoren noch stärker konkretisieren

AG_3

1. Inhalte der kapitelspezifischen Fragen/Kriterien/Indikatoren

	Fragen	Kriterien	Indikatoren
Relevanz	relevant	relevant erachtet	zu viele ⇧ Flexibilität
		Definitionsprobleme für welche soft-facts bzw. hard-facts welche für Synthese relevant??? Verständnisprobleme	
	Land: Maßn- ansatz, EU: Progammansatz		
Voll-ständigkeit / Ergänzungen	Landschaft sinnvoll (+/-) : ökon. Aspekte	Tierschutz Bodenaspekte: Humus Sonderstandorte Luft	

AG_4

1. Inhalte der kapitelspezifischen Fragen/Kriterien/Indikatoren
⇨ Fragen Beantworten stand im Vordergrund

	Fragen	Kriterien	Indikatoren
Relevanz	÷ ja, alle	÷ nur teilweise	÷ ja (teilweise) ÷ schwierig zu erheben, zu detailliert ÷ unterschiedliche Herangehensweisen:
Voll-ständigkeit	÷ ja		÷ **Österreich:** Indikatorenkatalog nur als Vorschlagsliste genutzt ÷ **Sachsen-Anhalt:** Keine Ergänzungen (Flurbereinigung)
Ergänzungen		÷ in Ausnahme-fällen ergänzt	÷ **6-Länder:** Ergänzt v.a. bei indirekte Wirkungen ÷ **Bayern:** Realisierungsindikatoren ergänzt

AG_5

1. Inhalte der kapitelspezifischen Fragen/Kriterien/Indikatoren

	Fragen / Kriterien	Indikatoren
Relevanz	**Forstliche Ressourcen (Vorräte/Kohlenstoffspeicherung)** - grundsätzlich relevant - nicht immer Teil des Zielsystems des Landes **Sozio-ökonomische Effekte (Beschäftigung/Einkommen)** - grundsätzlich relevant - nicht immer Teil des Zielsystems des Landes **Schutzfunktionen** - relevant	- fast alle schwierig zu beantworten/ herzuleiten - sehr speziell formuliert - je nach Angebot der Maßnahmen können einzelne Indikatoren nicht verwendet werden - generell: Angaben über die Lage der Flächen in Schutzgebieten liegen nicht vor
Ergänzungen	**Erstaufforstungen** - regionale Verteilung **Sonstige** - aus Zeitgründen nicht möglich	

AG_1

2. Beantwortung der kapitelspezif. Fragen/Kriterien/Indikatoren

	Problempunkte	Lösungen
Daten/ Repräsen-tativität	- unzureichendes Begleitsystem, insbesondere EU-Monitoring - geringe Motivation zur Daten-erhebung und -bereitstellung - Repräsentativität der Daten nicht gegeben - hoher Zeitaufwand für Daten-sammlung	- Konzeption eines zusätzlichen Begleitsystems - ad-hoc-Erhebungen - Erhöhung des Umfangs der Daten
Methodik	- kein mit/ohne-Vergleich möglich	- Vorher/nachher-Vergleich - Fallstudien
Referenz-situation	- keine Vergleichsgruppen	- evt. normativ

AG_2

2. Beantwortung der kapitelspezif. Fragen/Kriterien/Indikatoren

	Problempunkte	Lösungen
Daten/ Repräsen-tativität	- Repräsentativität häufig nicht gegeben - Datenlücken - kurze Zeitreihen	- stärkere Betonung von Fall-studien statt massenstatistischer Auswertungen
Methodik	- Auswahl der Vergleichs-gruppen (Mit-Ohne-Vergleich) - Vergleich verschiedener Rechtsformen	- Aussagen relativieren und durch Fallstudien verstärkt ergänzen
Referenz-situation	- häufige Änderung der Fördervoraussetzungen - Abgrenzungsprobleme (betrieblich/regional/sektoral)	- Kontinuität in der Förderung - Sonderauswertungen der Statistik - sektoraler Vergleich grundsätzlich problematisch
. . .		

AG_3

2. Beantwortung der kapitelspezif. Fragen/Kriterien/Indikatoren

	Problempunkte	Lösungen
Daten/ Repräsen-tativität	Monitoringsystem absolut ungeeignet (hohe Aggregationsebene) Förderdaten: unterschiedliche Datenqualität (Inhalt/Datenhaltung) Abwägung zwischen Datentiefe (Fläche) und Kosten - Umweltdaten: allgemeine Lücken, Monotoring abiotische Werte (--)	
Methodik	Eine eindeutige Beziehung zwischen Maßnahmen und Ressourcenwirkung herzustellen: - Vermeidung der Doppelzählung von Förderflächen bzw. ↑Transparenz - Flächenaddition als Aussage über Effizienz unzureichend - Treffsicherheit/räumliche Aquivalenz/Wichtung - Bewertung: negative U-wirkung????	
Referenz-situation	Gute fachliche Praxis: theoretische Referenz, bei Nutzung von Praxisdaten ggf. diese nicht abgebildet Mit-ohne Vergleich.: alle potentiellen TN nehmen auch teil Vorher-Nachher: nur bei Neueinsteiger, Problem (!): VO 2078/1999	
. . .		

AG_4

2. Beantwortung der kapitelspezif. Fragen/Kriterien/Indikatoren

	Problempunkte	Lösungen
Daten/ Repräsen-tativität	÷ Vorhandensein Einzelprojektdaten	÷ teilweise Nutzung von Bewilligungs-daten, sammeln von Einzelprojektdaten im Land
Methodik	÷ Nettoeffekte/Mitnahme-effekt ÷ Erfassung indirekter Wirkungen	÷ nur in Ansätzen dargestellt, Mischung quantitativ/qualitativ, Artikulieren! Annahmen transparent machen! ÷ Baden-Württemberg: Nutzen beschreiben, in Euro abschätzen ÷ andere: Wirkungen beschreiben
Referenz-situation	÷ Vergleichsgruppen nicht zu finden	÷ Vergleichsgruppen nur bei einzelbe-trieblicher Förderung, ansonsten nicht ÷ Eigener Indikatorenkatalog
Indikatoren	÷ komplizierte Kriterien/ Indikatoren, hohe Auf-gliederung nicht leistbar	÷ mehr Beschreibung, zusätzliche Indikatoren ÷ Aufgliederung nicht abgearbeitet

AG_5

2. Beantwortung der kapitelspezif. Fragen/Kriterien/Indikatoren

	Problempunkte	Lösungen
Daten/ Repräsen-tativität	- vorhandene **Sekundärdaten** oft ungeeignet, reine Outputdaten. Erheblicher Aufwand bei nach-träglicher Erhebung. - **Primärdaten**: „von der Stichprobe zum Fallbeispiel“	- standardisierte Auswertung von Antragsdaten - Fallbeispiele - Unterstützung durch andere Institutionen - Monitoringsystem in Österreich
Methodik	- Hochrechnung - Befragung (taktisches Verhalten) - Fallbeispiele: sehr zeitaufwendig	
Referenz-situation	- nicht vorhanden - kann durch Befragung nicht hergestellt werden	
. . .		

AG_1

3. Bewertungsergebnisse (qualitativ/ quantitativ) und Handlungsempfehlungen

A) Zu welchen Fragen/Kriterien gibt es Ergebnisse?

- kaum Ergebnisse
- teilweise Ergebnisse zu einzelnen Indikatoren
- detaillierte Analyse sprengt den Rahmen

B) Wie belastbar sind diese?

- sehr vage, wenig belastbar
- meist qualitativ
- zum jetzigen Zeitpunkt kaum Nachweise über beabsichtigte Effekte möglich (time lag, Datenverfügbarkeit)
- Wirkungszuordnung bei Mehrfachförderung schwierig
- Mitnahmeeffekt kaum identifizierbar

C) Handlungsempfehlungen

- Entflechtung der Förderung
- Präzisere Maßnahmenziele

AG_2

3. Bewertungsergebnisse (qualitativ/ quantitativ) und Handlungsempfehlungen

A) Zu welchen Fragen/Kriterien gibt es Ergebnisse?

- vorwiegend zum Einkommensbeitrag
- Grundlagen für Aufrechterhaltung der landwirtschaftlichen Nutzung
- Kontextindikatoren für Erhalt lebensfähiger Gesellschaftsstrukturen
- Umweltindikatoren erlauben nur bedingt kausale Zusammenhänge und sind wenig spezifisch

B) Wie belastbar sind diese?

- zumeist geringe Belastbarkeit
- hohe Sensibilität je nach Einkommens und Bezugsgröße sowie Betriebsgruppe
- kaum Aussagen für NE

C) Handlungsempfehlungen

- generelle Überprüfung der benachteiligten Gebiete
- verbesserte Datengrundlage für NE
- Verschneidung von InVeKoS und Testbetriebsdaten

AG_3

3. Bewertungsergebnisse (qualitativ/ quantitativ) und Handlungsempfehlungen

A) Zu welchen Fragen/Kriterien gibt es Ergebnisse?

- zu allen,
  Ausnahme: Nachbarschaftsbeziehung erst möglich, wenn InVeKos-GIS

B) Wie belastbar sind diese?

- alle Aussagen zu Flächenberechnung sind innerhalb der gewählten Methodik belastbar, jedoch Synthese ????

C) Handlungsempfehlungen

- wurden ausgesprochen und diskutiert

AG_4

3. Bewertungsergebnisse (qualitativ/ quantitativ) und Handlungsempfehlungen

A) Zu welchen Fragen/Kriterien gibt es Ergebnisse?

- wenig Ergebnisse, vorläufiger Charakter
- erst wenige abgeschlossene Projekte
- Output-, Befragungsergebnisse (diese auch belastbar)
- administrative Umsetzung als Schwerpunkt in Deutschland
- in Österreich zu allen relevanten Fragen belastbare Ergebnisse (Ausnahme Einkommen)

C) Handlungsempfehlungen

- nur wenig Empfehlungen insgesamt (bezogen auf Maßnahmenausgestaltung)
- Empfehlungen zur administrativen Umsetzung
- an uns: Fragensystem weiter anpassen

AG_5

3. Bewertungsergebnisse (qualitativ/ quantitativ) und Handlungsempfehlungen

A) Zu welchen Fragen/Kriterien gibt es Ergebnisse?

- quantitative Schätzungen/Ergebnisse zu fast allen Fragestellungen, zu denen Flächendaten vorliegen und verwendet werden können (Durchschnittswerte oder Hochrechnungen)
- verbale Beschreibung von Wirkungsketten insbesondere bei sozio-ökonomischen Ergebnissen

B) Wie belastbar sind diese?

- belastbar wie Durchschnittswerte oder Hochrechnungen
- sozio-ökonomische Ergebnisse eher weniger belastbar

C) Handlungsempfehlungen

- Angaben zur Lage der Förderfläche
- Controlling für den Schutzwald umsetzen
- Gebietskulissen ?

AG_1

4. Welche Grundlagen wurden für die Ex-post-Evaluierung geschaffen?

A) Begleit- und Monitoringsysteme (EU und sonstige)

- ergänzende Datenerhebungssysteme durch Evaluatoren geschaffen (⇨ EU- Monitoringsystem unzureichend)

B) Methodik

- teilweise, sehr unterschiedliche Aussagen
- konzeptionelle Grundlagen geschaffen (Indikatoren, Analyseinstrumente)

AG_2

4. Welche Grundlagen wurden für die Ex-post-Evaluierung geschaffen?

A) Begleit- und Monitoringsysteme (EU und sonstige)

- Schaffung von methodischen Voraussetzungen für ex-post-Analysen
- EU-Monitoringsystem wenig brauchbar
- teilweise durch Verschnitt von Datenquellen aussagekräftiges Monitoring für einzelne Länder
- Konzepte für Sonderauswertungen für Agrarstatistik wurden erarbeitet

B) Methodik

- Fallstudien verstärkt mit einbeziehen
- Verschnitt verfügbarer Daten
- Sensibilitätsanalysen

AG_3

4. Welche Grundlagen wurden für die Ex-post-Evaluierung geschaffen?

A) Begleit- und Monitoringsysteme (EU und sonstige)

- Konsens: EU-Monitoring für Eval absolut ungeeignet:
  - zu hohe Aggregation
  - Anträge statt Bewilligung

B) Methodik

- Eval-methodik wurde def., diese lässt sich fortführen
  - Definition u. Abstimmung der Zielsysteme
  - z.T. Primärdatenerfassung angelegt, die fortgeführt wird: z.B. SA
  - methodische Grundlagen geschaffen/angeraten für z.B. Umweltdaten (fachliches Monitoring)

AG_4

4. Welche Grundlagen wurden für die Ex-post-Evaluierung geschaffen?

Nicht diskutiert

AG_5

4. Welche Grundlagen wurden für die Ex-post-Evaluierung geschaffen?

A) Begleit- und Monitoringsysteme (EU und sonstige)

- Erfassungssysteme angeschoben
- Antragsdaten werden aufgearbeitet
- Zahlstellen wurden zusammengeführt (in Österreich)
- Vorgaben sollten erhalten werden

B) Methodik

- Hochrechnungen / Durchschnittswerte
- Musterfragebögen
- Fallbeispiele können weiter verfolgt werden (Langfristigkeit der forstlichen Umsetzung)
- Methodikaustausch zwischen den Evaluatoren

AG_1

5. Programmplanungsdokumente als Grundlage für die Kapitelbewertung

A) Interventionslogik des Landes

- oberflächlich nachvollziehbar
- SWOT-Analyse sehr global
- stärker regional differenzierte Ansätze notwendig

B) Zielbeschreibung- und Quantifizierung

- unpräzise Ziele
- kaum Quantifizierung (gleichzeitig birgt Quantifizierung Probleme)

C) Ex-ante-Evaluierung

- sehr großer Zeitmangel bei der Erstellung
- meist unzureichend oder fehlend
- Verknüpfung der ex ante-Bewertung mit dem Programmantrag ist problematisch

D) Empfehlungen für die zukünftige Programmplanung

- Programmansatz im Integrationsanspruch zurückfahren
- präzisere Ziele

AG_2

5. Programmplanungsdokumente als Grundlage für die Kapitelbewertung

A) Interventionslogik des Landes

- Programmplanungsdokumente analytisch gut
- SWOT-Analyse nicht ausreichend nach Gebietstypen differenziert

B) Zielbeschreibung- und Quantifizierung

- in wenigen Ländern Ziele hinreichend quantifiziert
- Nutzung der Handlungsspielräume durch Landesprogramme
- Ziele gewichten

C) Ex-ante-Evaluierung

- nur in wenigen Ländern extern
- unzureichende Berücksichtigung im Zuge des Genehmigungsverfahrens

D) Empfehlungen für die zukünftige Programmplanung

- neue Legitimierung der AZ (Gebietsabgrenzung, AZ nicht nach Kassenlage)
- Vereinheitlichung der zu fördernden Anbaukulturen

AG_3

5. Programmplanungsdokumente als Grundlage für die Kapitelbewertung

A) Interventionslogik des Landes

- Ideal nach Interventionslogik: SWOT - Programmplanung - ex-ante
- Realität: hoher politischer Einfluß nicht zwingend auf SWOT aufbauend
  - fehlende Maßnahmen
  - fehlende operationelle Ziele, kaum Wirkungsziele

B) Zielbeschreibung- und Quantifizierung

- Zielsysteme mussten nachgearbeitet werden

C) Ex-ante-Evaluierung

- lag i.d.R. vor
- völlig unterschiedliche Qualität
- kritischere Genehmigung z.T. gewünscht

D) Empfehlungen für die zukünftige Programmplanung

- besseres timing
- Maßnahmenblatt bei Programmerstellung als Checkliste, keine langen Erläuterung über potentielle Wirksamkeit des ökologischen Landbaus

AG_4

5. Programmplanungsdokumente als Grundlage für die Kapitelbewertung

Nicht diskutiert

AG_5

5. Programmplanungsdokumente als Grundlage für die Kapitelbewertung

A) Interventionslogik des Landes

- keine stringente/neue Interventionslogik
- Strategie/Maßnahmen neuen Anforderungen nicht angepasst
- andererseits: Interventionslogik wurde von der EU im Rahmen der Programmgenehmigung akzeptiert

B) Zielbeschreibung- und Quantifizierung

- Ziele sind beschrieben, aber überwiegend unrealistisch
- Output (z.B. Fläche) ist quantifiziert, aber überwiegend unrealistisch
- keine Gebietskulissen und keine Zielgruppen

C) Ex-ante-Evaluierung

- sehr wenig verwendbar, nur vereinzelt konkrete Vorgaben

D) Empfehlungen für die zukünftige Programmplanung

- Rahmenbedingungen für die Erstaufforstungen attraktiver machen
- Maßnahmen mit bestimmter Intention müssten ausgelagert werden
- stärkere sozio-ökonomische Ausrichtung
- Vereinfachung der Abwicklung

AG_1

6. Evaluationskultur in den Ländern

A) Verhältnis Auftraggeber - Evaluator

- knappe zeitliche Vorgaben für die Evaluation
- Auftragsvergabe und Finanzierung ausschließlich durch EU-KOM?

B) Lernprozess oder „Hofberichterstattung“?

- geringes Interesse der Auftraggeber an formativer Evaluation
- teilweise Einflussnahme der Auftraggeber auf Berichterstellung
- Evaluation überwiegend als Erfüllung der KOM-Verpflichtung
- kaum Diskussionsinteresse der Auftraggeber nach Abgabe des Berichtes

AG_2

6. Evaluationskultur in den Ländern

A) Verhältnis Auftraggeber - Evaluator

- **ze**ntrale Evaluation: Konsensfindung schwierig
- unterschiedliche Erwartungen an AZ
- Parallelbewertung durch Programmbewerter
- Ergebnisverwertung im Endbericht teilweise nicht bekannt
- Evaluation als Prozess durch Länder nicht gewährleistet

B) Lernprozess oder „Hofberichterstattung“?

- Chancen einer Effizienzsteigerung teilweise nicht genutzt
- keine Aussagen zur qualitativ besseren Ausgestaltung der AZ im Ländervergleich
- Informations- und Datenfluss teilweise verbesserungsbedürftig
- mangelnde Evaluationskultur / teilweise Berichterstattung als Pflicht und nicht als Lernprozess

AG_3

6. Evaluationskultur in den Ländern

A) Verhältnis Auftraggeber - Evaluator

- allgemeines konstruktives Verhältnis
- Verhaltenskodex als gutes Beispiel bzw. als Instrument einführen/ definieren/ einhalten
    - Grundlage sachliche Bewertung
    - keine Einflussnahme auf Ergebnisse
    - Erfolge und Kritikpunkte anführen
- ständiger Austausch Auftraggeber/Evaluator - begleitender Prozess

B) Lernprozess oder „Hofberichterstattung“?

- Evaluierungskultur in Deutschland noch in den Kinderschuhen, dafür Erfolge
    - wird von allen Gruppen als Lernprozess bzw. Chance gesehen
    - Auftraggeber wünsche sich sehr konkrete Empfehlungen mit Umsetzungsanweisungen

AG_4

6. Evaluationskultur in den Ländern

A) Verhältnis Auftraggeber - Evaluator

- Zusammenarbeitsstrukturen mit viel Austausch
- insgesamt: hohe Zufriedenheit mit Zusammenarbeitsverhältnis, kooperative Haltung
- grundsätzlichere Probleme/Anregungen:
  - Land sucht Evaluator aus, aber was wäre eine Alternative? Mitsprache der EU?
  - Steuerung der Evaluation über Folgeaufträge, Informationen, ...
  - ist ein Evaluatorenwechsel überhaupt möglich?
  - EU könnte mehr vermitteln, zusätzliche Instanz
  - Evaluatoren haben schon viel „Macht“

B) Lernprozess oder „Hofberichterstattung“?

- Lernprozess ja, aber nur im Land
- Umsetzung von Empfehlungen? Stärkere Einbindung der EU hierbei?
  Aber: EU hat eigene Interessen

AG_5

6. Evaluationskultur in den Ländern

A) Verhältnis Auftraggeber - Evaluator

- a
- a
- a
- a
- a
- a

B) Lernprozess oder „Hofberichterstattung“?

- a
- a
- a
- a
- a
- a