

QUALITY ASSESSMENT FORM¹

Ex-post evaluation of social legislation in road transport and its enforcement

Study contract no. MOVE/D3/2014-256

Final report

DG/Unit MOVE D3

Official(s) managing the evaluation : Ewa Ptaszyńska, Alicja Hübner

Evaluator/contractor Ricardo-AEA (UK) with its consortium partners: TRT (IT) and Milieu (BE)

Assessment carried out by:

Inter-service steering group composed of representatives of: DG MOVE, DG EMPL, DG JUST, DG SG, LS

Scale for scoring: **Poor** **Satisfactory** **Good** **Very good** **Excellent**

Date of the Quality Assessment

¹ Refer to the [‘Guide on Scoring the Criteria’](#) for how to assess each criterion.

(1) RELEVANCE

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

SCORING

Good

Arguments for scoring: The study covers all aspects outlined in the terms of reference, including the scope, key issues, timeframe, target groups and geographical area. The final report provides required information enabling an overall assessment of the functioning of the road transport social legislation and its enforcement regime. However, not all evaluation questions have been provided with comprehensive and clear answers, which was mainly due to limitations in gathering reliable and full-fledged data or subjective character of certain aspects such as those related to assessing working conditions and satisfying needs of the sector. For instance, the need for and magnitude of exemptions and derogations from the current rules are not sufficiently presented. Where quantitative analysis was not possible due to data limitations the contractor provided qualitative analysis instead allowing to answer the evaluation question concerned. The unintended positive and negative effects of the legislation have also been identified. The assessment took into account changes in the problems and needs compared to the situation at the start of the intervention. The evaluation added value to existing policy knowledge.

(2) APPROPRIATE DESIGN

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

SCORING

Very Good

Arguments for scoring: The methodology is appropriate and clearly described. It allows for gathering variety of data and information from different sources and using different tools, for comparing, triangulation, cross-checking and analysing the gathered information in the context of evaluation questions. Design of the evaluation was adequate to the intervention. Contractor adjusted its data collections methods accordingly to the limitations revealed in the study. Limitations and weaknesses of the methods are clearly spelled out in the report.

(3) RELIABLE DATA

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

SCORING

Very good

Arguments for scoring:

The contractor has identified various available information and sources throughout desk and field research. Relevant literature and previous studies have been reviewed and existing monitoring systems were used. Data collection rationale is explained; and it is coherent with the design of the study. Tools and means used to collect and process data (e.g. tailor-made surveys, interviews and case studies) were: selected in relation to criteria specified in the inception phase; 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.

However, there is a problem of insufficient quantitative data for assessing at least certain effects of the legislation in force. In particular quantitative data concerning the period preceding the entry into force of the legislation is unavailable. This does not allow for establishing a thorough evidence-based baseline scenario, which in consequence makes it difficult to quantify the developments observed as a result of the adoption of the legislation. The data collection limitations have been explained by the contractor and where relevant the qualitative information and assessments have been gathered and extrapolation and assumption-based calculations have been performed. Available data in general can thus be considered reliable.

(4) SOUND ANALYSIS

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

SCORING

Good

Arguments for scoring: The data are systematically analysed and where appropriate triangulation and extrapolation is made to cross-check the reliability of data and to address the incompleteness of available statistics. The analysis is focused on the relevant cause/effect relations or on correlations between analysed aspects of the legislation and their relevant factors. However, where causation or correlation is not established the additional analysis is not performed. The analysis uses appropriate quantitative and qualitative techniques, although their use is limited by the unavailability of certain data. The report provides a logical argumentation and presents sufficiently evidence and limitations of the analysis. The context and the inputs from important stakeholders are well taken into accounts.

(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?

SCORING

Good

Arguments for scoring: The findings follow logically the preceding analysis and data, are credible and corroborate the existing knowledge. In the absence of quantitative data the contractor made use of stakeholders' opinions and other qualitative data available. However, the latter technique, by its subjective nature, could not lead to unbiased and evidence-based findings. The differences and contradictions between stakeholders' opinions and available data have been explained and carefully assessed when drawing the main findings.

(6) VALID CONCLUSIONS

Are conclusions non-biased and fully based on findings?

SCORING

Good

Arguments for scoring: The conclusions stem logically from findings and are based on impartial judgment. They are satisfactorily addressed to the evaluation questions and orderly presented. However, some conclusions are not definitive leaving the room for further questions about impacts.

(7) HELPFUL RECOMMENDATIONS

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

SCORING

Good

Arguments for scoring: The recommended improvements to the legislative framework follow logically the analysis performed and conclusions established. Recommendations are realistic, prudent and potentially useful. Some of them are made on the basis of stakeholders' views and qualitative assessment, where quantitative data and analysis were not available.

(8) CLARITY

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

SCORING

Good

Arguments for scoring: The report describes the legislation being evaluated, its context, the evaluation purposes, methodology, contextual limitations, market and sector developments, findings, analysis, conclusions and recommendations in a structured manner. There are several cross-references between the sections of the report, which does not facilitate the reading. This aspect as well as the significant length of the report is partly justified by the number and the level of details and complexity of the evaluation questions posed. Key messages are summarized and highlighted. The report contains a concise executive summary which includes general conclusions and recommendations.

OVERALL ASSESSMENT

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?
- 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?

The final evaluation report is of overall good quality. The evaluation fulfills contractual obligations and the contractor followed meticulously the Commission's comments and suggestions. Some inconsistencies appear in the report. The findings and conclusions are reliable and all limitations concerning the data availability and subsequently the comprehensiveness of analysis have been explained and duly justified. The report demonstrates appropriate results based on triangulation and extrapolation of available data. The information in the report substantiates and enriches the existing knowledge and is useful for designing further intervention.

Given the contextual and contractual constraints encountered:

- What lessons can be learned from the evaluation process?

The evaluation questions should be designed in a simpler and more straightforward manner taking account of the (un-)availability of data and subjective nature of certain aspects of evaluation.

GUIDE ON SCORING THE CRITERIA

This list of indicators aims at helping to score each criterion and it can also assist in the process of developing the argumentation underpinning the score.

The indicators may be adapted according to the specificities of each evaluation and some indicators may be omitted and others added when appropriate.

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 while the concurrent accomplishment of those at the end of the list may suggest a higher score).

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*).

