



**EUROPEAN  
COMMISSION**

# **DG MARKT Guide to Evaluating Legislation**

## **ANNEXES**

Brussels  
March 2008

## **Table of contents**

<b>Annex 1 - Internal Control Standard 14 and the Evaluation Standards .....</b>	<b>64</b>
<b>Annex 2 - Evaluation Charter in DG MARKT .....</b>	<b>69</b>
<b>Annex 3 - Evaluation Background and Theory.....</b>	<b>73</b>
Background .....	73
Evaluation Theory .....	74
1. <i>What is an evaluation of legislation?</i> .....	74
2. <i>Why is evaluation theory different for legislation?</i> .....	74
3. <i>Where to find the relevant information for the intervention logic</i> .....	77
4. <i>The role of impact assessment and evaluation in Activity Based Management</i>	77
<b>Annex 4 - Template of DG MARKT evaluation mandate.....</b>	<b>79</b>
<b>Annex 5 - Case Study for evaluating legislation: example of the evaluation mandate for Services Directive .....</b>	<b>81</b>
<b>Annex 6 - Quality Assessment Form .....</b>	<b>87</b>
1.    Guide for organising the quality assessment.....	87
1.1 <i>What are the main aims, stages and actors involved in the QA process?</i> .....	87
1.2 <i>How to fill in the quality assessment form?</i> .....	88
1.3 <i>How to do the scoring?</i> .....	88
1.4 <i>Publishing the quality assessment report?</i> .....	88
2.    Quality Assessment Form .....	90
3.    Guide on scoring the criteria .....	94
<b>Annex 7 - Concept/Issue Mapping.....</b>	<b>98</b>
1.    Who should / can do it ?.....	98
2.    How best to use this tool ? .....	98
<i>Step 1 : List relevant issues / impacts, etc.</i> .....	98
<i>Step 2 : Arranging relevant issues / impacts into groupings</i> .....	100
<i>Step 3 : Plotting groupings into similarity matrix</i> .....	101
<i>Step 4 : Merging individual similarity matrices into a single matrix</i> .....	101
<i>Step 5 : Translating the matrix into a 'point map' to select the clusters of impacts ..</i>	103
<i>Step 6 : Rating impacts and calculating average scores for clusters</i> .....	104
<b>Annex 8 - Social Network Analysis .....</b>	<b>106</b>
<b>Annex 9 - Checklist for dissemination strategy .....</b>	<b>111</b>
<b>Annex 10 - Action Plan .....</b>	<b>112</b>

**Figures**

Figure 1 : Evaluation Charter - work flow of key evaluation stages and actors..... 72

Figure 2 : Attempt to use a 'linear evaluation model' for legislation..... 75

Figure 3 : Cause-effect logics of implementing legislation..... 76

Figure 4 : Impact assessment vs. evaluation in a broader context..... 78

Figure 5 : Example of issue listing ..... 99

Figure 6 : Example of initial response grouping ..... 100

Figure 7 : Example of an "individual similarity matrix" reflecting issues plotted together ..... 101

Figure 8 : Example of a" group similarity matrix" – reflects views of a 3 person evaluation team 102

Figure 9 : Example of a point map ..... 103

Figure 10 : Example of a cluster map..... 104

Figure 11 : Example of point rating map..... 104

Figure 12 : Example of a cluster analysis..... 105

Figure 13 : The causal relationships examined by social network analysis ..... 106

Figure 14 : Example of why the MS action is extracted in the intervention logic ..... 106

Figure 15 : Example of why the MS action is extracted in the intervention logic even when clustered ..... 107

Figure 16 : Example of social network analysis illustrating how MS action has affected EU measures..... 109

Figure 17 : Example: applying the social network analysis back to an intervention logic ..... 110

**Tables**

Table 1 : Differences between an impact assessment and evaluation ..... 78

Table 2 : Example of open-ended question ..... 99

Table 3 : Average rating per cluster ..... 105

Table 4 : Measure by Member State matrix for social network analysis ..... 108

## Annex 1 - Internal Control Standard 14 and the Evaluation Standards

### Extract from SEC(2007)1341 - Revision of the Internal Control Standards and Underlying Framework

**ICS 14. Evaluation of Activities:** Evaluations of expenditure programmes, legislation and other non-spending activities are performed to assess the results, impacts and needs that these activities aim to achieve and satisfy.

#### **REQUIREMENT**

- Evaluations are performed in accordance with the guiding principles of the Commission's evaluation standards. Corresponding evaluation baseline requirements are applied for retrospective evaluations (interim, final and ex-post) while prospective evaluations (ex-ante and impact assessments) follow the relevant specific guidelines.

(Main references: Commission communication "Responding to Strategic Needs: Reinforcing the use of evaluation SEC (2007) 213; DG Budget's Guide for ex-ante evaluation; Impact assessment guidelines issued by SG)

#### **TIPS FOR ASSESSING CONTROL EFFECTIVENESS**

*Reminder: When assessing control effectiveness, two fundamental questions should be asked:*

- (1) Considering the DG's specific activities and risks, are the current control arrangements sufficient?*
- (2) Do the control arrangements work as intended in practice?*

Set out below are some questions management may want to consider when assessing control effectiveness with regard to this particular ICS:

- *Are evaluation activities appropriately organised and resourced to meet their purposes?*
- *Are evaluation activities planned in a transparent and consistent way so that relevant evaluation results are available in due time for operational and strategic decision-making and reporting needs?*
- *Does the evaluation design provide clear and specific objectives, and appropriate methods and means for managing the evaluation process and its results?*
- *Do evaluation activities provide reliable, robust and complete results? Are the evaluation reports used by management in practice, i.e. do they have a real impact on the DG's decision-making or the policy and legislative proposals prepared by the DG? If not, why?*
- *Are evaluation results communicated in such a way that they ensure maximum use of the results and that they meet the needs of decision-makers and stakeholders?*

## **EVALUATION STANDARDS**

(ANNEX II of SEC(2007)213)

### **Context and purpose**

Evaluation involves a judgement of interventions according to their results, impacts and needs they aim to satisfy. It is a systematic tool which provides a rigorous evidence base to inform decision-making and contributing to making Commission activities more effective, coherent, useful, relevant and efficient. Evaluation also enhances transparency, learning and accountability. To achieve this, the Commission's evaluation standards aim to ensure relevant and timely evaluations of high quality and that evaluation results are communicated to decision-makers and other relevant stakeholders in a clear and transparent manner to facilitate the use of evaluation results.

In light of the above objectives, the standards are grouped into five categories:

- Resources and organisation of evaluation activities (A1-A3),
- Planning evaluation activities (B1-B5),
- Designing evaluations (C1-C3),
- Conducting evaluations (D1-D5), and
- Dissemination and utilisation of evaluation results (E1-E5).

The standards are expressed as a set of guiding principles. For each guiding principle, a number of baseline requirements (forming an integral part of the standards) have been defined which should contribute to achieving compliance with the overriding principle. Meeting the baseline requirements will hence be important, but not necessarily sufficient, to ensure full compliance with the guiding principles.<sup>1</sup>

The standards are an integral part of the Commission's Internal Control Standard n°23 on evaluation, which means that they are binding and that the way they are implemented may be audited on this basis.

### **Scope**

The standards apply to Commission evaluations of policy instruments such as expenditure programmes, legislation and other non-spending activities.<sup>2</sup> The standards are binding upon all DGs and Services of the Commission with activities that affect entities outside the European institutions (e.g. organisations, companies and citizens).

The standards also apply where a DG performs evaluation of internal policies or service provision. However, additional organisational structures are not necessarily needed in these cases. The key issue is to clarify *who* is responsible for *what* and it is the responsibility of the Director General to consider the most appropriate way of organising evaluation activities in accordance with their needs.

---

<sup>1</sup> The implementation of the baseline requirements will normally need to be complemented by additional measures, such as developing and implementing good practices or the various actions set out in the present Communication.

<sup>2</sup> Separate evaluations of individual projects financed under programmes are not subject to these standards. However, project evaluations required by specific provisions, for example pilot projects, are covered by the standards.

The standards apply to the different temporal types of evaluations. However, whilst the *guiding principles* for designing and conducting evaluations and dissemination and utilisation of evaluation results apply to all types of evaluation, the corresponding *baseline requirements* refer only to retrospective or combinations of retrospective evaluations (interim, final and ex-post). By contrast, purely prospective evaluations (ex-ante and impact assessments) must be carried out in accordance with DG Budget's guide for ex-ante evaluation<sup>3</sup> or the Commission's Impact Assessment Guidelines to ensure adequate quality.

Moreover, the standards apply irrespective of the nature of the author of the evaluation, i.e. to both internal and external evaluations (and combinations thereof).

## A) RESOURCES AND ORGANISATION OF EVALUATION ACTIVITIES

**Evaluation activities must be appropriately organised and resourced to meet their purposes.**

1. Each Directorate General must have an evaluation function with a clearly defined responsibility for co-ordinating and monitoring evaluation activities of the Directorate General (from the planning of evaluations until their dissemination and use), promoting quality of evaluation and organisational learning, and assisting the central services in the implementation of the Commission Evaluation Policy.
2. Each Directorate General must ensure that human and financial resources are clearly identified and proportionately allocated for evaluation activities to be carried out.<sup>4</sup>
3. Each Director General must clearly define the tasks, responsibilities, organisation and procedures for all actors involved in planning, designing and conducting evaluations, and disseminating and using evaluation results.

## B) PLANNING EVALUATION ACTIVITIES

**Evaluation activities must be planned in a transparent and consistent way so that relevant evaluation results are available in due time for operational and strategic decision-making and reporting needs.**

1. An annual evaluation plan and an indicative multi-annual evaluation programme are to be prepared by the evaluation function in consultation with the other units in the Directorate General and integrated in the Annual Management Plan.
2. The multi-annual evaluation programme must be drawn up on the basis of the life cycle of the interventions, the operational and strategic decision-making needs of the Directorate General, general requirements for evaluation, and any specific requirement for evaluation as set out in the legal base of the intervention.
3. All activities addressed to external parties must be periodically evaluated in proportion with the allocated resources and the expected impact.

---

<sup>3</sup> The existing ex-ante guidelines will be updated and developed to be more complementary to the impact assessment guidelines (action 12.1).

<sup>4</sup> Especially in the SPP cycle within the APS and AMP exercises.

4. The timing of evaluations must enable the results to be fed into decisions on the design, renewal, modification or suspension of activities.
5. All relevant services (in particular the evaluation function, SPP/policy planning co-ordinators, IA co-ordinators and key operational units) must contribute to or be consulted on the annual evaluation plan and the indicative multi-annual evaluation programme.

### C) DESIGNING EVALUATIONS

**Evaluation design must provide clear and specific objectives, and appropriate methods and means for managing the evaluation process and its results.**

1. Save in duly justified cases, a steering group must be set up for each evaluation to advise on the terms of reference, support the evaluation work and take part in assessing the quality of the evaluation at the appropriate regularity; its composition must be adjusted to the specific needs and circumstances of each evaluation and the evaluation function must be advised thereon.
2. Terms of reference must be established for each external evaluation and a corresponding document/mandate must be established for each internal evaluation, which must at least specify the following points: purpose and objectives, key questions, scope, expected outputs, deadlines, and quality criteria.<sup>5</sup>
3. Issues of relevance to all services concerned must be considered for the terms of reference.

### D) CONDUCTING EVALUATIONS

**Evaluation activities must be conducted to provide reliable, robust and complete results.**

1. The evaluation must be conducted in such a way that the results are supported by evidence and rigorous analysis.
2. All actors involved in evaluation activities must comply with principles and rules regarding conflict of interest.
3. Evaluators must be free to present their results without compromise or interference, although they should take account of the steering group's comments on evaluation quality and accuracy.
4. The final evaluation reports must as a minimum set out the purpose, context, objectives, questions, information sources, methods used, evidence and conclusions.

---

<sup>5</sup> The evaluation questions should reflect the following evaluation issues whenever relevant: effectiveness, efficiency/cost-effectiveness, relevance, coherence, sustainability, utility and/or community added value, and where relevant the contribution to broader strategic objectives. Additional evaluation issues may also have to be added to the terms of reference.

5. The quality of the evaluation must be assessed on the basis of the pre-established criteria throughout the evaluation process and the quality criteria must as a minimum relate to relevant scope, appropriate methods, reliable data, sound analysis, credible results, valuable conclusions and clarity of the deliverables.

## E) DISSEMINATION AND UTILISATION OF EVALUATION RESULTS

**Evaluation results must be communicated in such a way that it ensures the maximum use of the results and that they meet the needs of decision-makers and stakeholders.**

1. The evaluation results must be examined by the services concerned, who must outline the actions they propose to take towards the formulation, planning and/or revision of the relevant interventions, in accordance with procedures set out by the Director General (cf. standard A1).
2. Evaluation results must be communicated effectively to all relevant decision-makers and other interested stakeholders/parties.
3. The evaluation results must be made publicly available<sup>6</sup> and targeted summary information should be prepared to facilitate communication to the general public.
4. The evaluation function must promote the use of evaluation in decision-making and organisational learning by ensuring that policy implications and lessons learnt from (and across) evaluations are synthesised and disseminated.
5. The use of the evaluation results must be regularly monitored by the evaluation function.

---

<sup>6</sup> Unless a case for confidentiality can be made under one of the exceptions provided for in article 4 of Regulation 1049/2001 of the European Parliament and the Council, 30 May 2001



# Annex 2 - Evaluation Charter in DG MARKT

## Introduction

In its commitment towards achieving better regulation, simplification and joined-up policy-making, DG MARKT recognises the importance of more systematic evaluation of all types of EU action. By providing key data, justification and sound knowledge that allows policy- and decision-makers to plan, design, implement and manage complex EU policies with ambitious objectives, evaluation is an integral component for DG MARKT's better regulation objectives. The following Charter and its accompanying innovative Guide to Evaluating Legislation demonstrate DG MARKT's commitment to evaluating its policies and activities.

## Why do we need an Evaluation Charter?

The new Communication on Evaluation (SEC[2007]213) reinforces the Commission's evaluation framework and clearly sets out evaluation's role within the context of better regulation and informed policy-making. One requirement of this Communication is that all DGs establish an Evaluation Charter to ensure the implementation of the revised evaluation standards within their respective procedures. To this end, the purpose of this MARKT Evaluation Charter is to

*"clearly define the tasks, responsibilities, organisation and procedures for all actors involved in planning, designing and conducting evaluations, and disseminating and using evaluation results" (Evaluation Standard A3).*

Detailed advice can be found in the guidance, but generally speaking ALL evaluation projects must be carried out according to the principles of this Charter.

## WHAT is an evaluation?

Evaluation is *"the judgement of interventions according to their results, impacts and needs they aim to satisfy"*. In other words, evaluation is an evidence-based assessment of how well actions (e.g. legislation) have achieved their stated objectives, by looking at both the positive and negative aspects of unexpected as well as expected effects through a range of direct and indirect impacts.

## WHY do evaluation?

In an ever-changing environment, evaluation can help identify what elements of an action have matched expectations and remain valid, as well as where needs/problems have evolved, providing a key knowledge-base to improve future policy- and decision-making.

## WHEN is an evaluation done?

Evaluation is carried out at pertinent stages in a policy's life-cycle. Previously the legal base has not always stipulated the precise evaluation requirements. However, since 2002, proposals should include evaluation articles, clearly setting out both the timing of evaluations and their requirements (building on the relevant section of the accompanying Impact Assessment). Where a review of policy objectives is required, the evaluation must be scheduled sufficiently in advance to provide the relevant information.

## HOW is an evaluation conducted?

An evaluation is spread out over several key stages, each of which must receive the appropriate attention. The following stages must be respected, regardless of whether an evaluation is conducted by external contractors, DG MARKT staff or as a mixture involving both.

<b>Planning</b>	exactly what will need to be evaluated and the relevant criteria and indicators must be considered from the outset of a proposal and must be clearly outlined in its impact assessment. Subsequently, careful planning must be <u>carried out sufficiently in advance</u> to identify and provide adequate resources (including staff, time and money), thereby ensuring that evaluation results are readily available to feed into policy revision or decision making in a timely manner.
-----------------	--

<b>Design</b>	an <b>evaluation mandate</b> must be drawn up, defining the context, purpose, scope and justification of the evaluation along with an overview of the policy objectives and expected achievements. It will include the evaluation questions, success criteria, data sources, and if appropriate, terms of reference for any external study. The mandate structures and focuses the evaluation, facilitating good quality analysis that can be integral to better policy making.
<b>Manage/Conduct study</b>	evaluation must be carried out in a way which ensures a high quality end product within the given timeframe. Each evaluation will be <u>managed</u> by a designated desk officer from the operational unit working in close cooperation with B.2. A <u>Steering Group</u> will also be created, to ensure that the quality of the analysis is adequately balanced and unbiased. <u>Quality Assessments</u> will be carried out in accordance with established criteria.
<b>Dissemination</b>	in the interests of transparency and accountability, evaluation results must be communicated to the widest possible audience. A <b>dissemination strategy</b> will establish what should be communicated to whom.
<b>Exploitation</b>	it is important that the <u>results</u> of evaluations <i>are used</i> both in policy making and a broader sense. Evaluation results will also feed into the strategic planning & programming cycle at the appropriate juncture e.g. Annual Policy Strategy (APS), Activity Statements (PDB), Annual Activity Reports (AAR), ....
<b>Follow-up</b>	the evaluation's recommendations and any resulting action(s) must be clearly outlined in an <b>action plan</b> to be implemented in an appropriate time-frame. Bi-annual progress reports from the operational unit should be presented until the action plan is completed. These will form the basis of B2's annual report to Directors.

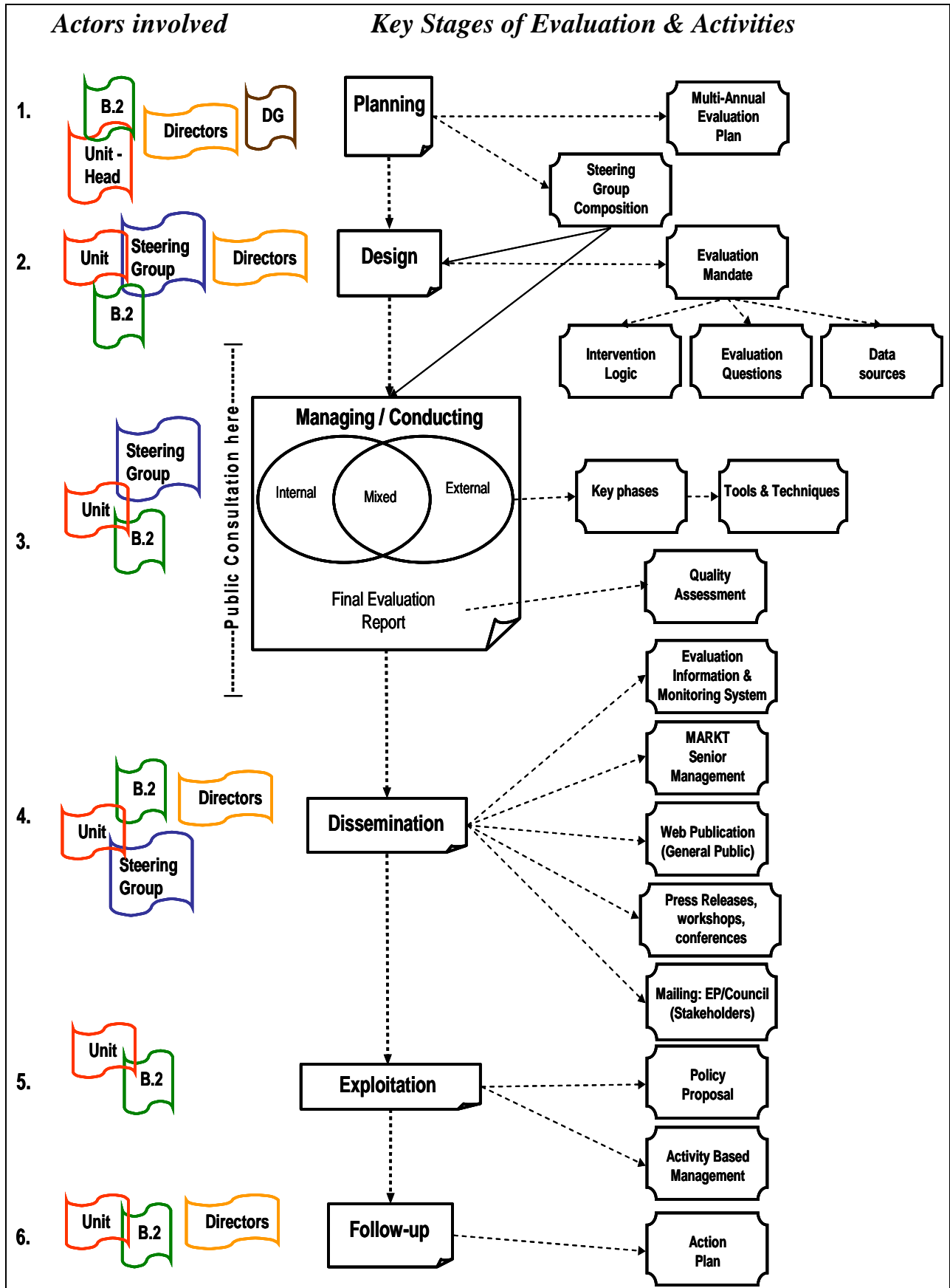
#### WHO does WHAT, WHEN?

WHO	ROLE (WHAT)	STAGE (WHEN)
Director General	<ul style="list-style-type: none"> <li>Approves MARKT evaluation plan as part of the Annual Management Plan</li> </ul>	Planning
Directors	<ul style="list-style-type: none"> <li>In the event of cross-cutting evaluations, Directors decide on the appropriate allocation of work between units</li> <li>Approve the evaluation mandate</li> <li>Approve final report, dissemination strategy, and follow-up action plans</li> </ul>	Planning  Design  Dissemination and Follow-up
Operational Unit - Head	<ul style="list-style-type: none"> <li>Ensures proper timely planning, including resource request</li> <li>Designates a desk officer to supervise the evaluation (to work in close cooperation with B2)</li> <li>Sets up a steering group of appropriate interested parties</li> <li>Feeds evaluation results into strategic planning and programming (APS, AMP, AAR, PDB) and other exercises</li> </ul>	Planning    Exploitation
Operational Unit - Desk Officer	<ul style="list-style-type: none"> <li>Building on the relevant section of the impact assessment that accompanied the legislative proposal, prepares the evaluation mandate (with assistance from B2)</li> <li>Drafts terms of reference for any external assistance required coherent with the above mandate</li> </ul>	Design

WHO	ROLE (WHAT)	STAGE (WHEN)
	<ul style="list-style-type: none"> <li>• Manages/conducts the evaluation (working in close cooperation with B2)</li> </ul>	Conduct/Manage
	<ul style="list-style-type: none"> <li>• Establishes a Quality Assessment for elements of the evaluation undertaken by external contractors (if relevant)</li> <li>• Reviews the relevant data input for the EIMS database (DG BUDG)</li> <li>• Drafts the dissemination strategy and ensures the evaluation findings are presented in a usable form (i.e. no technical jargon) to the appropriate audience</li> <li>• Drafts and ensures implementation of a follow-up action plan and provides the bi-annual progress report</li> </ul>	Dissemination  Follow-up and Exploitation
Evaluation Function - Unit B2	<ul style="list-style-type: none"> <li>• Draws up the evaluation plan for the DG</li> <li>• Advises the Operational Unit on the contents of the evaluation mandate</li> <li>• Provides input to terms of reference for any external assistance required</li> <li>• Assists the operational unit with managing/conducting the evaluation</li> <li>• Establishes a Quality Assessment for elements of the evaluation undertaken by internal staff (if relevant)</li> <li>• Coordinates with DG BUDG on evaluation issues</li> <li>• Inputs necessary data to EIMS database (DG BUDG)</li> <li>• Feeds evaluation results into strategic planning and programming (APS, AMP, AAR, PDB) and other exercises</li> <li>• Advises on the dissemination strategy</li> <li>• Reports to Directors annually on the implementation of action plans</li> </ul>	Planning  Design  Conduct / Manage  Exploitation  Dissemination  Follow-up
Steering Group	<ul style="list-style-type: none"> <li>• Assists with drafting the mandate and any associated terms of reference, facilitating work of the evaluator with information, contacts, etc.</li> <li>• Monitors progress and quality of the evaluation study</li> <li>• Contributes to the quality assessments</li> <li>• Assists with dissemination, if appropriate</li> </ul>	Design  Conduct / Manage  Dissemination

J. Holmquist  
Director General

Figure 1 : Evaluation Charter - work flow of key evaluation stages and actors



## Annex 3 - Evaluation Background and Theory

### Background

The evaluation of activities, particularly spending programmes<sup>7</sup>, is a well established discipline at the Commission. For a long time, evaluation has been a useful way to assess and provide evidence for sound and efficient management. With the arrival of Activity Based Management (ABM) in 2000, evaluation became the main tool for judging how well interventions were achieving their objectives and determining any necessary improvements. Two Communications<sup>8</sup> provided a strong framework for evaluation, defining rules and standards and embedding evaluation as a key Commission tool for evidence-based decision-making. Spending programmes in particular, have relied on evaluation<sup>9</sup> results to produce an evidence-based judgement for budgetary decision-making.

As ABM has evolved over the last decade, so too has evaluation. Evaluation has expanded from a method for judging whether programmes and policies had achieved their objectives into an essential learning and development tool. Today, evaluation also provides key data, justification and sound knowledge allowing policy- and decision-makers to plan, design, implement and manage complex EU policies with ambitious objectives. As a result, there is a need for more systematic evaluation of all types of EU action as an integral component in achieving better regulation, simplification and joined-up policy-making.

Although Impact Assessment<sup>10</sup> is now an accepted means of ensuring that new proposals are based on more informed policy-making, evaluating the results of the policy instruments used (e.g. legislation) is a relatively new and developing field. Even though it is a less common practice, assessing whether and to what extent the expected impacts of current legislative instruments have borne true is equally valuable, if not more so, to policy- and decision-making. To maximise the benefits of evaluation, it should/must always have a specific purpose – be that to provide key information upon which future decisions can be made (e.g. revision) and/or provide solid evidence legitimising decisions taken (e.g. to the citizen or European Parliament). As long as the fundamental requirements of evaluation are met, it may be combined with other exercises (e.g. reviews) when it makes sense to do so.

---

<sup>7</sup> Spending programmes are specifically defined measures that aim to fund, from the EU budget, development across the EU. Certain policy areas, such as AGRI, REGIO and RTD have long traditions in dispersing such funds and therefore, evaluating whether the funds have been spent wisely through these programmes.

<sup>8</sup> SEC (2000)1051 and C(2002)5267.

<sup>9</sup> This "traditional" type of evaluation will be referred to throughout this document as programme evaluation.

<sup>10</sup> COM (2002)276.

## Evaluation Theory

### 1. What is an evaluation of legislation?

The guide itself sets out the basics of what an evaluation is - an **evidence-based assessment** of how well actions have achieved their stated objectives. An evaluation of legislation is a factual analysis that looks at the extent to which its implementation is addressing the issue/problem initially identified.

This helps us to see not just whether our actions are (or are becoming) successful, but will also show where, if at all, the process can be improved; highlight good practices for others to learn from; and/or point out whether it has achieved more or less than expected.

### 2. Why is evaluation theory different for legislation?

While **spending programmes** tend to fit within the linear model of traditional evaluation theory (Figure 2, Section 1.2 of the guide) quite well, it being generally easier to establish for their evaluation :

- a) what they are expected to achieve (**objective**) and the resources available (**inputs**) within a fixed period,
- b) what tangible and measurable outputs (action) and **results** came about,
- c) what **beneficiaries and affected parties were involved**, and
- d) whether money is being spent wisely.

In the case of legislation, which often deals with "concepts" or "principles", the cause-effect logic is somewhat more complex. There are many more potentially diverging paths and possible interactions, both positive and negative, which nevertheless, (hopefully!) lead towards achieving the desired change/s (**objective**).

In evaluating legislation, the focus tends to be on how these many diverging interactions affect the outcome (or not) of the issue/problem it was aiming to address and on the wider impact of the legislation in an EU (and possibly global) context. Trying to fit this within the linear model of traditional evaluation theory would create an impossible picture to decipher:

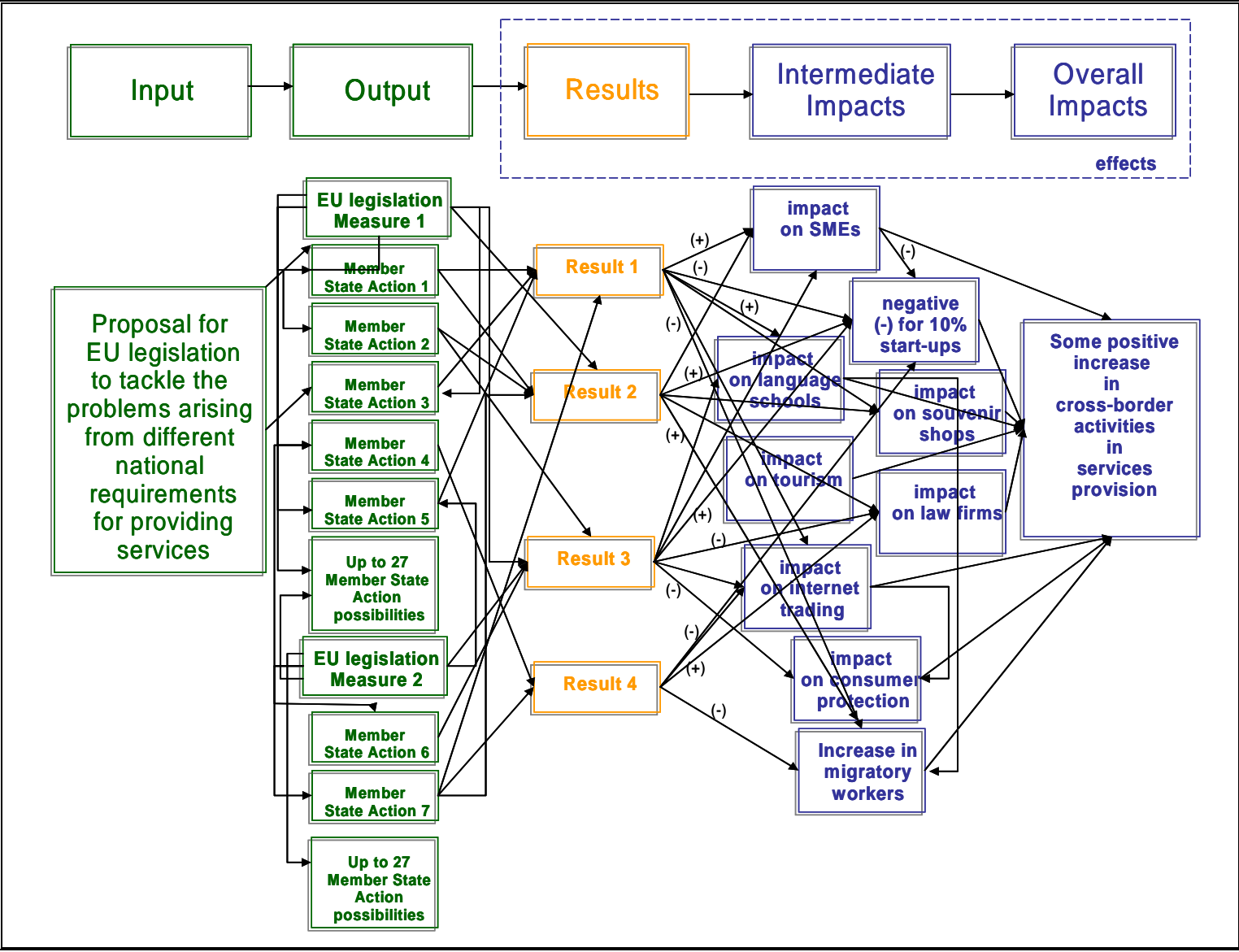


Figure 2 : Attempt to use a 'linear evaluation model' for legislation

Looking at the illustration in Figure 2, it becomes apparent that there is a need for a more workable model that can capture the complexity of these variables, particularly in relation to the effects. A cause-effect model was developed for legislation expanding the standard logic model to capture the elements specific to legislation without losing the important intervening relationships:

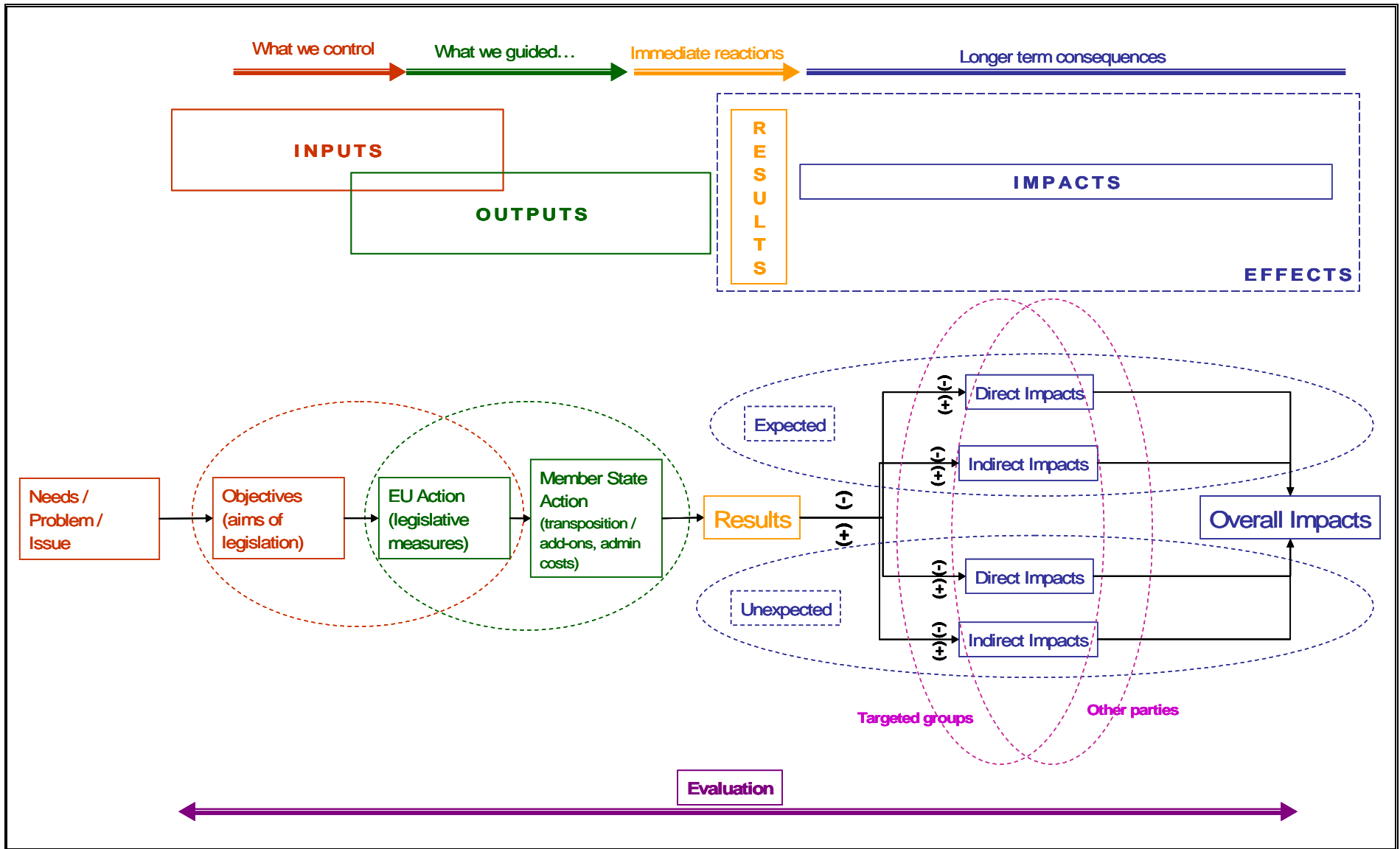


Figure 3 : Cause-effect logics of implementing legislation



The elements specific to legislation that needed attention in developing an appropriate cause-effect logic starts, as can be seen from Figure 3 above, with the fact that EU legislation can be seen both in terms of an "input" (i.e. a resource towards meeting EU policy objectives) and an "output" (i.e. a product upon which Member States react towards meeting EU policy objectives). This dual role is now better reflected. The next element that needed capturing was the **Member State action**, which can potentially be quite diverse (for Directives) and this in turn can significantly affect the successful achievement of the required objectives and the extent of the positive impacts expected, even potentially magnifying the negative impacts. While we cannot say we have total control over this "output" as suggested by Figure 2, Section 1.2 of the guide, legislation nevertheless provides certain limitations within which Member State choices are made. Finally, notwithstanding that evaluation of other EU actions (spending programmes included) also examine the full range of impacts, evaluation of legislation – due to its very nature - does however, focus more on the area of **impacts of all kinds**.

### **3. Where to find the relevant information for the intervention logic**

While good starting point is the Impact Assessment (if it exists), do take account of any changes that may have been made to the legislation during adoption, as the intervention logic is meant to represent how legislation is expected to function as of its adoption and entering into force. Obviously not all legislation being evaluated at present will have an Impact Assessment, but over time, this should become standard. Where there is no Impact Assessment it should be possible to identify these key factors from other documents and/or discussions with staff involved in the area.

- Implementation Reports will also be a key source for mapping legislation according to the above model. Depending on the type of instrument, how Member States transpose legislation can significantly alter the final outcomes. While some legislative instruments (such as Regulations) are strictly binding and enter into force according to the definition set by the EU, others, especially Directives, may include options or derogations that can infer up to 27 variations on transposition, giving Member States generally much more freedom as to how transposition may be carried out.

### **4. The role of impact assessment and evaluation in Activity Based Management**

Both Impact Assessment (IA) and Evaluation are important tools in the Activity Based Management (ABM), Better Regulation (BR) and policy cycles. The planning elements of ABM (i.e. the Annual Policy Strategy and Annual Management Plan) establish the issues for which IA and evaluation findings will be needed. The resulting findings must then feed back into the appropriate decision-making stages of ABM, which equally responds to the BR agenda.

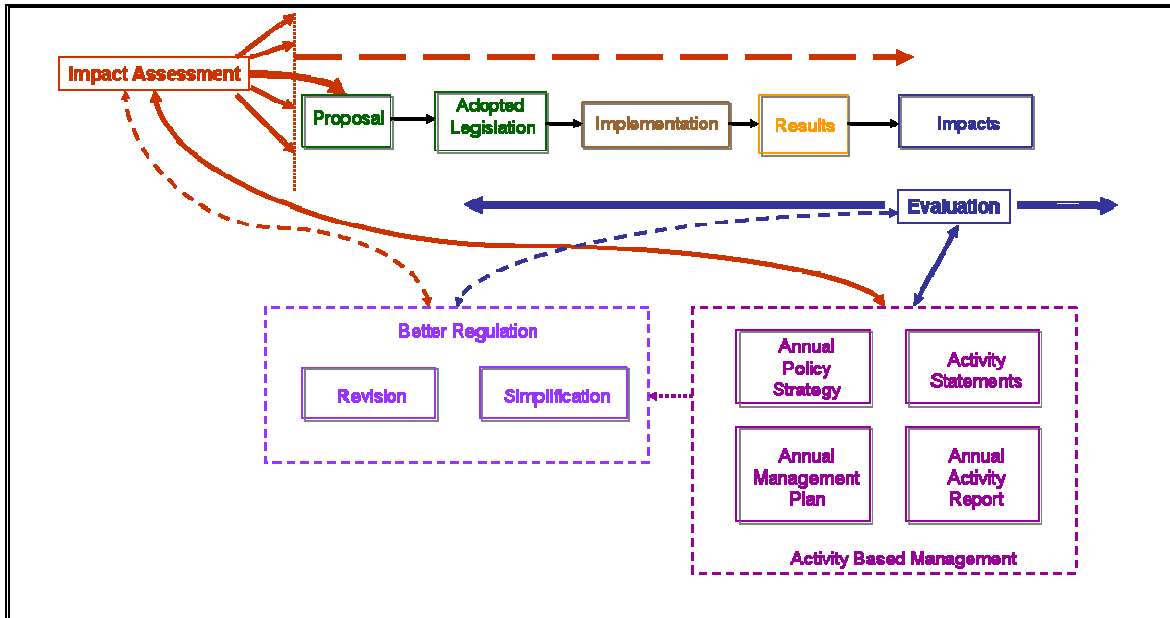
Although in practice the boundaries between these two tools are somewhat blurred, the main differences are presented in the table below:

**Table 1 : Differences between an impact assessment and evaluation**

<p>an Impact Assessment:</p> <ul style="list-style-type: none"> <li>- is forward-looking (prospective / "ex ante")</li> <li>- analyses various options to solve the problem/issue</li> <li>- tries to anticipate likely impacts of the selected option</li> </ul>	<p>an Evaluation:</p> <ul style="list-style-type: none"> <li>- is principally backward-looking (retrospective / "ex-post") although can also look forward</li> <li>- verifies the extent to which the chosen option has really fulfilled the expectations</li> <li>- analyses the actual impacts that have occurred</li> </ul>
---	--

In DG MARKT, this places Impact Assessment firmly in the pre-proposal part of the policy cycle and Evaluation in post-implementation, as illustrated in Figure 4:

**Figure 4 : Impact assessment vs. evaluation in a broader context**



The forward-looking part of an evaluation (i.e. how the measure concerned could be improved) can often provide a sound structure for an Impact Assessment to build upon. It can help identify the nature and extent of a problem and provide first indications as to how these may be solved. In turn, an Impact Assessment should clearly define criteria and expectations for future monitoring and evaluation of the chosen option being proposed – this will ensure that there is an appropriate framework for the evaluation to provide a sound analysis based on solid key data.

## Annex 4 - Template of DG MARKET evaluation mandate

<b>Title of the evaluation (subject)</b>	
<b>Type of the evaluation<sup>1</sup>:</b>	<input type="checkbox"/> prospective <input type="checkbox"/> retrospective <input type="checkbox"/> both
<b>Organisation<sup>2</sup>:</b>	<input type="checkbox"/> internal <input type="checkbox"/> external <input type="checkbox"/> mixed
<b>Timing<sup>3</sup>:</b>	Start date: _____      Completion date: _____
<b>Contact person(s):</b>	Desk Officer (☎99999) / Back up person (☎99999)
<b>Evaluation Function:</b>	Your evaluation correspondent in B.2 (☎99999)
<b>Budgetary Resources<sup>4</sup>:</b>	amount: 000 000 € (+ABB N° & ABB description)
<b>Human Resources:</b>	(N° of Person/Months)
<b>Purpose &amp; objectives of the evaluation:</b>	
<i>define the general objectives and intended use of the evaluation</i>	
<b>Justification:</b>	<input type="checkbox"/> obligatory <input type="checkbox"/> voluntary
<i>describe the legal requirement for carrying out the evaluation</i>	
<b>Description of the action to be evaluated:</b>	
<i>basic information about the legislation that includes objectives of the legislation; the measures it sets out; expected results and impacts.</i>	
<b>Scope of the evaluation:</b>	
<i>Coverage of the evaluation: all/parts of legislation; timescale, i.e. first 5 years; geographical, i.e. all EU-27</i>	
<b>Intervention Logic:</b>	
<i>Insert the drawing of the intervention logic<sup>5</sup></i>	
<b>Key questions of the evaluation:</b>	
<i>Relevance; Effectiveness; Efficiency; Distributional effects; Acceptability; Consistency; EU added-value</i>	

Title of the evaluation (subject)			
<b>Data Sources:</b>			
Category	Evaluation Questions	Data Type	Source
<b>Stakeholders:</b> <i>(in each case indicate who)</i>			
Within Commission:			
Member State authorities:			
Business, trade, NGOs, etc:			
Other:			
<b>Steering Group:</b> <i>(in each case indicate who)</i>			
<input type="checkbox"/> DG internal steering:			
<input type="checkbox"/> With participation of other Commission services (recommended):			
<input type="checkbox"/> Steering including Member State's representatives (optional):			
<input type="checkbox"/> Steering including external stakeholders (optional):			
<b>Milestones, deadlines:</b>			
a) Designation of the Steering Group:	__/__/__	e) Quality Assessment	__/__/__
b) Validation of the Mandate (including ToR)	__/__/__	f) Summary of Key Findings & Recommendations	__/__/__
c) Selection of Proposals (if external evaluation)	__/__/__	g) Dissemination Strategy	__/__/__
d) (Draft) Final Report	__/__/__	h) Follow-up Action Plan	__/__/__
<b>Other relevant information/remarks:</b>			

<sup>1</sup> 90% evaluations in MARKT are both.

<sup>2</sup> Majority of evaluations tend to be mixed.

<sup>3</sup> Start date = work on completing mandate; completion date = dissemination of results.

<sup>4</sup> ABB N° & Descriptions: 12.020100 - Implementation and development of Internal Market, 12.010211.04 Studies and Consultations (please check with A1 if not sure).

<sup>5</sup> Or attach as separate document if necessary for legibility.

## Annex 5 - Case Study for evaluating legislation: example of the evaluation mandate for Services Directive

Example of the evaluation mandate as it should be completed.

<b>Case Study – Design of Evaluation for Authorisation Schemes EVALUATION MANDATE</b>	
<b>Evaluation of Authorisation Schemes (part of Services Directive)</b>	
<b>Type of the evaluation:</b> <input type="checkbox"/> prospective <input checked="" type="checkbox"/> retrospective <input type="checkbox"/> both	
<b>Organisation:</b> <input type="checkbox"/> internal <input checked="" type="checkbox"/> mixed <input type="checkbox"/> external	
<b>Timing:</b>	Start date: January 2011                      Completion date: June 2012
<b>Contact person(s):</b>	Desk Officer: Tom Jones, Z3 (☎99991) / Fred Smith, Z3 (☎99992)
<b>Evaluation Function:</b>	Your evaluation correspondent in B.2 Bill Wiggins (☎90009)
<b>Budgetary Resources:</b>	<i>amount: 300 000 € (ABB N° 12.020100. -Implementation and development of Internal Market)</i>
<b>Human Resources:</b>	<i>24 Person/Months</i>
<b>Purpose &amp; objectives of the evaluation:</b> <i>define the general objectives and intended use of the evaluation</i>	
<p>The evaluation should provide a snap shot of the current situation allowing comparisons to be drawn both with the situation prior to adoption and that presented in the 2008 report (stipulated by Article 39(1)).</p> <p>It will assess the progress that has been made in applying the content of the legislation and provide indications as to whether the desired objectives are being achieved.</p> <p>It should assess if any changes or additional measures are necessary. In particular, it should identify any parties who are not taking action to implement appropriate measures.</p>	
<b>Justification:</b> <input checked="" type="checkbox"/> obligatory <input type="checkbox"/> voluntary <i>describe the legal requirement for carrying out the evaluation</i>	
<p>Directive 2006/123/EC (hereafter termed the "Services Directive") foresees in the Review Clause (article 41) the completion of a first review by the end of 2011 and every three years thereafter. Due to the scope of the review, an evaluation has been considered more appropriate for the needs of the decision-making process.</p>	
<b>Description of the action to be evaluated:</b> <i>basic information about the legislation that includes objectives of the legislation; the measures it sets out; expected results and impacts</i>	
<p>Authorisation schemes are procedures which oblige a service provider or recipient to obtain from a competent authority a formal decision, or implied decision, concerning access to a service activity or permission to carry out said activity. Under the Services Directive all authorisation schemes must adhere to the following principles:</p> <ul style="list-style-type: none"> <li>• a scheme should be non-discriminatory;</li> <li>• schemes must be justified and in the public interest;</li> <li>• the objectives of such a scheme cannot be achieved by means of less restrictive measures.</li> </ul>	

**Case Study – Design of Evaluation for Authorisation Schemes  
EVALUATION MANDATE**

**Evaluation of Authorisation Schemes (part of Services Directive)**

Articles 9-13 of the Services Directive provide guidance on when an authorisation scheme is permitted; how it should be defined and operated; the restricted number of reasons when the duration of an authorisation may be granted for a limited period; and when and how selection procedures may be used in cases where only a limited number of authorisations are allowed. Once a provider has obtained such an authorisation it should be able to access/provide the service. Article 14 sets out requirements that are prohibited under such authorisation schemes and Article 15 establishes the specific requirements which are to be evaluated for authorisation schemes.

*Problems:* The main problems cited by service providers in relation to the issue of establishment are: unnecessary, burdensome and discriminatory requirements linked to the granting of authorisations; lengthy, costly and non-transparent procedures.

*Target Population:* Services affected by authorisation schemes, authorities administering such schemes

N.B. the Services Directive covers only services performed for an economic consideration. Services of General Interest are not covered, nor are certain Services of General Economic Interest (e.g. transport, postal). Other excluded services include: Financial services, electronic communications networks and services, transport services, healthcare, audiovisual services, gambling, social services and taxation. (See Directive for details).

*Objectives:* In practical terms the provisions in the Services Directive relating to authorisations are intended to ensure that within the EU, all authorisation schemes for the creation of establishments are justified, having clear, transparent, objective and non-discriminatory criteria. Existing schemes should be screened to evaluate whether they meet these standards. Where appropriate existing schemes should be removed or replaced by less restrictive means such as simple notifications. In this way, businesses should have a clearer understanding of what is required of them and gain better access to markets, in a more predictable and timely manner.

Member States also have to abolish a number of restrictions on establishment, such as nationality requirements or "economic needs tests" (which require businesses to commission market studies, often costing hundreds of thousands of Euros, to prove to the authorities that they will not "destabilise" local competition).

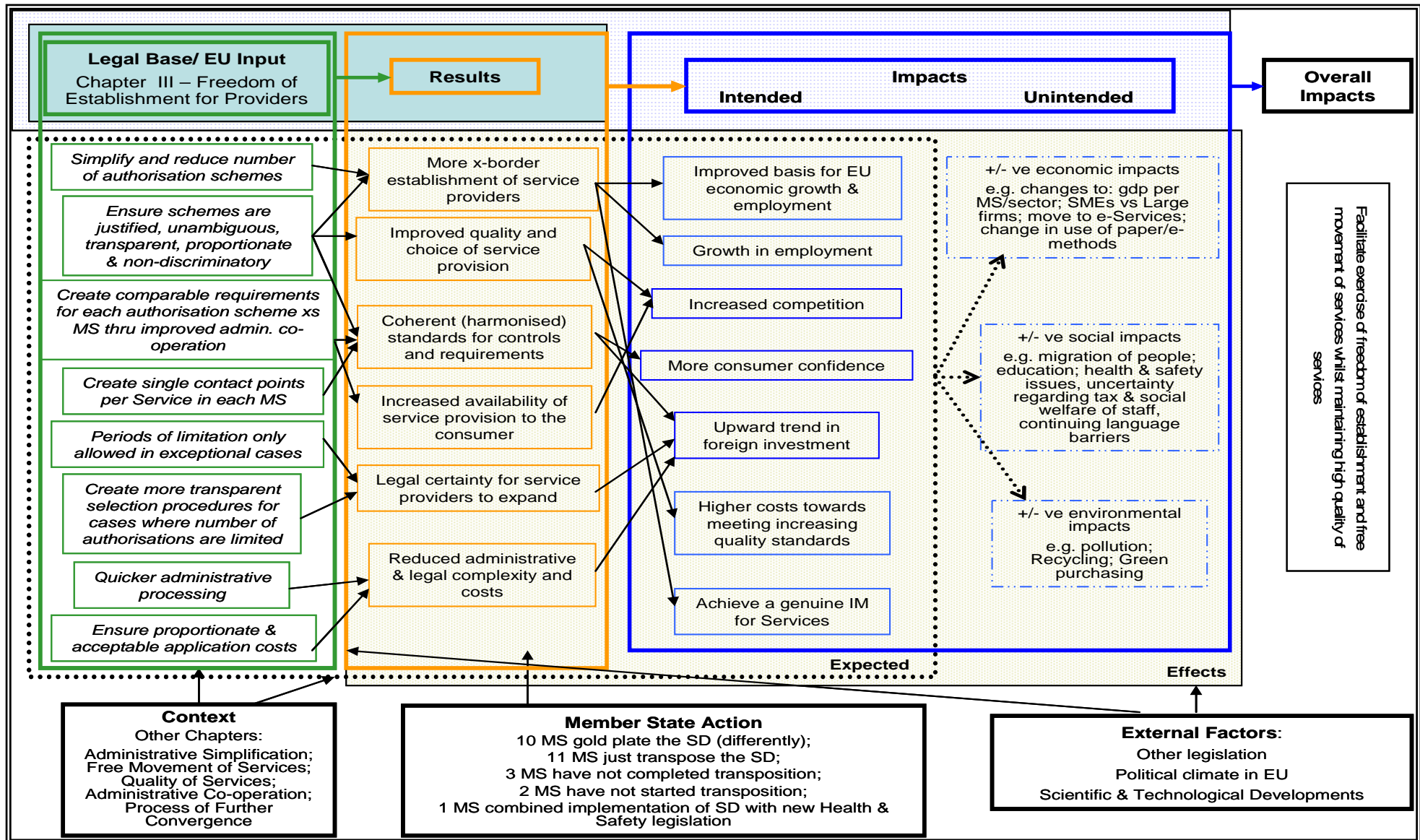
*Outputs, expected results and impacts:* 1) simpler, non-discriminatory and justified authorisation schemes; 2) Easier and quicker access to markets in other MS leading to more cross border provision of services; 3) Improved Internal Market delivering Economic benefits to EU 4) Freedoms of establishment and movement facilitated

**Scope of the evaluation:**

*Coverage of the evaluation: all/parts of legislation; timescale, i.e. first 5 years; geographical, i.e. all EU-27*

The evaluation focuses on the Articles relating to Authorisation in the Services Directive and the effects this legislation has had in the EU 27 during the period January 2007 to December 2010. It should provide an overview of the situation in all relevant sectors, with particular emphasis on 3 major sectors (e.g. construction and real estate, legal advice and consultation, tourism and leisure).

**Intervention Logic:**



## Evaluation of Authorisation Schemes (part of Services Directive)

### Key questions of the evaluation:

*Relevance; Effectiveness; Efficiency; Distributional effects; Acceptability; Consistency; EU added-value*

#### Relevance:

- ◆ Are the principles set out for authorisation schemes still appropriate today?
- ◆ Should other principles be defined?

#### Effectiveness:

- ◆ Which authorisation schemes have been changed, abolished or maintained as a result of the implementation of the directive?
- ◆ What different measures have been implemented at national/sector level and how has these affected results?
- ◆ To what extent have the above changes (positive or negative) contributed to the objective of the Services Directive in upholding the Freedom of Establishment?
- ◆ If there are still problems with authorisation procedures, what are they and to what extent are these problems related to how the Directive has been interpreted or implemented?

#### Efficiency:

- ◆ Could improvements have been made to authorisation schemes by any other measure – was legislation in this form necessary?
- ◆ What were the costs of screening the existing authorisation schemes?
- ◆ What were the costs of changing authorisation schemes?
- ◆ To what extent are current schemes simpler to administer and apply for, quicker and cheaper? And what is the potential for, and cost of, further simplification?

#### Distributional effects:

- ◆ How have the observed costs/benefits been shared (equally/proportionally) across the various groups involved, in particular, service providers, consumers, new market entrants, expanding businesses, employees, different geographic levels/countries, economic sector?
- ◆ What measures have been introduced to combat any undesirable effects?

#### Acceptability:

- ◆ What has been the reception from stakeholders – how does it differ between different stakeholders?
- ◆ Are there differences according to sector?

#### Consistency:

- ◆ Have the changes to authorisation schemes, both positive and negative, contributed to the Single Market objectives of creating cross-border opportunities without conflicting with the specific objectives of the authorisation schemes themselves?
- ◆ In what way have such changes influenced the global objectives of authorisation schemes?
- ◆ What steps have been taken to maximise synergies for authorisation schemes or minimise the impacts of conflicting objectives?

#### EU Added Value:

- ◆ To what extent could the changes brought about by the Services Directive have been achieved by national measures only?



Evaluation of Authorisation Schemes (part of Services Directive)			
Data Sources:			
Category	Evaluation Questions	Data Type	Source
Relevance	Are the principles set out for authorisation schemes still appropriate today? Should other principles be defined?	<ul style="list-style-type: none"> <li>An overview of the situation in EU – a "state of play";</li> <li>A summary of the situation as was at time of drafting legislation</li> </ul>	<ul style="list-style-type: none"> <li>Impact Assessment;</li> <li>Implementation Reports;</li> <li>Academic Literature;</li> <li>case law</li> <li>MS authorities;</li> <li>Infringement unit;</li> <li>National judiciaries;</li> <li>European Court of Justice</li> </ul>
	How has the number of authorisation schemes changed as a result of the implementation of the directive? Have some schemes been deleted?	<ul style="list-style-type: none"> <li>Statistics on authorisation schemes</li> <li>Qualitative data on the number of the authorisation scheme per MS and sector</li> </ul>	<ul style="list-style-type: none"> <li>from the relevant authorities</li> </ul>
Effectiveness	Are current schemes simpler to administer and apply for; quicker and cheaper?	----	----
	To what extent are the changes (positive or negative) attributable to the Authorisation articles?	----	---
	If there are still problems with ....?	.....	.....
	What different ...?	....	
	Have ..?		
Efficiency			
Distributional effects			
Acceptability			
Consistency			
EU value-added			
<b>Stakeholders:</b> <i>(in each case indicate who)</i>			
Within Commission:		DGs represented in the Steering Group	
Member State authorities:		IMAC members responsible for implementing the Services Directive	

Business, trade, NGOs, etc:	<p><u>Business organisations from the EU and national levels – general ones as well as sectoral where authorisations schemes play particularly important role, e.g.:</u></p> <ul style="list-style-type: none"> <li>• UNICE - Confederation of European Business</li> <li>• EuroCommerce</li> <li>• ETOA – European Tour Operators Association</li> <li>• Eurochambers</li> <li>• FIEC - European Construction Industry Federation</li> <li>• CIAA - Confederation of the food and drink industries of the EU</li> <li>• RICS - Royal Institution of Chartered Surveyors (real estate)</li> </ul> <p><u>Consumer organisations, e.g.:</u></p> <ul style="list-style-type: none"> <li>• EUROCOOP - European Community of Consumer Cooperatives</li> <li>• ANEC - European Consumer Voice in Standardisation</li> <li>• BEUC - The European Consumers Organisation</li> </ul> <p><u>NGOs:</u></p> <ul style="list-style-type: none"> <li>• European Citizen Action Service, ECAS</li> </ul> <p><u>Trade unions, e.g.:</u></p> <ul style="list-style-type: none"> <li>• UNI-Europa</li> <li>• ETUC/CES - European Trade Union Confederation</li> </ul>		
Other:			
<b>Steering Group:</b> <i>(in each case indicate who)</i>			
<input type="checkbox"/> DG internal steering: <input checked="" type="checkbox"/> With participation of other Commission services (recommended): COMP, EMPL, ENTR, MARKT, SANCO, SG <sup>11</sup> <input type="checkbox"/> Steering including Member State's representatives (optional): <input type="checkbox"/> Steering including external stakeholders (optional):			
<b>Milestones, deadlines:</b>			
a) Designation of the Steering Group	01/01/2011	e) Quality Assessment	31/07/2012
b) Validation of the Mandate (including ToR)	28/02/2011	f) Key Findings & Recommendations	15/06/2012
c) Selection of Proposals (if external evaluation)	15/06/2011	g) Dissemination Strategy	30/06/2012
d) (Draft) Final Report	30/05/2012	h) Follow-up Action Plan	15/09/2012
<b>Other relevant information/remarks:</b>			

<sup>11</sup> Even though this is a recommended option, in this case it would be useful to have a broader Steering Group e.g. to include member of the IMAC advisory committee, or some key stakeholder. In this case the Commission services would be the core SG (e.g. having a final say on all contractual issues) but other members could be involved in some parts, esp. in the data collection phase, validation, etc. It always brings more transparency to the evaluation process and makes your stakeholders take active part in it.

## Annex 6 - Quality Assessment Form

The quality assessment framework consists of a quality assessment form, a guide on scoring the criteria and an explanatory guide for organising the quality assessment process.

1. **The guide for organising the quality assessment** Process provides information on the main stages of the assessment process and on the purpose of each rubric of the Quality Assessment Form. Moreover, it explains the key options concerning publication of the quality assessment.
2. **The quality assessment form** should be used in order to ensure that the assessment is based on the relevant quality criteria and to keep the assessment focused on the quality of the actual report.
3. **The guide on scoring the criteria** explains a set of indicators relevant for each of the 8 quality criteria.

### 1. Guide for organising the quality assessment

#### 1.1 What are the main aims, stages and actors involved in the QA process?

Assessing the quality of an evaluation can help to:

- encourage the evaluator(s) to conform to the terms of reference and to professional standards;
- ensure transparency about the quality of the evaluation to all potential users;
- enable organisational learning about the evaluation process.

Accordingly, the assessment must be part of the evaluation process right from the start, i.e., not just added on once to the final report becomes available.

Firstly, the terms of references must from the onset make (external) evaluators aware of the quality assessment criteria (*standard C4*) that will be applied and about the possibility that the assessment will be made public. The criteria should hence be attached or hyperlinked to the terms of reference.

Throughout the evaluation process, the official(s) managing the evaluation and the steering group should alert the evaluator (*standard C7*) about potential quality problems.

A preliminary quality assessment should be prepared based on the draft final report and the results of this exercise must be communicated and explained to the evaluator(s) so they are offered the possibility of improving the work (clarity of drafting; conclusions/recommendations that aren't sufficiently underpinned by data and analysis etc).

The actual quality assessment is prepared once the final report becomes available. The assessment is carried out by the official(s) managing the evaluation in cooperation with the steering group (*standard C7*). To be credible (cf. in particular the *aim of transparency* concerning the quality) it is important to differentiate responsibility in managing the intervention, the evaluation process and quality assessment.

### 1.2 How to fill in the quality assessment form?

The first box of the Quality Assessment Form should be filled-in with information that identifies the object being evaluated, the evaluator and the participants in the assessment exercise.

For *each criterion*, a first section provides space for the score and the arguments for it. The second section is optional and is meant to provide contextual information (e.g. lack of data or of access to data base) about the evaluation process and contractual constraints (e.g. resources)<sup>12</sup>. The aim of this section is to make the scoring independent from these constraints. Besides, it can be used to collect information for organisational learning about the evaluation process. The quality assessment exercise concerns how the *conclusions* and *recommendations* were obtained and presented, but must not compromise the evaluators ability to present their results.

The *overall assessment* at the end of the form summarises key elements of the overall quality as detailed in the eight preceding criteria. It will be helpful to potential readers of the quality assessment if the overall assessment highlights the consequences of this for different types of use (contract management; accountability and internal (partial) use of the findings).

This rubric also contains space for synthesising the information on contextual and contractual constraints, so general lessons can be drawn for improving future evaluations. As a caveat, it should be noted that the purpose of the assessment form is to assess the quality of the evaluation report not the evaluation process.

### 1.3 How to do the scoring?

The Guide on Scoring the Criteria provides a set of *indicators* for each criterion. The aim is to facilitate a correct and consistent scoring, and to help the assessor(s) to develop a comprehensive and coherent argumentation to underpin the score given for each criterion.

The indicators are, roughly speaking, presented in order of importance (i.e., those at the start of the list are crucial even for a moderate score), and they may not all be relevant to specific evaluation

The application of the indicators to a particular quality assessment should be adapted to the specificities of the evaluation. The weight attached to, or attention paid to each indicator should be proportionate to their relevance for the evaluation being assessed, and additional indicators may be necessary to capture the specificity of the evaluation.

### 1.4 Publishing the quality assessment report?

The directorate general managing the evaluation decides on whether the quality assessment should be published. Publication, along with the evaluation report, can reinforce some of the purposes of the Quality Assessment exercise. In general, publication offers the advantages of:

- ensuring transparency about the reliability of the evaluation results to external users and stakeholders  
(*a generalised policy of publishing all assessments, regardless of the quality of the evaluations, may also safeguard against external stakeholders who expect routine implementation of recommendation*).

---

<sup>12</sup> This does not concern the context in which the object being evaluated is implemented, only the context in which the evaluation is carried out.

- encouraging the evaluator to conform to the terms of reference and to professional standards.

However, such publication also reinforces the need to maintain a consistent application of the quality criteria between different evaluations and over time.

Publication may involve the full quality assessment report; parts of it (in particular criteria 5: '*credible findings*', 6: '*valid conclusions*', and probably 7: '*helpful recommendations*') or publication of a global statement on the quality.

The evaluation function should in any case make the quality assessment available to all relevant internal stakeholders.

## 2. Quality Assessment Form

**Title of the evaluation ...** <sup>13</sup>

- Draft final report...
- Final report .....

**DG/Unit** .....

- Official(s) managing the evaluation .

**Evaluator/contractor** .....

---

**Assessment carried out by<sup>(\*)</sup>:**

- Steering group .....
- Evaluation Function.....
- Other (please specify).....

(\*) Multiple crosses possible

**Date of the Quality Assessment** .....

<b>(1) RELEVANCE</b>					
<i>Does the evaluation respond to information needs, in particular as expressed in the terms of references?</i>					
<b>SCORING</b>	<b>Poor</b>	<b>Satisfactory</b>	<b>Good</b>	<b>Very Good</b>	<b>Excellent</b>
	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
<b>Arguments for scoring:</b>					
If relevant: <b>Contextual</b> (such as deficient terms of references) and <b>contractual constraints</b> (such as lack of time, insufficient resources)					

<sup>13</sup> Refer to the [‘Guide on Scoring the Criteria’](#) for how to assess each criterion.

## (2) APPROPRIATE DESIGN

*Is the design of the evaluation adequate for obtaining the results needed to answer the evaluation questions?*

SCORING	Poor	Satisfactory	Good	Very Good	Excellent
	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Arguments for scoring:

If relevant: **Contextual** (unexpected issues) and **contractual constraints** (such as lack of time and resources)

## (3) RELIABLE DATA

*Are data collected adequate for their intended use and have their reliability been ascertained?*

SCORING	Poor	Satisfactory	Good	Very Good	Excellent
	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Arguments for scoring:

If relevant: **Contextual** (such as lack of data or access to data base) and **contractual constraints** (such as lack of time and resources)

## (4) SOUND ANALYSIS

*Are data systematically analysed to answer evaluation questions and cover other information needs in a valid manner?*

SCORING	Poor	Satisfactory	Good	Very Good	Excellent
	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Arguments for scoring:

If relevant: **Contextual** and **contractual constraints** (such as lack of resources and time)

### **(5) CREDIBLE FINDINGS**

*Do findings follow logically from and are justified by, the data/information analysis and interpretations based on pre-established criteria and rational?*

<b>SCORING</b>	<b>Poor</b>	<b>Satisfactory</b>	<b>Good</b>	<b>Very Good</b>	<b>Excellent</b>
	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Arguments for scoring:

If relevant: **Contextual and contractual constraints**

### **(6) VALID CONCLUSIONS**

*Are conclusions non-biased and fully based on findings?*

<b>SCORING</b>	<b>Poor</b>	<b>Satisfactory</b>	<b>Good</b>	<b>Very Good</b>	<b>Excellent</b>
	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Arguments for scoring:

If relevant: **Contextual and contractual constraints**

### **(7) HELPFUL RECOMENDATIONS**

*Are areas needing improvements identified in coherence with the conclusions? Are the suggested options realistic and impartial?*

<b>SCORING</b>	<b>Poor</b>	<b>Satisfactory</b>	<b>Good</b>	<b>Very Good</b>	<b>Excellent</b>
	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Arguments for scoring:

If relevant: **Contextual and contractual constraints**



## (8) CLARITY

*Is the report well structured, balanced and written in an understandable manner?*

SCORING	Poor	Satisfactory	Good	Very Good	Excellent
	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Arguments for scoring:

If relevant: **Contextual and contractual constraints**

## OVERALL ASSESSMENT OF THE FINAL EVALUATION REPORT

**Is the overall quality of the report adequate, in particular:**

- Does the evaluation fulfil contractual conditions?
- Are the findings and conclusions of the report reliable, and are there any specific limitations to their validity and completeness?
- Is the information in the report potentially useful for designing intervention, setting priorities, allocating resources or improving interventions?

**Given the contextual and contractual constraints encountered:**

- What lessons can be learned from the evaluation process?

### 3. Guide on scoring the criteria

#### 1. - RELEVANCE

*Does the evaluation respond to information needs, in particular as expressed in the terms of references?*

This criterion concerns how well the evaluation responds to the terms of references.

- The evaluation report deals with and responds to the evaluation questions
- A justification was provided if any evaluation questions wasn't answered
- The scope covers the requested periods of time, geographical areas, target groups, parts of budget, regulations, etc
- Limitations in scope are discussed and justified
- Effects on other policies, programs, groups, areas etc are considered
- Unintended effects are identified
- The evolution of the intervention is taken into account possible changes in the problems and needs compared to the situation at the start of the intervention have been addressed
- The evaluation broaden the scope or enlighten the approaches in the policy cycle
- The evaluation add value to existing policy knowledge
- Other

#### 2. - APPROPRIATE DESIGN

*Is the design of the evaluation adequate for obtaining the results needed to answer the evaluation questions?*

This criterion concerns the inception phase. It operationalises and possibly complements the terms of reference. In some cases, because of unforeseen events, it may also relate to a subsequent reorientation of parts of the evaluation work.

- The rationale of the program, cause-effect relations, outcomes, policy context, stakeholder interests, etc have been studied and taken into account in design the evaluation
- The evaluation method chosen is coherent with evaluation needs and requests
- The method is clearly and adequately described, in enough detail for the quality to be judged. It is described to the extent that the evaluation can be replicated
- Information sources and analysis tools are adequate for answering the evaluation questions
- Judgement criteria to help answer the evaluation question were pre-defined
- Weaknesses of the selected method are pointed out along with potential risks
- Other methodological alternatives are considered; their pros and cons are explained
- Research design has been validated with experts or relevant stakeholders if appropriate (e.g. experts on related policies, specific evaluation know-how)
- Ethical issues are properly considered (confidentiality of sources of information, potential harms or difficulties of participation of stakeholders, etc)
- Other

### 3. - RELIABLE DATA

#### *Are data collected adequate for their intended use and have their reliability been ascertained?*

This criterion concerns the relevance and correctness of both primary and secondary data.

- Available information and sources are well identified
- Relevant literature and previous studies have been sufficiently reviewed
- Existing monitoring systems were used
- Data and information are free from factual or logic errors; data gathered are correct and sufficient
- Data collection rationale is explained; and it is coherent with the design of the study
- The quality of existing or collected data was checked and ascertained
- The amount of qualitative information and quantitative data is balanced and appropriate for a valid and reliable analysis
- Tools and means used to collect and process data (e.g. surveys, case studies, expert groups, etc...) were: selected in relation to criteria specified in the inception phase; complete and suitable for answering the evaluative questions; adequately used as to guarantee the reliability and validity of results
- Tools and data collection limitations (missing coverage, non-participation or non-attendance of selected cases) are discussed and explained.
- Correcting measures have been taken to avoid any potential bias and/or their implications
- Other

### 4. - SOUND ANALYSIS

#### *Are data systematically analysed to answer evaluation questions and cover other information needs in a valid manner?*

This criterion refers to the correct interpretation of data and to the adequacy of the method applied.

- There is a clear, solid and coherent deductive analysis (e.g. controlled comparison, experimental research, inferential statistics, etc...)
- The analysis is well focussed on the most relevant cause/effect relations and influences underlying the program logic, and alternative explanations have been considered
- The analysis uses appropriate quantitative or qualitative techniques, suitable to the evaluation context
- Cross checking of findings has taken place. The analysis relies on two or more independent lines of evidence
- Explanatory arguments are explicitly (or implicitly) presented
- The context (historical, socio-economic, etc...) is well taken into account in the analysis
- The report reflects an appropriate range of stakeholders consulted
- Inputs from important stakeholders are used in a balance way
- The limitations of the analysis and exceptions to general explanations or evidences were identified, discussed and transparently presented
- Other

## 5. - CREDIBLE FINDINGS

*Do findings follow logically from and are justified by, the data/information analysis and interpretations based on pre-established criteria and rational?*

This criterion concerns the coherence of the findings with the preceding analysis and data.

- Judgements are based on transparent criteria
- Findings are supported by evidence originating from sound analysis
- Generalisations or extrapolations, when made, are justified (e.g., through the sampling or selection of cases)
- Findings corroborate existing knowledge; differences or contradictions with existing know-how are explained
- Stakeholder opinions were considered and reflected when appropriate
- Main findings are replicable
- Limitations on validity are pointed out; trade-offs between internal and external validity are identified and discussed
- Results of the analysis reflect an acceptable compromise of the perceptions of stakeholders and those described by figures and facts observed and estimated
- Other

## 6. - VALID CONCLUSIONS

*Are conclusions non-bias and fully based on findings?*

This criterion concerns the extent to which conclusions logically stem from findings and are based on impartial judgement.

- Conclusions are properly addressed to the evaluation questions and other information needs
- Conclusions are coherently and logically substantiated by evaluation findings
- There are no relevant conclusions missing according to the evidences presented
- Conclusions are interpreted in relation to the policy context
- Conclusions are free of personal or partisan considerations; potential influence of values and interests of the evaluation team in the research method and outcome are openly discussed
- Conclusions are orderly presented and related (categorised, ranked, priorities, sequence)
- Controversial issues are presented in a fair and balanced manner
- Other

## 7. - HELPFUL RECOMMENDATIONS

*Are areas needing improvements identified in coherence with the conclusions? Are the suggested options realistic and impartial?*

This criterion concerns the soundness and realism of the recommendations

- Recommendations stem logically from conclusions
- Plausible options for improvements are identified
- Recommendations covers all relevant main conclusions
- They are realistic, impartial, and potentially useful
- Relations among recommendations are taken into account (e.g. priority ranking, sequencing,

etc)

- Recommendations provide certain guidance for action planning
- Where feasible, cost of recommendations were estimated
- Other

## 8. - CLARITY

### *Is the report well structured, balanced, and written in an understandable manner?*

This criterion concerns to the clarity of the presentation and the appropriateness of the content of the evaluation.

- The content of the report describes the policy being evaluated, its context, the evaluation purposes, contextual limitations, methodology, findings, etc in a neat and well structured manner
- The report is well structured and signposted to guide and facilitate reading
- Key messages are summarised and highlighted
- There is a clear presentational linked sequence among data, interpretation and conclusions
- The report includes a relevant and concise executive summary, which includes main conclusions and recommendations in a balance and impartial manner
- Specialised concepts were used only when necessary and were they clearly defined
- Tables, graphs, and similar presentational tools are used to facilitate understanding; they are well commented with narrative text
- the length of the report (excluded appendices) is proportionate (good balance of descriptive and analytical information)
- Detailed information and technical analysis are left for the appendix; information overload is avoided in the report
- The report provides a proper focus of truly relevant issues
- Written style and presentation is adapted for the various relevant target readers; the evaluator show awareness of potentially different needs and interests
- Other

## OVERALL ASSESSMENT OF THE REPORT

The overall assessment of the evaluation report is not a self-standing criterion. Instead it summarises key elements and consequences of the eight preceding criteria. Moreover, the overall assessment needs to consider the concerns of the potential users of each specific evaluation:

- Does the evaluation fulfil contractual conditions? (*certain internal users*);
- Are the findings and conclusions reliable, and are there any specific limitations to their validity and completeness? (*most internal and external users*)
- Notwithstanding intrinsic weaknesses, is the information in the report -or parts of it- a useful input for designing intervention, setting priorities, allocating resources or improving interventions? (*certain internal users*)

## Annex 7 - Concept/Issue Mapping

As discussed in the main section of the guide, in order to get the most useful results from an evaluation, the analysis may need to be focused on specific elements or issues. To date, this would have been done by studying what happened in a number of 'representative' Member States and trying to extrapolate the results to other countries. Such a choice was often made intuitively without examining in detail any particular issues that could be further examined. While this works well for specifically defined measures such as programmes, there are too many interlinked factors in legislative measures such as Directives to make this a reliable method from which to extrapolate findings. Due to the complexity of legislation, identifying key aspects on which to focus from the intervention logic is not so obvious. **To help select areas where the evaluation may focus its enquiry**, chapter 3.3.2. of **the main guide suggests using Concept/Issue Mapping and Social Network Analysis** (see Annex 8).

By providing an alternative perspective of the intervention logic, **these tools capture the multi-dimensional aspects of legislation and can draw out strands that may be worth closer inspection**, rather than focusing in-depth enquiry on a particular geographic or demographic selection and extrapolating it to the wider world.

By establishing similarities/differences between entities, this tool is used for identifying clusters and types within various elements of the Intervention Logic (such as common types of Member State action; several impacts relating to cross-border activity or administrative cooperation). A fairly straight-forward technique, it involves clustering similar topics together and rating each cluster according to its perceived importance.

### 1. Who should / can do it ?

This analysis can be carried out by the operational unit (OU) with or without the assistance of B2. Equally, if consultants are being used, they could be tasked with this piece of work. If such an exercise is conducted before finalising the technical annex (for external evaluations) or the mandate (for internal/mixed evaluations), it may be worth encouraging the Steering Group to participate.

### 2. How best to use this tool ?

The following steps illustrate the best way to use this tool:

#### ***Step 1 : List relevant issues / impacts, etc.***

As a first step, using available literature and an initial stakeholder consultation (e.g. with relevant authorities who have worked on the transposition of the legislation in the Member States) try to establish a global picture of all the issues which are relevant and group them under useful "headings". For example, you could gather: all the ways in which Member States transposed the legislation, including all gold-plating measures; or all potential impacts regardless of importance.

Any survey provided to selected stakeholders (covering as many as appropriate) should be structured in a focused way whilst still enabling respondents to propose extra issues (done

by asking open-ended questions).<sup>14</sup> An added bonus here would be to ask them at the same time to rate the importance they attach to the various issues (do this using a simple rating scale such as 1-3 or 1-5, where you want to allow them a middle choice; use a 1-4 scale when you want to force them to decide between the two extremes).

**Table 2 : Example of open-ended question**

Please provide 3 impacts that you consider this Directive has had and rate the importance you attach to each:	1-5 (1=least important; 5 = most)
1.	
2.	
3.	

Once all the responses have been received, the impacts identified should be amalgamated into one long (numbered) list. At this stage, **all** the impacts provided should be included in the list – no items should be removed due to duplication/similarity.

To illustrate this, consider the following possible impacts resulting from the implementation of the Services Directive (suggested by respondents to an "imaginary" consultation).

**Figure 5 : Example of issue listing**

N°	Impact	Rating
1	increased employment	5
2	lower set-up costs	2
3	increased migration	3
4	improved consumer confidence	2
5	increased standards	1
6	higher quality of services	3
7	create jobs in health sector	4
8	less administrative red-tape	2
9	increased market share	4
10	need to recruit outside own territory	5
11	larger customer base	5
12	availability of qualified staff from over-seas	3
13	need for more nurses	2
14	improved choice of professional services	1
15	fewer consumer complaints	3
16	upward trend in immigration	1
17	less complex legal issues	3
18	more competition from foreign companies	4
19	increased labour costs to ensure quality	2
20	loss of customers to cheaper foreign competitor	2
21	better quality services	2
22	reduced prices for consumers	4
23	better choice of professional services	4
24	costs in meeting improved standards	2
25	increased investment required for market expansion	5
26	lack of appropriately qualified staff	1
27	easier and cheaper set-up administration	2
28	rising employment costs	3
29	uncertainty regarding tax & social welfare issues re: staff	4
30	continuing language barriers re: staff and customers	1

The respondents also ranked the importance of each impact giving scores from 1 to 5 (where 1=least important; 5 = most). These ratings will be used further in step 6.

Outcome of step 1: an exhaustive list of possible impacts, relevant issues, etc.<sup>15</sup>

<sup>14</sup> A simple tool for on-line consultation using open ended questions available at the Commission is IPM.

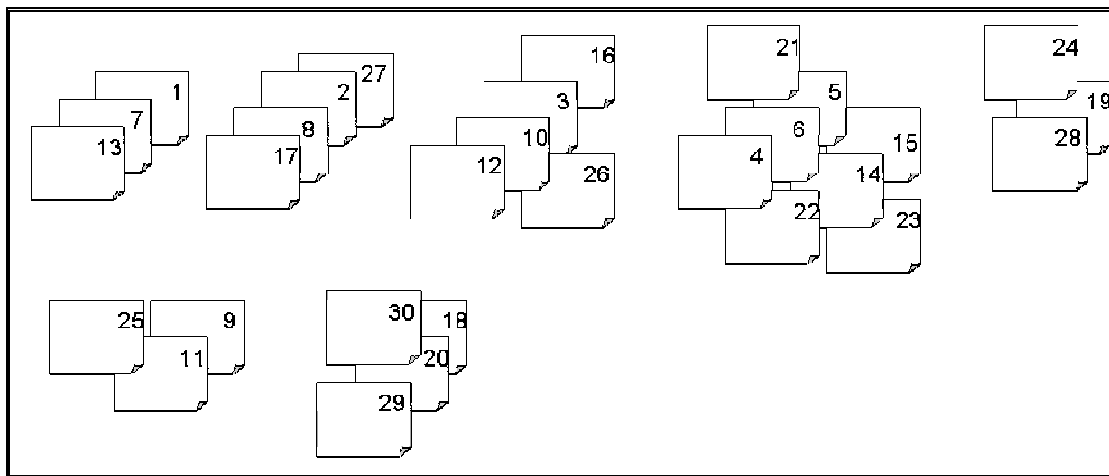
<sup>15</sup> In this case, in order to facilitate the exercise, we limited the number of impacts to 30 but in reality this list might be much longer depending on what the respondents in a stakeholder consultation consider important.

## Step 2 : Arranging relevant issues / impacts into groupings

Once all the issues have been gathered they then need to be grouped together according to their similarities. This works best if a number of individuals participate<sup>16</sup>, first creating their own groups, and then discussing their suggestions, confirming similarities and ascertaining any ambiguities. The exercise can be repeated as many times as necessary to organise similar ideas into "concept clusters". In this way large numbers of issues (e.g. impacts) can be reduced to say, 5 manageable clusters for evaluation.

So continuing our example, each evaluation team member arranges the list into what they consider to be similar response groupings. One member, let's call him/her: team member "A" grouped the responses as follows:

**Figure 6 : Example of initial response grouping**



S/he decided that the three impacts:

- ⇒ Impact 1: increased employment
- ⇒ Impact 7: create jobs in health sector
- ⇒ Impact 13: need for more nurses

could be grouped together, as they all deal with employment and

- ⇒ Impact 3: increased migration
- ⇒ Impact 10: need to recruit outside own territory
- ⇒ Impact 12: availability of qualified staff from over-seas
- ⇒ Impact 16: upward trend in immigration
- ⇒ Impact 26: lack of appropriately qualified staff

make up another group relating to migration. Interestingly, another member of the evaluation team (team member B) thought that the issues regarding overseas staff were linked to employment, so s/he created one group from all these impacts rather than two.

Outcome of step 2: there will be a different set of groupings for each team member, which may or may not be alike; for a team composed of 3 evaluators we will have up to 3 different groupings of impacts. The number of groupings will vary depending on the number of team members.

<sup>16</sup> E.g. for internal evaluation it can be done separately by each member of the Steering Group, or for the external ones, every member of an evaluation team does his/her own groupings.



**Step 3 : Plotting groupings into similarity matrix**

Now, each team member should plot their own results into a similarity matrix (a square grid where the long list of impacts are listed both horizontally and vertically). A "1" is placed in the row/column intersection of two impacts which have been placed in the same group; if the impacts are not in the same group, a "0" should be entered. The mapping of each impact with itself will always be "1", creating a diagonal line of "1"s cutting the matrix in half.

Figure 7 shows the similarity matrix created from the groups identified by team member A in stage 2.

**Figure 7 : Example of an "individual similarity matrix" reflecting issues plotted together**

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	
1	1						1						1																		
2		1						1									1												1		
3			1							1		1					1										1				
4				1	1	1							1	1									1	1							
5				1	1	1							1	1								1	1	1							
6				1	1	1							1	1								1	1	1							
7	1						1						1																		
8		1						1									1												1		
9								1	1	1																1					
10		1							1	1	1					1										1					
11									1	1	1															1					
12			1							1	1	1					1										1				
13	1						1					1																			
14				1	1	1							1	1									1	1							
15				1	1	1							1	1								1	1	1							
16			1							1		1															1				
17		1						1				1					1												1		
18																		1		1									1	1	
19																		1		1				1					1		
20				1	1	1							1	1					1	1		1	1	1					1	1	
21				1	1	1							1	1					1	1		1	1	1					1	1	
22				1	1	1							1	1					1	1		1	1	1					1	1	
23				1	1	1							1	1					1	1		1	1	1					1	1	
24																		1						1							
25								1	1	1															1		1				
26			1						1	1	1					1											1				
27	1							1		1							1												1		
28																		1		1				1						1	
29																			1	1		1								1	1
30																				1	1									1	1

1-30 = list of 30 impacts gathered from stakeholder consultation

Outcome of step 3: for a team composed of 3 evaluators we will have 3 different (or not) individual similarity matrices.

**Step 4 : Merging individual similarity matrices into a single matrix**

In this step, the individual similarity matrices should now be condensed into a single matrix to establish the team's view of the common/related elements in the long list. As some responses may fall into potentially different groups, this should lead to a discussion of the various categorisations, using the most popular views to help place any ambiguous responses. This may need to be discussed by the evaluation team to find a consensus on ambiguous issues. Figure 8 below shows such a 'merged' matrix into which the results of 3 individual matrices were introduced.

How to read this matrix:

- The values on the diagonal intersections (e.g. impact1 mapped to impact 1, impact 2 mapped to impact 2, etc) will equal the number of members in the team whose matrices are being merged, i.e. in this example, 3.
- where all 3 evaluators grouped the same 2 impacts together, the value will also be 3 (e.g. impacts number 4 (improved consumer confidence) and number 5(increased standards ) were placed together by all 3 evaluators, so the value in '5 down, 4 across' is 3 as is the value in "4 down, 5 across").
- Value 2 inserted in a given field indicates that 2 individual evaluators grouped those impacts together, and so on.

**Figure 8 : Example of a " group similarity matrix" – reflects views of a 3 person evaluation team<sup>17</sup>**

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	
1	3	0	1	0	0	0	3	0	0	1	0	1	3	0	0	1	0	0	0	0	0	0	0	0	0	1	0	0	0	0	
2	0	3	0	0	0	0	0	3	1	0	1	0	0	0	0	3	0	0	0	0	0	0	0	0	0	0	3	0	0	0	
3	1	0	3	0	0	0	1	0	0	2	0	3	1	0	0	3	0	0	0	0	0	0	0	0	0	2	0	0	0	0	
4	0	0	0	3	3	3	0	0	0	0	0	0	0	2	3	0	0	0	0	0	3	2	2	0	0	0	0	0	0	0	
5	0	0	0	3	3	3	0	0	0	0	0	0	0	2	3	0	0	0	0	0	3	2	2	0	0	0	0	0	0	0	
6	0	0	0	3	3	3	0	0	0	0	0	0	0	2	3	0	0	0	0	0	3	2	2	0	0	0	0	0	0	0	
7	3	0	1	0	0	0	3	0	0	1	0	1	3	0	0	1	0	0	0	0	0	0	0	0	0	1	0	0	0	0	
8	0	3	0	0	0	0	0	3	1	0	1	0	0	0	0	0	3	0	0	0	0	0	0	0	0	0	3	0	0	0	
9	0	1	0	0	0	0	0	1	3	0	3	0	0	0	0	0	1	0	0	0	0	0	0	0	0	2	0	1	0	0	
10	1	0	2	0	0	0	1	0	0	3	0	2	1	0	0	2	0	0	1	0	0	0	0	0	0	3	0	1	0	0	
11	0	1	0	0	0	0	0	1	3	0	3	0	0	0	0	0	1	0	0	0	0	0	0	0	2	0	1	0	0	0	
12	1	0	3	0	0	0	1	0	0	2	0	3	1	0	0	3	0	0	0	0	0	0	0	0	0	2	0	0	0	0	
13	3	0	1	0	0	0	3	0	0	1	0	1	3	0	0	1	0	0	0	0	0	0	0	0	0	1	0	0	0	0	
14	0	0	0	2	2	2	0	0	0	0	0	0	0	3	2	0	0	0	0	0	2	1	3	0	0	0	0	0	0	0	
15	0	0	0	3	3	3	0	0	0	0	0	0	0	2	3	0	0	0	0	0	3	2	2	0	0	0	0	0	0	0	
16	1	0	3	0	0	0	1	0	0	2	0	3	1	0	0	3	0	0	0	0	0	0	0	0	0	2	0	0	0	0	
17	0	3	0	0	0	0	0	3	1	0	1	0	0	0	0	0	3	0	0	0	0	0	0	0	0	0	3	0	0	0	
18	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	3	0	3	0	0	0	1	1	0	0	0	2	2	
19	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	3	0	0	0	0	2	0	1	0	3	0	0	
20	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	3	0	3	0	0	0	1	1	0	0	0	2	2	
21	0	0	0	3	3	3	0	0	0	0	0	0	0	2	3	0	0	0	0	0	3	2	2	0	0	0	0	0	0	0	
22	0	0	0	3	3	3	0	0	0	0	0	0	0	2	3	0	0	0	0	0	3	3	2	0	0	0	0	0	0	0	
23	0	0	0	2	2	2	0	0	0	0	0	0	0	3	2	0	0	0	0	0	2	1	3	0	0	0	0	0	0	0	
24	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	2	1	0	0	0	3	1	0	0	0	2	1	1
25	0	0	0	0	0	0	0	0	2	0	2	0	0	0	0	0	0	1	0	1	0	0	0	1	3	0	0	0	1	1	
26	1	0	2	0	0	0	1	0	0	3	0	2	1	0	0	2	0	0	1	0	0	0	0	0	0	3	0	1	0	0	
27	0	3	0	0	0	0	0	3	1	0	1	0	0	0	0	0	3	0	0	0	0	0	0	0	0	0	3	0	0	0	
28	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	3	0	0	0	0	0	2	0	1	0	3	0	0	
29	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	0	2	0	0	0	1	1	0	0	0	3	3	
30	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	0	2	0	0	0	1	1	0	0	0	3	3	

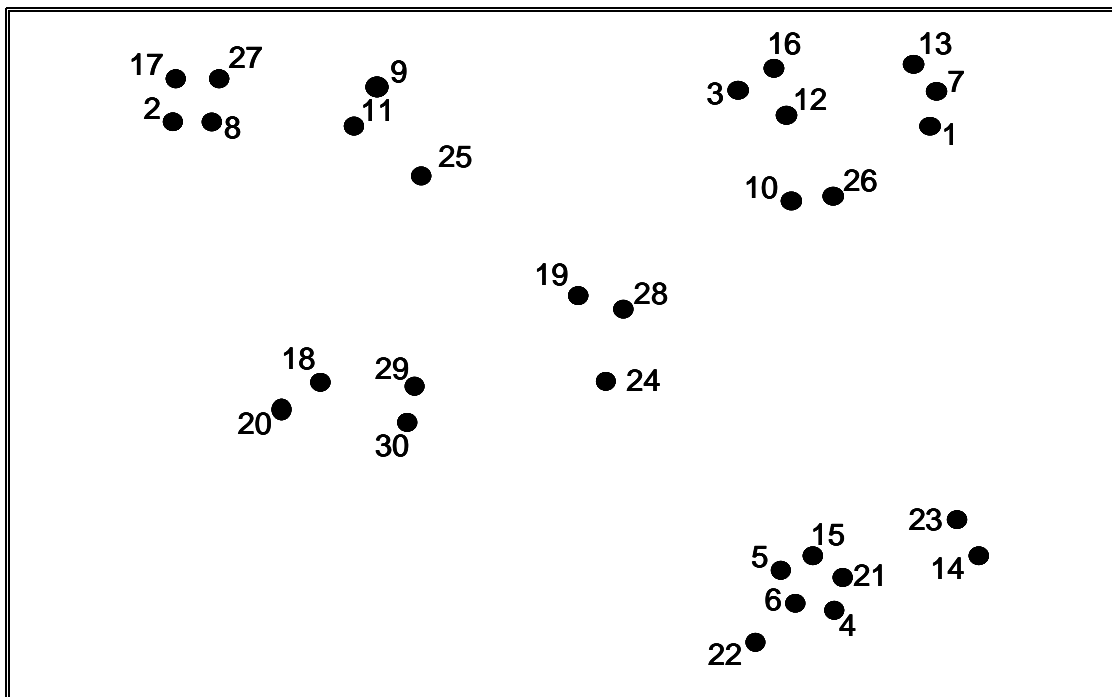
Outcome of step 4: Matrix showing definitive groupings common to an evaluation team.

<sup>17</sup> Note: the matrix is symmetrical along the diagonal, so the row total for a given impact will always equal the column total for the same impact. The maximum number in a given cell is equal to the number of evaluators in the team. Individual row/column totals are unlikely to be equal as they reflect the team's opinion of which impacts can be grouped together.

**Step 5 : Translating the matrix into a 'point map' to select the clusters of impacts**

This matrix should then be plotted in a point map by placing each issue close to those it was matched up with. The larger the number in the matrix the closer these impacts should be to each other reflecting the fact that they were placed together more frequently. For example, in Figure 9, items that scored 3 in the matrix are plotted within a 1cm diameter of each other, those that scored 2 within a 2 cm radial, those that scored 1 were plotted 3 cm radius and those that scored 0 were plotted outside this maximum radius. This should help to establish clusters of points that can be labelled with a common theme (similarity).

**Figure 9 : Example of a point map**

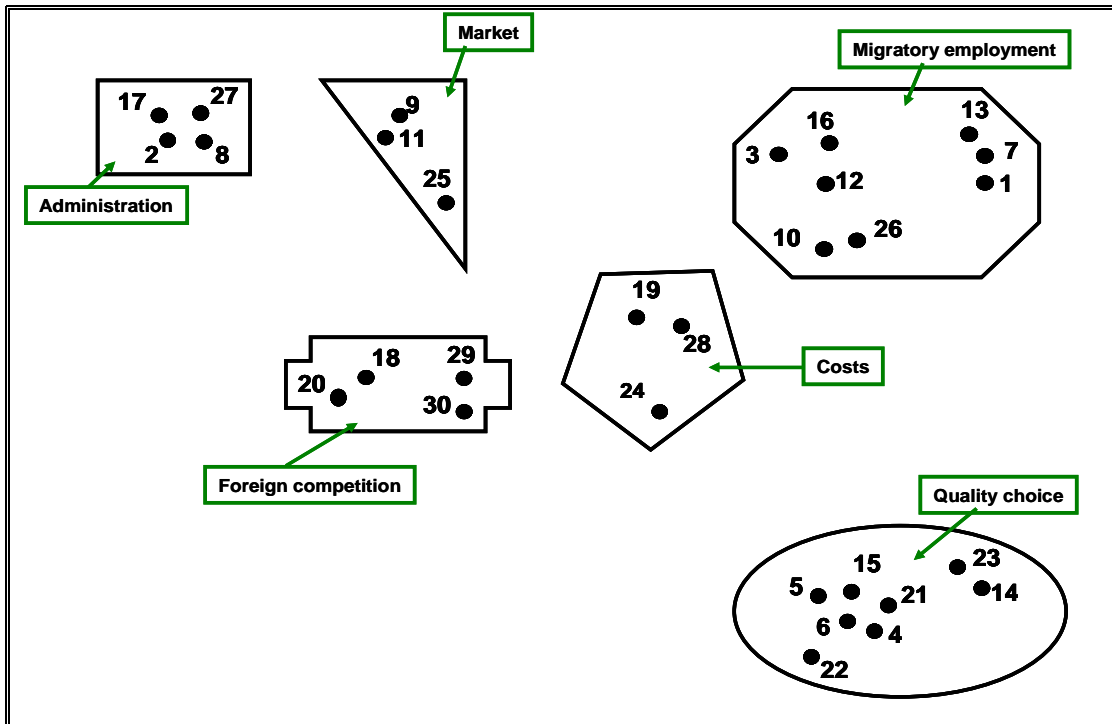


Based on the closeness of the points to each other, it should now be possible to select appropriate clusters and label them. (N.B. this can be a fairly subjective task!). For example one of the clusters groups the following impacts:

- ⇒ 2 - lower set-up costs,
- ⇒ 8 - less administrative red-tape,
- ⇒ 17 - less complex legal issues and
- ⇒ 27 - easier and cheaper set-up administration,

which could then be labelled "administration".

Figure 10 : Example of a cluster map

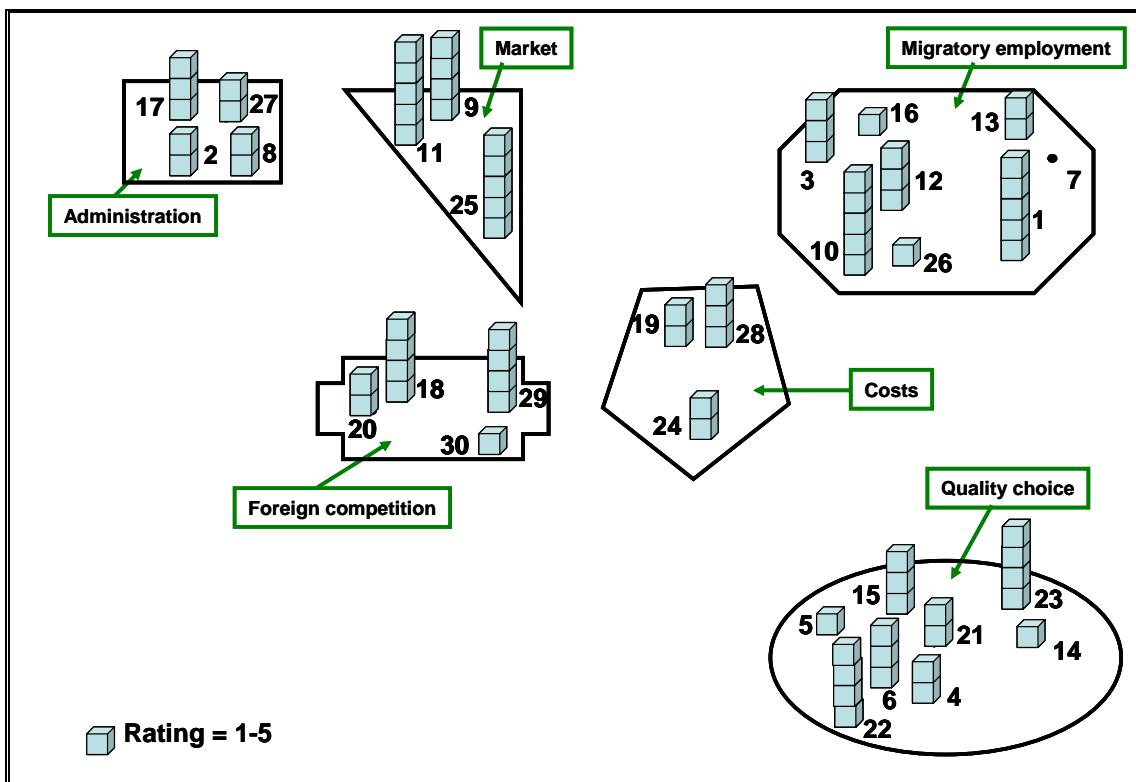


Outcome of step 5: Thematic clusters of grouped impacts.

**Step 6 : Rating impacts and calculating average scores for clusters**

Each impact was also rated by the respondent during the 1<sup>st</sup> step (listing main impacts / relevant issues, etc). So for example according to Figure 5, impact number 17 (less complex legal issues) was given 3 in the scale from 1 (least important) to 5 (most). These ratings for each impact should now be plotted in the point map (see Figure 11).

Figure 11 : Example of point rating map



As a final step, it is now possible to calculate the average rating scores given to each cluster to establish the perceived importance of each grouping.

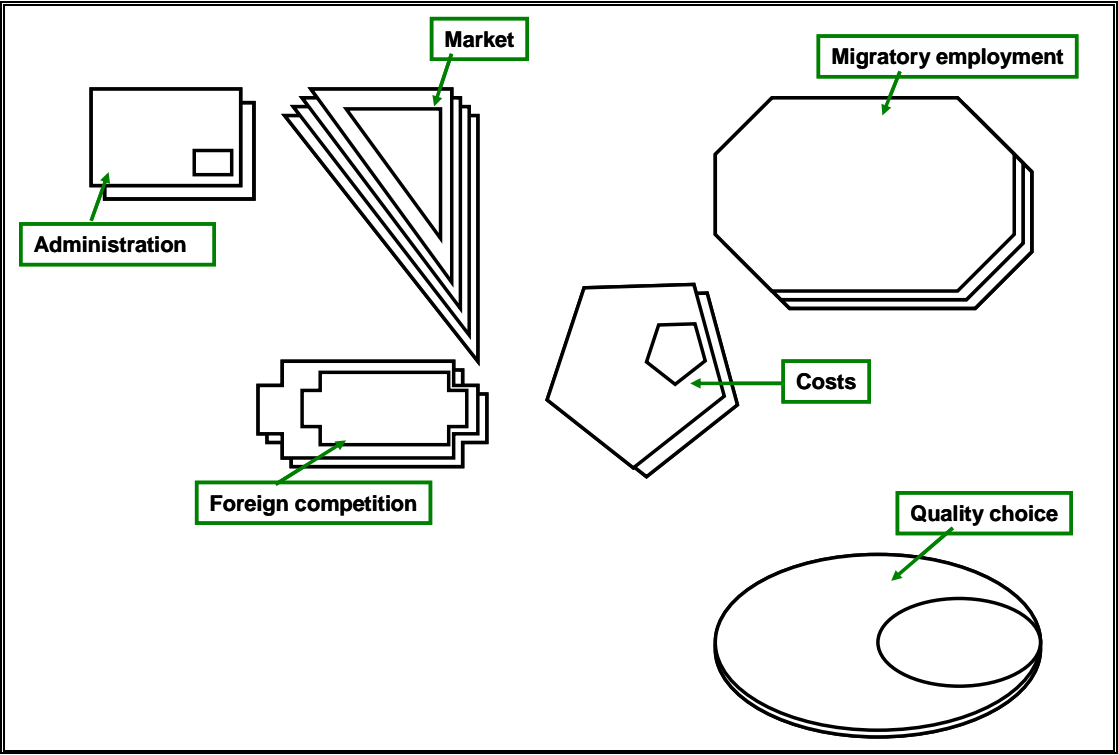
For example, "administration": impacts 2, 8, 17 and 27 have been placed close together with ratings of 2, 2, 3 and 2 respectively; hence the average for their cluster is  $(2+2+3+2)/4 = 2.25$ . The table below shows the calculated average rating for all the clusters identified in our example.

**Table 3 : Average rating per cluster**

Clusters :	Average ratings:
Administration	2.25
Market	4.67
Migratory employment	3.00
Costs	2.33
Foreign competition	2.75
Quality choice	2.50

If you want, this can also be depicted in the cluster analysis map:

**Figure 12 : Example of a cluster analysis**



As a result of this exercise, complex intervention logics may be adjusted to reflect the clustering identified. By using the label/theme in the logic model, it should be possible to establish a clearer view of the more important strands requiring focused examination, thus reducing the complexities of legislative measures to more manageable categories for evaluation. So in practical terms our example shows that the most important issue that will need special focus during the evaluation exercise is a "market" cluster (the highest rating of 4.67) that includes impacts like increased market share, larger customer base and increased investment required for market expansion. Other issues identified as important include concern migratory employment and foreign competition.

Outcome of step 6: A list of areas / issues to be focused on during evaluation.

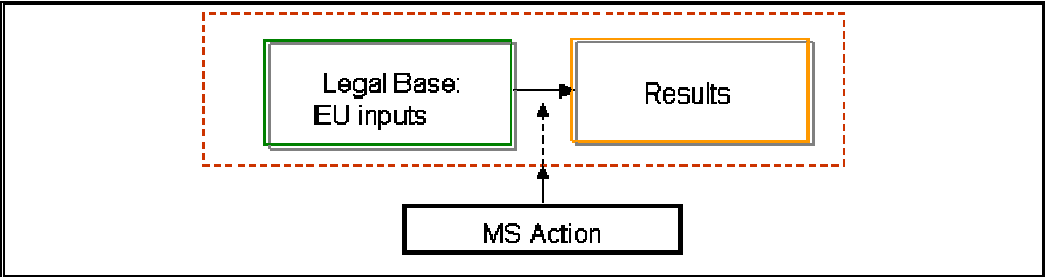
# Annex 8 - Social Network Analysis

Expanding on the logical frameworks, this tool can be particularly effective in understanding complex causal relationships. For example, it may help clarify the links between Member State action and the wide ranging impacts, positive & negative, on various groups, while also taking external factors into account. This may be particularly true in the case of directives or soft law. It can help identify and rate the causal links establishing a better understanding of their nature and the consequence their interaction has on the cause-effect. It may make subsequent choice of case studies or decisions on where to focus further in-depth enquiry easier.

## Practical usage of social network analysis

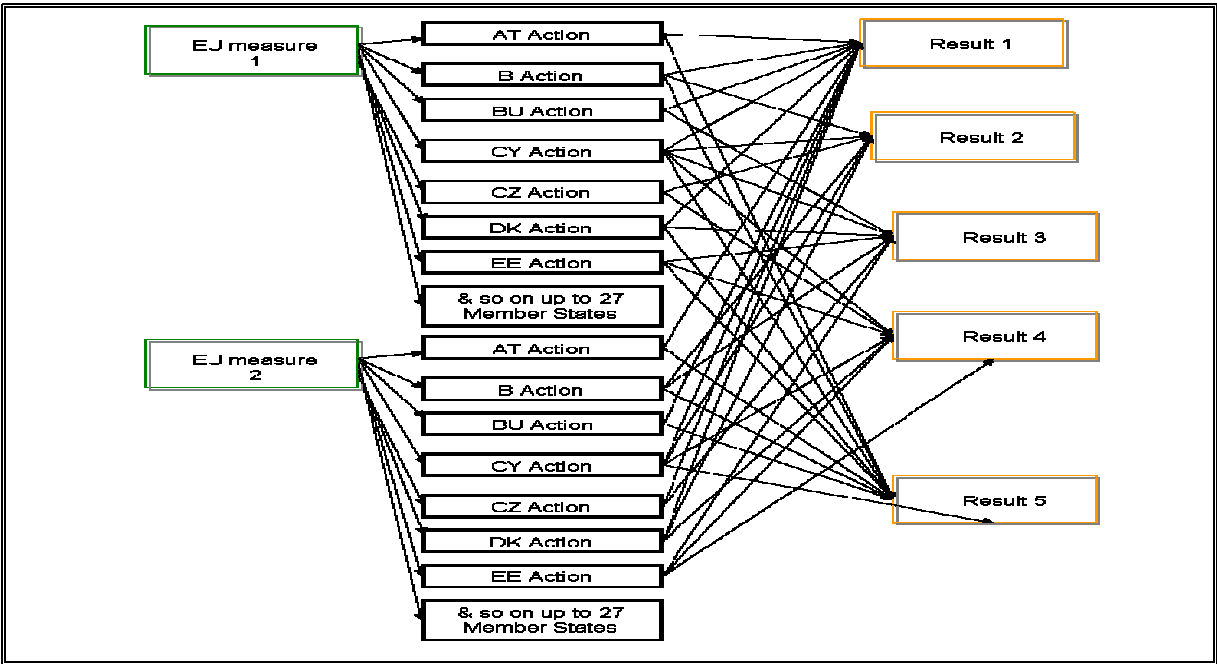
One of the major complexities in evaluating legislation (particularly so-called "soft-law") as opposed to programmes is that the relationship between the measures of EU legislation and the results they anticipated can often be significantly altered by the manner in which Member States decided to implement them (as explained in the guide - Stage 2, Section i) on Intervention Logic).

**Figure 13 : The causal relationships examined by social network analysis**



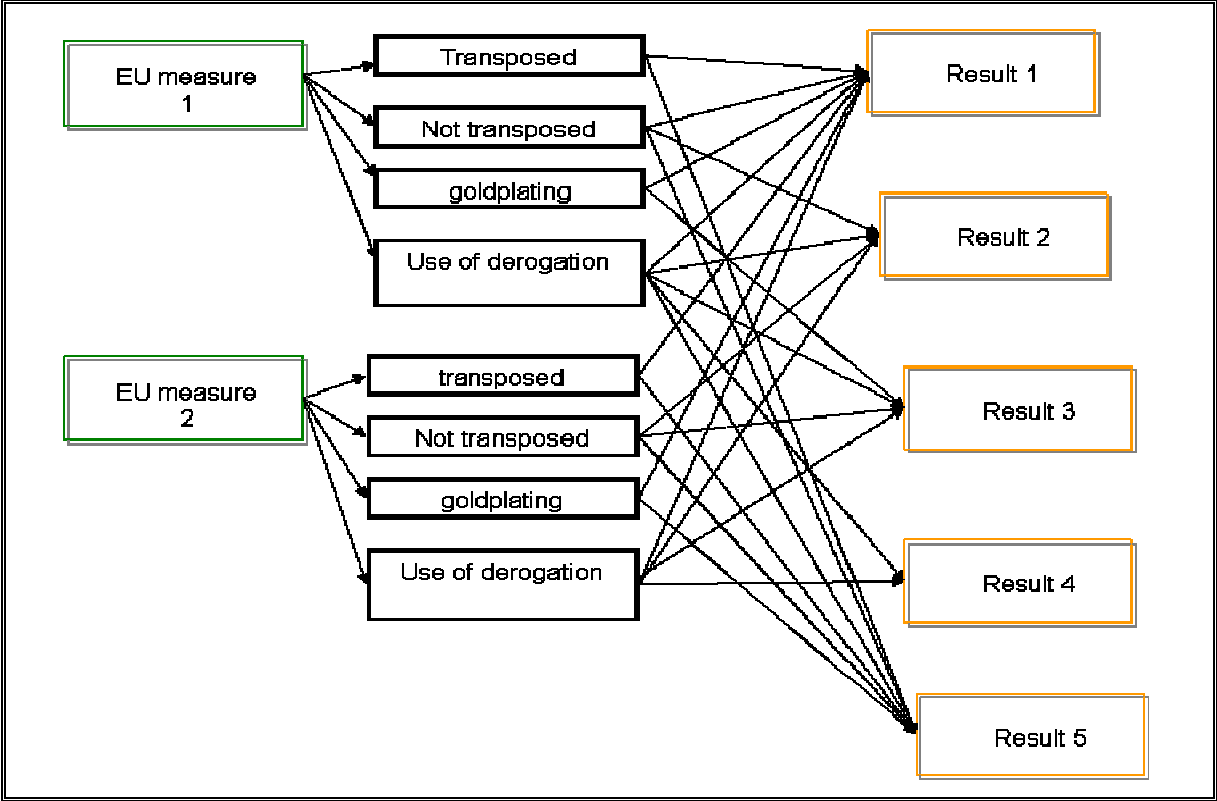
It is important to understand HOW Member State action can deflect the cause-effect arrow – but this needs to be done in a "manageable" way. If this were to be included in the Intervention Logic directly, it would no longer be possible to follow the logic as the model would be confused by all the different actions (boxes and arrows!):

**Figure 14 : Example of why the MS action is extracted in the intervention logic**



Clustering the Member State action - for example, into 4 thematic areas (e.g. transposed, not transposed, gold-plating & use of derogation) - can help make the intervention logic more manageable, but the diagram still remains superficial in terms of the evaluative information it provides.

**Figure 15 : Example of why the MS action is extracted in the intervention logic even when clustered**



However, using the established clusters (such as the 4 types of MS action suggested for the example above) to analyse the relationship arrows can provide you with a more informative picture against which you may view strands for more in-depth enquiry or where the evaluation should focus.

By plotting a matrix of measures by 27 Member States using the clusters (probably identified in the concept mapping, especially if the issue of gold-plating is also broken down further into common clusters), it is then possible to map out the strength of the relationship between the measure and the MS action. It is expected that such information will be available either in Implementation Reports or if not, may have to be attained through a targeted preliminary consultation/questionnaire.

In the table below measures A to F are examples of various measures that could be undertaken by a Member State as a result of implementing a Directive. A value of 1 in the table means that a given measure was simply transposed, 2 indicates not transposed, 3 means gold-plated and 4 that the derogation option was used.

**Table 4 : Measure by Member State matrix for social network analysis**

	Measure A	Measure B	Measure C	Measure D	Measure E	Measure F
Member States	prohibited Requirements	conditions for granting authorisation	conditions for establishing authorisation schemes	selection from amongst several candidates	authorisation procedures	duration of authorisation
AT	1	1	1	1	1	1
BE	1	1	1	1	1	1
BG	2	2	2	2	2	2
CY	1	2	1	2	2	2
CZ	1	3	3	3	3	1
DK	1	3	3	3	1	1
EE	1	3	3	3	1	1
FI	1	3	3	3	3	3
FR	1	3	3	3	1	1
DE	1	3	3	3	1	1
GR	1	3	3	3	3	1
HU	1	3	3	3	3	1
IE	4	4	4	4	1	1
IT	2	2	2	2	2	2
LV	1	1	1	1	1	1
LT	1	1	1	1	1	1
LU	1	3	3	3	3	1
MT	1	3	3	3	3	1
NL	1	3	3	3	3	1
PL	1	1	1	1	1	1
PT	1	3	3	3	3	1
RO	1	3	3	3	3	3
SI	1	3	3	3	3	2
SK	1	2	2	1	2	2
ES	1	3	3	3	3	1
SE	1	3	3	3	3	1
UK	3	3	3	3	3	3
<b>Summary:</b>						
transposed	23	5	6	6	10	19
not transposed	2	4	3	3	4	5
gold-plated	1	17	17	17	13	3
use of derogation	1	1	1	1	0	0
<b>Legend:</b> 1 transposed; 2 not transposed; 3 gold-plated; 4 use of derogation						

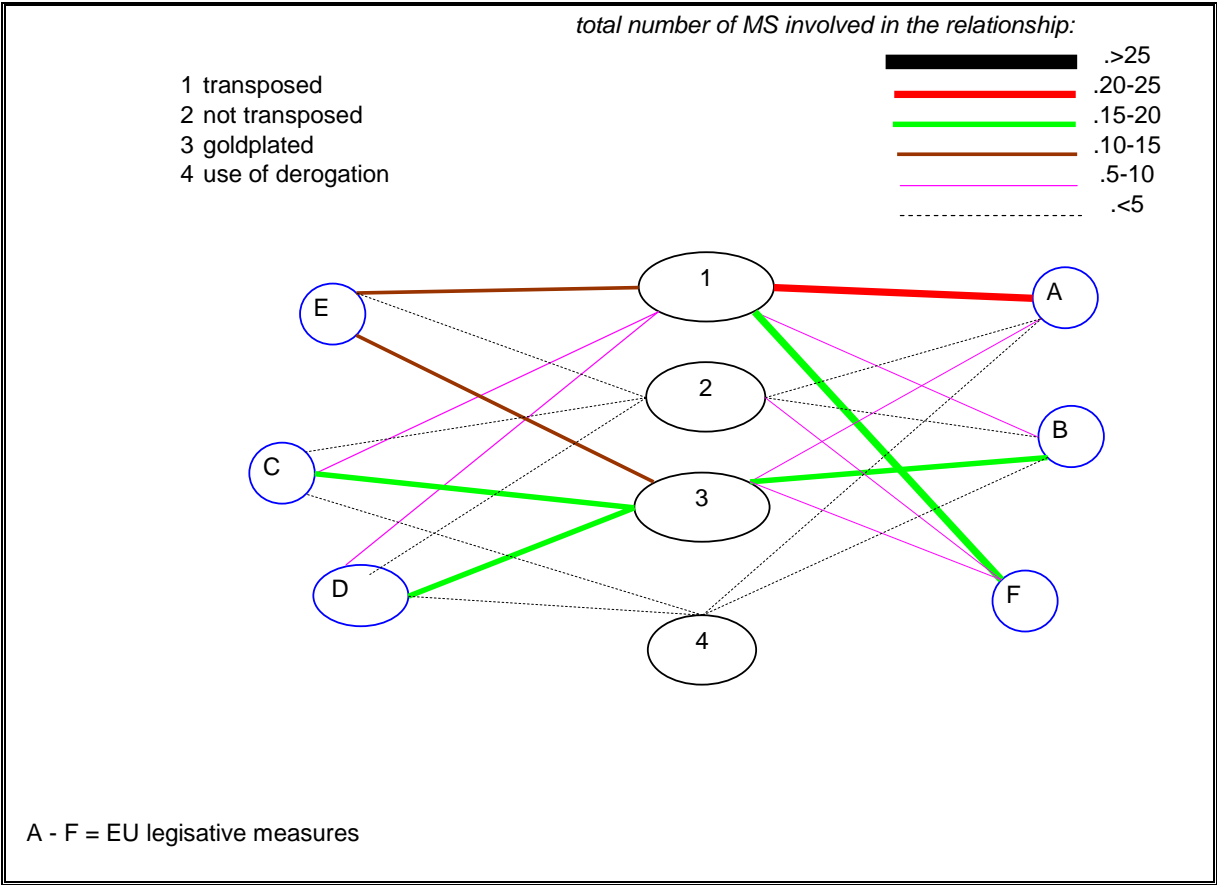
In this hypothetical example, the table above already shows that Measures B, C and D have been gold-plated by a large number of Member States (all score 17), while Measures A and F have more often been simply directly transposed (scores 23 and 19 respectively).

From this it may be worth focusing further investigation as to *how* Measures B, C and D have been gold-plated and the effect on the anticipated results. It also confirms that the evaluation will only need to focus on verifying whether Measures A and F actually did produce the expected effects as these will have been implemented (in the majority) as intended. Measure E was equally either transposed or gold-plated, and therefore may be a good source for a case study comparison.



This is perhaps more easily represented in the network analysis diagram below:

**Figure 16 : Example of social network analysis illustrating how MS action has affected EU measures**



To plot this diagram, the 6 measures (A-F) and 4 possible transposition behaviours (1-4) are set out and the linked by patterned or coloured lines representing the number of MS which have adopted particular transposition behaviour relative to that measure. For example, Measure E has been transposed (1) by a total of 10 MS, so E and 1 are linked by a brown line (representing 10-15 MS); Measure C has been "gold-plated" (3) by 17 MS in total, so C and 3 are linked with a green line (representing 15-20 MS).

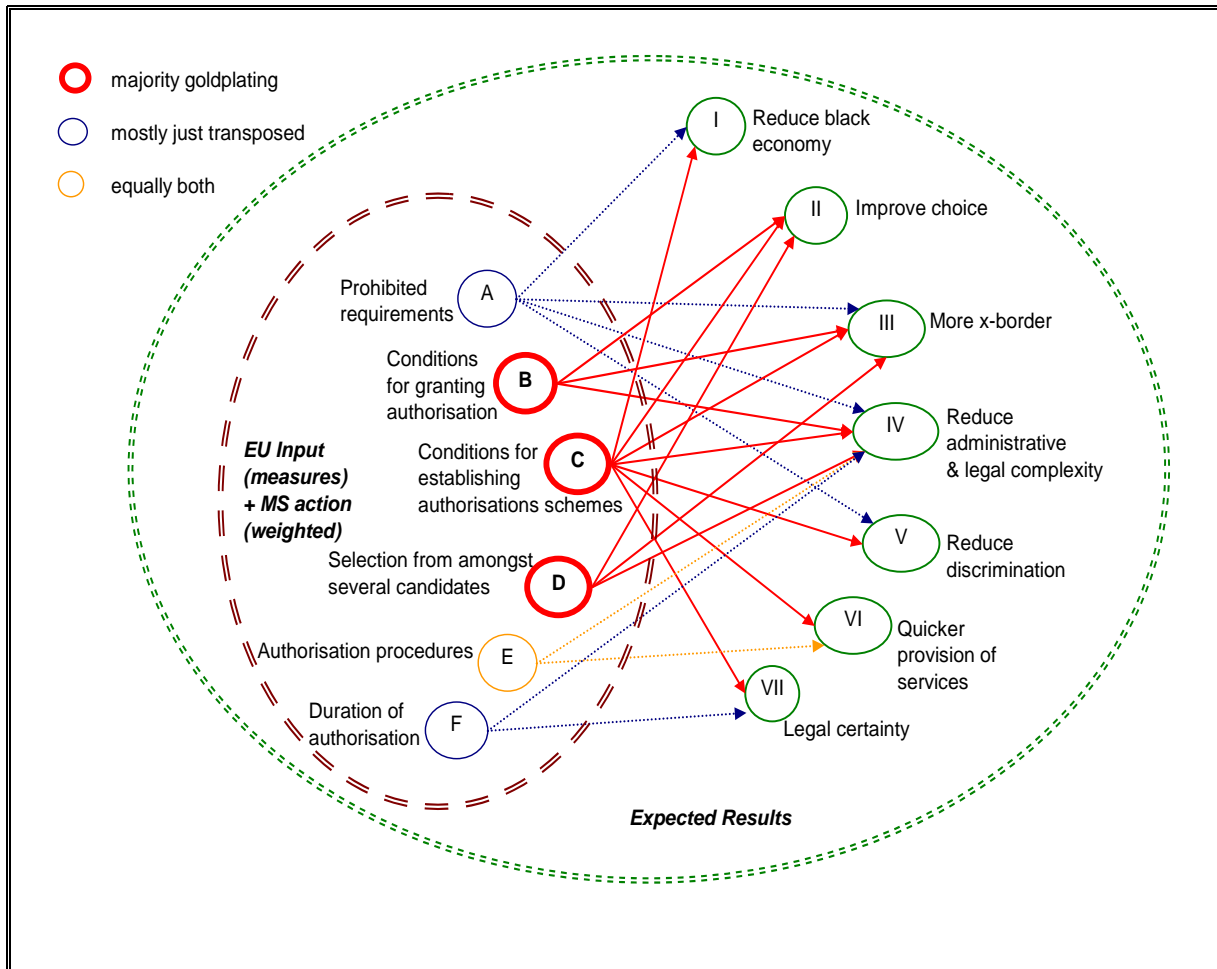
It is now possible to see the obvious results of measures that were mainly gold-plated or transposed, without losing the context of the other transposition behaviour.

If this information is now transferred back to the intervention logic, the logic starts to take on a 3-dimensional picture, where the specific threads become more readily identifiable and the additional elements can be also considered.

Having established from the Social Network Analysis above, that transposition behaviour 2 (not transposed) and 4 (use of derogation) are of minor importance in this instance (all liaisons equal less than 5), one can attribute three transposition behaviour patterns to the causal links in the intervention logic: majority gold-plating, mostly just transposed and equally both.

By applying these patterns (or weightings) to the relevant part of the intervention logic, the element relating to MS action that was initially extracted to facilitate building a workable representation, can now be reintroduced without complicating the flow of cause-effect:

**Figure 17 : Example: applying the social network analysis back to an intervention logic**



This can be an excellent aide for structuring interview guides, wider surveys, etc. helping focus data collection on key issues. This technique can be used again to explore other relationship links across the intervention logic where a better understanding of the complex interlinking of elements is required, principally for example, in the analysis of positive and negative impacts on a variety of target groups.

## Annex 9 - Checklist for dissemination strategy

The following checklist includes the main tasks to be done after an evaluation has been completed. The tasks mentioned below are only those obligatory steps for each evaluation but of course the target groups should be tailored to each study and the list below should be completed for each evaluation.

The table includes tasks related not only to a dissemination strategy of an evaluation study itself but to the dissemination of other related documents as well, e.g. quality assessment, action plan, etc.

<b>TASK</b>	<b>WHO ?</b>	<b>DONE ?</b>
Send final report to participants of the Steering committee	OU	<input type="checkbox"/>
Send a note including main conclusions and recommendations to Director General with copy to other Directors	OU	<input type="checkbox"/>
Publish final report (in PDF format) on the external website of DG MARKT	Webmaster A4	<input type="checkbox"/>
Publish final report (in PDF format) link on the internal website of DG MARKT	Secretariat of B2	<input type="checkbox"/>
Send executive summary to main stakeholders / contributors to the findings (mentioning the website where full report is available)	OU	<input type="checkbox"/>
Send executive summary to EP, EU Ombudsman, ...	OU	<input type="checkbox"/>
Upload evaluation findings, conclusions and recommendations together with relevant documents and reports into the EIMS database (Evaluation Information and Monitoring System managed by DG BUDG)	OU (through B2)	<input type="checkbox"/>
Send Final Quality assessment to DG BUDG and unit B.2	OU	<input type="checkbox"/>
Prepare Action Plan on the recommendations of the final report with provisional calendar and comments on actions proposed	OU	<input type="checkbox"/>
Send Action Plan to hierarchy, B.2 and A1	OU	<input type="checkbox"/>
...		<input type="checkbox"/>

OU – Operational Unit

## Annex 10 - Action Plan

### ACTION PLAN

Evaluation of \_\_\_\_\_

<b>Finding</b>	<b>Recommendation</b>	<b>Concrete actions to be taken and provisional calendar</b>	<b>Follow up (after 1 year)</b> <i>(Were planned actions implemented? If yes, when? If not, why not? ...)</i>