



Fiche outil 7

BàO - Evaluation

Grille qualité rapport provisoire & rapport final

MEAL Unit – 2020

1. Pourquoi et comment apprécier la qualité d'une évaluation ?

Pourquoi ?

Il convient d'apprécier la qualité de l'évaluation pour vérifier si l'évaluation satisfait les exigences en matière de qualité (ces exigences sont déclinées en p2).

Quand ?

Ce document doit être communiqué à l'évaluateur avant qu'il entame la rédaction du rapport provisoire, afin qu'il puisse s'autoévaluer

Lors de la remise du rapport provisoire de l'évaluateur au COPIL, ce dernier utilise cette grille qualité pour apprécier le rapport et faire ses retours

Comment ?

L'analyse qualité d'une évaluation ne se limite pas à la qualité du rapport final, elle dépend de la façon dont on a su gérer la **qualité de l'ensemble du processus évaluatif (cf. FO_1 : processus évaluatif)** et concerne donc l'ensemble des acteurs.

Toutefois, **le rapport final et sa synthèse** étant les résultantes de l'ensemble du processus évaluatif, il est nécessaire que ceux-ci soient conformes à certains critères.

La Commission Européenne propose 9 critères pour apprécier la qualité d'un rapport final (cf. page 2). Les critères qualité doivent être présentés à l'évaluateur en amont (cf Pack Evalueur) afin qu'il puisse les prendre en compte lors de la rédaction du rapport.

- Pour savoir comment apprécier les résultats de la grille, voir en page 3
- Pour une guidance détaillée de chacun des 9 aspects, voir page 4

2. Grille synthétique d'appréciation de la qualité des travaux d'évaluation

Inacceptable Acceptable Bon Excellent

1. **La satisfaction des demandes** : L'évaluation traite-t-elle de façon adéquate les demandes d'information formulées par les commanditaires et correspond-elle aux TDR ?
Pourquoi ? :

2. **La pertinence du champ** : La raison d'être du projet, ses réalisations, ses résultats sont-ils étudiés dans leur totalité, y compris ses interactions avec d'autres politiques et ses conséquences imprévues ?
Pourquoi ? :

3. **L'adéquation de la méthodologie** : La conception de l'évaluation est-elle adaptée et adéquate pour fournir les résultats nécessaires (avec leurs limites de validité) pour répondre aux principales questions évaluatives ?
Pourquoi ? :

4. **La fiabilité des données** : Les données primaires et secondaires collectées ou sélectionnées sont-elles adaptées ? Offrent-elles un degré suffisant de fiabilité par rapport à l'usage attendu ?
Pourquoi ? :

5. **La solidité de l'analyse** : L'analyse des informations quantitatives et qualitatives est-elle conforme aux règles de l'art, complète et adaptée afin de répondre correctement aux questions évaluatives ?
Pourquoi ? :

6. **La crédibilité des résultats** : Les résultats découlent-ils logiquement et sont-ils justifiés par l'analyse des données et par des interprétations basées sur des hypothèses explicatives présentées avec soin ?
Pourquoi ? :

7. **Des conclusions justifiées** : Les conclusions sont-elles claires? Sont-elles fondées sur des résultats crédibles?
Pourquoi ? :

8. **Des recommandations impartiales** : Les recommandations sont-elles justes, non-biaisées par des considérations personnelles ou partisans et sont-elles assez détaillées pour être concrètement mises en œuvre ?
Pourquoi ? :

9. **La clarté du rapport** : Le rapport décrit-il le contexte et le but du programme/projet évalué ainsi que son organisation et ses résultats de façon telle que les informations fournies soient aisément compréhensibles ?
Pourquoi ? :

Attention :

L'appréciation synthétique doit tenir compte des contraintes qui pèsent sur l'évaluation et l'équipe qui l'a réalisée. Il faut donc nuancer la première conclusion à laquelle on est parvenu en appliquant strictement les critères de qualité. Par exemple, un rapport peut être jugé peu adéquat non en raison de déficiences liées aux travaux, mais parce que les TdR étaient eux-mêmes irréalistes ou que le contexte s'est dégradé. Il est fréquent de constater que les ressources allouées et le temps imparti à l'équipe d'évaluation limitent la portée et la robustesse des conclusions.

Comment apprécier les résultats

La qualité générale du rapport découle de la notation accordée à chacun des 9 critères :

1. A partir de 3 « inacceptables », le rapport doit être considéré comme inacceptable.
2. Lorsque 2 critères ou plus ne sont pas atteints, il est possible de demander à l'évaluateur de reprendre certaines parties ; ce cas doit être prévu et inscrit dans les clauses du contrat.
3. Lorsque l'évaluation est jugée "inacceptable" ou "excellente" pour un critère donné, l'appréciation devra être étayée par au moins deux exemples argumentés.
4. Lorsque l'évaluation est jugée "bonne" pour un critère donné, l'appréciation devra être appuyée par un exemple ou une référence explicative.

Mode de notation par critère :

Critère 1 : La satisfaction des demandes

- **Inacceptable** : Certaines questions des TDR ont été traitées de manière inadéquate ou n'ont été que partiellement abordées. De trop nombreuses questions des TDR n'ont pas été abordées ou n'ont été que partiellement abordées.
 - **Acceptable** : Les demandes formulées dans les TDR ont reçu une réponse de manière correcte. En particulier, les questions évaluatives ont été traitées de manière satisfaisante.
 - **Bon** : Le rapport d'évaluation a bien exposé une vue d'ensemble de la manière dont les objectifs annoncés ont été atteints et a clarifié la logique d'intervention. Le rapport d'évaluation a été au-delà des demandes du cahier des charges et a abordé d'autres sujets d'intérêt.
 - **Excellent** : Les questions abordées couvrent non seulement les demandes des TDR mais replacent l'évaluation dans un cadre beaucoup plus général en liaison avec les autres politiques communautaires, nationales ou locales liées.
-

Critère 2 : La pertinence du champ d'évaluation

En général le champ d'évaluation a trois composantes : le champ temporel, le champ géographique et le champ réglementaire (en particulier les groupes cibles concernés).

- **Inacceptable** : Deux des trois champs sont mal ou insuffisamment traités. L'un des 3 champs d'évaluation est insuffisamment ou mal traité.
 - **Acceptable** : Les 3 champs, temporel, géographique et réglementaire sont correctement pris en compte. Les principaux effets non attendus ont été relevés.
 - **Bon** : Au-delà des 3 champs concernés, l'évaluation s'est intéressée aux interactions du projet avec d'autres politiques aux niveaux local, national ou communautaire. Tous les effets non attendus ont été traités.
 - **Excellent** : En plus des remarques sur le niveau bon, le rapport a systématiquement étudié en détail les effets non attendus du projet.
-

Critère 3 : L'adéquation de la méthodologie

- **Inacceptable** : Il n'existe aucune stratégie d'évaluation et les choix méthodologiques apparaissent souvent en inadéquation par rapport aux résultats recherchés. A la lecture du rapport d'évaluation, il apparaît que des choix méthodologiques ont été faits mais ils n'ont été ni explicités ni défendus.
 - **Acceptable** : La stratégie d'évaluation est clairement explicitée et est effectivement appliquée au cours de l'étude. Les choix méthodologiques sont adéquats pour répondre aux TDR.
 - **Bon** : Les limites inhérentes à la stratégie d'évaluation ont été clairement précisées et les choix méthodologiques ont été discutés et défendus par rapport à d'autres options.
 - **Excellent** : l'évaluateur fait une analyse critique de sa stratégie globale et choix méthodologiques et indique les avantages/inconvénients d'alternatives méthodologiques.
-

Critère 4 : La fiabilité des données

Ce critère ne juge pas de la validité intrinsèque des données disponibles mais de la manière dont le consultant a trouvé les données et comment il les a utilisées.

- **Inacceptable** : Les données utilisées primaires ou secondaires sont manifestement biaisées par

des méthodes de collecte inadaptées ou mal mises en œuvre (par exemple échantillons ou études de cas mal sélectionnés) ou fournissent des renseignements inutilisables.

- **Acceptable** : Les sources de données quantitatives et qualitatives sont identifiées. La fiabilité des données primaires et secondaires a été testée et discutée par le consultant. Les méthodes de collecte ont été clairement explicitées et sont adaptées aux informations recherchées.
 - **Bon** : Les données ont été systématiquement croisées à travers des sources ou des méthodes de recherche indépendantes les unes des autres. Les limites de validité des données et des méthodes de collecte des données sont clairement exposées.
 - **Excellent** : Tous les biais découlant des informations fournies sont analysés et corrigés par des méthodes reconnues.
-

Critère 5 : La solidité de l'analyse

- **Inacceptable** : 2 des 3 éléments (Cf. ci-dessous : méthode d'analyse, relations causales, comparaisons) sont mal traités.
 - **Acceptable** : Les méthodes d'analyse des données quantitatives et/ou des données qualitatives sont faites rigoureusement suivant des méthodes reconnues et pertinentes par rapport aux types de données analysées. Les comparaisons (par exemple : avant/après, bénéficiaires/non bénéficiaires ou contrefactuelles) sont réalisées de manière adaptée.
 - **Bon** : Les méthodes d'analyse sont explicitées et leur limite de validité précisée. Les relations causales entre une mesure et les différents effets sont explicitées. Les limites de validité des comparaisons effectuées sont indiquées.
 - **Exceptionnel** : Tous les biais d'analyse (à travers les 3 éléments) ont été systématiquement analysés et présentés avec leur conséquence sur la limite de validité de l'analyse.
-

Critère 6 : La crédibilité des résultats

Ce critère est objectivement le plus difficile à juger.

- **Inacceptable** : Les résultats de l'analyse apparaissent très peu crédibles. Le texte contient des affirmations qui ne sont pas étayées. Les extrapolations faites et les généralisations de l'analyse ne sont pas pertinentes.
- **Acceptable** : Les résultats produits par l'analyse apparaissent fiables et équilibrés, notamment au vu du contexte dans lequel le programme est évalué. Les hypothèses interprétatives et les extrapolations faites sont acceptables. Les résultats reflètent un compromis acceptable entre la réalité décrite par les données et les faits constatés ou estimés et la réalité du programme/projet telle qu'elle est perçue par les acteurs et les bénéficiaires.
- **Bon** : Les limites des hypothèses interprétatives et des extrapolations faites sont explicitées et discutées. Les effets propres aux mesures évaluées sont isolés des effets dus au contexte et aux contraintes dans lesquels elles sont appliquées. L'équilibre entre la validité interne (absence de biais au sein de la méthode) et la validité externe (représentativité des résultats) est satisfaisant.
- **Exceptionnel** : Les déséquilibres entre la validité interne et la validité externe des résultats sont systématiquement analysés et leurs conséquences sur l'étude d'évaluation explicitées. Les effets contextuels ont été isolés et ont pu être démontrés grâce à des indicateurs pertinents. Les biais apportés dans le choix des hypothèses interprétatives et dans les extrapolations faites sont analysés et leurs conséquences explicitées.

Critère 7 : La validité des conclusions

Ce critère ne juge pas de la valeur intrinsèque des conclusions mais de la manière dont les conclusions ont été atteintes.

- **Inacceptable** : Les conclusions ne s'appuient pas sur une analyse pertinente et rigoureuse. Les conclusions se basent sur des données non prouvées. Les conclusions sont partiales car elles reflètent plus les a priori de l'évaluateur que l'analyse des faits.
 - **Acceptable** : Les conclusions découlent de l'analyse. Les conclusions sont argumentées par des faits et des analyses facilement identifiables dans le reste du rapport. Les limites de validité des conclusions sont indiquées.
 - **Bon** : Les conclusions sont discutées au regard du contexte dans lequel l'analyse a été faite. Les limites de validité des conclusions sont explicites et argumentées.
 - **Excellent** : Les conclusions sont hiérarchisées, elles sont faites en rapport avec la globalité du programme évalué et elles tiennent compte des relations de ce programme avec le contexte dans lequel il se situe, en particulier en tenant compte des autres programmes ou politiques publiques affectant ce programme particulier.
-

Critère 8 : L'utilité des recommandations

Ce critère ne juge pas de la valeur intrinsèque des recommandations mais de leur pertinence par rapport à la manière dont a été conduite l'étude et notamment au regard des conclusions.

- **Inacceptable** : Les recommandations sont déconnectées des conclusions. Les recommandations sont partiales car elles reflètent de façon prépondérante les points de vue de certains acteurs ou de certains bénéficiaires ou elles reflètent la pensée propre de l'évaluateur en référence à un système de valeur socio-économique et un objectif par rapport au programme étudié.
 - **Acceptable** : Les recommandations découlent logiquement des conclusions. Les recommandations sont impartiales.
 - **Bon** : En plus du référentiel précédent, les recommandations sont hiérarchisées et elles sont présentées sous forme d'options d'actions possibles.
 - **Excellent** : Outre le référentiel du niveau bon, les recommandations sont testées et les limites de validité de celles-ci sont indiquées.
-

Critère 9 : La clarté du rapport

- **Inacceptable** : Absence de résumé. Rapport illisible et/ou de structure désordonnée. Absence de chapitre de conclusions (et de recommandations).
- **Acceptable** : Le rapport est lisible facilement et la structure du rapport est logique ou reflète les demandes des lignes directrices. Le résumé court reflète le rapport. Les concepts spécialisés et les démonstrations techniques sont présentés en annexe avec des références claires dans le corps du texte.
- **Bon** : Le corps du rapport est court et concis, de lecture fluide. La structure du rapport est mémorisable facilement. Le résumé est clair et présente de façon équilibrée et impartiale les conclusions et recommandations principales.
- **Excellent** : Le rapport se lit "comme un roman" et sa structuration est d'une logique inattaquable. Le résumé est opérationnel en lui-même.