

Die Evaluierungspolitik der Europäischen Kommission. Eröffnungsreferat der 8. Jahrestagung der DeGEval am 12.-14. Oktober 2005 in Essen

Werner Stengg

Europäische Kommission, Brüssel

1. Einführung

Evaluation kann in der Europäischen Union auf eine lange Tradition verweisen. Dies gilt insbesondere für die Regionalpolitik, wo Strukturfonds-Programme seit vielen Jahren Gegenstand regelmäßiger und rigoroser Bewertung sind.¹ Aber auch in anderen Politikbereichen, die mit erheblichen Ausgaben verbunden sind (wie etwa in der Forschungspolitik oder der Entwicklungszusammenarbeit), konnte sich Evaluation schon früh etablieren.

Im Jahr 1996 beschloss die Europäische Kommission, ihre Evaluierungspolitik auf sämtliche Politikbereiche auszuweiten (allerdings beschränkt auf ausgabenrelevante Aktivitäten), was unter anderem bedeutete, dass alle Generaldirektionen ihre eigenen Evaluierungsreferate (oder kleinere organisatorische Einheiten) einrichten mussten. Seit 2000 existiert schließlich die Verpflichtung, auch Tätigkeiten *außerhalb* des traditionellen Anwendungsbereiches der ‚Ausgabenprogramme‘ zu evaluieren (wie etwa Politiken oder Gesetzgebungsakte).

Im Jahr 2004 belief sich die Zahl der von der Kommission in Auftrag gegebenen (oder selbst durchgeführten) Evaluationen auf mehr als 140 – eine Zahl, die jene Evaluationen, die in den Mitgliedstaaten im Rahmen der Regionalpolitik durchgeführt werden, nicht einschließt. Der Schwerpunkt lag nach wie vor auf ausgabenrelevanten Interventionen, wenn auch ein deutlicher Trend zur vermehrten Evaluation von ‚Politiken‘ und Gesetzgebungsakten zu erkennen war. Mit Ausnahme von Ex-ante-Evaluationen wurden fast alle Evaluationen von externen Beratern durchgeführt. Die Durchschnittskosten dieser Evaluationen beliefen sich im Jahr 2004 auf € 188.000.

¹ Dies erfolgt auf der Grundlage einer Partnerschaft mit den Mitgliedstaaten, die für die Ex-ante-, Halbzeit- oder laufende Bewertung verantwortlich zeichnen, während die Kommission für thematische und Ex-post- Bewertungen zuständig ist. Ausführliche Informationen dazu sind auf der Evaluierungs-Webseite der Generaldirektion für Regionalpolitik zu finden: http://europa.eu.int/comm/regional_policy/sources/docgener/evaluation/evaluation_de.htm.

Die nachstehenden Ausführungen geben einen Überblick über die organisatorischen und rechtlichen Rahmenbedingungen der Evaluationspolitik der Kommission. Besonderes Augenmerk wird dabei auf die Fragestellung gelegt, inwieweit diese Rahmenbedingungen die Art der zu erwartenden Nutzung der Evaluierungsergebnisse beeinflussen. Kapitel 4 untersucht die Bestrebungen der Kommission, bereits im Designstadium von Interventionen deren Wirkfähigkeit und spätere Evaluierbarkeit zu gewährleisten. Zum Abschluss werden aktuelle Herausforderungen aufgezeigt und mögliche künftige Schwerpunkte skizziert.

2. Organisatorische Rahmenbedingungen

Die Evaluationspolitik der Europäischen Kommission ist durch eine *dezentralisierte Organisationsform* gekennzeichnet. Es obliegt den einzelnen Generaldirektionen, entsprechende Evaluierungskapazitäten aufzubauen, Evaluationen im Einklang mit den bestehenden rechtlichen Anforderungen zu planen, durchzuführen (oder an externe Berater zu vergeben) und für die Verbreitung und Verwendung der Evaluationsergebnisse zu sorgen.

Die Kommission hat aber auch einige Mechanismen vorgesehen, um die Kohärenz ihrer Evaluierungsaktivitäten zu gewährleisten:

- Verbindliche *Evaluationsstandards* stellen Mindestanforderungen an die Organisation, Planung, Durchführung, Verbreitung und Verwendung von Evaluationen.
- Ein *Evaluierungsnetzwerk* (welches die Evaluierungsreferate aller Generaldirektionen umfasst) sorgt für den regelmäßigen Austausch von Informationen und bewährten Verfahren. Es dient auch als Pool an Expertise, auf den zurückgegriffen werden kann, wenn es etwa gilt, Antworten auf methodologische Herausforderungen zu finden.
- Im Zentrum dieses Netzwerks steht das *Evaluierungsreferat der Generaldirektion Haushalt*, das die rechtlichen Rahmenbedingungen vorgibt (einschließlich der bereits erwähnten Evaluationsstandards), den Austausch von Informationen und bewährten Verfahren fördert (im Rahmen des Evaluierungsnetzwerks, aber auch durch die Organisation themenspezifischer Work Shops, etc.), methodologische Unterstützung anbietet (durch die Erstellung von Leitfäden, die Teilnahme an ‚steering groups‘, etc.), die Einhaltung der rechtlichen Mindestanforderungen überwacht (vor allem bei Kommissionsvorschlägen mit budgetären Auswirkungen), die kommissionsweit anfallenden Evaluierungsberichte archiviert (und durch eine Datenbank zugänglich macht), und die anderen europäischen Institutionen regelmäßig über Entwicklungen der Evaluierungspolitik der Europäischen Kommission informiert (d.h. über wichtige Resultate, Entwicklungen der Evaluierungskapazitäten, geplante Evaluationen, etc.).
- Das *Generalsekretariat* der Kommission koordiniert und unterstützt die Aktivitäten der operativen Dienste auf dem Gebiet des Folgeabschätzungsverfahrens.

Das Hauptmerkmal dieser dezentralisierten Struktur liegt darin, dass das Instrument der Evaluation in den operativen Kontext eingebettet ist, in dem der jeweilige Eva-

luierungsgegenstand situiert ist. Ein zentral gesteuertes ‚one-size-fits-all‘ Konzept erschiene angesichts der höchst unterschiedlichen Kontexte, in denen die verschiedenen Generaldirektionen agieren, nur schwer realisierbar: Die Europäische Forschungspolitik stellt Evaluation vor andere Herausforderungen und Fragestellungen als die Evaluation von Binnenmarktgesetzgebung. Politikbereiche, in denen die Europäische Union über stark ausgeprägte Kompetenzen verfügt (z.B. die Wettbewerbspolitik) haben andere Evaluierungsschwerpunkte als solche, in denen die Rolle der Kommission weitgehend auf Koordinationstätigkeiten beschränkt ist (wie etwa im Bereich Justiz und Inneres). Ausgabenprogramme, die direkt von der Kommission implementiert werden (z.B. im Bereich der Aus- und Weiterbildung), unterscheiden sich maßgeblich von jenen, die primär auf nationaler und regionaler Ebene umgesetzt werden (wie etwa im Bereich der Strukturfonds).

Die organisatorische Eingliederung der Evaluation in die operativen Dienste der Kommission führt zu einem größeren Verständnis des Evaluierungsgegenstandes sowie des dazugehörigen politischen und sozioökonomischen Kontexts. Dies erleichtert den Auftraggebern von Evaluationen die Aufgabe, sinnvolle *terms of reference* zu erstellen (da in relativ kurzer Zeit realistische und relevante Evaluierungsfragen identifiziert werden können), sowie die Qualität der von externen Beratern geleisteten Arbeit zu beurteilen.

Erhöhte *ownership* ist ein weiterer Vorteil dieses dezentralisierten Modells. Eine Evaluation, die von einer zentralen Stelle (z.B. der Generaldirektion für Haushalt) in Auftrag gegeben wird, könnte von den Programmverantwortlichen als reines ‚Kontrollinstrument‘ verstanden werden und zu einer reflexhaften Abwehrhaltung führen. Je größer die (organisatorische) Nähe des Auftraggebers zum Evaluierungsgegenstand, desto eher wird Evaluation als Managementinstrument empfunden: Die subjektiv wahrgenommene Chance, ‚besser zu werden‘, drängt die Befürchtung, ‚bestraft zu werden‘, in den Hintergrund. Evaluation ist innerhalb der Europäischen Kommission daher dort am erfolgreichsten (und am wenigsten umstritten), wo formative Ziele (z.B. die Verbesserung von Design oder Umsetzung von Interventionen) und Lernprozesse im Mittelpunkt stehen.

Auf Kritik stößt das dezentralisierte Modell hingegen, wenn das Ziel der Rechenschaftslegung im Mittelpunkt steht. ‚Programmnähe‘ wird dann nicht mehr als Vorteil empfunden, sondern als Indikator für die mangelnde Unabhängigkeit der Evaluation und die dadurch verringerte Glaubwürdigkeit ihrer Resultate. Wenn gleich die starke Einbindung der Programmverantwortlichen in den Evaluierungsprozess deren Einfluss auf die Evaluationsergebnisse zweifellos erhöht, gibt es doch eine Anzahl von Faktoren, die eine übergebührliche Einflussnahme durch die Programmverantwortlichen erschweren: Fast alle Zwischen-, End- und Ex-post-Evaluationen werden von externen Beratern durchgeführt, deren Unabhängigkeit in den Evaluierungsstandards der Kommission festgeschrieben ist: *„The evaluator’s independence in his/her work must be respected and the evaluation results must not be interfered with”*.² Der Prozess selbst wird häufig nicht von den Programmverantwortlichen gesteuert, sondern von den Evaluierungsreferaten der betreffenden

2 Die Evaluierungsstandards sind auf folgender Webseite erhältlich: http://europa.eu.int/comm/budget/evaluation/Key_documents/regureq_de.htm.

Generaldirektionen. Zudem werden für die meisten Evaluationen Lenkungsgruppen (*„steering groups“*) eingesetzt, in die auch Vertreter anderer Generaldirektionen (wie zum Beispiel der Generaldirektion Haushalt) eingeladen werden.

Ein unbestreitbarer Nachteil des dezentralen Modells liegt allerdings darin, dass strategische, bereichsübergreifende Themen nicht ausreichend Berücksichtigung finden. Während die von einzelnen Generaldirektionen zu verantwortenden Aktivitäten regelmäßig auf deren Relevanz, Wirtschaftlichkeit und Wirksamkeit überprüft werden, werden Fragestellungen, die sich über mehrere Generaldirektionen erstrecken, nur sehr sporadisch aufgegriffen. Bereichsübergreifende Evaluationen können wertvolle Einsichten liefern, wenn die Aktivitäten verschiedener Generaldirektionen zu denselben übergeordneten Zielen beitragen (z.B. ‚Lissabon-Prozess‘, nachhaltige Entwicklung, etc.), oder sich an dieselbe Zielgruppe richten (z.B. KMUs)³.

3. Rechtliche Rahmenbedingungen

Die wichtigsten Rechtsgrundlagen

Die rechtlichen Vorgaben in Bezug auf Umfang, Zeitpunkt und Zweck der Evaluierungstätigkeiten der Europäischen Kommission sind sehr ambitioniert. So verlangt die derzeit gültige, kommissionsinterne Mitteilung zur Evaluation, dass grundsätzlich ‚alle Aktivitäten‘ der Kommission evaluiert werden sollten, egal ob sie Budgetmittel beanspruchen oder nicht. Natürlich gilt der Grundsatz der Proportionalität, mit dem gewährleistet werden soll, dass sich Evaluation auf jene Bereiche konzentriert, wo ihr potenzieller Nutzen am größten ist: *„Umfang, Häufigkeit und Zeitpunkt der Evaluierung [sollen] auf die aus dem Entscheidungsprozess erwachsenden Anforderungen, die Laufzeit und Art der jeweiligen Maßnahme sowie auf die verfügbaren Mittel abgestimmt werden“*⁴.

Diese allgemeinen Erfordernisse werden in spezifischeren Rechtsakten weiter konkretisiert. Dies gilt vor allem für die Haushaltsordnung für den Gesamthaushaltsplan der Europäischen Gemeinschaften, welche festlegt, dass all jene Programme und Tätigkeiten regelmäßig zu evaluieren sind, *„die mit erheblichen Ausgaben verbunden sind“*.⁵ Die Durchführungsbestimmungen zur Haushaltsordnung enthalten detailliertere Bestimmungen zu Zweck, Inhalt und Zeitpunkt von Ex-ante-, Zwischen- und Ex-post-Evaluationen.⁶ Ein weiteres Beispiel für sektorspezifische Bestimmungen sind die bereits erwähnten Strukturfonds-Verordnungen.

3 Zum KMU-Thema wurde im Jahr 2003 eine ‚Strategische Evaluierung der finanziellen Unterstützung für kleine und mittelgroße Unternehmen‘ durchgeführt: http://europa.eu.int/comm/budget/evaluation/Key_documents/evalguide_study_de.htm.

4 Mitteilung an die Kommission, 26/07/2000, ‚Ergebnisorientierung: Verstärkte Evaluierung von Kommissionsmaßnahmen‘, S. 5, http://europa.eu.int/comm/budget/evaluation/pdf/sec20001051_de.pdf.

5 Verordnung N° 1605/2002 des Rates vom 25. Juni 2002, Artikel 27(4) <http://www.cc.cec/budg/lex/en/legis/rf2002/index.htm>.

6 Verordnung N° 2342/2002 der Kommission vom 23. Dezember 2002, Artikel 21.

In Ergänzung zu der in der Haushaltsordnung bereits vorgesehenen ‚Ex-ante-Evaluation‘ (die für ausgabenrelevante Vorschläge verpflichtend ist), verfasste die Europäische Kommission im Jahr 2002 eine Mitteilung zum Folgeabschätzungsverfahren (‚*impact assessment*‘). Seit dem Jahr 2004 werden solche Verfahren für alle wichtigen ‚Initiativen‘ (d.h. Vorschläge, die im Arbeitsprogramm der Kommission aufscheinen) durchgeführt, um die zu erwartenden ökonomischen, sozialen und umweltbezogenen Wirkungen verschiedener Politikoptionen zu analysieren (mehr dazu in Kapitel 4).

Der Evaluierungsgegenstand

‚Programme‘ stehen nach wie vor im Zentrum der Evaluierungstätigkeit der Europäischen Kommission, wobei unter Programmen zeitlich begrenzte Maßnahmenbündel zu verstehen sind, die mit Ressourcen ausgestattet und auf bestimmte Ziele hin ausgerichtet sind.

Die Haushaltsordnung und die Mitteilung zur Evaluation verlangen allerdings auch die Evaluation von ‚Tätigkeiten‘ (‚*activities*‘) im Sinne der Organisation des Haushaltsplans nach Tätigkeitsbereichen (‚*activity based budgeting*‘). So haben alle Generaldirektionen ihre diversen Aktivitäten in eine geringe Zahl von ‚Tätigkeiten‘ zusammengefasst (üblicherweise zwischen zwei und fünf), für die entsprechende strategische Ziele und Indikatoren festgelegt wurden. In Analogie zur obigen Programmdefinition stellen diese Tätigkeiten also ein Bündel von Maßnahmen dar, die auf gemeinsame Ziele hin ausgerichtet sind. Ihr hohes Aggregationsniveau führt zu einem beträchtlichen Grad an Heterogenität: Tätigkeiten können einzelne Ausgabenprogramme beinhalten, aber auch eine Vielzahl von Politiken oder Gesetzgebungsakten, die keinerlei budgetären Ausgaben verursachen.

Aufgrund dieser Komplexität und Heterogenität ist die Verpflichtung, ‚alle Tätigkeiten‘ zu evaluieren, nicht wörtlich zu verstehen: Wo es nicht möglich oder sinnvoll ist, eine Tätigkeit in einer einzigen globalen Evaluation zu erfassen, steht es den Generaldirektionen frei, diese in ausreichend homogene ‚Sub-Aktivitäten‘ zu untergliedern und diese getrennt zu evaluieren. In solchen Einzelevaluationen sollte aber der Gesamtkontext nicht vernachlässigt werden, sodass eine spätere Zusammenfassung der einzelnen Evaluationsergebnisse ausreichend Auskunft über den Zielerreichungsgrad und die Kohärenz der übergeordneten Tätigkeit zu erteilen vermag.

Über die Evaluation von einzelnen Programmen oder Tätigkeiten hinausgehend, unterstreicht die Mitteilung zur Evaluation aber auch die Notwendigkeit, bereichsübergreifende, horizontale Themen zu evaluieren.

Die Evaluation einzelner ‚Projekte‘ steht hingegen nicht im Zentrum der Evaluierungspolitik der Kommission, nicht zuletzt, weil viele Projekte von Seiten der Mitgliedstaaten umgesetzt und auch evaluiert werden. Evaluationen von Programmen, die von der Kommission selbst implementiert werden, erfassen in der Regel nur eine Stichprobe der durch das Programm finanzierten Projekte. Spezifische Ausnahmeregelungen existieren jedoch in bestimmten Politikbereichen (z.B. für Großprojekte, die im Rahmen der Kohäsionspolitik durchgeführt werden) und für bestimmte Projekttypen (vor allem ‚Pilotprojekte‘, deren Evaluation Aufschlüsse über die Sinnhaftigkeit der Fortsetzung dieser Projekte in Programmform erteilen soll).

Die Evaluierungsinhalte

Das Hauptaugenmerk liegt üblicherweise auf der ‚Wirksamkeit‘ (*effectiveness*) des jeweiligen Evaluierungsgegenstandes, das bedeutet, auf der Fragestellung, in welchem Maße die kurz-, mittel- und (soweit erfassbar) langfristigen Ziele der betreffenden Maßnahme erreicht wurden.

Ein weiterer Schwerpunkt liegt auf der Evaluation der ‚Wirtschaftlichkeit‘ (*efficiency*) – dem Verhältnis zwischen intendierten Outputs und Resultaten zu den für ihre Realisierung eingesetzten Mitteln. Dies gilt insbesondere für Maßnahmen, die Budgetmittel in Anspruch nehmen.

Neben den zentralen Begriffen der Wirksamkeit und Wirtschaftlichkeit stehen noch jene der ‚Relevanz‘, der ‚Kohärenz‘, der ‚Nachhaltigkeit‘ und des ‚Nutzens‘. Diesen wird je nach Art des Evaluierungsgegenstandes und Zeitpunkt der Evaluation unterschiedliche Bedeutung beigemessen. Eine Intervention ist *relevant*, wenn ihre Zielsetzungen im Einklang stehen mit den der Intervention zu Grunde liegenden Problemen oder Bedürfnissen. Sie ist *kohärent*, wenn ihre Interventionslogik schlüssig und widerspruchsfrei ist (‚interne‘ Kohärenz) und wenn sie nicht im Widerspruch zu anderen Interventionen steht, die ähnliche Ziele verfolgen oder ähnliche Zielgruppen ansprechen (‚externe‘ Kohärenz). Bei einer *nachhaltigen* (*sustainable*) Intervention dauern deren positive Effekte auch nach ihrer Beendigung an. Der tatsächliche *Nutzen* (*utility*) einer Intervention ergibt sich aus der Beantwortung der Frage, in welchem Verhältnis ihre *tatsächlichen* Wirkungen und Nebenwirkungen zu den zugrunde liegenden Problemen oder Bedürfnissen stehen.

Der Evaluierungszweck

Aus den bereits erwähnten Rechtsgrundlagen lässt sich schlüssig ableiten, dass der Hauptzweck der Evaluierungspolitik der Kommission darin liegt, Informationen zu liefern, die für den Entscheidungsfindungsprozess relevant sind. So sind etwa laut Haushaltsordnung (und deren Durchführungsbestimmungen) Evaluationen durchzuführen, *um die Beschlussfassung zu verbessern*. Für den konkreten Fall von Ausgabenprogrammen gilt es, die Evaluierungsergebnisse *„bei allen Beschlüssen über die Fortschreibung, Änderung oder Unterbrechung des Programms zu berücksichtigen“*⁷.

Gemäß der allgemeinen Mitteilung zur Evaluation dienen Evaluationen *„vor allem dazu, die Entwicklung der Gemeinschaftspolitik zu unterstützen und die Effizienz der Maßnahmen kontinuierlich zu verbessern“*. Evaluierungsergebnisse sollen in den *„Entscheidungsprozess auf Kommissionsebene eingehen, wenn Prioritäten festgesetzt und Ressourcen verteilt werden“*⁸.

In einem kürzlich verabschiedeten (nicht publizierten) Audit über die Evaluierungspolitik der Europäischen Kommission hat der Europäische Rechnungshof (anerkennend) festgestellt, dass Evaluation in den meisten Generaldirektionen zu einem *„established management tool“* geworden ist.

7 Durchführungsbestimmungen zur Haushaltsordnung, Artikel 21(2).

8 Mitteilung an die Kommission vom 26/07/2000, Seiten 3 und 9.

Der Evaluierungszeitpunkt

Evaluierungszeitpunkt und Evaluierungszweck sind eng miteinander verbunden. Wenn Evaluation als ‚management tool‘ verstanden wird und primär der ‚Verbesserung der Beschlussfassung‘ dienen soll, müssen Evaluationen zu einem Zeitpunkt durchgeführt werden, der es ermöglicht, dass entscheidungsrelevante Evaluierungsergebnisse zum richtigen Zeitpunkt vorliegen. Die operativen Generaldirektionen müssen, laut Mitteilung zur Evaluation, ihre Evaluierungstätigkeit so planen, „*dass die Evaluierungen für ihre eigenen politischen Entscheidungen genutzt werden können*“.⁹

Dies schlägt sich auch in der Programmevaluation nieder: Die Durchführungsbestimmungen zur Haushaltsordnung sehen für solche Evaluationen einen Zeitplan vor, „*der es ermöglicht, die Bewertungsergebnisse bei allen Beschlüssen [bezüglich] des Programms zu berücksichtigen*“.¹⁰

4. Maßnahmen zur Erhöhung der ‚Wirkfähigkeit‘ von Interventionen

In den bisherigen Ausführungen wurde primär darauf eingegangen, wie die Europäische Kommission die Wirkfähigkeit ihrer Aktivitäten ‚überprüft‘ (d.h. evaluiert). Um diese (‚ex post‘) Überprüfung zu erleichtern, ist es von großer Bedeutung, bereits zum Ex-ante-Zeitpunkt Maßnahmen zu setzen, die geeignet sind, die spätere Wirkfähigkeit von Aktivitäten zu erhöhen.

Für budgetwirksame Maßnahmen sieht die Haushaltsordnung schon seit einigen Jahren eine systematische Ex-ante-Evaluation vor.¹¹ Wesentlich größere öffentliche Aufmerksamkeit hat allerdings das im Jahr 2002 eingesetzte Folgeabschätzungsverfahren auf sich gezogen, das für alle wichtigen Initiativen durchgeführt werden muss (ob ausgabenwirksam oder nicht).¹²

Das Verfahren¹³ beginnt mit einer Problem- beziehungsweise Bedarfsanalyse, an die sich unmittelbar die Fragestellung anschließt, ob die EU-Ebene die am besten geeignete Ebene darstellt, dem identifizierten Problem oder Bedarf zu begegnen. Nach einem ersten Versuch¹⁴, realistische und relevante Ziele zu definieren, werden mögliche Alternativlösungen diskutiert. Hierbei gilt es, die gesamte Bandbreite möglicher Interventionsgrade in Betracht zu ziehen, wie etwa: die Möglichkeit, *nicht* tätig zu werden; die Förderung von Selbstregulierungsmechanismen (z.B. Selbstverpflichtungen von Seiten der Industrie); die Setzung von positiven

9 Mitteilung an die Kommission vom 26/17/2000, S. 9.

10 Durchführungsbestimmungen zur Haushaltsordnung, Artikel 21(2).

11 Durchführungsbestimmungen zur Haushaltsordnung, Artikel 21(1).

12 Ausführliche Informationen zum Folgeabschätzungsverfahren findet man auf folgender Webseite: http://europa.eu.int/comm/secretariat_general/impact/index_en.htm

13 Die nachstehenden Verfahrensschritte gelten sowohl für eine Ex-ante-Evaluation auf der Grundlage der Haushaltsordnung, als auch für das so genannte Folgenabschätzungsverfahren.

14 Der Zielsetzungsprozess ist ein iterativer Prozess, der sich über das gesamte Verfahren erstreckt.

oder negativen Anreizen, um marktwirtschaftliche Mechanismen in Gang zu setzen; die direkte Förderung von erwünschten Aktivitäten (z.B. durch Subventionen oder Ausgabenprogramme); bis hin zu klassischen Regulierungsmaßnahmen (d.h. Richtlinien oder Verordnungen).

Die viel versprechendsten Optionen werden schließlich auf der Grundlage ihrer zu erwartenden Effekte und Wirkungen miteinander verglichen. Bei der Analyse dieser Effekte und Wirkungen wird ein ‚*bottom-up*‘ Ansatz mit einem ‚*top-down*‘ Ansatz verknüpft: Zunächst wird üblicherweise eine ‚Programmtheorie‘ entwickelt (z.B. in Form einer Interventionslogik), um die erhofften Ursache-Wirkungszusammenhänge explizit zu machen. Da diese Interventionslogik tendenziell die Betrachtung ‚erwünschter‘ Effekte und Wirkungen begünstigt, wird in einem zweiten Schritt sichergestellt, dass mögliche unerwünschte Wirkungen und *trade-offs* nicht im Dunkeln bleiben: Anhand einer Check-Liste wird kritisch hinterfragt, ob die betreffende Option negative Auswirkungen auf eine Reihe ökonomischer, sozialer oder umweltbezogener Faktoren haben könnte.

Das Ziel dieser Vorgehensweise ist es nicht, dass nur Politiken vorgeschlagen werden können, die mit keinerlei *trade-offs* verbunden sind. Es gilt vielmehr, alle denkbaren *trade-offs* offen zu legen, sodass die politischen Entscheidungsträger ihre Entscheidung auf der Grundlage möglichst kompletter Information fällen können. Das Folgeabschätzungsverfahren ist also nicht als Substitut für den politischen Entscheidungsfindungsprozess zu verstehen, sondern lediglich als Entscheidungshilfe.

Ist die Entscheidung zugunsten einer bestimmten Option gefallen, werden die Ziele näher ausformuliert und entsprechende Indikatoren festgelegt. Danach gilt es, effektive Monitoring-Mechanismen einzurichten, um zu gewährleisten, dass alle Daten, die für das Management und die spätere Evaluation der Maßnahme erforderlich sind, erhoben werden können. Schließlich werden noch der Zeitpunkt, Inhalt und Zweck späterer Evaluationen festgelegt.¹⁵

Die konsequente Anwendung dieses Verfahrens sollte nicht nur dazu führen, dass qualitativ bessere Vorschläge gemacht werden; sie erleichtert auch die spätere Evaluation der Wirksamkeit der betreffenden Politiken.

5. Herausforderungen und zukünftige Entwicklungen

Das Ziel, in regelmäßigen Abständen ‚alle wichtigen Aktivitäten‘ der Europäischen Kommission zu evaluieren, ist sehr ambitioniert und muss daher differenziert betrachtet werden. Da Evaluationen beträchtliche Finanz- und Humanressourcen in Anspruch nehmen, müssen sie sich auf jene Bereiche konzentrieren, in denen sie den größten Mehrwert erbringen können – das heißt auf jene Maßnahmen und Po-

15 In vielen Berichten zu Folgeabschätzungsverfahren wird man die in diesem Absatz angeführten Fragestellungen nur im Ansatz beantwortet finden. Dies ist in der Regel dadurch begründet, dass Kommissionsvorschläge im weiteren Gesetzgebungsprozess häufig starken Veränderungen unterworfen werden, und die detaillierte Ausgestaltung von Zielen, Indikatoren und Monitoring-Systemen erst zu einem späteren Zeitpunkt erfolgt.

litiken, die die größten Auswirkungen auf den EU-Haushalt und/oder EU-Unternehmen und/oder EU-Bürger haben. Und sie sollten zu einem Zeitpunkt durchgeführt werden, der es ermöglicht, dass ihre Ergebnisse verwendet werden können.

Der selektive Einsatz von Evaluationen ist nicht nur aufgrund von Ressourcenüberlegungen erforderlich, sondern auch um die Akzeptanz des Instruments selbst nicht zu gefährden. Wenn Programmverantwortliche den Eindruck haben, dass es sich bei einer Evaluation lediglich um eine bürokratische Pflichtübung handelt, wird es nicht gelingen, das volle Potenzial dieser Evaluation auszuschöpfen.

Die Kommission sieht sich aber nicht nur mit der *Vielzahl* der (potenziell) zu evaluierenden Maßnahmen konfrontiert, sondern auch mit deren *Komplexität*, die interne wie externe Evaluatoren vor beträchtliche methodologische und logistische Schwierigkeiten stellt: EU-Politiken richten sich (derzeit) an fünfundzwanzig Staaten mit unterschiedlichen Kulturen und sozio-ökonomischen Kontexten, sie werden häufig von einer Vielzahl von Akteuren umgesetzt (Kommission, Mitgliedstaaten, Agenturen, etc.), und sie koexistieren häufig mit verwandten Politiken, die auf anderen Ebenen – ob national, regional oder lokal – mit vergleichbaren Zielsetzungen durchgeführt werden. Insbesondere der letztgenannte Aspekt erschwert die Zuordnung von (vermeintlichen) Wirkungen auf spezifische EU-Maßnahmen ungemein.

Methodologische Herausforderungen sind auch mit dem Ziel verbunden, Gesetze und Politiken verstärkt zu evaluieren. Während die klassische Programmevaluation auf eine lange Tradition zurückblicken kann, existieren noch verhältnismäßig wenige Erfahrungswerte für die Evaluation von Gesetzen und Politiken.

Die Kommission hat es sich auch zum Ziel gesetzt, die Relevanz ihrer Evaluierungsaktivitäten für ihren *strategischen* Entscheidungsfindungsprozess zu erhöhen. Die bisherigen Erfahrungen haben gezeigt, dass der Nutzen von Evaluation häufig auf die konkrete, evaluierte Maßnahme beschränkt bleibt: Evaluationsergebnisse werden nachweislich verwendet, um die Wirtschaftlichkeit und Wirksamkeit der betreffenden Maßnahme zu erhöhen. Wesentlich seltener leisten sie hingegen einen Beitrag, um die *übergeordneten* Ziele und Politiken, in die diese Maßnahmen eingebettet sind, zu überprüfen und zu verbessern. Es gilt also, einzelne Interventionen vermehrt in ihrem ‚weiteren Politik-Kontext‘ zu evaluieren, um Erkenntnisse zu gewinnen, die nicht nur von operationellem, sondern auch von strategischem Nutzen für die Kommission sind.

Bei allen Bestrebungen, das Potenzial von Evaluation als ‚management tool‘ maximal auszuschöpfen, darf aber nicht der Fehler begangen werden, Erwartungshaltungen zu wecken, die realistischweise nicht erfüllt werden können: Dies könnte negative Auswirkungen auf die Glaubwürdigkeit und Akzeptanz dieses Instruments haben. So bedarf es etwa einer realistischen Einschätzung, wie weit Evaluatoren in der Lage sind, die Wirkungen (ob erwünscht oder unerwünscht) von komplexen Politiken zu erfassen und diese ursächlich auf einzelne Interventionen zurückzuführen.

Es muss auch Sorge getragen werden, dass die Grenze zwischen Politikberatung und Politikgestaltung nicht überschritten wird. Evaluation soll Fakten in den Entscheidungsfindungsprozess einfließen lassen, den Entscheidungsträgern aber die politische Entscheidung – und damit Verantwortung – nicht abnehmen.