The Mid Term Evaluation in Objective 1 and 2 Regions

Growing Evaluation Capacity

Final Report

DG Regional Policy
Evaluation Unit – REGIO.C.2

November 2004

TABLE OF CONTENTS

1. SYNTHESE 4

1.1. LE ROLE DE L'EVALUATION A MI-PARCOURS 4
1.2. LE PROCESSUS DE L'EVALUATION A MI-PARCOURS 4
1.3. LES EVALUATEURS 4
1.4. LA QUALITE DES EVALUATIONS 5
1.5. LE COUT DES EVALUATIONS ET LEUR RAPPORT COUT-EFFICACITE 5
1.6. LA METHODOLOGIE SUIVIE 5
1.7. LES REPONSES AUX QUESTIONS D'EVALUATION 6
1.8. LA CONTRIBUTION AUX PRIORITES DE LISBONNE 7
1.9. L'UTILISATION DES EVALUATIONS 7
1.10. L'EVALUATION A MI-PARCOURS – POINTS FORTS ET POINTS FAIBLES 8
1.11. LA MISE A JOUR DE L'EVALUATION A MI-PARCOURS 9
1.12. LES DEFIS POUR L'AVENIR 9

1. ZUSAMMENFASSUNG 11

1.1 DIE ROLLE DER HALBZEITBEWERTUNG 11
1.2 DER PROZESS DER HALBZEITBEWERTUNG 11
1.3 DIE BEWERTER 11
1.4 QUALITÄT DER BEWERTUNGEN 11
1.5 KOSTEN UND KOSTENEFFIZIENZ 12
1.6 METHODISCHER ANSATZ 12
1.7 ANTWORTEN AUF DIE BEWERTUNGSFRAGEN 13
1.8 BEITRAG ZU DEN PRIORITÄTEN VON LISSABON 13
1.9 VERWENDUNG DER BEWERTUNGEN 14
1.10 DIE HALBZEITBEWERTUNG – STÄRKEN UND SCHWÄCHEN 15
1.11 AKTUALISIERUNG DER HALBZEITBEWERTUNG 15
1.12 HERAUSFORDERUNGEN FÜR DIE ZUKUNFT 16

1. SUMMARY 17

1.1. THE ROLE OF THE MID TERM EVALUATION 17
1.2. THE MID TERM EVALUATION PROCESS 17
1.3. THE EVALUATORS 17
1.4. QUALITY OF THE EVALUATIONS 17
1.5. COST AND COST EFFECTIVENESS 18
1.6. METHODOLOGIES 18
1.7. ANSWERS TO THE EVALUATION QUESTIONS 19
1.8. CONTRIBUTION TO LISBON PRIORITIES 19
1.9. USE OF THE EVALUATIONS 20
1.10. THE MID TERM EVALUATION – STRENGTHS AND WEAKNESSES 20
1.11. UPDATE OF THE MID TERM EVALUATION 21
1.12. CHALLENGES FOR THE FUTURE 21

2. INTRODUCTION 23

2.1. OBJECTIVES OF THIS REPORT 23
2.2. THE ROLE OF THE MID TERM EVALUATION 23
2.3. CONTENT OF THE REPORT 24

3. THE PROCESS 26

3.1. PARTNERSHIP WORKING – THE ROLES AND RESPONSIBILITIES OF ACTORS 26
3.1.1. Organisation of the Mid Term Evaluation 28
3.1.2. Timing 28
3.2. THE EVALUATORS 28
3.3. THE QUALITY OF THE MID TERM EVALUATIONS 29
3.4. COST OF THE EVALUATION AND COST EFFECTIVENESS 30

4. METHODOLOGIES 32

4.1. MIX OF METHODOLOGIES USED 32
4.1.1. Desk Research 32
4.1.2. Primary Research 33

4.1.3. Modelling Work ... *34*
4.2. METHODOLOGICAL CONCERNS ... 34

5. MID TERM EVALUATION FINDINGS ... 36

5.1. ANSWERS TO EVALUATION QUESTIONS ... 36
5.1.1. Limited Need for Strategy Shifts ... *36*
5.1.2. Effectiveness – Much Activity Underway after a Slow Start ... *37*
5.1.2.1. Macro-economic Impacts ... 37
5.1.2.2. Results and Outputs ... 39
5.1.3. Efficiency – A Start Made in the Analysis of Unit Costs ... *41*
5.1.4. Quantification of Objectives – an Urgent Need for Improvements ... *42*
5.1.5. Implementation Systems – Significant Improvements compared to the Past ... *43*
5.2. CONTRIBUTION TO LISBON PRIORITIES ... 45
5.2.1. Transport – Good progress in large Road and Rail projects; smaller scale projects behind schedule ... *45*
5.2.2. Knowledge Based Economy ... *46*
5.2.2.1. Research, Innovation, Information and Communication Technologies (ICT) – A growing priority with good results in Objective 2 regions; slower but definite progress being made in Objective 1 regions ... 46
5.2.2.2. Human Capital ... 47
5.2.3. Entrepreneurship and Financial Instruments –Performing well but affected by the global economic downturn ... *48*
5.2.4. Social Inclusion – Integrated actions are most effective but are resource intensive *49*
5.2.5. Environment and Sustainable Development – more consideration needed of the practical implementation of horizontal priorities ... *51*

6. USING THE EVALUATIONS ... 53

6.1. THE USERS OF THE MID TERM EVALUATIONS ... 53
6.2. AMENDED FINANCIAL ALLOCATIONS ... 54
6.3. CHANGES TO IMPLEMENTATION SYSTEMS ... 54
6.4. PUBLICATION OF REPORTS ... 55
6.5. FURTHER ACTIONS PLANNED ... 55
6.6. FACTORS MILITATING AGAINST USE OF THE MID TERM EVALUATIONS ... 56

7. CONCLUSIONS ... 57

7.1. THE PROCESS ... 57
7.2. THE EVALUATORS ... 57
7.3. QUALITY OF THE EVALUATIONS ... 57
7.4. COST AND COST EFFECTIVENESS ... 58
7.5. METHODOLOGIES ... 58
7.6. ANSWERS TO THE EVALUATION QUESTIONS ... 59
7.7. CONTRIBUTION TO LISBON PRIORITIES ... 59
7.8. USE OF THE EVALUATIONS ... 60
7.9. THE MID TERM EVALUATION – STRENGTHS AND WEAKNESSES ... 60
7.10. UPDATE OF THE MID TERM EVALUATION ... 61
7.11. CHALLENGES FOR THE FUTURE ... 61
7.11.1. Further Strengthening Evaluation Capacity ... *61*
7.11.2. Improving Monitoring Systems ... *62*
7.11.3. Creating a Platform for the Exchange of Experience ... *62*

1. SYNTHESE

Les évaluations à mi-parcours des programmes financés par les Fonds structurels ont été achevées fin 2003 conformément aux dispositions prévues par le règlement n° (CE) 1260/1999 du Conseil. L'examen à mi-parcours des programmes est à présent en cours ou terminé. Il convient donc de faire le point sur les conditions de déroulement de cet exercice sur les résultats globaux de l'évaluation à mi-parcours, afin d'en tirer les enseignements nécessaires pour l'avenir et d'envisager sur cette base les éventuels transferts de bonnes pratiques.

Le présent document relatif aux évaluations à mi-parcours des objectifs 1 et 2 a été préparé par l'unité d'évaluation de la Direction Générale Politique Régionale et Cohésion. Il a été discuté avec les États Membres lors d'un séminaire qui s'est tenu en octobre 2004 à Bruxelles. Il a ensuite été finalisé à la lumière de ces débats et des informations complémentaires que ce séminaire a pu apportées quant à l'utilisation des évaluations.

1.1. Le rôle de l'évaluation à mi-parcours

Les prescriptions relatives à l'évaluation à mi-parcours sont exposées à l'article 42 du règlement du Conseil. Elles ont été complétées par un document de travail[1] de la Commission élaboré en concertation avec les États membres en 2000. Le but de l'évaluation à mi-parcours était de faire le point sur les réalisations effectives puis au vu d'un examen des développements intervenus dans le cadre du programme, de mettre en évidence les éventuelles modifications à apporter au programme afin d’en optimiser l'impact à plus long terme. Le document de travail a apporté les orientations nécessaires quant au processus d’élaboration de l'évaluation à mi-parcours lui-même et à ses huit composantes essentielles.

1.2. Le processus de l'évaluation à mi-parcours

L'évaluation à mi-parcours est un processus complexe auquel participent la Commission, l'État membre, l'autorité de gestion, le comité de suivi et l'évaluateur. Deux aspects du processus de l'évaluation à mi-parcours peuvent particulièrement être considérés comme des points forts: la qualité de la programmation et la mobilisation du partenariat.

1.3. Les évaluateurs

Le marché de l'évaluation est un marché relativement limité et l'évaluation à mi-parcours n'a pas permis d’élargir significativement le nombre des évaluateurs. Les évaluations ont été principalement effectuées par des sociétés de conseil spécialisées dans l'évaluation, parfois en association avec des centres de recherche universitaires, et l’on note également un nombre croissant d'évaluations réalisé par des consortiums d'évaluateurs. Le caractère impératif du délai réglementaire fixé pour la présentation de toutes les évaluations à mi-parcours a mis sous pression le marché de l'évaluation. Certains évaluateurs, qui avaient pris en charge plusieurs évaluations, se sont ainsi retrouvés dépassés par la masse de travail impliquée.

[1] Document de travail n° 8 relatif à l'évaluation à mi-parcours des interventions structurelles.

1.4. La qualité des évaluations

Les évaluations sont dans l’ensemble d'une qualité supérieure à celles effectuées par le passé et d’excellents exemples ont pu être relevés. L'examen qualitatif des évaluations par la Commission a permis établir que pour les deux tiers d’entre elles, elles ont été de bonne voire d'excellente qualité. Cette amélioration qualitative est directement imputable à la compétence croissante des responsables de l’évaluation au sein des autorités de gestion. Ces dernières ont globalement affecté à l'évaluation plus de ressources qu’au cours des exercices précédents.

Les principales faiblesses relevées sont liées à l'ampleur des tâches à accomplir, au grand nombre de questions évaluatives auxquelles il a fallu répondre et à la complexité de l'évaluation des programmes les plus importants des. Les évaluations les moins performantes sont la plupart du temps le fait d'évaluateurs surmenés en raison du nombre trop important de projets d'évaluations pris en charge et de l'insuffisance des ressources qu’ils ont pu y consacrer. Quelques lacunes méthodologiques ont également limité la qualité de certaines évaluations.

1.5. Le coût des évaluations et leur rapport coût-efficacité

Plus de 35 M€ ont été dépensés pour les évaluations à mi-parcours des objectifs 1 et 2, avec un coût moyen de 172.000 euros par évaluation (près de 240 000 euros pour les programmes de l'objectif 1 contre 97 000 euros pour ceux de l'objectif 2). Le coût des évaluations a représenté en moyenne 0,08% des dotations annuelles totales, dont 0,07% pour l'objectif 1 et 0,09% pour l'objectif 2. Le coût moyen est plus faible pour l'objectif 1, étant donné la taille plus étendue des programmes concernés.

Le rapport coût-efficacité dépend des modalités d’utilisation des évaluations comme facteurs d'amélioration de la qualité et de la pertinence des programmes, processus qui est encore en cours. La Commission considère dès à présent que les ressources affectées à l'évaluation à mi-parcours ont été adaptées et globalement plus élevées qu'auparavant, mais que le rapport coût-efficacité s’avère encore limité du fait d’un démarrage lent ou tardif de certains programmes, du grand nombre de questions évaluatives et de l'ampleur de certains programmes. Cela a conduit à une approche parfois trop superficielle conjuguée à quelques faiblesses méthodologiques.

1.6. La méthodologie suivie

En général, les évaluations à mi-parcours sont le fruit de la combinaison de diverses méthodes allant de la recherche documentaire, au travail de terrain en passant, pour ce qui concerne les plus grands programmes, à des exercices de modélisation macroéconomique. Des améliorations significatives sont apparues dans la qualité et la rigueur des évaluations. Pour un grand nombre d'évaluations, la principale difficulté a été liée au fait que le système de contrôle et de suivi ne fonctionnait pas encore parfaitement et que les données administratives relatives la mise en oeuvre du programme étaient encore limitées. La recherche de terrain a consisté en des consultations, des enquêtes, des groupes de réflexion et des études de cas.

L’évaluation des Fonds structurels est complexe et les évaluateurs sont encouragés à recourir à des méthodologies appropriées. A l’avenir, certaines questions méthodologiques devront être traitées:

- En ce qui concerne **la recherche primaire**, les évaluations devraient obtenir davantage d'informations quant à la mise en oeuvre du programme «sur le terrain». Bien que l'importance de certains programmes puisse rendre cette observation difficile pour toutes les mesures, il est envisageable que seules les priorités essentielles fassent l'objet d'investigations concrètes et poussées.

- En ce qui concerne **l'analyse des données**, certaines évaluations à mi-parcours ont été beaucoup trop descriptives et pas assez synthétiques. L'analyse menée par les évaluateurs doit être plus synthétique et menée à un niveau plus stratégique. Les résultats du programme global doivent rassembler les données administratives issues de la mise en oeuvre, de la recherche primaire et de l'analyse des changements intervenus dans l'environnement global.

- **La présentation des rapports** doit encore être améliorée. Beaucoup de rapports sont trop longs, notamment du fait de l'abondance de données descriptives non synthétisées, allant à l'encontre de leur pertinence et de leur utilité.

Les évaluateurs doivent développer leurs capacités dans ces domaines, tandis que les prescripteurs doivent avoir à leur égard des exigences plus élevées.

1.7. Les réponses aux questions d'évaluation

Les évaluations visaient à répondre à un certain nombre de questions évaluatives que l'on peut résumer ainsi :

- Dans presque tous les cas, les évaluateurs ont conclu que les **stratégies** adoptées dans les programmes **étaient pertinentes**.

- Il apparaît dans la plupart des cas **prématuré d'évaluer l'efficacité des programmes**, étant donné leur démarrage lent ou tardif. Cela est à mettre en relation avec le système de contrôle et de suivi qui a souvent nécessité une révision des indicateurs et de leurs objectifs.

- Il est également **trop tôt pour mesurer l'impact des programmes**, bien que les évaluations aient recommandé dans certains cas la révision des objectifs tirés de l'expérience des premières années de mise en oeuvre. La modélisation macroéconomique entreprise pour de grands CCA laisse à penser que l**es programmes semblent en mesure** d'atteindre leurs objectifs globaux.

- L'efficacité n'a le plus souvent été définie qu'en termes d'absorption financière et non pas en termes de coût unitaire de réalisation ou de résultats obtenus. **L'absorption financière s'est clairement améliorée** au regard du passé, notamment du fait de l'application de la règle N+2.Les évaluations ont mis à jour **l'urgente nécessité d'améliorer les systèmes de quantification des objectifs.**

- **Les systèmes de mise en oeuvre** ont rempli les conditions réglementaires **toute en s’améliorant sensiblement** par rapport au passé.

1.8. La contribution aux priorités de Lisbonne

Bien que les évaluations à mi-parcours n'aient pas eu pour objectif d'évaluer la contribution des Fonds structurels aux priorités de Lisbonne, certains rapports d'évaluation fournissent à cet égard des éléments intéressants:

- En ce qui concerne **le transport et l'accessibilité**, les évaluations font apparaître des progrès intéressants en matière de grands projets routiers et ferroviaires. Les projets de développement des ports, aéroports ainsi que d’autres initiatives locales de transport sont également en bonne voie tout en évoluant à un rythme plus lent.
- Les Fonds structurels soutiennent **l'économie fondée sur la connaissance** dans le cadre de la coopération entre les entreprises et les organismes de recherche, du développement des groupes d'entreprises et des centres de recherche, des 'investissements pour l'accès aux réseaux à large bande, du développement des stratégies régionales d'innovation et de la formation des chercheurs, ainsi que des projets de recherche appliquée. Beaucoup de projets novateurs bénéficient d'une aide, certains nécessitant plus de temps qu'initialement prévu pour se concrétiser.
- **L'esprit d'entreprise** est encouragé par le biais des aides au démarrage, aux PME et aux grandes entreprises innovatrices, au développement des zones d’activités commerciales, à un soutien aux activités de conseil et à la création de structures de capital-risque. Plusieurs de ces interventions, très liées à l’évolution de la demande, connaissent des progrès plus lents que prévus, notamment du fait du ralentissement global de la conjoncture économique européenne au cours des premières années de programmation.
- La promotion **de l'inclusion sociale** est une dimension essentielle pour la majorité des programmes de l'objectif 1 et de l'objectif 2 et la plupart des évaluations constatent que les résultats sont en la matière satisfaisants ou supérieurs au but visé. Les actions intégrées s'avèrent les plus efficaces, bien que fortement consommatrices de ressources.
- **L'environnement** était une priorité horizontale pour les Fonds structurels et les évaluations donnent un aperçu des méthodes susceptibles de favoriser l’intégration de cette priorité conformément aux principes du **développement durable**. Les Fonds structurels financent directement des actions environnementales, comme la revitalisation des sites industriels, le développement des sources d'énergie renouvelables ou celui des équipements destinés à l'approvisionnement en eau douce ou à l'évacuation des déchets, ainsi que celui des zones protégées. En général, les projets de type traditionnel sont plutôt financés en début de programme alors que les actions innovantes se développent plutôt dans un second temps.

1.9. L’utilisation des évaluations

La première utilisation de l'évaluation à mi-parcours concernait l'examen à mi-parcours et l’analyse de la pertinence des programmes. Les évaluations ont été principalement utilisées par les autorités de gestion, les comités de suivi et les organismes de mise en oeuvre. D'une façon générale, les évaluations n'ont pas suffisamment favorisé le débat public, à quelques exceptions près, principalement dans les régions de l'objectif 1, où les Fonds structurels sont mieux connus et engagent des sommes plus importantes.

Les changements de dotations financières dans les programmes ont majoritairement découlé de préoccupations en termes d'absorption, bien que la contribution des évaluations ne soit pas non plus négligeable en matière de redéfinition des objectifs. Les évaluations ont eu une nette incidence sur l'amélioration des systèmes de mise en oeuvre, particulièrement sur le développement des systèmes d'indicateurs, la mise en oeuvre des priorités horizontales et l'amélioration des critères de sélection des projets.

Deux facteurs-clés ont influencé positivement l'utilisation des évaluations à mi-parcours:

- **la qualité des évaluations –** les évaluations de meilleure qualité s'appuyant sur des méthodologies plus rigoureuses ainsi que la présentation claire des résultats ont permis de formuler des recommandations concrètes et précises plus à même d'être utilisées dans la pratique ;
- **un partenariat actif lors de l'évaluation** – lorsque l'autorité de gestion et le comité de suivi étaient intéressés par les résultats de l'évaluation et qu'ils se sont activement impliqués dans le processus, apportant les informations nécessaires tout en demandant des explications aux évaluateurs dans le respect de leur indépendance, on a constaté une réelle et concrète utilisation concrète des évaluations.

1.10. L'évaluation à mi-parcours – points forts et points faibles

L'opération d'évaluation à mi-parcours témoigne de la capacité d'évaluation croissante des États membres. Si beaucoup d'améliorations ont été constatées par rapport aux évaluations précédentes, des possibilités d'amélioration subsistent néanmoins.

Tels sont les points forts de l'évaluation à mi-parcours:

- son **organisation et sa planification**,
- la mobilisation du **partenariat**,
- l'**allocation des ressources –** à la fois financières et humaines pour conduire le processus,
- la **capacité d'évaluation accrue** de la part des évaluateurs comme des prescripteurs, ce qui a débouché sur des évaluations de meilleure qualité.

Les principaux points faibles observés dans les rapports sont les suivants:

- la **rigidité du délai a conduit à** une évaluation prématurée dans certains cas qui a mis inutilement sous forte pression le marché restreint de l'évaluation,
- les **exigences trop larges en matière d'évaluation**, qui ont chercher à couvrir toutes les mesures et toutes les questions d'évaluation, de la pertinence de la stratégie à l'efficacité de sa mise en oeuvre,
- les **lignes directrices de l'évaluation n'ont pas toujours été suffisamment adaptées** aux besoins des régions ou des secteurs,
- certains évaluateurs n'ont pas produit un travail d'une **qualité** suffisante tandis que certaines autorités de gestion n'ont pas montrer suffisamment d'exigences à leur égard,

- En termes de conclusions pour l'avenir, la Commission a proposé que l'on remplace l'évaluation à mi-parcours au cours de la prochaine période de programmation par un programme d'évaluation permanent sur tout ou partie des programmes, établi en fonction des besoins locaux et nationaux. Cela permettrait de tirer le meilleur parti des points forts de l'évaluation à mi-parcours, tout en remédiant aux points faibles liés aux rigidités et à l'étendue de son champ d'application.

1.11. La mise à jour de l'évaluation à mi-parcours

Une mise à jour de l'évaluation à mi-parcours est requise d'ici la fin 2005 et la Commission a récemment publié des orientations quant à sa teneur. Cette actualisation arrive à un moment important, juste avant que ne commence la planification pour la période de programmation suivante, constituant également une préparation importante pour l'évaluation ex post. La Commission a proposé que l'élément essentiel de cette mise à jour soit une analyse des réalisations et des résultats atteints jusque là, tout en laissant le champ libre à toute question évaluative particulière, laissée à l'appréciation des autorités de gestion en fonction des besoins régionaux ou sectoriels.

1.12. Les défis pour l'avenir

Après avoir analysé le processus, la méthodologie, le contenu et l'utilisation des évaluations à mi-parcours et établi ses points forts et ses points faibles, nous pouvons, en conclusion, déterminer un certain nombre de défis pour l'avenir, en ce qui concerne l'évaluation et les Fonds structurels d'une manière plus générale.

- **Le renforcement de la capacité d'évaluation:** Les États membres qui ont entrepris l'évaluation à mi-parcours ont démontré une capacité d'évaluation croissante qui a débouché sur la production de rapports de bien meilleure qualité. Toutefois, ces résultats contrastés restent à conforter– à la fois du point de vue des prescripteurs de l'évaluation que de celui des évaluateurs eux-mêmes:
  - Il est nécessaire dans certains cas **d'augmenter le nombre et la qualité des ressources humaines consacrées à l'évaluation**.
  - En ce qui concerne les **évaluateurs, des efforts accrus** sont nécessaires **en matière méthodologique**, dans le souci de mieux rendre compte des résultats des programmes. Les propositions relatives à la possibilité de mener des évaluations plus ciblées devraient permettre à l'avenir l'élaboration de méthodologies plus spécifiques.
  - La capacité d'évaluation est encore à un stade précoce de son développement dans la plupart des États membres ayant récemment adhéré à l'U.E. En l'espèce, le défi consiste à **tirer parti de l'expérience** des autres États membres afin d'agir à la fois sur la capacité des administrations et sur celle des évaluateurs.
- **Amélioration des systèmes de contrôle:** le besoin est évident d'améliorer les systèmes de contrôle et de suivi qui doivent devenir de véritables instruments de gestion, fournissant en temps réel toutes les données physiques relatives à la mise en oeuvre des programmes. Conçu comme des outils d'aide à la décision, ils en deviendront gages de programmes de meilleure qualité. De bons systèmes de contrôle et de suivi sont donc essentiels quant à la capacité des autorités publiques régionales et nationales et de la Commission européenne elle-même à mieux rendre compte, à mieux expliquer et à plus clairement démontrer la valeur ajoutée de la masse considérable des ressources investies par le biais des Fonds structurels. Il est

nécessaire d'améliorer ces systèmes de contrôle et de suivi afin de les focaliser davantage sur les indicateurs-clés et de faire en sorte que les données régulièrement collectées puissent rapidement être prises en compte par les décideurs et mises à la disposition des évaluateurs.

- **Création d'une plate-forme pour l'échange d'expérience:** À la suite de l'élargissement de l'UE, il existe désormais un large éventail de capacités d'évaluation au sein des États membres. Cette capacité peut encore s'enrichir au travers d'un fructueux échange d'expérience. La Commission prendra prochainement des mesures afin d'encourager cette mise en réseau et proposera l'an prochain de développer le GUIDE de l'évaluation du développement socio-économique figurant sur le WEB (www.evalsed.info) de façon à en faire un véritable site interactif favorisant de tels échanges d'expérience.

1. ZUSAMMENFASSUNG

Die Halbzeitbewertungen aller Strukturfondsprogramme wurden, wie in der Verordnung (EG) Nr. 1260/1999 des Rates verlangt, bis Ende 2003 abgeschlossen. Die Halbzeitüberprüfung der Programme in den Mitgliedstaaten ist im Gange bzw. ebenfalls bereits abgeschlossen. Die Zeit ist daher gekommen, den Prozess und die Ergebnisse der Halbzeitbewertung zu betrachten und zu überlegen, welche Lehren gezogen wurden und wie künftig an bewährte Praktiken in diesem Bereich angeknüpft werden kann.

Dieses Papier wurde vom Referat Bewertung der GD Regionalpolitik erstellt und betrifft die Halbzeitbewertungen für die Ziele 1 und 2. Es wurde mit den Mitgliedstaaten im Oktober 2004 in Brüssel auf einem Seminar erörtert. Im Anschluss an das Seminar wurde das Papier unter Berücksichtigung dieser Erörterungen und weiterer Informationen darüber, welcher Gebrauch von den Bewertungen gemacht wurde, fertig gestellt.

1.1 Die Rolle der Halbzeitbewertung

Die Anforderungen an die Halbzeitbewertungen wurden in Artikel 42 der Verordnung (EG) Nr. 1260/1999 des Rates festgelegt und durch ein Arbeitspapier der Kommission[2], das im Jahr 2000 in Konsultation mit den Mitgliedstaaten erarbeitet wurde, ergänzt. Der Zweck der Halbzeitbewertung bestand darin, die bisherigen Fortschritte festzuhalten und auf dieser Grundlage sowie anhand einer Untersuchung der im Programmumfeld stattgefundenen Entwicklungen zu ermitteln, welche Änderungen gegebenenfalls vorgenommen werden sollten, um die langfristigen Auswirkungen des Programms zu maximieren. Das Arbeitspapier enthielt Anleitungen für den Prozess der Halbzeitbewertung sowie deren acht Kernelemente.

1.2 Der Prozess der Halbzeitbewertung

Die Halbzeitbewertung war ein komplexer Prozess, an dem die Kommission, der Mitgliedstaat, die Verwaltungsbehörde, der Begleitausschuss und der Bewerter beteiligt waren. Zwei Aspekte dieses Prozesses können als besondere Stärken angesehen werden: die Qualität der Planung und der positive Beitrag der Partnerschaft.

1.3 Die Bewerter

Der Bewertungsmarkt ist relativ begrenzt, und die Zahl der Bewerter hat im Zuge der Halbzeitbewertung nicht signifikant zugenommen. Die Bewertungen wurden in den meisten Fällen von Bewertungsunternehmen durchgeführt; teilweise waren Wissenschaftler mitbeteiligt, und eine steigende Zahl von Bewertungen wurde von Konsortien von Bewertern vorgenommen. Die einheitliche, für die Übermittlung aller Halbzeitbewertungen geltende Frist sorgte auf dem Bewertungsmarkt für Engpässe. Einige Bewerter, die mehrere Bewertungen übernommen hatten, waren damit überfordert.

1.4 Qualität der Bewertungen

Die Bewertungen waren von höherer Qualität als in der Vergangenheit mit einigen herausragenden Beispielen. Die Kommission prüfte die Qualität der Bewertungen und

[2] Arbeitspapier Nr. 8 über die Halbzeitbewertung der Strukturfondsinterventionen

bescheinigte zwei Dritteln von diesen eine gute bis ausgezeichnete Qualität. Diese Verbesserung ist auf einen Anstieg der Kapazitäten unter den Bewertern und den Verwaltungsbehörden zurückzuführen. Die Verwaltungsbehörden stellten für die Bewertung mehr Ressourcen zur Verfügung, als dies bei den vorangegangenen Bewertungen der Fall gewesen war.

Die Grenzen, an die die Bewertungen stießen, hängen mit dem Umfang der Aufgabe zusammen, da eine große Zahl von Bewertungsfragen zu beantworten waren und einige sehr umfangreiche und komplexe Programme bewertet wurden. Die schwächere Qualität einiger Bewertungen geht darauf zurück, dass der Bewerter zu viele Bewertungen übernommen hatte, damit überfordert war und nicht genügend Ressourcen für alle Bewertungen einsetzte. Außerdem wurde die Qualität einiger Bewertungen durch Schwächen bei der Methodik beeinträchtigt.

1.5 Kosten und Kosteneffizienz

Für die Halbzeitbewertungen für die Ziele 1 und 2 wurden insgesamt über 35 Mio. EUR aufgewendet mit durchschnittlich 172 000 EUR pro Bewertung (knapp 240 000 EUR für die Ziel-1-Programme und 97 000 EUR für Ziel 2). Die Kosten der Bewertungen machten im Schnitt 0,08% der Gesamtheit der jährlichen Mittelzuweisungen aus (0,07% für Ziel 1 und 0,09% für Ziel 2). Der niedrigere Prozentsatz bei Ziel 1 war angesichts des Umfangs vieler der größeren Ziel-1-Programme zu erwarten.

Die Kosteneffizienz hängt davon ab, wieweit die Bewertungen zur Verbesserung der Qualität und Relevanz der Programme beitragen, ein Prozess, der noch nicht abgeschlossen ist. Die Kommission hat den Eindruck, dass die für die Halbzeitbewertungen eingesetzten Mittel angemessen waren und höher lagen als frühere Mittelzuweisungen, während die Kostenwirksamkeit durch den späten oder schleppenden Start einiger Programme, die hohe Zahl von Bewertungsfragen und den großen Umfang einiger Programme (was in einigen Fällen einen eher oberflächlichen Ansatz zur Folge hatte) sowie durch Schwächen bei der Methodik beeinträchtigt wurde.

1.6 Methodischer Ansatz

Den Halbzeitbewertungen lag im Allgemeinen eine geeignete Kombination von Schreibtischrecherchen, Recherchen vor Ort und – bei umfangreicheren Programmen – makroökonomischer Modellierung zugrunde. Qualität und Stringenz der Bewertungen haben sich merklich verbessert. Bei vielen Bewertungen stellte sich das Problem, dass das Begleitsystem noch nicht gut funktionierte und nur in begrenztem Unfang administrative Daten zur Durchführung des Programms vorlagen. Die Recherchen vor Ort umfassten Konsultationen, Umfragen, Zielgruppenkontakte und Fallstudien.

Die Bewertungen im Rahmen der Strukturfonds sind ein komplexes Unterfangen, und die Bewerter sollten zur Entwicklung und Verbesserung von geeigneten Methoden ermutigt werden. Bei künftigen Bewertungen ist auf folgende die Methodik betreffende Aspekte zu achten:

- Was die **eigenen Untersuchungen** anbelangt, so sollten sich die Bewertungen auf die konkrete Realität der Programmdurchführung stützen. Auch wenn es angesichts des Umfangs einiger Programme schwierig sein könnte, dies bei allen Maßnahmen durchzuziehen, sollten doch zumindest die großen Schwerpunkte Gegenstand von Recherchen vor Ort sein.

- Zur **Analyse** ist zu sagen, dass einige Halbzeitbewertungen ausführliche Beschreibungen anstelle einer mehr synthetischen Analyse enthielten. Die Bewerter sollten anstreben, die Leistung des Gesamtprogramms auf strategischer Ebene synthetisch zu analysieren und dabei administrative (Begleit)daten von der Basis oder eigene Untersuchungen mit einer Analyse der Veränderungen im Gesamtumfeld zu verbinden.
- Die **Präsentation der Berichte** muss verbessert werden. Viele Berichte waren – häufig aufgrund von ausführlichen Beschreibungen – zu lang, was ihren Nutzen beeinträchtigte.

Die Bewerter müssen ihre Kapazitäten in diesen Bereichen ausbauen, während die Auftraggeber von den Bewertern die Einhaltung hoher Standards fordern müssen.

1.7 Antworten auf die Bewertungsfragen

Die Bewertungen sollten Antworten auf eine Reihe von Bewertungsfragen liefern. Diese Antworten lassen sich wie folgt zusammenfassen:

- In nahezu allen Fällen kamen die Bewerter zu dem Schluss, dass die **Strategien** der Programme **weiterhin angemessen** waren.
- In vielen Fällen war es für eine **Beurteilung der Effektivität noch zu früh**, da die Programme spät oder schleppend angelaufen waren. Eine Schwierigkeit bereitete das Begleitsystem, bei dem häufig eine Überarbeitung der festgelegten Indikatoren sowie von deren Zielvorgaben erforderlich war.
- Für eine **Messung der Auswirkungen** war es ebenfalls **zu früh**, wenngleich in den Bewertungen in einigen Fällen eine Überarbeitung der Zielvorgaben aufgrund der Erfahrungen aus den ersten Durchführungsjahren empfohlen wurde. Die für die großen GFK vorgenommene makroökonomische Modellierung lässt erkennen, dass die **Programme auf dem Wege sind,** die Gesamtziele zu erreichen.
- Effizienz wurde häufiger als Inanspruchnahme der Mittel und nicht als Kosten je Output oder Ergebnis definiert. Die **Inanspruchnahme der Mittel** hat sich – bedingt durch die Erfordernisse der „N+2“-Regel - gegenüber früher **deutlich verbessert.**
- Die Bewertungen ergaben, dass die **Systeme zur Quantifizierung der Ziele dringend verbessert** werden müssen.
- **Die Durchführungssysteme** entsprechen den Anforderungen der Verordnungen und haben sich gegenüber der Vergangenheit **erheblich verbessert.**

1.8 Beitrag zu den Prioritäten von Lissabon

Obwohl die Halbzeitbewertungen nicht zum Ziel hatten, den Beitrag der Strukturfonds zu den Prioritäten von Lissabon zu bewerten, ermöglichen einige Bewertungsergebnisse doch interessante Einblicke:

- Im Bereich **Verkehr und Erreichbarkeit** wurden in den Bewertungen gute Fortschritte bei den großen Straßen- und Schienenprojekten festgestellt. Die Projekte zum Ausbau von Häfen und Flughäfen und die Initiativen zugunsten des örtlichen Verkehrs schreiten ebenfalls voran, wenngleich häufig weniger rasch.
- Die Strukturfonds fördern die **wissensbasierte Wirtschaft** über die Zusammenarbeit zwischen Forschungsinstituten und Unternehmen, die

Entwicklung von Unternehmensclustern und Forschungszentren, Investitionen in den Breitbandzugang, die Entwicklung von regionalen Innovationsstrategien, die Ausbildung von Forschern sowie über Projekte im Bereich der angewandten Forschung. Es werden zahlreiche innovative Projekte und Ansätze unterstützt, deren Aufbau in einigen Fällen mehr Zeit in Anspruch nahm als ursprünglich erwartet.

- **Der Unternehmergeist** wird gefördert durch Zuschüsse für neu gegründete Unternehmen, kleine, innovative und große Unternehmen, die Entwicklung von Unternehmensparks, Beratungsunterstützung und die Errichtung von besonderen Wagniskapitalfonds. Viele dieser Interventionen sind nachfragegesteuert, und die Tatsache, dass diese Interventionen langsamer voranschreiten als erwartet, hängt mit der allgemeinen Konjunkturabschwächung in den ersten Jahren der Programmlaufzeit zusammen.

- Die Förderung der **sozialen Eingliederung** ist eine wichtige Komponente der meisten Ziel-1- und Ziel-2-Programme, und die meisten Bewertungen zeigen, dass die Ergebnisse zufrieden stellend sind oder die Zielvorgaben sogar übertreffen. Integrierte Maßnahmen erweisen sich als am effizientesten, sind aber ressourcenintensiv.

- Die **Umwelt** bildet für die Strukturfonds einen horizontalen Schwerpunkt, und die Bewertungen ermöglichen Einblicke, wie die praktische Einbeziehung dieser Priorität verstärkt und das Konzept der **nachhaltigen Entwicklung** in den Programmen verwirklicht werden kann. Darüber hinaus fördern die Strukturfonds direkte Umweltmaßnahmen wie z.B. die Revitalisierung von Industriegeländen, die Entwicklung erneuerbarer Energiequellen, Anlagen für die Trinkwasserversorgung oder Abfallentsorgung sowie die Schaffung von Schutzgebieten. Die Projekte von eher herkömmlicher Art schreiten im Allgemeinen planmäßig voran, während innovativere Maßnahmen eine längere Anlaufzeit haben.

1.9 Verwendung der Bewertungen

Die Halbzeitbewertungen wurden in erster Linie für die Halbzeitüberprüfung und für die Anpassung der Programme verwendet und zwar vor allem von den Verwaltungsbehörden, den Begleitausschüssen und den Durchführungsstellen. Im Allgemeinen zogen die Bewertungen keine öffentliche Debatte nach sich, wobei es allerdings Ausnahmen gab (vor allem in den Ziel-1-Regionen, wo die Strukturfonds ein stärkeres Profil haben).

Änderungen der Mittelzuweisungen innerhalb der Programme wurden in erster Linie aus Gründen der Mittelinanspruchnahme vorgenommen, wobei die Bewertungen allerdings zu diesen Entscheidungen beigetragen haben. Die Bewertungen hatten eine erhebliche Wirkung, indem die Durchführungssysteme verbessert (insbesondere Ausbau der Indikatorsysteme), die horizontalen Schwerpunkte umgesetzt und die Kriterien für die Auswahl der Projekte verbessert wurden.

Zwei Schlüsselfaktoren hatten einen positiven Einfluss auf die Verwendung der Halbzeitbewertungen:

- die **Qualität der Bewertungen:** Bewertungen von höherer Qualität mit einer stringenten Methodik, einer klaren Präsentation der Ergebnisse und präzisen Empfehlungen wurden mit größerer Wahrscheinlichkeit auch verwendet;

- eine **aktive, an der Steuerung der Bewertung mitwirkende Partnerschaft:** Dort, wo die Verwaltungsbehörde und der Begleitausschuss an den Resultaten der Bewertung interessiert waren und am gesamten Prozess mitwirkten, indem sie Informationen bereitstellten und an die Bewerter - unter Achtung von deren Selbständigkeit - hohe Anforderungen stellten, wurden die Bewertungen in der Regel auch verwendet.

1.10 Die Halbzeitbewertung – Stärken und Schwächen

Die Halbzeitbewertungen lassen einen Anstieg der Bewertungskapazitäten in den Mitgliedstaaten erkennen. Trotz großer Fortschritte gegenüber früheren Bewertungen kann die Bewertung der Strukturfonds aber weiter verbessert werden. Die ermittelten Stärken der Halbzeitbewertung betreffen folgende Punkte:

- **Organisation und Planung**,
- Beitrag der **Partnerschaft**,
- **Einsatz der** - finanziellen und personellen - **Mittel** für die Verwaltung der Bewertung,
- eine **Zunahme der Bewertungskapazitäten** auf dem Markt und in den Behörden, die zu Bewertungen von höherer Qualität geführt hat.

Die ermittelten Schwächen betreffen folgende Punkte:

- der **starre zeitliche Rahmen,** der dazu führte, dass die Bewertungen in einigen Fällen zu früh vorgenommen wurden und auf dem Bewertungsmarkt Engpässe entstanden;
- die **zu breit gefassten Bewertungsanforderungen**, die sich auf alle Maßnahmen erstreckten und acht verschiedene komplexe Bewertungsfragen von der Strategie über die Effizienz bis hin zur Durchführung umfassten;
- die für die Bewertung geltende **Leistungsbeschreibung**, die bisweilen **nicht ausreichend** auf die Bedürfnisse der Regionen oder Sektoren **abgestimmt** war;
- die nicht ausreichende **Qualität** der von einigen Bewertern abgelieferten Arbeit, wobei es die Verwaltungsbehörden versäumten, den Prozess aus der Nähe zu steuern und eine bessere Qualität zu verlangen.

Mit Blick auf die Zukunft hat die Kommission vorgeschlagen, die Halbzeitbewertung im kommenden Programmplanungszeitraum durch eine fortlaufende Bewertung zu ersetzen, die sich je nach Bedarf auf alle Programme oder einen Teil von diesen erstreckt. Damit könnte an die Stärken der Halbzeitbewertung angeknüpft werden, während die durch Inflexibilität und den Umfang bedingten Schwächen ausgeräumt würden.

1.11 Aktualisierung der Halbzeitbewertung

Die Halbzeitbewertung muss bis Ende 2005 aktualisiert werden. Die Kommission hat kürzlich Leitlinien für den Inhalt dieser Aktualisierung herausgeben. Diese wird zu einem wichtigen Zeitpunkt - kurz vor Beginn der Planung für den nächsten Programmplanungszeitraum – erfolgen und ist zugleich eine wichtige Vorbereitung für die Ex-post-Bewertung. Die Kommission hat vorgeschlagen, sich bei der Aktualisierung auf eine Analyse der bisherigen Outputs und Resultate zu konzentrieren, zugleich aber nach Maßgabe der regionalen oder sektoralen Bedürfnissen auch Raum für spezielle Bewertungsfragen zu lassen.

1.12 Herausforderungen für die Zukunft

Nach dieser Analyse der Prozesse, der Vorgehensweisen, des Inhalts und der Verwendung der Halbzeitbewertungen und der Ermittlung von deren Stärken und Schwächen können abschließend eine Reihe von künftigen Herausforderungen im Zusammenhang mit der Bewertung und den Strukturfonds identifiziert werden.

- **Weiterer Ausbau der Bewertungskapazitäten:** In den Mitgliedstaaten, die die Halbzeitbewertung durchgeführt haben, ist eine Zunahme der Bewertungskapazitäten zu erkennen, wodurch sich die Qualität der Berichte verbessert hat. Die Leistungen waren jedoch unterschiedlich, und die Kapazitäten müssen weiter ausgebaut werden – sowohl bei denjenigen, die die Bewertungen in Auftrag geben, als auch bei denen, die sie ausführen:
    - In einigen Fällen sind der **Umfang und die Qualität der für die Verwaltung der Bewertung eingesetzten Humanressourcen zu verstärken**.
    - Was die **Bewerter** anbelangt, so müssen **verstärkte Anstrengungen zur Entwicklung von geeigneten Vorgehensweisen** unternommen werden, um die Rechenschaftspflicht zu gewährleisten und mehr Einblick in die Leistung der Programme zu gewinnen. Die Vorschläge für eine künftig gezieltere Ausrichtung der Bewertungen dürften die Entwicklung von spezifischen Vorgehensweisen bei den Bewertungsthemen ermöglichen.
    - In vielen Mitgliedstaaten, die der EU kürzlich beigetreten sind, sind die Bewertungskapazitäten gerade erst im Aufbau begriffen. Hier gilt es, **an die Erfahrungen anderer Mitgliedstaaten anzuknüpfen** und Maßnahmen zu treffen, um die Bewertungskapazitäten sowohl in den Behörden als auch auf dem Bewertungsmarkt auszubauen.

- **Verbesserung der Begleitsysteme:** Die Begleitsysteme müssen eindeutig verbessert werden, damit sie zu einem wirklichen Managementinstrument werden, das frühzeitig sachdienliche physische Daten zur Durchführung der Programme liefert und zur Verbesserung der Beschlussfassung und der Programmqualität beiträgt. Gute Begleitsysteme sind wesentlich zur Gewährleistung der Rechenschaftspflicht, damit die regionalen und nationalen Behörden und die Europäische Kommission die sehr erheblichen in die Strukturfondsprogramme investierten Mittel rechtfertigen und den damit erzielten Mehrwert belegen können. Die Begleitsysteme sind zu verfeinern, indem der Schwerpunkt auf Schlüsselindikatoren gelegt und die regelmäßige Sammlung von Daten gewährleistet wird, die von den Entscheidungsträgern geprüft und den Bewertern zugänglich gemacht werden.

- **Schaffung einer Plattform für den Erfahrungsaustausch:** Mit der Erweiterung der EU gibt es nun eine breite Palette von Bewertungskapazitäten in den Mitgliedstaaten, die durch den Austausch von Erfahrungen verstärkt werden können. Die Kommission wird Maßnahmen zur Förderung der Vernetzung treffen und im nächsten Jahr vorschlagen, den im Internet abrufbaren „GUIDE to the evaluation of socio-economic development“ (Leitfaden für die Bewertung der sozioökonomischen Entwicklung - www.evalsed.info) zu einer interaktiven Website auszubauen, die solche Tätigkeiten erleichtern kann.

1. SUMMARY

The mid term evaluations of all Structural Fund programmes were completed by the end of 2003 as required by Council Regulation 1260/1999 and the mid term review of programmes is underway or complete in Member States. It is appropriate, therefore, to pause and take stock of the process and results of the mid term evaluation and consider what has been learnt and how good practice can be built on in the future.

This paper has been prepared by the Evaluation Unit of the DG for Regional Policy and concerns the mid term evaluations of Objectives 1 and 2. It was discussed with the Member States at a seminar in October 2004 in Brussels. Following the seminar, the paper was finalised in light of the debate and further information on how the evaluations were used.

1.1. The Role of the Mid Term Evaluation

The requirements for the Mid Term Evaluation were set out in Article 42 of Council Regulation 1260/1999. These requirements were complemented by a Commission working paper[3] which was drawn up in 2000 in consultation with Member States. The aim of the mid term evaluation was to report on what had been achieved to date and, on that basis and on the basis of an examination of developments in the programme environment, to identify any changes which should be made in order to maximise the longer term impact of the programme. The Working Paper provided guidance on the process for the mid term evaluation and its eight core elements.

1.2. The Mid Term Evaluation Process

The mid term evaluation was a complex process involving the Commission, the Member State, the managing authority, the monitoring committee and the evaluator. Two aspects of the process for the mid term evaluation can be identified as particular strengths: the quality of planning and the positive contribution of partnership.

1.3. The Evaluators

The evaluation market is relatively limited and the mid term evaluation did not see a significant expansion of the number of evaluators. Evaluations were mostly undertaken by evaluation consultancy companies, with some involvement of academics and growing evidence of evaluations being undertaken by consortia of evaluators. The single legal deadline for the submission of all mid term evaluations created pressure on the evaluation market. Some evaluators who took on a number of evaluations were over-stretched.

1.4. Quality of the Evaluations

The evaluations were of a higher quality than those undertaken in the past with some excellent examples. The Commission's examination of the quality of the evaluations resulted in two thirds being rated as being of good or excellent quality. This increase in quality results from a growing capacity among evaluators and managing authorities.

[3] Working Paper No. 8 on the Mid Term Evaluation of Structural Interventions

Managing authorities allocated more resources to the evaluation than had been the case in previous evaluation exercises.

Limitations of the evaluations related to the scale of the task, with a large number of evaluation questions to be answered and some very large complex programmes being evaluated. Some weaker evaluations resulted from evaluators being over-stretched in undertaking too many evaluation projects and not devoting sufficient resources to all projects. Methodological weaknesses also limited the quality of some evaluations.

1.5. Cost and Cost Effectiveness

Over €35 million was spent on the Objective 1 and 2 mid term evaluations with an average cost per evaluation of €172,000 (nearly €240,000 for Objective 1 programmes and €97,000 for Objective 2). The cost of the evaluations represented on average 0.08% of total annual allocations, 0.07% for Objective 1 and 0.09% for Objective 2. The lower percentage cost for Objective 1 is to be expected, given the scale of many of the larger Objective 1 programmes.

Cost effectiveness depends on the usefulness of the evaluations in improving the quality and relevance of programmes, a process which is still underway. The Commission's perception is that the resources allocated to the mid term evaluation were appropriate and higher than previous allocations, but that cost effectiveness was limited by the late or slow start of some programmes, the large number of evaluation questions and the large scale of some programmes leading to in some cases a rather superficial approach, and methodological weaknesses.

1.6. Methodologies

The mid term evaluations in general were based on an appropriate mix of desk research, primary research and, for larger programmes, some macro-economic modelling. Significant improvements were evident in the quality and rigour of the evaluations. A difficulty for many of the evaluations was that the monitoring system was not yet functioning well and there was limited administrative data on the implementation of the programme. Primary research undertaken included consultations, surveys, focus groups and case studies.

Structural Fund evaluations are complex and evaluators should be encouraged to develop and improve appropriate methodologies. The following methodological concerns need to be addressed in future evaluations :

- As regards **primary research**, evaluations should be informed by the "on-the-ground" reality of the implementation of the programme. While the scale of some programmes might make this difficult across all measures, at least the major priorities should be subject of fieldwork.
- In relation to **analysis**, some mid term evaluations contained extensive descriptive material rather than more synthetic analysis. Evaluators should aim to analyse synthetically the performance of the overall programme at a strategic level, bringing together bottom up administrative (monitoring) data or primary research and analysis of changes in the global environment.
- The **presentation of reports** needs to be improved. Many reports were extremely long, often because of the extensive descriptive material. This militated against their usefulness.

Evaluators need to develop their capacities in these areas, while those who commission evaluations must demand high standards from the evaluators.

1.7. Answers to the Evaluation Questions

The evaluations aimed to answer a number of evaluation questions and we can synthesise the answers as follows:

- In nearly all cases the evaluators concluded that the **strategies** adopted in the programmes **were still appropriate**.
- It was **too early to assess effectiveness** in many cases, given the late or slow start-up of programmes. A difficulty was presented by the monitoring system which often required revision of the definition of indicators and their targets.
- It was likewise **too early to measure impact**, although evaluations did in some cases recommend the revision of targets based on the experience of the early years of implementation. Macro economic modelling undertaken for large CSFs suggests that **programmes are on course** to achieve their global objectives.
- Efficiency was more often defined as financial absorption rather than the cost per output or result achieved. **Financial absorption is clearly improved** compared to the past, driven by the requirements of the N+2 rule.
- Evaluations found an urgent **need for improvement of the systems of quantification of objectives**.
- **Implementation systems** were found to fulfil the regulatory requirements and to be **significantly improved** on the past.

1.8. Contribution to Lisbon Priorities

Although the mid term evaluations did not have an objective of evaluating the contribution of the Structural Funds to the Lisbon priorities, some evaluation findings provide interesting insights:

- As regards **transport and accessibility**, the evaluations find good progress being made in large road and rail projects. Projects to develop ports, airports and local transport initiatives are also progressing but often at a slower rate.
- The Structural Funds are supporting the **knowledge based economy** through co-operation between research institutes and businesses, the development of business clusters and research centres, investment in broad-band access, the development of regional innovation strategies and the training of researchers, as well as applied research projects. Many innovative projects and approaches are being supported with some taking longer to establish than originally anticipated.
- **Entrepreneurship** is supported through grant aid to start-up, small, innovative and large enterprises, the development of business parks, consultancy support and the creation of special venture funds. Many of these interventions are demand led and slower progress than anticipated is linked to the global economic slow-down in the early years of the programmes.
- The promotion of **social inclusion** is an important dimension in the majority of Objective 1 and 2 programmes and most evaluations find that performance is satisfactory or above target. Integrated actions are found to be the most effective, although these are resource intensive.

- The **environment** was a horizontal priority for the Structural Funds and the evaluations provide insights into how to enhance the practical integration of the priority and how the concept of **sustainable development** can be made real in the programmes. The Structural Funds also support direct environmental actions including the revitalisation of industrial sites, the development of renewable energy sources and the development of facilities for fresh water supply or waste disposal, as well as the development of protected areas. In general, the more traditional types of projects are proceeding according to plan with more innovative measures taking longer to start up.

1.9. Use of the Evaluations

The primary use of the mid term evaluation was in relation to the mid term review and adaptation of programmes. The evaluations were used mainly by managing authorities, monitoring committees and implementing bodies. Generally, the evaluations did not stimulate public debate, although there were some exceptions, mainly in Objective 1 regions where the Structural Funds have a higher profile.

Changes to financial allocations in programmes were primarily driven by absorption concerns, although the evaluations did contribute to decisions in this regard. The evaluations had a strong impact in improving implementation systems, particularly the further development of systems of indicators, implementing horizontal priorities and improving project selection criteria.

Two key factors influenced positively the use of the mid term evaluations:

- The **quality of the evaluations** – higher quality evaluations that had rigorous methodologies and clear presentation of results and that made precise recommendations were more likely to be used;
- An **active partnership involved in managing the evaluation** – where the managing authority and monitoring committee were interested in the outcome of the evaluation and participated throughout the process in providing information and challenging evaluators while respecting their independence, the evaluations tended to be used.

1.10. The Mid Term Evaluation – Strengths and Weaknesses

The mid term evaluation exercise demonstrates growing evaluation capacity among Member States. While much improved compared to previous evaluations, there is still scope to further improve Structural Fund evaluation. We can identify the following strengths of the mid term evaluation:

- Its **organisation and planning**,
- The contribution of the **partnership**,
- The **allocation of resources –** both financial and human to manage the process,
- **Increased evaluation capacity** in the market and in the administrations leading to higher quality evaluations.

The weaknesses identified in the report are:

- The **rigid deadline** leading to the evaluation being too early in some cases and creating pressure on the evaluation market,

- The **requirements of the evaluation being too broad**, covering all measures and eight different and complex evaluation questions, ranging from strategy to effectiveness to implementation,
- Evaluation **terms of reference sometimes not sufficiently adapted** to the needs of regions or sectors,
- Some evaluators not delivering sufficiently high **quality** work, with Managing Authorities not managing the process closely enough and demanding better quality.

In terms of proposals for the future, the Commission has proposed that the mid term evaluation should be replaced in the next programming period with a programme of ongoing evaluation of all or parts of the programmes depending on needs. This would allow the strengths of the mid term evaluation process to be built upon, but would address the weaknesses related to rigidities and scope.

1.11. Update of the Mid Term Evaluation

An update of the mid term evaluation is required by the end of 2005 and the Commission has recently produced guidance on its content. The update will be produced at an important time, just before planning starts for the next programme period, and it will also be an important preparation for the ex post evaluation. The Commission has proposed that the core element of the update should be an analysis of outputs and results achieved to date, but that there should be scope for particular evaluation questions to be addressed, depending on regional or sectoral needs.

1.12. Challenges for the Future

Having analysed the processes, methodologies, content and use of the mid term evaluations and concluded on the strengths and weaknesses of the exercise, we can in conclusion identify a number of challenges for the future for evaluation and the Structural Funds.

- **Further Strengthening Evaluation Capacity:** The Member States which undertook the mid term evaluation demonstrate a growing evaluation capacity which has led to the production of better quality reports. However, performance was variable and there is a need to further strengthen capacity – both among those who commission evaluation and those who carry them out:
  - In some cases there is a need **to increase the amount and quality of human resources devoted to managing evaluation**.
  - As regards the **evaluators, increased efforts to develop appropriate methodologies** are needed to ensure accountability and to provide greater insights into the performance of programmes. Proposals for more targeted evaluations in the future should permit the development of methodologies specific to the issues being evaluated.
  - Evaluation capacity is at an early stage of development in many of the Member States which have recently joined the EU. Here, the challenge is **to build on experience** of other Member States and to take action to develop capacity both within the administrations and in the marketplace of evaluators.
- **Improving Monitoring Systems:** There is a clear need to improve monitoring systems so that they become a genuine management instrument – providing timely and meaningful physical data on the implementation of programmes and

contributing to better decision making and higher quality programmes. Good monitoring systems are essential for accountability, in order to allow regional and national authorities and the European Commission account for and demonstrate the added value of the very considerable amount of resources being invested in Structural Fund programmes. There is a need for monitoring systems to be refined to concentrate on core indicators and to ensure that data are gathered regularly and considered by decision makers and are available to evaluators.

- **Creating a Platform for the Exchange of Experience:** With enlargement of the EU, there is now a wide range of evaluation capacity across the Member States and there is potential for that capacity to be enhanced through the exchange of experience. The Commission will take action to encourage networking and proposes in the next year to develop the web-based GUIDE to the evaluation of socio-economic development (www.evalsed.info) into a truly interactive site which can facilitate such activities.

2. INTRODUCTION

2.1. Objectives of this Report

The mid term evaluations of all Structural Fund programmes were completed on time by the end of 2003 as required by Council Regulation 1260/1999. The mid term review of programmes is underway in Member States with it already complete in some cases. It is appropriate, therefore, to pause and take stock of the process and results of the mid term evaluation and consider what has been learnt and how good practice can be built on in the future.

This paper was prepared by the Evaluation Unit of the Directorate General for Regional Policy and it focuses on the mid term evaluations of Objectives 1 and 2. The first draft of the paper provided an input for discussion with Member States on the experience and results of the mid term evaluation. This discussion took place at a seminar in October 2004 in Brussels, attended by members of the Technical Group on Evaluation[4] as well as some evaluators and regional authorities who presented good practice examples. Following the seminar, the paper was finalised in light of the debate and further information on how the evaluations were used.

2.2. The Role of the Mid Term Evaluation

The requirements for the Mid Term Evaluation were outlined in Article 42 of Council Regulation 1260/1999. These requirements were complemented by a working paper[5] which was drawn up by the Commission in 2000 in consultation with the Member States through the Technical Group on Evaluation.

The Regulation states that the role of the mid term evaluation is, along with ex ante and ex post evaluation, to appraise the impact of programmes with respect to the objectives of the Structural Funds and to analyse their effects on specific structural problems. Effectiveness was to be measured on the basis of the impact on the strengthening of economic and social cohesion and the impact of the individual priorities in the programmes. The Working Paper outlines the role for the mid term evaluation in more specific terms, taking account of its place in the evaluation life cycle and the contribution it could make at the point in time in which it took place:

> "The mid term evaluation is not an end in itself but a means to improve the quality and relevance of programming. It provides an opportunity to identify reorientations to the programming which may be needed to ensure the achievement of the original objectives."
>
> "The overall aim of the mid term evaluation is to assess the initial results of the various forms of assistance and to make recommendations for any changes needed to ensure that they achieve their objectives."

---

[4] The Technical Group on evaluation comprises representatives of all Member States. It meets two or three times each year. Through this Group, the Commission consults with the Member States on EU-wide evaluations and the development of guidance for Structural Fund evaluation.

[5] Working Paper No. 8 on the Mid Term Evaluation of Structural Interventions

The twofold aims of the mid term evaluation are therefore clear. It aimed to report on what had been achieved to date and, on that basis and on the basis of an examination of developments in the programme environment, to identify any changes which should be made to the programme content or implementation system in order to maximise the longer term impact.

The Working Paper provides guidance on the process for the production of the mid term evaluation and its core elements. The core elements were extensive, comprising:

- Analysis of previous evaluation results;
- Analysis of the continuing validity of the ex ante evaluation's analysis of the strengths, weaknesses and potential of the State, region or sector;
- Assessment of the continuing relevance and the consistency of the strategy;
- Quantification of objectives – outputs, results and impacts;
- Evaluation of effectiveness and efficiency to date and expected socio-economic impacts;
- Quality of implementation and monitoring arrangements;
- Results for performance reserve indicators.

Article 42 of Council Regulation 1260/1999 – The Mid Term Evaluation

(1) Mid-term evaluation shall examine, in the light of the ex-ante evaluation, the initial results of the assistance, their relevance the extent to which the targets have been attained. It shall also assess the use made of financial resources and the operation of monitoring and implementation.

(2) Mid-term evaluation shall be carried out under the responsibility of the managing authority, in cooperation with the Commission and the Member State. It shall cover each Community support framework and each assistance. It shall be carried out by an independent assessor, be submitted to the Monitoring Committee for the Community support framework or assistance concerned in accordance with Article 35(3), and then sent to the Commission, as a general rule three years after adoption of the Community support framework or assistance, and no later than 31 December 2003, with a view to the revision referred to in Article 14(2).

(3) The Commission shall examine the relevance and quality of the evaluation on the basis of criteria defined beforehand by the Commission and the Member State in partnership, with a view to reviewing the assistance and allocating the reserve referred to in Article 44.

(4) As a continuation of mid-term evaluation, it shall be updated for each Community support framework and assistance and completed no later than 31 December 2005 in order to prepare for subsequent assistance operations.

2.3. Content of the Report

The first section of this report deals with the process of the evaluation, including the roles of the various actors involved and an assessment of the quality of the evaluations. This is followed by a section on the methodologies used. The following substantial section deals with the findings of the report, both in terms of the answers to the main evaluation questions for the mid term evaluation and an analysis of how Structural Fund programmes are contributing to the achievement of the Lisbon priorities. This section is based on a review of a sample of 65 evaluations across all Member States most of which were assessed by the Commission as being of good quality. The section on the use of the evaluations is primarily based on contributions from the Member States. Finally, some conclusions are drawn on the mid term evaluation process and results and some thoughts for the future are raised for consideration and debate.

3. THE PROCESS

The mid term evaluation involved a wide range of partners, extending from beneficiaries who were interviewed for the evaluation, through the evaluator, the steering group, the monitoring committee, the managing authority, the Member State and, finally, to the European Commission. Each stakeholder had a different role and responsibility in the process, but all contributed to the quality of the process and, ultimately, the quality of the product – the final evaluation report. This section analyses the role and responsibilities of the different actors and also examines the quality and cost of the evaluations.

3.1. Partnership Working – the Roles and Responsibilities of Actors

The mid term evaluation was carried out on the basis of the partnership – with Regulation 1260/1999 specifying a role for the Commission, the Member State, the managing authority, the monitoring committee and the evaluator. In addition, the beneficiary was involved. The experiences and views of beneficiaries are important since it is their needs and the needs of their regions which programmes aim to address. In most cases, steering committees were set up, as recommended by the Commission. They usually comprised a sub-committee of the monitoring committee and their role was to guide the evaluation process from the establishment of terms of reference to the finalisation of the evaluation. The table below outlines the range of interacting roles and responsibilities discharged by the different partners involved in the evaluation.

Partnership worked at several different levels during the mid term evaluation process and in different ways depending on national and regional practice, budgets involved and the scale of the programmes. The role of the steering groups varied across programmes and countries and they were not always involved from the beginning of the process. In general, the contribution of the steering groups resulted in higher quality and more relevant evaluations and enhanced ownership and involvement in the evaluation process by the partnership. That said, in some cases there seemed to be some attempt to steer the evaluation too much towards particular conclusions and recommendations. The presentation of the report to the monitoring committee broadened the involvement of the partnership in the process. In some cases (e.g., Greece, Ireland, Wales), there was keen media and political interest in the findings of the evaluation.

The evaluation process was also characterised by a strong partnership with the European Commission. The Commission provided an overall framework for the evaluation in consultation with the Member States. It participated in the steering groups and focused particular efforts on the terms of reference and the inception and draft reports. In this way it aimed to stimulate improvements in the evaluation at significant moments in the process.

Roles of the Different Partners/Stakeholders in the Mid Term Evaluation	
Partner/Stakeholder	**Role**
Commission	➢ To provide overall guidance; ➢ To give views on the terms of reference, inception report and draft final report; ➢ To participate on the steering committees; ➢ To examine the draft final report and feed back views to Managing Authorities; ➢ To assess the quality of the final reports; ➢ To use the evaluations in considering proposals for the allocation of the performance reserve and the mid term review.
Member State	➢ To provide overall guidance at a national level, including core terms of reference in some cases; ➢ To co-operate in the carrying out of the evaluation; ➢ To organise the assessment of quality, in some cases.
Managing Authority	➢ To manage the mid term evaluation process under its core responsibility; ➢ To develop draft terms of reference; ➢ To organise the selection of evaluators and award a contract; ➢ To manage the evaluation; ➢ To provide information to the evaluator, in terms of documentation, monitoring information and views and opinions; ➢ To co-ordinate the work of the steering committee and the role of the monitoring committee in contributing to the mid term evaluation; ➢ To submit the evaluation to the monitoring committee and then the Commission; ➢ To facilitate the process of consideration of evaluation findings in the context of the mid term review.
Monitoring Committee	➢ To agree the membership of the steering committee; ➢ To agree the terms of reference (in some cases); ➢ To comment on interim outputs from the evaluation (in some cases); ➢ Members to give interviews to the evaluators on the performance of the programme; ➢ To consider the evaluation findings in the context of the mid term review.
Steering Committee	➢ To approve the terms of reference; ➢ To select the evaluator; ➢ To provide guidance and information to evaluators on content and methodologies based on knowledge of the programme; ➢ To review interim and final reports and provide feedback to ensure their quality; ➢ To assure the quality of interim and final reports.
Evaluator	➢ To undertake the evaluation in line with the terms of reference and taking account of the guidance and information provided by the steering committee and the managing authority, but maintaining an objective view in drawing conclusions and making recommendations.
Beneficiary	➢ To provide information to the evaluator on how the programme is impacting on him/her or his/her business. This may be in response to a written questionnaire, a telephone interview or a face to face interview or through a group interview.

3.1.1. Organisation of the Mid Term Evaluation

The mid term evaluation process was a complex and wide-ranging exercise and this was particularly the case in large Objective 1 regions with Community Support Frameworks, the evaluation of which built on the findings of Operational Programme evaluations. The quality of the organisation of the mid term evaluation process contributed to higher quality of the final reports.

Core terms of reference and guidance were drawn up at Member State level in a number of countries (e.g., Austria, France, Italy, Ireland and UK). These interpreted the Commission's guidance into a national context and contributed to ensuring the evaluations were well designed. The only drawback to the provision of core terms of reference was that in some cases regions did not adapt them sufficiently to meet their needs and at the end of the process it became clear that additional evaluation questions relevant to the region or sector which should have been added were not included.

3.1.2. Timing

In its guidance, the Commission emphasised the need for sufficient time to be devoted to the mid term evaluation exercise. In most cases, planning started appropriately early and evaluators were in place in time to allow them to produce high quality work. Delays occurred in Greece and Luxembourg (Objective 2). In Luxembourg this was due to the late start and small scale of the programme and a good quality evaluation was produced in a short time. In Greece intensive work once the evaluators were in place ensured that harmful impacts on the quality of the work were minimised.

3.2. The Evaluators

In total approximately 170 evaluation companies were involved in the evaluation of Objective 1 and 2 programmes. In many cases consortia of evaluators were established. This is particularly the case for Germany, which had the largest number of evaluators (28) involved in the mid term evaluation process, compared to the number of evaluations (19). By contrast, in Sweden one evaluation company undertook all 6 evaluations while another company carried out two evaluations concerning management issues for ESF measures under Objective 1. The largest number of evaluations undertaken by one evaluation company was 10 in Spain, but in several cases this company worked in consortium with other companies. In Portugal one evaluation company was involved in 9 evaluations while another was involved in 7, but in both cases the companies were in consortium with each other and/or other companies. In France, one company undertook 6 evaluations. For Germany, Belgium, Greece, Italy and the UK, the maximum number of evaluations undertaken by one company was 4. It was less than this in Austria, Denmark, Finland, Ireland, Luxembourg and the Netherlands.

In terms of involvement in evaluations across national boundaries, there is some limited evidence of transnational involvement with evaluators from Ireland, Austria, UK and Denmark involved in the German evaluations, Swiss evaluators undertaking one Austrian evaluation, a Belgian evaluator undertaking the Luxembourg evaluation a Spanish evaluator involved in one Portuguese evaluation, and a German evaluator involved in the Irish CSF evaluation.

An innovative approach to the mid-term evaluation based on a transnational exchange of experience was demonstrated in the Objective 2 regions of Nordrhein-Westfalen and the East and West of Scotland. Building on long term co-operation links, the regional authorities decided to launch a comparative study on the mainstreaming of horizontal issues in these Objective 2 programmes. The objective was to stimulate debate on experience of implementation of the Structural Funds and to strengthen the learning process. The results of the study were discussed in a workshop of representatives of the three regions.

Overall, the mid term evaluation process does not demonstrate a significant increase in the size of the marketplace of evaluators. There is evidence that there was extreme pressure on that market with the 31 December 2003 deadline for all mid term evaluations while at the same time some evaluators were also working on ex ante evaluations for the new Member States. Those responsible for managing the evaluations had to insist on obtaining quality work as scheduled, particularly from evaluators engaged in a large number of evaluations, some of whom appeared to be over-stretched. This underlines the importance of evaluation capacity in the management of evaluation being available in the managing authorities to ensure the quality of the process. It also highlights a need for evaluators to ensure that they meet professional standards for evaluation in carrying out Structural Fund evaluations.

3.3. The Quality of the Mid Term Evaluations

Article 42(3) of Regulation 1260/1999 required the Commission to examine the quality and relevance of the mid term evaluations on the basis of criteria defined beforehand by the Commission and the Member State. The criteria used were those developed under the MEANS[6] programme of methodological guidance on evaluation of the Structural Funds. In order to contribute to raising the quality of the evaluations, the Commission decided to undertake quality assessments of Objective 1 and 2 reports at draft stage where they were available and to give feedback to managing authorities in time for evaluators to improve the quality of the final reports. The following table demonstrates the very considerable improvement made in the evaluations between draft and final report stage.

Quality Assessments – Objective 1 and 2 Reports

	Draft Reports %	**Final Reports %**
Good/Excellent	34	65
Acceptable	48	35
Unacceptable	18	0
Total	100	100

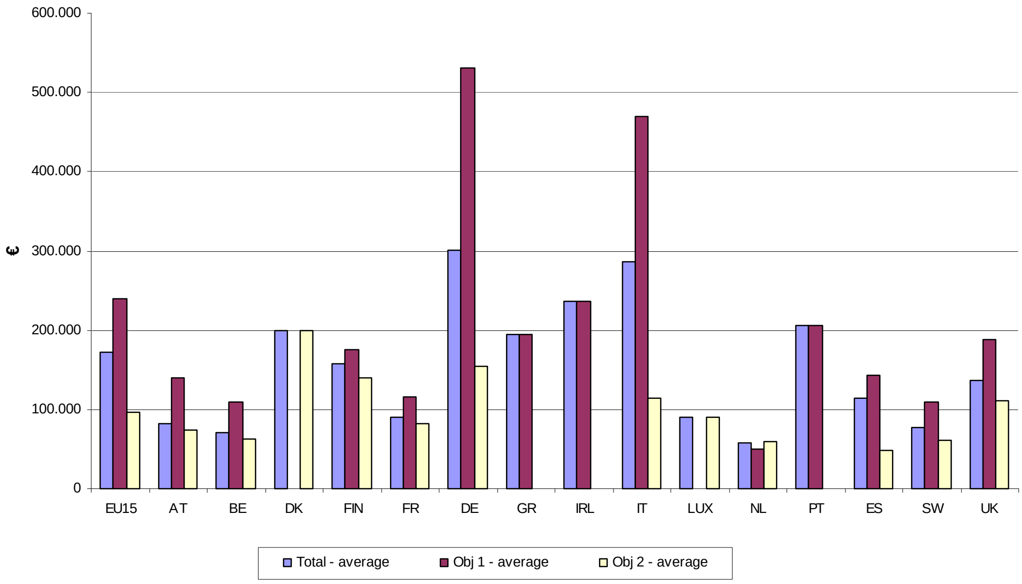
[6] European Commission, (1999). *Evaluating Socio-Economic Programmes. Volumes 1-6.* Luxembourg: Office for Official Publications of the European Communities. (The MEANS collection has been updated and a new GUIDE to the Evaluation of Socio-Economic Development can now be accessed through the website of the Directorate General for Regional Policy or directly at www.evalsed.info).

The final outcome, with nearly two thirds of the evaluation reports assessed as being good or excellent, is positive. At draft stage, the high number of unacceptable reports was due to many being unfinished. To be judged acceptable, a report needed to address all the key evaluation questions. Good or excellent evaluations did so in more depth and based on a greater degree of primary research. The end result, with the significant improvement in quality of the reports between draft and final stage demonstrates contribution of the partnership between the Commission, the managing authorities and the steering groups in the management of the evaluations.

3.4. Cost of the Evaluation and Cost Effectiveness

Very significant resources were allocated to the mid term evaluation exercise, which contributed to its higher quality compared to the past. In total, Objective 1 and 2 evaluations cost over €35 million with an average cost of €172,000 per evaluation (nearly €240,000 for Objective 1 evaluations and €97,000 for Objective 2 evaluations). There was a very large range in costs, with the lowest cost of an evaluation €27,000 and the highest at €906,000. Some Member States allocated more than others to the mid term evaluation, with the cost of the entire mid term evaluation exercise highest in Italy, followed by Germany, Greece, the UK, Spain and France. Figure 1 below presents the average costs of the mid term evaluations broken down by Member State and by Objective (excluding CSF evaluations).

Figure 1
Average Costs of Evaluations broken down by Member State and Objective

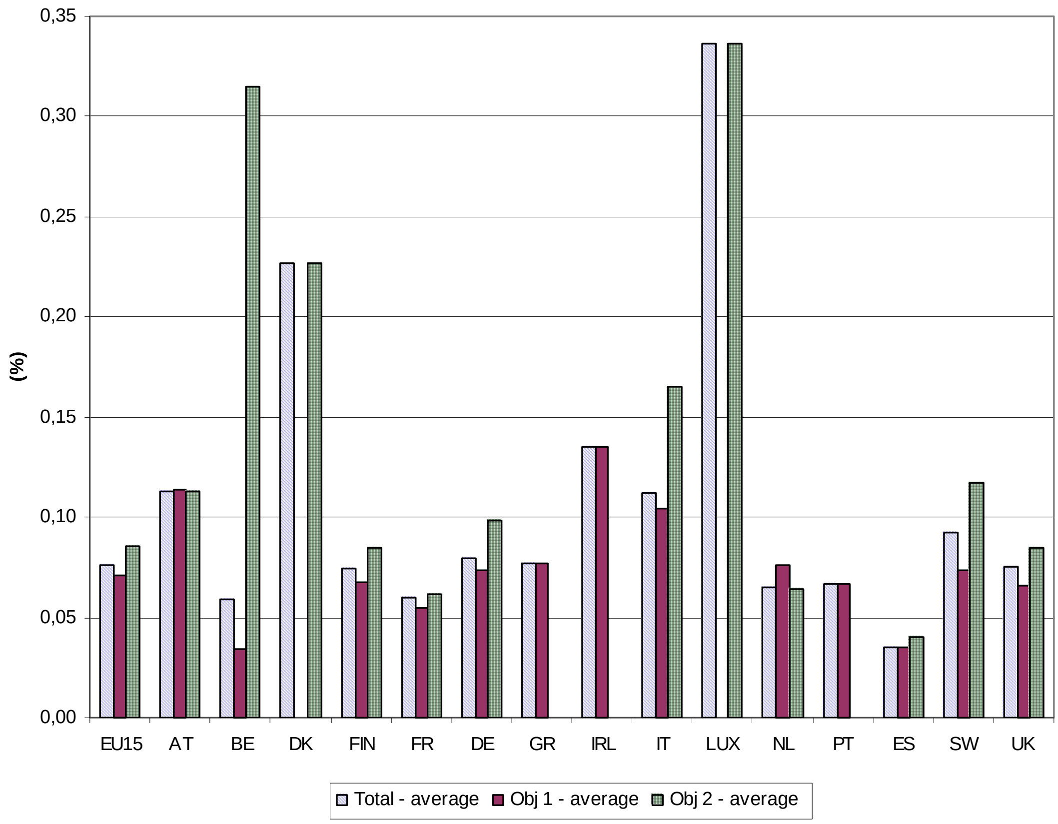


There can be no hard and fast rules about the most appropriate cost of an evaluation. The new GUIDE to the evaluation of socio economic development suggests that the cost will be generally less than 1% of the total annual cost of the programme being evaluated. For very large Structural Fund programmes, a lower percentage is appropriate and this is particularly the case for CSF evaluations which cover all programmes in a country or macro region.

The mid term evaluations for the 1994-1999 period cost significantly less than 0.1% of the total annual cost. For the 2003 evaluations of Objective 1 and 2 programmes, the average cost (excluding CSF evaluations) was 0.076% of total annual cost of programmes, with the average cost for Objective 1 evaluations 0.071% and for Objective 2 evaluations 0.091%. The lower percentage cost of Objective 1 evaluations is to be expected, given the scale of many of the larger Objective 1 programmes. The highest percentage costs for the Belgian, Danish and Luxembourg Objective 2 programmes reflect the small scale of those programmes.

There was no correlation between the cost of the evaluations and the Commission's assessment of their quality. The size of the programme was a more important factor in determining the budget for the evaluation, although here too significant variation was evident across Member States.

Figure 2
Average Costs of Evaluations as % of Total Annual Programme Costs by Member State and Objective



It is difficult to assess the cost effectiveness of the mid term evaluation process. It is undoubtedly true that a very significant amount of resources was allocated to the mid term evaluation. The ultimate cost effectiveness relates to the usefulness of the evaluations in improving the quality of the programmes and whether the evaluations provided added value to managing authorities and to Member States and the Commission in pointing to possible improvements to programmes and in accounting for what had been achieved to date. This is an issue to which we will return in our conclusions following our analysis of methodologies used, the content of the evaluations and their use.

4. METHODOLOGIES

As demonstrated in the previous section, the scale of programmes to be evaluated and the scale of the budget allocated to the evaluations varied considerably. The scale and complexity of the programme and the budget for the evaluation clearly influence the methodology. In general, the evaluations were based on a mix of desk research and some fieldwork. For larger Objective 1 programmes or CSFs, some macro economic modelling was carried out. In this section we analyse the methodologies used and highlight some concerns which need to be addressed in the future. The analysis is based on a sample of 88 evaluations.

4.1. Mix of Methodologies Used

As noted above, the mix of methodologies used included desk research, some primary research and some modelling work. The mix was required, given the wide range of questions posed in the mid term evaluation, which included evaluation of:

- developments in the programme environment,
- the implementation mechanisms of the programme,
- the performance of the individual measures, the priorities and the programme as a whole in achieving their objectives.

These different elements required different techniques and approaches from the evaluators in order to answer the questions posed.

Some excellent evaluations had a mix of methodologies specifically designed to answer the evaluation questions and to build on the available information. For example, the Merseyside evaluation, in addition to desk research and key stakeholder consultations, was based on 128 project reviews explicitly designed to provide a “reality check” on the views of implementing bodies and two beneficiary surveys targeted at businesses and individuals.

In general, the examination of implementation systems took on greater importance than foreseen in the Commission’s working paper because of the late or slow start-up of many programmes, leading to limited information on outputs and results on which to base judgements on the performance of programmes. This focus of the evaluations on implementation systems was particularly notable in Italy and France.

4.1.1. Desk Research

All mid term evaluations had an important element of desk research. This research included analysis of previous evaluation reports and studies and the updating and analysis of recent trends in national statistics. This was done extensively in most cases, although it was notable how few evaluators concluded that changes in the programme environment had any implications for the programme strategy, despite the global economic slowdown.

An important element of the desk research was analysis of the monitoring data provided by the managing authorities. The monitoring data was a crucial starting point for the evaluator who had to examine outputs and results achieved to date and compare these to what was planned. Ideally, primary research should then be designed to explain the results being achieved and analyse the extent to which needs are being met.

Unfortunately, for a number of reasons this type of analysis was not possible in many evaluations:

- Some programmes were late or slow in starting and had very limited outputs and no results to report. In these cases evaluators, appropriately, focused particularly on implementation systems, including project selection and indicators.
- Some monitoring systems were not functioning properly and were unable to provide the evaluator with the information required. In these cases there seems to be a problem in channelling information from the project level into the monitoring system. This raises questions about the efficacy of monitoring, when it does not permit access to timely information on what is happening in the programmes. For the mid term evaluations, the unavailability of monitoring information presented some evaluators with a serious challenge. Some responded by undertaking a greater amount of primary research. Others, unfortunately, did not respond to the challenge and weak evaluations were produced as a result.

Where monitoring information was available, even if incomplete, the quality of the analysis varies. In the best cases, (e.g., Steiermark, Austria Objective 2, Hainaut, Belgium Objective 1) the monitoring data was analysed at programme and priority level as well as at measure level, with comparisons made to targets and to financial absorption. As a result of this analysis and the primary research, recommendations were made, where necessary, for more relevant indicators and more realistic targets. Less satisfactory were the evaluations which contained measure by measure or sub-measure by sub-measure information, only sometimes with performance compared to target, and, crucially, the entire mass of information not analysed at a priority or programme level in the interests of assessing overall programme performance.

4.1.2. Primary Research

The extent of primary research varied considerably across evaluations. In all cases structured interviews and consultations were held, although it was not always clear that interviews included stakeholders not directly responsible for programme management and implementation.

The methodologies of 88 mostly good mid term evaluations were examined for this report. Since the absence of primary research was one of the main reasons for lower ratings of the quality of the evaluations, it should be noted that the analysis in this section contains an over-positive picture of the extent of primary research undertaken for the Objective 1 and 2 mid term evaluations.

Approximately 57% of evaluations examined included some survey work, although the extent of this varied considerably. There was more survey work in Objective 2 evaluations (65% of our sample) compared to Objective 1 evaluations (49%). Surveys were undertaken with beneficiaries, project managers and sometimes the wider community. Some programmes had little activity to report and therefore there was no point in undertaking beneficiary surveys at this stage. This was the case in a number of Italian Objective 2 programmes for example. In general, surveys were undertaken in smaller programmes with fewer measures. For large Objective 1 programmes with a very wide range of types of measures, beneficiary surveys were often not undertaken. While the cost of survey work across such extensive programmes might have been prohibitive, evaluators sometimes undertook targeted surveys for priority measures.

Germany and the UK were notable for the amount of survey work incorporated in evaluations of large programmes.

Focus groups or workshops were also used in some evaluations, where the evaluators brought together a range of beneficiaries or other stakeholders for a facilitated discussion of particular issues. Such groups were used in 42% of the sample of evaluations examined and were used slightly more in Objective 1 evaluations compared to Objective 2 evaluations.

Case studies are another approach which allows the evaluator to examine a limited number of projects in detail to obtain information on the functioning of the programme. The case study methodology was used in approximately 51% of evaluations, although it is not always clear that the case studies were more than project visits. Case studies were used more frequently in Objective 1 than Objective 2 evaluations.

In 21% of the evaluations examined, focus groups or case studies were used as an alternative to more extensive survey work and this was particularly the case for Objective 1 evaluations (31% of those examined). In 44% of cases, focus groups and case studies complemented survey work.

4.1.3. Modelling Work

Section 5.1.2.1 below presents the results of the macro economic modelling undertaken in the CSF evaluations for the Cohesion countries, Germany and Italy as well as for the Hainault Objective 1 phasing out region in Belgium. In addition to this work, a number of the evaluations of large regional programmes in Germany and Spain undertook macro economic modelling, often adapting the national model to the regional context.

4.2. Methodological Concerns

Evaluations of Structural Fund programmes are complex, involving the evaluation of wide-ranging programmes in a multi-stakeholder partnership context. The list of issues to be examined in the mid term evaluation was extensive while the time and budget available for the exercise were finite. Taking account of the complexity of the task, it can be concluded that significant improvements were evident in the quality and rigour of the evaluations undertaken for the mid term compared to previous comparable evaluations. Further improvements could be made, however, and the following methodological concerns can be identified:

- In some cases, there was limited primary research which was focused only on those responsible for implementing the programme. An evaluation of a Structural Fund programme should incorporate the “on-the-ground” reality and use this to explain the findings of desk research and results of interviews with managing authorities and implementing bodies. Even where programmes were slow or late in starting up, once projects are selected, the perceptions of beneficiaries can provide useful information on the functioning of the programme.

- In many cases monitoring data was weak. Some programmes were late starting and had only limited data gathered on activity, while for others the monitoring system is not yet functioning and providing information on physical performance to the monitoring committee and to the evaluators for this exercise. This created problems for evaluators in reporting on what has actually been achieved to date.

- Linked to the unavailability of monitoring data but also to the strong focus on financial performance due to the demands of N+2, some evaluators over-emphasised financial analysis with a limited focus on what was actually being achieved by the programmes. This led to recommendations which were purely finance driven, rather than reviewing the strategic implications of such re-allocations. A further conceptual concern was that many evaluators defined "efficiency" as the absorption of funds, rather than the unit cost of achieving outputs and results.

- Evaluations often did not produce sufficiently operational conclusions and recommendations and some did not clearly differentiate between the two. Conclusions should be specific and clearly based on evidence presented while recommendations should be practical and actionable.

- A final point relates to the length of evaluations and in some cases their lack of analytical content. Many of the evaluations were too long and their length militated against the dissemination of their findings. There is a need for evaluators to develop a capacity for more synthetic judgement and to focus more on analysis rather than including very long descriptions of measures and sub-measures.

These points need to be addressed by evaluators in their future evaluation work. However, it is also important that managing authorities are aware of these potential weaknesses. In relation to data availability, managing authorities need to work with evaluators to establish how best to source objective information on which to base evaluations. They also need to demand and insist on high quality analysis, clear and practicable conclusions and recommendations and concise reports.

5. MID TERM EVALUATION FINDINGS

This chapter presents an analysis of the findings of the mid term evaluations based on two different approaches. In the first place, the answers to the key evaluation questions are analysed, focusing on the appropriateness of strategies, the achievements to date and implementation issues. In the second part, a thematic analysis is undertaken, focusing on the relevant themes of the Lisbon Agenda. The aim here is to seek to draw conclusions on what the mid term evaluations tell us about progress in relation to these priorities.

The analysis is based on a sample of 65 mid term evaluations which the Commission judged to be of good quality. These evaluations were selected on the basis that the higher quality evaluations would be more likely to conclude on the evaluation questions and to provide insights into developments in programmes in relation to particular themes. The evaluations covered both Objective 1 (49%) and Objective 2 (51%), large and small programmes and regional and sectoral programmes.

5.1. Answers to Evaluation Questions

The answers to evaluation questions are based on a synthesis of the conclusions in relation to each of the key components of the evaluation.

5.1.1. Limited Need for Strategy Shifts

In nearly all cases, the evaluators concluded that the strategies were still appropriate and that there was no justification for radical changes. The global downturn in the world economy was analysed across the different programmes, but the conclusion was that the programmes were still appropriate to meet the development challenges. Just two mid term evaluation out of our sample of 65 (the Science, Technology and Information programme in Portugal and the Highlands and Islands in UK) concluded that the initial strategy adopted in the programming documents should be extensively revised.

Some adjustments or refinements to strategies for the second half of the programming period were recommended:

- The Austrian evaluators found that the SWOT analyses were still valid and appropriate to meet the objectives of the programmes, but a stronger focus should be placed on gender mainstreaming in the Burgenland programme. Some changes should also be made to the weighting of individual priorities and measures in the Oberösterreich programme. Evaluators of the Steiermark programme recommended that in further programme implementation more attention should be paid to international aspects (the enlargement of EU), better co-ordination with Interreg and the effective implementation of R&D measures.

- The French evaluations found that in most cases the strategies were broadly appropriate. Recommendations mainly focus on a stronger emphasis on employment and on preventive measures for local employment in rural areas or those in industrial decline (Bourgogne). A greater emphasis is proposed to strengthen competitiveness and develop ICT (Alsace, Bourgogne).

- The German CSF evaluation found that the overall strategy was still relevant. The mid term evaluations emphasized the need for greater concentration on important priorities with decisions on actions to be funded to be made on the basis of the potential for a contribution to long term growth. All Objective 1 programmes need to be fine-tuned in 2004 to reflect new labour market legislation.

- The Irish evaluations examined the under-absorption of allocations for enterprise and research in particular and suggested a shift towards transport infrastructure, further strengthening the emphasis on this area.

- The Netherlands evaluations reported that the overall the strategy and the associated priorities remain appropriate. They take due account of the relevant policy framework (European, national and regional). The evaluators concluded that the Oost Nederland SPD strategy must now prove that it can operate successfully under current favourable conditions.

- The Portuguese evaluations recommended a better articulation with the European Employment Strategy, a greater focus on the modernisation of the services sector and public administration, and a stronger effort to diffuse innovative public policies and to develop training programmes for SMEs.

- The Spanish evaluations recommended some modifications as a result of the "Prestige" catastrophe as well as the adaptation of programmes to focus more on research, development and innovation.

- The UK evaluations in general found strategies to be still sound. For Merseyside, the evaluators suggested that opportunities for using Objective 1 resources should be considered with a view to securing a stronger culture of innovation and technology transfer, with an increased focus on the demand side. There is also a recommendation for a closer alignment between the East of England Objective 2 programme and the Regional Economic and Regional Social Strategies, with regard to strategic aims, funding arrangements and project selection procedures.

5.1.2. Effectiveness – Much Activity Underway after a Slow Start

Effectiveness is considered in relation to macro economic impacts and outputs and results achieved. The macro-economic impacts are reviewed for the large CSFs which undertook macro economic modelling.

5.1.2.1. Macro-economic Impacts

One of the objectives of the mid-term evaluations is to provide feedback on the likely achievement of the projected impacts of Structural Funds interventions. Where the scale of intervention is large enough, macro-economic models can be used to estimate economic effects. The approach to model building varies greatly from evaluation to evaluation. As a consequence, the results cannot be compared. Not only the underlying assumptions are diverse (for example, on the timing of payments or treatment of private funds), but also the aim of the analysis differed from updating the model of the ex ante evaluation to determining the effect of actual expenditures between 2000 and 2002. Furthermore, estimations were made with different time horizons (usually up to 2008 or 2010). Although the sophistication of the models and the scope of simulations vary, a general conclusion can be drawn that Member States have progressed in terms of the quality of the analysis.

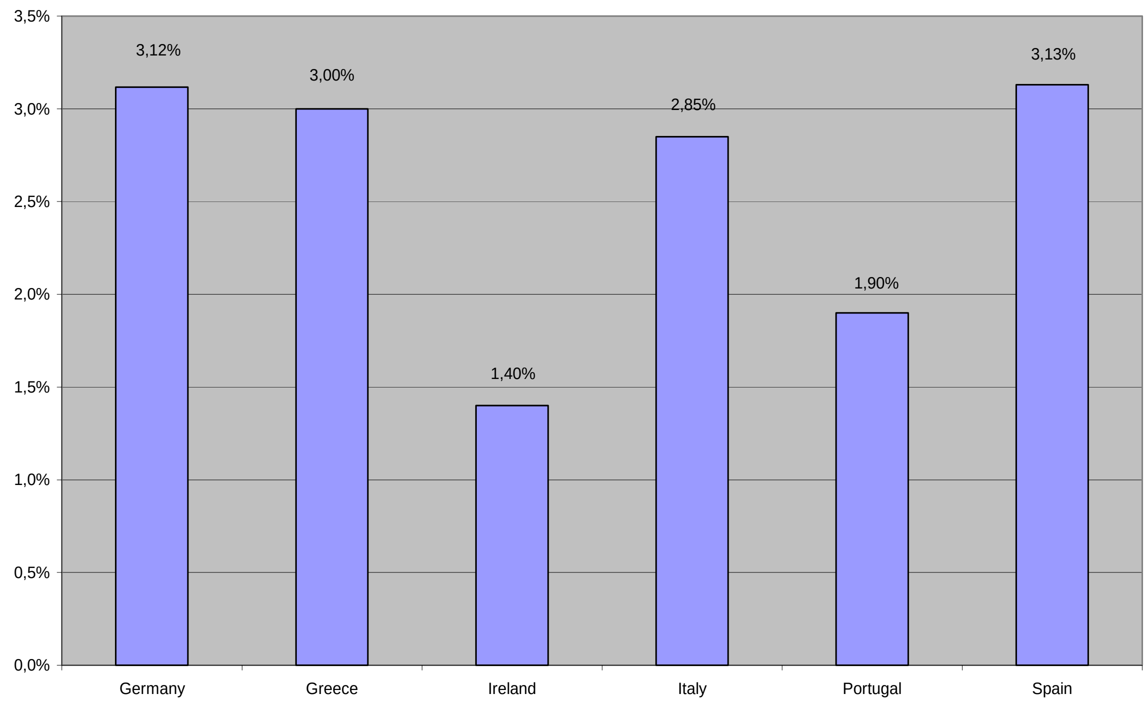
Most models distinguish between supply and demand side effects. The distinction is important, as demand side effects, which result from the fact that the immediate financial “inflow” increases demand, tend to be short-lived and cease when the flow of funds stop. At the same time supply side effects, which derive from the structural adjustment of the economy, develop over a longer period but are more permanent. The capability to assess short and long term impacts helps us to better determine the sustainability of Structural Fund interventions.

A specific problem of macroeconomic modelling arises in countries where the whole country is not eligible for assistance under Objective 1 (Germany, Italy, and Spain). The evaluators addressed this problem in different ways: in Germany, the geographic scope is East Germany; in Spain, the impact of CSF was analysed at a national level; while in Italy, the effects were assessed for Southern Italy and Italy as a whole.

In the mid-term evaluations only minor changes were proposed in financial allocations, and none of the models ran simulations to make estimations of the impact of possible changes.

Figure 3

Impact of EU funding on GDP level (average 2000-2006, deviation from baseline scenario)



Greece: 2000-2008
Ireland: based on expenditures 2000-2002
Italy: average of low and high externality scenarios
Portugal: based on expenditures 2000-2003

The impact of Structural Fund assistance is illustrated in Figure 3 above. The Structural Fund contribution was assessed for the total financial allocation for 2000-2006, except for Ireland and Portugal. Ireland and Portugal assessed the effects based on actual expenditure. The other four countries aimed to determine the expected impact of Structural Fund assistance based on available information. In Greece and in Italy, the time horizon runs to 2008 while in Germany and in Spain the period under analysis extends to 2010.

The models used are, in general, more sophisticated than those used in the ex ante evaluations. The Italian model, for example, underwent a series of changes. The Italian authorities reviewed the evaluation made ex ante using a more elaborated model and incorporated changes in global environment. The Portuguese authorities developed a new model called HERPOR, which is able to give more reliable results. In the mid-term evaluation of the Irish NDP, the macroeconomic modelling was complemented by several studies, which increased the reliability of the analysis by providing better estimation on some parameters.

Assessing macro-economic impact in small areas – Hainaut, Belgium

For small areas, specific models can be designed to estimate long-term impact on growth. A model was developed for the Objective 1 phasing out region Hainaut (Belgium), which provides valuable insight into how structural adjustment takes place.

The difference in GDP between the “with phasing out assistance” and “without phasing out assistance” is not great (0.3% in the short, 0.07% in the long run). However, looking at other variables, an increase in industrial productivity of (2.7% and 1.7% in the short and long run respectively) and an employment shift from industry to the tradable services sector can be noted as signs of structural adjustment.

5.1.2.2. Results and Outputs

Achievements measured in terms of *financial* performance were presented in all mid term evaluations, although recommendations were sometimes weak on actions to take in relation to under- or over-spends. Analysis of *physical* performance presented problems in many Member States. In a number of cases, the monitoring system provided little or only partial information on achievements, as was discussed in section 4.1.1 above.

It is clear that a substantial amount of activity has already taken place under the 2000-2006 period, although the full impact of that activity will only be evident over time. Below we outline some of the outputs and results achieved according to the evaluations. The outputs represent examples of achievements and are not generalisable across all Objective 1 and 2 programmes.

- In Austria, the overall performance of the programmes is considered good, although figures vary across priorities and measures.

- In Denmark, the mid term evaluation found that expected effects in the form of the total forecast number of jobs created is relatively high. The approved projects are forecast to create 2,833 jobs, which represents 143% of the target.

- In Finland, implementation of the programmes has proceeded well in most programmes and good quantitative results have been gained. In Eastern Finland, 15,900 jobs had been created, representing 90% of the target and 3,200 businesses had been created (105% of target). In Northern Finland, 6,700 new jobs had been created, 62% of the target, and over 900 new companies had started up (94% of target). For Objective 2, 8,300 new jobs and 1,000 companies were created in

Southern Finland (55% and 20% of target respectively) while 16,000 jobs and 2,200 companies were created in Western Finland (46% and 22% of targets).

- In Germany Objective 1, macro economic modelling in the CSF mid term evaluation estimates that the level of GDP will be 2.6% to 3.4% higher by 2006 as a result of the Structural Fund investment and the Structural Fund interventions will lower the rate of unemployment by approximately 2%. In the Nordrhein-Westfalen Objective 2 region, 38% of the resources are committed and 14% are paid. 20,000 new jobs have been created, 20% of the target, while 46,000 jobs have been saved, well above the target of 29,000. In Rheinland-Pfalz, also Objective 2, 2,066 new jobs had been created, 36% of the target and 385 jobs had been saved. In R&D and business start-ups, 70 new jobs had been created and 213 jobs saved.

- In Spain the CSF evaluation estimates that the rate of annual growth was 0.41% higher in 2000 to 2002 as a result of CSF expenditures. Improved accessibility has been achieved through 476 km of roads/motorways built or improved, 173 km of rail built or improved. In addition, 810 km of energy distribution networks have been constructed and 250 km of gas pipeline built. Support for the productive environment has resulted in 4,600 SMEs being assisted with a leverage effect on investment estimated at approximately €12.2 billion. A contribution to the development of the knowledge society is evident in 1,503 research centres and 48,199 researchers supported and the installation of 26,864 ADSL lines. Approximately 7 million beneficiaries have been supported by the ESF, of which 57% received training.

- In Portugal, the evaluation of the Economy Operational Programme states that 38,000 jobs have been created in SMEs. The gas distribution network has been extended by nearly 3,000 km and 25,000 consumers have converted to natural gas. With the support of the Agriculture and Rural Development Operational Programme, nearly 2,300 young farmers were newly installed. Around 150,000 active workers were involved in training programs promoted by the Operational Programme for Employment, Training and Social Development.

- In Sweden, in the two Objective 1 programmes, completed projects have created almost 5,000 jobs, supported the establishment of 1,390 SMEs and the training of over 18,000 people. The two Objective 2 programmes supported the creation of more than 7,500 jobs and 2,400 SMEs.

- In UK, the aggregate data suggest that progress in terms of predicted outputs is broadly in line with the ambitions of the West Wales and the Valleys as well as East of England Programmes, and the so far limited evidence in terms of actual outputs does not show major discrepancies from these predictions. Nevertheless, the mid term evaluations for the South Yorkshire, South West of England and Yorkshire and Humber demonstrate that the outputs achieved across the Priorities are relatively low.

Overall, we can conclude that a significant amount of activity is now underway in Objective 1 and 2 regions. However, due to the late start-up of some programmes and weaknesses in the indicators, target setting and data collection systems, it is not possible to draw conclusions on aggregate performance across programmes and Member States.

5.1.3. Efficiency – A Start Made in the Analysis of Unit Costs

Working Paper No 8 recommended that mid term evaluations should examine the efficiency of the form of assistance, in terms of the outputs and results achieved compared to the inputs, primarily the resources deployed. Most of the mid term evaluations understood efficiency in terms of financial absorption (commitment and payments). Out of our sample of 65 mid term evaluations, the examination of unit costs was undertaken in only 21 evaluations of programmes in Germany, France, Greece, Ireland, Italy, Netherlands, Portugal, Sweden and UK.

- In Germany, some consideration was given to the opportunity cost of interventions (for example support for enterprises versus unemployment benefits) in the CSF evaluation report. The evaluations for the Berlin Objective 1 and 2 programmes examine the efficiency of ERDF measures and unit cost per participant in ESF interventions. Efficiency of EAGGF interventions is measured in terms of invested EAGGF resources per job and per hectare of land.

- In Greece, an examination of efficiency in terms of unit cost was undertaken by all evaluators. Most evaluations analysed unit costs in relation to outputs and results achieved at measure level and included a qualitative analysis of the factors differentiating planned and actual costs.

- In France, evaluations for Guyane and Martinique examined the unit costs at the level of certain measures.

- In Ireland, the evaluation of the CSF performed the calculation of unit costs for around 100 cost categories (indicators) from the two regional programmes. Detailed analysis of cost inflation was undertaken for infrastructural investments. Efficiency was also examined in the evaluation of the programme for the Borders, Midland and Western Region.

- In the Netherlands, the evaluation for the Oost-Nederland programme dealt with this issue using three approaches: ordinary cost effectiveness, multiplier effects or ERDF monies in terms of leverage of private resources, and the efficiency of effects on employment opportunities.

- In Spain, a comprehensive study of efficiency with detailed unit cost presentation as well as comparison against unit costs in other sectors of the economy was carried out by the evaluator of the Andalusia Objective 1 programme.

- In Sweden, detailed analysis of the unit cost per job created and new enterprise established was undertaken at measure level.

- In the United Kingdom, the Merseyside mid term evaluation examined unit costs for the public cost per net job, per million pounds net additional GVA and net GVA per net job created. The study also compared the planned (SPD) unit cost values with the observed expenditure and impacts, broken down by priority and as a total.

Some evaluators stated that that since the programme aims to achieve objectives, effectiveness is more an issue than efficiency (mid term evaluation for Productive Sector OP in Ireland). Others stated that the late start of the programme or the lack of reliable monitoring data were the reasons not to examine the efficiency in terms of the outputs and results achieved compared to the inputs.

5.1.4. Quantification of Objectives – an Urgent Need for Improvements

The quantification of objectives is an area where most evaluations found room for considerable improvement, although there were some very good examples of quantification of objectives feeding into an effective monitoring system (e.g., Germany - Nordrhein-Westfalen). In some cases, (e.g., Greece, France, Spain) baselines were missing which made evaluation work difficult. Recommendations to revise and simplify the indicator system were made in Belgium, Finland, France, Germany and Greece, while in Denmark some additional indicators were recommended. In Portugal the evaluators recommended further changes to improve the indicator system which had already been simplified during 2003.

In many cases, inadequate definition of indicators was identified as a problem and recommendations were made to address this issue. Particular problems arose with impact indicators, especially job effects (defining gross and net jobs or jobs created and safeguarded).

The mid term evaluations in Austria, Finland, France, Germany and UK highlighted difficulties with the indicators for horizontal priorities, which often did not have objectives, were difficult to measure or not monitored.

Some mid term evaluations, for example, the evaluation of the Hainaut Objective 1 programme in Belgium and the Berlin Objective 1 programme in Germany, reported the problems related to the monitoring of ESF interventions, especially possible double counting.

A further issue arises where performance was far in excess of targets, suggesting that the target setting exercise was flawed at least in some areas (e.g., Austria, Finland, the Netherlands, Spain, Sweden). The explanation may be extremely good performance, but the evaluators do not always analyse the reasons for such over-performance. The fact that it is often accompanied by expenditure patterns in line with target suggests that in these cases targets were set too low. In Spain, the evaluators have recommended changes to targets to set them at more realistic levels.

In the UK, the quantification of objectives was generally of a good quality since most programmes were based on a methodology developed during the last programming period[7]. This work provided a good starting point and the mid term evaluations have recommended further improvements to bring targets more in line with regional and programme specificities. Evaluators for the West Wales and the Valleys Programme in the UK found that significant progress had been made in developing clear definitions of the key indicators and in rationalizing the “hierarchy” of targets.

[7] EKOS Limited, (1998): *Baselines and Quantification in the United Kingdom Objective 2 Programmes*

5.1.5. *Implementation Systems – Significant Improvements compared to the Past*

As noted earlier, the evaluations provided extensive material on implementation systems and made many recommendations to improve and streamline practice. However, in nearly all cases, the evaluations found that implementation systems respect the formal requirements and had improved considerably compared to the past, with more rigorous project selection procedures, better financial control and evaluation and improving monitoring systems. In some cases (e.g., Greece, Finland, UK) systems were found to be over-complex and recommendations were made to simplify project selection or other decision making procedures. The following summarises some of the key findings related to implementation:

- The mid term evaluations in Austria found that in comparison with the previous programming period the implementation systems have been simplified, streamlined and made more transparent. Nevertheless, the evaluation for Burgenland found that responsibilities at priority level were not sufficiently clearly defined in contrast to programme and measure level.

- In Denmark, the evaluators concluded that regional structures have been successfully created and that they operate effectively in implementing the programme. The evaluators found that administrative structures and the distribution of work between regional and central levels were satisfactory. Strengthened guidance and communication were recommended in the areas of audit, control and network establishment.

- In Finland, the evaluators considered the implementation of the programmes to be successful. Criticism was directed towards the fragmented nature of the programme structure, a heavy management model at national level (East Finland) and the use of monitoring system (West Finland).

- In France, the findings of the mid term evaluations varied. The evaluators considered that the implementation and control systems have been improved compared with the previous programming period and perform satisfactorily in Martinique, Aquitaine, Alsace and Haute Normandie. Although the management structures fulfil the legal requirements in the Guyane programme, the evaluators considered that weaknesses in implementation resulted from a shortage of manpower.

- In Germany, the evaluators found that overall comprehensive, transparent and well-functioning management systems have been established. However, various shortcomings need to be addressed in different programmes. More competitive project selection processes and better co-ordination between different monitoring systems were recommended for Berlin programmes. Further monitoring system improvements were recommended for the Sachen-Anhalt, Bayern and Schleswig-Holstein programmes. Recommendations were made for improved integration between the authorities responsible for different funds in Rheinland-Pfalz.

- While the existence of an integrated monitoring system was identified as a positive element for Greece, the system for collecting and processing information was found to be weak although progress has been made compared to the past. The evaluations concluded that a continuing emphasis is needed on this area to bring about further improvements. In the evaluation of the Attica programme, the evaluators concluded that the managing authority complied largely with regulatory obligations. Strengthening the capacity of the managing authority to provide support to final beneficiaries and improvement of the Integrated Management System were identified as areas for particular attention.

- The implementation systems were found to be adequate in the Italian mid term evaluations. A revision of project selection criteria was recommended in Sardenia and Bolzano. According to the evaluators, the monitoring system functions well in Abruzzo, Toscana, Emilia-Romagna, but a need for improvement was identified in Sardenia and Piemonte

- The evaluation of the CSF in Ireland identified three principal concerns for implementation: project selection and prioritisation, cost estimation and control, and programme management.

- In the evaluation of the Oost-Nederland programme, the implementation structures were found to be reasonably satisfactory and efficient. According to the evaluators, the selection criteria and procedures work well and, compared to previous periods, communication and publicity has progressed substantially. Room for improvement was identified in partnership structures, coherence between relevant committees and the efficiency of the monitoring system.

- In Portugal, the mid term evaluations recommended a stronger co-ordination role in managing the CSF and a more strategic role in managing and monitoring the CSF and OPs. They also suggested the revision of project selection criteria and an increase in the human resources involved in managing the Science, Technology and Innovation programme.

- In Spain overall, programming, co-ordination, monitoring and evaluation have been assessed as adequate for an efficient implementation of the programmes. Nevertheless, excessive bureaucracy and scarcity of human and material resources were identified as areas where the further improvements are needed in the Andalucia and Galicia programmes.

- In Sweden, the programmes fulfil the requirements of financial and administrative management. The evaluations suggested the introduction of special requirements for larger projects.

- In the United Kingdom, the mid term evaluations considered project development, appraisal and selection systems to be substantially improved compared to the previous programming period. Recommendations were made for further improvements. A significant rationalisation of programme management was recommended in the North West England programme while improvements in ESF arrangements were deemed necessary in the East of England programme. The UK evaluations demonstrate a strong partnership basis for programming and implementation, which is complemented by an appropriate degree of top-down co-ordination.

5.2. Contribution to Lisbon Priorities

It is important to note that evaluation of the contribution of the Structural Funds to progress towards meeting the Lisbon objectives was not one of the tasks of the mid term evaluation. Structural Fund programmes were mostly drawn up before the Lisbon European Council in 2000 and the focus of the evaluations was therefore on how well programmes were meeting their own objectives. That said, much Structural Fund investment is targeted at sectors which are important for the achievement of the Lisbon objectives. Therefore, in this section of the report, we identify some relevant findings which can give insights on the progress being made in these important areas.

5.2.1. Transport – Good progress in large Road and Rail projects; smaller scale projects behind schedule

Of the 65 evaluations analysed in detail, investment in transport infrastructure was included in 49. The 49 evaluations were of 26 (53%) Objective 1 programmes with the remainder Objective 1 programmes. Not surprisingly investment in transport infrastructure was a higher priority in Objective 1 programmes.

In Objective 1 regions, transport infrastructure has a particularly high priority in the four cohesion countries and in Germany and Italy. In general, in these countries the transport measures are progressing well with high level of commitments. Road and rail projects are advancing at a faster rate than projects to improve ports and airports and local transport initiatives. In Spain, this priority has the highest absorption rates of all priorities. In Greece some regions present a similar picture. In Ireland, absorption rates are also high, although the evaluators conclude that cost inflation means that the outputs and results to be achieved will be below target.

For smaller Objective 1 regions, the focus tends to be on smaller scale infrastructure, such as the development of port facilities, access to particular industrial areas and the development of public transport in urban areas. These projects often demonstrate a slower start-up than the large-scale road and rail projects cited above – many of which were carried over from the previous planning period, in terms of planning or development of earlier stages.

In Objective 2 regions, the main transport investments concern the integration of different modes of transport, the improvement of facilities at ports and airports and some road and rail building and improvement. The last were particularly evident in Spain where good progress was being made. Slow start up of many programmes meant that in other countries, progress is behind target.

5.2.2. *Knowledge Based Economy*

5.2.2.1. Research, Innovation, Information and Communication Technologies (ICT) – A growing priority with good results in Objective 2 regions; slower but definite progress being made in Objective 1 regions.

All the mid term evaluations examined address the issue of research, innovation and ICT and where programmes do not include this area as a priority, the evaluations make recommendations to develop measures or actions to respond to this important Community priority.

The following types of actions are supported by the Structural Funds in Objective 1 and 2 regions:

- Co-operation between research institutes and businesses,
- Development of business clusters,
- Investments to increase broadband access in businesses, educational institutions and public administrations,
- Connection to regional and international networks, and the development of internet platforms,
- Development of research centres, including their construction and/or equipment,
- Applied research projects, including environmental research,
- Development of regional innovation strategies,
- Training of researchers.

Information society is also present as a horizontal theme in many programmes.

In Objective 1 regions, the mid term evaluations in general find problems with slow start-up of measures and consequently little progress achieved in terms of outputs or results. This is particularly notable in Greece, France, Ireland and Italy. In Spain, while the CSF evaluation identifies low absorption rates, some regional programmes are performing quite well in this domain (e.g., Andalucia). German Objective 1 regions seem to be demonstrating good performance in this sector as are some of the smaller Objective 1 regions, including Burgenland in Austria, East Finland, Norra Norrland in Sweden.

Andalucia Objective 1 – Spain – R&D measures performing well

Priority 2 in the Integrated Operational Programme for Andalusia contains measures to foster research and development in the public and private sector in the region. The priority has performed well in financial terms (40%), although the level of financial execution varied considerably between measures. The physical performance of the priority has been very satisfactory and in many measures exceeded 100% for the 2000-2006 period. During 2000-2002, 86 research, development and innovation projects, and 2,109 research groups have been co-financed, 23 research centres have been supported, 3,266 networks have been created in public administration services, and 1,878 teams have been established to expand the network of environmental information

Burgenland Objective 1 – Austria – Good Evidence of Outputs and Results

Research, technology and innovation is one of the priorities of the programme, focusing on development of business clusters, competence centres, co-operation projects between research institutes and businesses and innovative services. The priority shows good progress. By mid-term, 3 co-operation projects between businesses and research institutes were complete, 4 new technology centres were-created and 1 extended, 11 clusters were developed, 9 new production/processing methods developed and the priority contributed to creation of 350 new jobs and saved 192 jobs in R&D. In addition, the Information technology, telecommunication-nets and applications measure focuses on connection to regional and international data networks, database, development of internet platforms. Here 8 co-operation projects were implemented, 14 new jobs created, 8 new products or processing methods developed.

In the UK, there is a varied performance across Objective 1 regions, with good progress in South Yorkshire and the Highlands and Islands, though in the latter case expenditure levels remain low because of problems in encouraging final beneficiaries to take advantage of the scheme. In Merseyside, the evaluators recommend increased effort towards innovation and technology transfer. In West Wales and the Valleys, difficulties are being experienced in implementing ICT as a horizontal theme.

In Objective 2 regions the evaluations in general find good progress being made with fewer difficulties than those experienced in Objective 1 regions. Slow take-up of measures is evident in France, Italy and the UK, but a more positive picture is evident in Austria, Germany, Denmark, Spain, Finland and Sweden.

Some of the slow progress can be attributed to the global economic slow-down and the consequent lack of businesses wishing to participate in the available schemes. Many of the measures proposed are, appropriately, demand-led and this has led to activity levels below those forecast. Managing authorities and monitoring committees clearly need to devote some consideration to how to promote a more active involvement of the business sector in these kind of activities in their regions.

A particular issue which merits further consideration is how information society, e-Europe and ICT in general function as horizontal themes. In many programmes, this issue has been included as a horizontal theme, but its impact to date seems limited and there are evident difficulties in its implementation. In France, the impact of the programmes on ICT is found to be weak and in Aquitaine the evaluation recommends the development of a regional strategy, including the creation of a specific measure and a focus on the development of ICT infrastructure. In the UK some of the evaluations suggest that more resources should be targeted towards making the cross-cutting priorities more “real”; this would include the development of working groups and the incorporation and monitoring of the issue as a project selection criterion.

5.2.2.2. Human Capital

The development of human capital in a knowledge based economy was not explicitly included in many of the programmes and therefore does not feature in the evaluations. In mono-fund Objective 2 programmes (e.g., Italy, Netherlands), there was a particular absence of this issue. The East Netherlands evaluation found that the nationally based Objective 3 programme limited the potential for complementarity with the regional programme. In the UK Objective 2 programmes, the focus was more on social inclusion (see below) than human capital development in the knowledge based economy.

Science, Technology and Information Objective 1 – Portugal – Good Progress being made

C'est l'axe "former et qualifier" qu'avec une dotation financière de €371M (€347M pour la mesure 1 et €24M pour la 2) couvre cette priorité. Les objectifs de cet axe sont: promouvoir la formation avancée de ressources humaines, en particulier pour la formation post-grade en visant une convergence rapide pour les valeurs moyens des qualifications avancées en ressources humaines que se vérifient dans l'UE, et stimuler l'insertion des ressources humaines au niveau de post grade dans les entreprises et instituions de I&D

La mesure « formation avancée » a financé jusqu'au 31/12/2003 6868 bourses pour :

- effectuer le doctorat (4464)
- le post-doctorat (972) et
- autres types de bourse pour une licence sabbatique, pour développer une carrière scientifique, stages en organisations scientifiques et/ou technologiques des scientifiques invités

La mesure «appui à l'insertion des docteurs et professeurs dans les entreprises et institutions de I&D» prévoyait l'insertion de 1500 chercheurs dans les entreprises à la fin de 2006

Well integrated programmes with human capital investments complementing those discussed above were not general. Where they did exist, the performance of human capital investments tended to mirror that of the research, innovation and ICT measures, with significant delays evident in some cases. There was more potential for such integrated regional development programmes in Objective 1 regions.

5.2.3. Entrepreneurship and Financial Instruments –Performing well but affected by the global economic downturn

Measures to promote entrepreneurship have been at the heart of regional policy for many years and it would be expected that the performance of these measures in terms of implementation should be relatively good. This is the case in general for the evaluations analysed and some recommend increasing allocations to this area. The slowdown in the global economy is cited in some cases as reducing demand for the support provided under the programmes. Serious delays in getting innovative measures under this heading up and running are evident in Greece and parts of Italy (e.g. Puglia). In these cases, evaluators often recommend a more proactive approach to identify businesses or organisations which can assist or benefit from this priority.

The main types of actions supported include the following:

- Direct grant aid to start-up, small, innovative or large enterprises,
- The development of business parks,
- Consultancy support with strategy development, marketing, export, etc.,
- The creation of special venture funds.

It is too early to measure the impact of the entrepreneurship and financial engineering measures of the Structural Fund programmes. However, the evaluators in Mecklenburg-Vorpommern (Germany Objective 1) have already tracked the survival rates of businesses supported. In relation to venture capital funds, these have been set up in a variety of countries, although in the UK and Germany there is a suggestion in some cases of saturation of the market and close monitoring will be needed to ensure that there is a real need for these interventions. The German CSF evaluation found a "jungle" of SME support measures with approximately 500 in existence, clearly suggesting a need for some rationalisation.

> *East Finland Objective 1 – Entrepreneurship measures a great success*
>
> One of the most successful elements in Eastern Finland Objective 1 Programme has been the creation of new entrepreneurship and innovation as well as the development of business activity. By June 2003, 3,444 projects and 2,532 new enterprises had been started.
>
> Finnvera loans are a new financing instrument in the Finnish Structural Fund programmes and the evaluators consider them very important in achieving the targets set for new enterprises and jobs. The loans granted during the first four years of the programme have contributed to the creation of some 500 new businesses.
>
> The importance of the programme to the competitiveness of the region is undeniable. In all areas, development has not been totally satisfactory, however. Within the region covered by Eastern Finland Objective 1 Programme, rural SME trends have been negative especially in remote areas. The evaluators note that the trend is a signal not only for the future, but also for the current programme period. It should therefore be explored whether the programme framework would permit the targeting of SME support measures specifically to such areas of decline.

> *North East of England Objective 2 – UK – A new model for business support*
>
> Priorities 1 and 2 provide support for new and existing businesses in the North East. The urgent need to invigorate the region's business base is reflected in the allocation of nearly 70% of the Programme's EU resources to these two Priorities.
>
> A new, customer-focused model for business support was due to be introduced in April 2004. The Business Brokerage Model will place purchasing power in the hands of businesses, provide them with the information they require in order to become discerning purchasers of business support, and improve knowledge gathering and market penetration on the part of the 'broker' organisation by releasing them from the delivery of business support services.
>
> The introduction of the business brokerage model will have profound implications for the delivery of Priorities 1 and 2. Business support activity in the region will move away from 'projects' to be tailored to the needs of individual businesses. This is likely to affect the draw down of committed EU funds and the gathering of monitoring and evaluation data. Area Business Brokerage Partnerships are being finalised and the role of business development agencies will become clear over the coming months. The scale of resources required under the brokerage model is still unclear, as is the extent to which the model will address all or part of the objectives/actions contained in Priorities 1 and 2.

5.2.4. Social Inclusion – Integrated actions are most effective but are resource intensive

In total 72% of the evaluations examined were of programmes which had a social inclusion dimension. Several Objective 2 programmes did not incorporate a specific social inclusion focus which tended to be handled separately in these cases through nationally organised Objective 3 programmes. A variety of kinds of actions were supported, including:

- ESF supported measures to support the integration of people with disabilities, immigrants, disadvantaged women, disadvantaged youth and ex prisoners. These include training interventions and integrated packages of support designed to meet the needs of the individual.
- Community development, including locally based projects which provide childcare, crime prevention initiatives and meeting points for groups with special needs,

• Physical improvement of urban neighbourhoods in decay.

Approximately two thirds of evaluations examined suggested that performance was satisfactory or above target. A good example of the kind of intervention which evaluators found to work well is the “pathways” approach of Merseyside Objective 1 in the UK. By combining actions for employment with social inclusion and lifelong learning, a comprehensive and integrated policy response is provided to the economic disparities in the region. The evaluators conclude that the initiative is performing well and they stress the often innovative and integrative nature of activities. A further example of a range of measures provided in an Objective 2 programme is provided by Catalonia in Spain.

For the remaining one third of evaluations examined, these measures demonstrated delays or a lack of demand leading to performance below expectations. In France, both in Objective 1 and 2 regions delays in implementation are evident, with recommendations for reallocations of resources in Guyane and Martinique. In Finland, the evaluators examined this issue closely. For East Finland, the evaluators are critical that the programme has not met needs in relation to the prevention of social exclusion. They recommend a greater effort by regional authorities to map potential project operators who could contribute pro-actively to the struggle against social exclusion. The role of mediating organisations needs to be further developed and strengthened. In the West Finland Objective 2 programme, the evaluators consider that too few projects have started in this area. Local groups and third sector organisations have not been able to start their own projects due to a lack of financing and therefore most projects implemented have been sponsored by the public sector. The Swedish Vastra Objective 2 programme evaluation also makes strong recommendations in this regard.

Cataluna Objective 2 – Spain – A range of measures for social inclusion

Although social inclusion was not a priority theme in the Catalan programme, several measures respond to the challenge of social inclusion under the Local and Urban Development Priority (5), and one measure under “Improvement of competitiveness and development of productive sector” Priority 1 (measure 1.8 “Strengthening of stability and adaptability in employment”). Under Priority 1, 44,800 persons, including nearly 800 immigrants benefited from employment projects.

Among other actions under Priority 5 were construction of social infrastructure, encouraging and supporting local development initiatives, creating centres of professional training and local development, as well as supporting local initiatives aiming at employment creation. Around 4,000 people annually used the social infrastructure in 2000-2002.

In the UK in the Yorkshire and Humber region, the evaluators note the difficulty of mapping Objective 2 activities onto the National Action Plan priorities which makes it difficult to comment on how well the programme is performing.

Vastra Objective 2 – Sweden – Specific recommendations to enhance social *inclusion focus*

As regards equal opportunities, the issue is seen as somewhat marginal and not decisive in selecting projects. The evaluators recommend a systematic analysis of the consequences a project has in terms of equal opportunities and that equal opportunity descriptions in larger projects should be followed up and evaluated.

As regards integration of immigrants, the evaluators found a lack of support for this issue. It is widely seen as marginal and does not receive adequate human or financial resource inputs. The evaluators recommend training for responsible desk officers and other decision makers.

The evaluation also recommends the more active development of indicators for social integration. The possibility of projects choosing a neutral or negative impact, i.e., "the project is neutral as regards integration" or "the project has a negative influence on integration" should be removed and the indicator system should be changed accordingly. Exactly what is meant by "this project improves the integration situation" should be clearly defined. The evaluators recommend that care should be taken to ensure the approach is facilitative rather than imposed in order to ensure that support for the concept is obtained.

5.2.5. Environment and Sustainable Development – more consideration needed of the practical implementation of horizontal priorities

The environment was a horizontal priority in nearly all programmes and nearly all mid term evaluations examined its implementation and potential impact, except the evaluations of some sectoral programmes (e.g., information society or human resources development) which do not address the issue. The concept of sustainable development was not fully developed at the time of the agreement of the 2000-2006 programmes; therefore, in most cases the evaluations examine the extent to which environmental objectives have been achieved.

As well as being a horizontal priority, the environment is the focus of direct action in many Structural Fund programmes, with measures concerned with a range of actions, including:

- Revitalisation of industrial sites
- The development of renewable energy sources
- Fresh water supply
- Sewage treatment facilities
- Farm waste initiatives
- Air pollution control and emission reduction
- Development of protected areas which can include soil regeneration measures, the reintroduction of plant and animal species, regeneration of beaches, etc.
- Renovation and maintenance of cultural heritage.

There is varied performance across programmes in terms of the implementation of these measures. In general, a trend can be identified of more "standard" environmental projects proceeding according to plan, but delays being experienced in the more innovative types of measures, particularly in relation to renewable energy.

Assessing the performance of the environment as a horizontal priority is a more complex exercise. Where there is a strong direct environmental element to programmes, they tend to place less emphasis on environment as a horizontal priority. Many of the evaluations present the experiences of programmes in assessing if projects have a positive, negative or neutral impact on the environment. Where this approach is used, the percentage of projects with a positive impact on the environment seems to be at levels between 20% and 35%. Some of the evaluations go further, however, and provide interesting analysis on the reality of the extent to which environment is effectively implemented as a horizontal priority. They note confusion in definitions of what a “positive impact on the environment” actually means. In some cases the issue is not incorporated into project selection or is incorporated too late, while in others there is only a requirement for the project to state whether the project has a positive, negative or neutral impact without any requirement for further details. The deeper level of analysis of this issue is particularly notable in evaluations of programmes in Finland, Sweden and the UK.

In the UK, the Merseyside Objective 1 evaluator notes that the environment priority as well as the other two cross cutting themes were not cascaded down to priority or measure level and built into commitments required of projects. In South Yorkshire, also Objective 1, the evaluator finds that the horizontal priority is well embedded into all priorities, but the primary research yielded more complex results. While project managers gave low scores for project contribution to the environment and relevance to South Yorkshire, almost two thirds of all stakeholders interviewed reported that it remained a “completely relevant” theme. The evaluators suggest two reasons for this divergence: there is a difference between strategic intent and delivery; and there is also a difficulty in translating environmental sustainability into tangible projects. In other Objective 1 programmes in the UK widely differing interpretations and understandings of the issue were found. For UK Objective 2 regions, while the evaluations were in general less critical of the implementation of the environment as a horizontal priority, issues of interpretation were also raised.

In Swedish Objective 1 and 2 regions, the evaluators found that environment was poorly integrated into the programme because it was introduced too late in the project application process. Recommendations are made for the development and dissemination of better guidance and the inclusion of the issue at an early stage in project selection so that it becomes a genuine project selection criterion. From the evidence of many of the evaluations, it seems that while the issue may be often included in the project application form, it is not actually taken into account when decisions are being made on the projects to fund.

The concept of sustainable development has been developed only recently and its translation into operational reality through Structural Fund programmes needs to be considered further, based on the findings of the mid term evaluations. For future programmes, it will be important to develop a shared understanding of how the sustainable development concept can add value to their implementation.

6. USING THE EVALUATIONS

The purpose of the mid term evaluation was to provide knowledge and analysis on the performance of programmes, with a view to improving their implementation and impact. The final stage in that process is therefore not the production of the report, but its use by managing authorities, monitoring committees, the Commission and indeed the wider public (including the political level).

The mid term review is the formal decision making process to adapt programmes, taking account of the analysis provided by the mid term evaluations. There are other means of using the evaluations, however, including the adaptation of implementation systems to make them more effective (which need not necessarily be reflected in the mid term review decision), disseminating information about the Structural Funds as part of a communication strategy and feeding debate on future directions for structural policies. An important role of all evaluations is to generate knowledge on the performance of public policies which can lead to better informed public debate.

The Commission is not and cannot be fully aware of the use made of evaluations, given that they were organised, managed and primarily used in the Member States. In order to complete the analysis of the use of the mid term evaluation, the Commission invited the Member States to provide their perspectives on the use and the usefulness of the mid term evaluations. This chapter is based on contributions from the Member States.

6.1. The Users of the Mid Term Evaluations

The main users of the evaluation were managing authorities, monitoring committees and implementing bodies. The broad range of interests represented on the monitoring committees, including social partners and various other interest groups, ensured a widespread awareness of the evaluation, at least among those involved in the programmes.

The primary way in which the mid term evaluations were used was in feeding into the mid term review. The fact that the mid term evaluation and mid term review had a legal basis, enshrined the requirement to use the evaluation and to consider its findings. In all cases, evaluation results were presented to the monitoring committee. This was followed by detailed proposals for changes to programmes on the basis of mid term evaluation findings which were prepared by the managing authority and considered by the monitoring committee before revised programmes were submitted to the Commission. Explicit linking of proposed changes to evaluation results was the practice in many cases (e.g., France, Portugal). In Greece some social partner organisations submitted proposals for changes to programmes based on the results of th evaluations. Evaluators had longer term contracts in Greece which did not end with the production of the mid term evaluation and evaluators participated in technical meetings and thus contributed to the formulation of proposals for the mid term review.

One of the important contributions of the mid term evaluation was that it facilitated consideration of other than purely financial issues. This was notable in Italy where the monitoring committees had to focus on what the programmes were actually achieving rather than issues of project selection or financial absorption. Even where the evaluations may not have been of the highest quality, they played a role in contributing to more strategic debate at the monitoring committees.

In Northern Ireland the evaluation of the PEACE programme combined the ex post evaluation of PEACE I which operated from 1995 to 1999 with the mid term evaluation of PEACE II. The evaluation was the central feature of a major public consultation exercise on the extension of PEACE II which was originally planned to end by the end of 2004. The evaluation helped to provide the basis for the extension of the programme.

In Finland, the evaluators arranged consultation seminars in the regions before the evaluations were finally completed and results, conclusions and recommendations were presented and discussed. In some seminars the media participated. This created interaction between the evaluators and those involved in implementing the programmes and enhanced both the quality of the reports and the process of consideration of evaluation findings.

6.2. Amended Financial Allocations

In some cases (e.g., Belgium, Finland, Sweden) changes to financial allocations did not occur, but in most cases there were changes. In Sweden, by the time the mid term evaluations were complete most of the resources had already been allocated. Elsewhere, the mid term review represented an important occasion to adapt programme structures to take account of measures which were under or over performing financially. Sometimes the decisions on re-allocations were influenced by broader policy considerations (e.g., Ireland) and financial considerations relating to progress in meeting expenditure targets were a determining factor for amended financial allocations in general throughout the EU.

6.3. Changes to Implementation Systems

In the majority of Member States, the mid term evaluations led to improvements in the system of indicators, streamlining them in many cases and sometimes suggesting new and more relevant indicators. A focus on the development of indicators for horizontal priorities was evident in France, Greece, Ireland, Sweden and Spain, among others. In Portugal, specific structures, “strategic integration platforms” are to be put in place to give more coherence and strategic direction to the implementation of horizontal priorities.

The other major area for improvement on the basis of the mid term evaluations was the project selection process, including the appraisal system. Adjustments were made in Finland, Italy, Northern Ireland and Scotland, among others.

Improving implementation issues was also important in the case of the PEACE programme where the changes were needed in the context of the extension of the programme.

Changes to the Implementation System in Wales

- The MTE reported little contact with projects once they have been given approval to proceed, The opportunity for learning from "best practice" is lost and problems with individual projects are not identified and addressed early enough. As a consequence WEFO established a separate team to mentor and support projects.
- The Objective 1 application form was changed to make it easier for applicants to present the information required
- Partnership structures were changed to streamline and simplify the process of considering applications
- A new process of considering the cross cutting themes was introduced, based on a risk-assessment. This provides a more robust system of identifying projects for full consideration by the Cross Cutting Themes Team

Sweden: An emphasis on identifying Strategic Projects and better implementation of the Horizontal Priorities

The Swedish evaluations recommended changes in two areas:

- Increased emphasis on **strategic projects**: In order to make the most of the remaining resources in the programme, activities should be carefully directed towards projects that show most promise in developing dynamic effects, which is more important than the importance of immediate results. The evaluators recommend how to identify such dynamic projects
- Working groups on **horizontal themes** have been set up to respond to the criticisms of the mid term evaluations and already activities are underway, including training courses for managing authorities and programme secretariats about the horizontal themes (the environment, gender equality and ethnic integration. 13 one-day courses in 5 locations have been run, attended by approximately 400 people. A national conference is planned for 2005 on integrating horizontal themes. An inventory of examples of good practice and working models and tools for integrating the horizontal themes is being compiled and work is underway on improving indicators.

6.4. Publication of Reports

In most Member States the mid term evaluations are freely available on web-sites. Where they are not, copies of the reports can be obtained from the managing authorities.

Dissemination of evaluation reports to the wider public can have a number of benefits. The UK authorities point out that availability of evaluations can help raise awareness of programmes and alert potential beneficiaries to the possibilities offered by the Structural Funds. The publication of reports can also contribute to public debate. Public debate on evaluation findings in parliaments, newspapers, etc., was evident in Finland, Greece, Ireland, Northern Ireland.

6.5. Further Actions Planned

In general further use of the mid term evaluations beyond the mid term review is not planned, although of course the knowledge gained from the evaluations will continue to influence the implementation of programmes and thinking for the future. However, the results of the mid term evaluations will be followed up by a number of specific ongoing evaluations in some Member States (e.g., Ireland, Austria, Spain, Sweden). While in some cases (e.g., Wales) recommendations could not be implemented within the timeframe and resources available for the current programming period, the learning gained will be reflected in proposals for the future. The mid term evaluation update will, of course, examine developments in relation to the key recommendations of the evaluations.

6.6. Factors militating against Use of the Mid Term Evaluations

The factors identified by the Member States which hindered the use of the evaluations are concerned both with design features of the entire evaluation process and weaknesses in the evaluations themselves.

The design of the mid term evaluations sometimes made it difficult to use. We have already noted the very broad scope of the mid term evaluation exercise which sometimes contributed both to the length of evaluations and a superficial approach both of which hampered use. In addition, while the partnership was a strength of the evaluation, gaining agreement among the partners on change to the programme was sometimes challenging. Some Member States also found the timescale extremely tight from completion of mid term evaluation to submission of mid term review proposals and this prevented systematic consideration of evaluation findings. The short time between production of the evaluation and the finalisation of mid term review proposals meant that “the mid term review swallowed without necessarily digesting the many interesting things that did come out of evaluations”.

Among the weaknesses in the evaluations were too many recommendations and/or vague recommendations which were insufficiently precise and concrete. The extreme length of some evaluations was also cited in a number of cases as militating against use.

7. CONCLUSIONS

In this final section, we draw conclusions on the process and content of the mid term evaluation and suggest how the lessons learnt can be reflected in the future.

7.1. The Process

Two aspects of the process for the mid term evaluation can be identified as particular strengths: the quality of planning and the positive contribution of partnership. As regards planning, the early availability of the Commission's guidance for the mid term evaluation was a good start which was built on in nearly all Member States. Good organisation and co-ordination across different evaluations in the same Member States was a general feature. Some Member States drew up core terms of reference which provided a good basis for the evaluations. The only weakness of this approach in some cases was that terms of reference were not sufficiently adapted to regional and sectoral needs by managing authorities and steering groups for individual programme evaluations.

The mid term evaluation was carried out in partnership which contributed to the relevance and quality of the evaluations. The role of the steering committees was particularly positive in involving the partnership throughout the process. The Commission participated actively in steering groups and in reviewing intermediate outputs of the evaluation process.

7.2. The Evaluators

The evaluation market is relatively limited and the mid term evaluation did not see a significant expansion of the number of evaluators. Evaluations were mostly undertaken by evaluation consultancy companies, with some involvement of academics and growing evidence of evaluations being undertaken by consortia of evaluators. There is some evidence that the involvement of universities and particular evaluation experts increased the quality of the evaluations with more rigorous and innovative methodologies used.

There is no doubt that the single legal deadline for the submission of all mid term evaluations of all Structural Fund programmes created pressure on the evaluation market and some evaluators who took on a number of evaluations were over-stretched. This underlines the importance of close monitoring of the evaluation process by Managing Authorities and the need to insist on high quality analytical work from the evaluators. For the evaluators, there is a need in some cases for closer attention to be paid to meeting professional standards for evaluation in relation to data gathering, analysis and presentation of results.

7.3. Quality of the Evaluations

The mid term evaluations were generally of a higher quality than those undertaken in the past with some excellent examples. The Commission's examination of the quality of the evaluations resulted in two thirds being rated as being of good or excellent quality. This increase in quality results from a growing capacity among evaluators but also among Managing Authorities. Managing Authorities allocated more resources to the evaluation than had been the case in previous evaluation exercises. This was both in terms of financial resources and the human resources involved in managing and steering the evaluation process.

Limitations of the evaluations usually related to the scale of the task, with a large number of evaluation questions to be answered and some very large complex programmes being evaluated. In other cases, the weaker evaluations related to evaluators being over-stretched in undertaking too many evaluation projects and not devoting sufficient resources to all projects. Methodological weaknesses, discussed below, also limited the quality of some evaluations.

7.4. Cost and Cost Effectiveness

Over €35 million was spent on the Objective 1 and 2 mid term evaluations with an average cost per evaluation of €172,000. The cost of the evaluations represented on average 0.08% of total annual allocations. The ultimate cost effectiveness depends on the usefulness of the evaluations in improving the quality and relevance of programmes. The Commission's perception is that the resources allocated to the mid term evaluation were appropriate and higher than previous allocations, but that cost effectiveness was limited by a number of factors, including the late or slow start of programmes, the broad nature of the evaluation and methodological weaknesses.

7.5. Methodologies

The mid term evaluations in general were based on an appropriate mix of methodologies, including desk research, primary research and for larger programmes some macro-economic modelling. Significant improvements were evident in the quality and rigour of the evaluations. A difficulty for many of the evaluations was that the monitoring system was not yet functioning well and there was limited administrative data on the implementation of the programme. As regards primary research, this included consultations, surveys, focus groups and case studies.

Structural Fund evaluations are complex exercises, especially those which cover the entire programme and evaluators should be encouraged to develop and improve appropriate methodologies. Methodological concerns arising from the mid term evaluations which need to be addressed in future evaluations concern the gathering of primary data, analysis and the presentation of findings:

- As regards primary research, evaluations should be informed by the "on-the-ground" reality of the implementation of the programme. While the scale of some programmes might make this difficult across all measures, at least the major priorities should be subject of fieldwork.
- In relation to analysis, some evaluations contained extensive descriptive material rather than more synthetic analysis. Evaluators should aim to analyse synthetically the performance of the overall programme at a strategic level, bringing together bottom up administrative (monitoring) data or primary research and analysis of changes in the global environment.
- Many reports were extremely long, often because of the extensive descriptive material. This militated against their usefulness.

Evaluators need to develop their capacities in these areas, while those who commission evaluations must demand high standards from evaluators. The Commission has already made these points to evaluators at a number of meetings of evaluation societies in the European Union, challenging evaluators to improve the quality and relevance of their work, building on the achievements which have been made to date. Member States also need to insist on high quality, working in partnership with evaluators to ensure that data collection systems are functioning and that evaluations are appropriately designed.

7.6. Answers to the Evaluation Questions

The evaluations aimed to answer a number of evaluation questions and we can synthesise the answers as follows:

- In nearly all cases the evaluators concluded that the strategies adopted in the programmes were still appropriate and only minor changes were in general proposed.
- It was too early to assess effectiveness in many cases, given the late or slow start-up of programmes. A difficulty was presented by the monitoring system which often required revision of the definition of indicators and their targets. The mid term evaluations have performed a useful function in reviewing the basis on which effectiveness will be judged, which should allow a robust basis for assessing results initially at the time of the mid term evaluation update and then in the ex post evaluation.
- It was likewise too early to measure impact, although evaluations did in some cases recommend the revision of targets based on the experience of the early years implementation of the programme. Macro economic modelling undertaken for large CSFs suggests that programmes are on course to achieve their global objectives.
- Efficiency was more often defined as financial absorption rather than the cost per output or result achieved. Financial absorption is clearly improved compared to the past, driven by the requirements of the N+2 rule. Some evaluations began an analysis of unit costs which can be built on in the future.
- Evaluations found an urgent need for improvement of the systems of quantification of objectives. The recommendations in this regard should contribute to a good basis for assessing the impact of programmes in the mid term evaluation update and the ex post evaluation.
- Implementation systems were found to fulfil the regulatory requirements and to be significantly improved on the past. In some cases complexities were identified and recommendations made to streamline systems. Recommendations were also made to refine project selection systems to make them more competitive and to incorporate the horizontal priorities in selection criteria.

7.7. Contribution to Lisbon Priorities

- As regards **transport and accessibility**, the evaluations find good progress being made in large road and rail projects. Projects to develop ports, airports and local transport initiatives are also progressing but often at a slower rate.
- The **knowledge based economy** is an important priority and the Structural Funds are supporting co-operation between research institutes and businesses, the development of business clusters and research centres, investment in broad-band access, the development of regional innovation strategies and the training of researchers, as well as applied research projects. Many innovative projects and approaches are being supported with some taking longer to establish than originally anticipated.

- The Structural Funds are supporting **entrepreneurship** through grant aid to start-up, small, innovative and large enterprises, the development of business parks, consultancy support and the creation of special venture funds. Many of these interventions are demand led and slower progress than anticipated is linked to the global economic slow-down in the early years of the programme.
- The promotion of **social inclusion** is an important dimension in the majority of Objective 1 and 2 programmes and the majority of evaluations find that performance is satisfactory or above target. Integrated actions are found to be the most effective, although these are resource intensive.
- The **environment** was a horizontal priority for the Structural Funds and the evaluations provide important insights into how to enhance the practical integration of the priority into the programmes and how the concept of **sustainable development** can be given practical reality in the programmes. The Structural Funds are also supporting direct environmental actions including the revitalisation of industrial sites, the development of renewable energy sources and the development of facilities for fresh water supply or waste disposal, as well as the development of protected areas. In general, the more traditional types of projects are proceeding according to plan with the more innovative measures taking longer to get off the ground.

7.8. Use of the Evaluations

Discussion with Member State representatives confirms that the primary use of the mid term evaluation was in relation to the mid term review and adaptation of programmes. The evaluations were used mainly by managing authorities, monitoring committees and implementing bodies. Generally, the evaluations did not stimulate public debate, although there were some exceptions mainly in Objective 1 regions where the Structural Funds have a higher profile.

Changes to financial allocations in programmes were generally driven by absorption concerns, although the evaluations did contribute to decisions in this regard. The evaluations generally had a strong impact in improving implementation systems, particularly the further development of systems of indicators, implementing horizontal priorities and improving selection criteria.

We can identify two key factors which influenced positively the use of the mid term evaluations:

- The quality of the evaluations – higher quality evaluations that had rigorous methodologies and clear presentation of results and that made precise recommendations were more likely to be used;
- The engagement of the managing authority and the monitoring committee – where the managing authority and monitoring committee were interested in the outcome of the evaluation and participated throughout the process in providing information and challenging evaluators while respecting their independence, the evaluations tended to be most useful.

7.9. The Mid Term Evaluation – Strengths and Weaknesses

The mid term evaluation exercise demonstrates growing evaluation capacity among Member States. While much improved compared to previous evaluations, there is still scope to further improve Structural Fund evaluation. As an overall conclusion, we can identify the following strengths of the mid term evaluation:

- Its organisation and planning,
- The contribution of the partnership,
- The allocation of resources – both financial and human to manage the process,
- Increased evaluation capacity in the market and in the administrations leading to higher quality evaluations.

The weaknesses identified in this report are:

- The rigid deadline leading to the evaluation being too early in some cases and creating pressure on the evaluation market,
- The requirements of the evaluation being too broad, covering all measures and eight different and complex evaluation questions, ranging from strategy to effectiveness to implementation,
- Evaluation terms of reference sometimes not sufficiently adapted to the needs of regions or sectors,
- Some evaluators not delivering sufficiently high quality work, with Managing Authorities not managing the process closely enough and demanding better quality.

In terms of proposals for the future, the Commission has proposed that the mid term evaluation should be replaced in the next programming period with a programme of ongoing evaluation of all or parts of the programmes depending on needs. This would allow the strengths of the mid term evaluation process to be built upon, but would address the weaknesses related to rigidities and scope.

7.10. Update of the Mid Term Evaluation

An update of the mid term evaluation is required by the end of 2005 and the Commission has recently produced its guidance on the content of the update. The update will be produced at an important time, just before planning starts for the next programme period, and it will also be an important preparation for the ex post evaluation. The Commission has proposed that the core element of the update should be an analysis of outputs and results achieved to date, but that there should be scope for particular evaluation questions to be addressed, depending on regional or sectoral needs.

7.11. Challenges for the Future

Having analysed the processes, methodologies, content and use of the mid term evaluations and concluded on the strengths and weaknesses of the exercise, we can in this final section identify a number of challenges for the future for evaluation and the Structural Funds.

7.11.1. Further Strengthening Evaluation Capacity

The Member States which undertook the mid term evaluation demonstrate a growing evaluation capacity which has led to the production of better quality reports. However, performance was variable and there is a need to further strengthen capacity – both among those who commission evaluation and those who carry them out:

- From the mid term evaluation exercise, we can see in general higher and appropriate financial allocations to evaluation. However, in some cases there is a need to increase the human resources devoted towards managing evaluation and to ensure that they are equipped with sufficient evaluation knowledge in order, where necessary, to challenge evaluators to improve the quality of their work.
- As regards the evaluators, increased efforts to develop appropriate methodologies are needed to ensure accountability and to provide greater insights into the performance of programmes. Proposals for more targeted evaluations in the future should permit the development of methodologies specific to the issues being evaluated.

Evaluation capacity is at an early stage of development in many of the Member States which have recently joined the EU. Here, the challenge is to build on the experiences of other Member States and to take action to build capacity both within the administrations and in the marketplace of evaluators.

7.11.2. Improving Monitoring Systems

There is a clear need to improve monitoring systems so that they become a genuine management instrument – providing timely and meaningful physical data on the implementation of programmes and contributing to better decision making and higher quality programmes. Good monitoring systems are essential for accountability, in order to allow regional and national authorities and the European Commission to account for and demonstrate the added value of the very considerable amount of resources being invested in Structural Fund programmes.

The Commission emphasised the issue of indicators in the drawing up of programmes in 2000 and 2001. It was also an important point in the negotiations on the programmes for the Member States which joined the EU in 2004. This emphasis led to the development of some large scale but unwieldy monitoring systems which in some cases three years after programme adoption did not provide useful data for decision makers or for evaluators. There is a need for monitoring systems to be refined to concentrate on core indicators and to ensure that data are gathered regularly and considered by decision makers and are available to evaluators.

7.11.3. Creating a Platform for the Exchange of Experience

With enlargement of the EU, there is now a wide range of evaluation capacity across the Member States and there is potential for that capacity to be enhanced through the exchange of experience. Discussions with those responsible for evaluation in national administrations suggest a willingness to engage in such activities. The Commission will take action to encourage such networking and proposes in the next year to develop the web-based GUIDE to the evaluation of socio-economic development (www.evalsed.info) into a truly interactive site which can facilitate such activities.