



Évaluation des démarches EPP de la MAP - Métaévaluation

SGMAP

Table des matières

1. Grille de métaévaluation	3
CONSTRUCTION DE LA GRILLE	3
GRILLE DE META EVALUATION	3
2. Méthode employée.....	5
QU'EST-CE QU'UNE META EVALUATION ?	5
DEFIS DE LA META EVALUATION	5
CONSTRUCTION D'UNE GRILLE D'ANALYSE	5
CHOIX DES REGISTRES, CRITERES ET INDICES DE QUALITE.....	6
MISE EN ŒUVRE ET EXPLOITATION	7
BIBLIOGRAPHIE INDICATIVE	7

1. Grille de métaévaluation

CONSTRUCTION DE LA GRILLE

La grille qualité est construite autour de 3 grands registres, 14 critères et 23 items d'analyse. Chaque item est lui-même décomposé en questions spécifiques.

GRILLE DE MÉTAÉVALUATION

Un processus d'évaluation est jugé de qualité lorsque...

Utilité

... Il prend en compte les usages qui pourront être faits de l'évaluation

ITEM 1 : LES MOTIFS DE L'ÉVALUATION ONT-ILS ÉTÉ EXPLICITÉS AU DÉMARRAGE DE L'ÉVALUATION ?

...Il permet de répondre à des besoins précis des parties prenantes

ITEM 2 : DES QUESTIONS ÉVALUATIVES ONT-ELLES ÉTÉ EXPLICITEMENT POSÉES ET PARTAGÉES AU DÉMARRAGE DE L'ÉVALUATION ?

...Le processus / Ses résultats sont conçus pour être utilisés dans l'évolution de l'action

ITEM 3 : DANS LE RAPPORT FINAL, Y A-T-IL AU MOINS UNE PROPOSITION IDENTIFIANT EXPLICITEMENT L'ACTEUR POUVANT PORTER LE CHANGEMENT PROPOSÉ ?

ITEM 4 : LES PRINCIPAUX DÉCIDEURS POLITIQUES SE SONT-ILS IMPLIQUÉS PERSONNELLEMENT DANS LE PROCESSUS D'ÉVALUATION AUX PRINCIPALES ÉTAPES DE L'ÉVALUATION ?

... Ses résultats arrivent à temps pour la prise de décision

ITEM 5 : UNE DATE SPÉCIFIQUE DE RENDU AVAIT-ELLE ÉTÉ FIXÉE ?

... Ses résultats sont diffusés, communiqués, etc.

ITEM 6 : L'ÉVALUATION A-T-ELLE ÉTÉ PUBLIÉE ?

Faisabilité

... Il précise le champ de l'évaluation de façon à être capable de répondre aux questions posées / aux enjeux de l'évaluation

ITEM 7 : LE CHAMP DE L'ÉVALUATION A-T-IL ÉTÉ EXPLICITÉ AU DÉMARRAGE DE LA DÉMARCHÉ ?

ITEM 8 : LES PARTIES PRENANTES DE LA POLITIQUE ÉVALUÉE EXTÉRIEURES AU CERCLE DE LA MISE EN ŒUVRE ONT-ELLES ÉTÉ EXPLICITEMENT IDENTIFIÉES EN DÉBUT DE DÉMARCHÉ ?

ITEM 9 : LA LOGIQUE D'INTERVENTION DU DISPOSITIF A-T-ELLE ÉTÉ EXPLICITÉE ?

... Les ressources affectées à l'évaluation (temps, budget) sont proportionnelles aux enjeux

ITEM 10 : QUEL EST LE COÛT ESTIMÉ DE L'ÉVALUATION ?

ITEM 11 : LA DURÉE DE LA MISSION EST-ELLE SUPÉRIEURE À 4 MOIS ?

... La méthode et les outils employés permettent de répondre aux questions et aux enjeux de l'évaluation

Il n'y a pas d'item dédié pour estimer l'atteinte de ce critère.

...Les personnes qui pilotent et réalisent l'évaluation ont les compétences nécessaires (sectorielles, évaluation) pour la faire

ITEM 12 : LE SGMAP A-T-IL ÉTÉ ASSOCIÉ AUX TRAVAUX D'ÉVALUATION ?

ITEM 13 : LA MISSION D'ÉVALUATION A-T-ELLE FAIT APPEL À UN EXPERT POUR COMPLÉTER LA MISSION ?

Fiabilité

... Les données sont collectées en suivant les standards méthodologiques de chaque outil employé (elles sont valides)

ITEM 14 : DES LIMITES OU DES OBSTACLES POTENTIELS À LA COLLECTE ET À L'ANALYSE ONT-ILS ÉTÉ REPÉRÉS AU DÉBUT OU PENDANT LA DÉMARCHE D'ÉVALUATION ?

ITEM 15 : L'ÉVALUATION A-T-ELLE EU RECOURS À DES PRESTATIONS D'ÉTUDE POUR LES OUTILS QUE L'ÉQUIPE NE MAÎTRISAIT PAS ? (À COCHER)

... Tous les points de vue légitimes sur le sujet ont été pris en compte

ITEM 16 : QUELS OUTILS DE COLLECTE ONT ÉTÉ EFFECTIVEMENT EMPLOYÉS LORS DE L'ÉVALUATION ?

ITEM 17 : Y A-T-IL UN COMITÉ DE PILOTAGE OU COMITÉ D'ÉVALUATION OUVERT ?

... Les données sont croisées de façon transparente, chaque hypothèse étant testée à charge et à décharge avant de donner lieu à constat (l'analyse est robuste)

ITEM 18*¹ : LA CONCLUSION PRINCIPALE S'APPUIE-T-ELLE SUR UNE ANALYSE SOLIDE ?

... Les conclusions s'appuient sur les constats du rapport

ITEM 19 : LE RAPPORT COMPREND-IL DES CONCLUSIONS PORTANT RÉTROSPECTIVEMENT UN JUGEMENT EXPLICITE SUR L'INTERVENTION ?

ITEM 20* : LA CONCLUSION PRINCIPALE S'APPUIE-T-ELLE SUR UNE ANALYSE AD HOC ?

... Les propositions s'appuient sur les conclusions et répondent suffisamment aux problèmes pointés par ces conclusions.

ITEM 21: LE RAPPORT PRÉSENTE-T-IL DES SCÉNARIOS D'ÉVOLUTION DE QUALITÉ ?

ITEM 22* : LA SUITE DE L'ÉVALUATION PEUT-ELLE ÊTRE RELIÉE AUX PROPOSITIONS ET CONCLUSIONS DE L'ÉVALUATION ?

¹ Les items suivis d'une astérisque s'appliquent à un échantillon du rapport uniquement.

2. Méthode employée

QU'EST-CE QU'UNE MÉTAÉVALUATION ?

Une métaévaluation peut être définie comme un processus cherchant à « décrire une activité évaluative au regard d'un ensemble d'idéaux à propos de ce qui constitue une bonne évaluation » (Stufflebeam, 1974).

Ce n'est pas une démarche de conformité vis-à-vis de normes qui auraient été fixées à l'avance : une métaévaluation permet de produire un jugement sur la qualité des évaluations, généralement afin de faire un bilan et, selon les cas, de rendre compte ou bien de tirer des leçons de l'expérience accumulée.

Une métaévaluation peut être aux deux extrêmes soit narrative (et dans ce cas généralement réalisée par des experts sur un petit nombre de cas ou bien au fil de l'eau), soit basée sur une grille d'analyse exclusivement fermée. Dans la mesure où l'évaluation de la démarche EPP de la MAP comprend également des études de cas en profondeur, c'est cette 2^e approche, qui facilite les généralisations, qui a été choisie.

DÉFIS DE LA MÉTAÉVALUATION

La réalisation d'une métaévaluation présente plusieurs défis :

- Il n'existe pas en évaluation de standards « opposables » comme dans le cas d'une norme, par exemple ; il faut donc choisir un ensemble de critères qui soient reconnus par la communauté évaluative, et pertinents en regard du corpus traité ;
- Les critères de qualité sont par nature de portée générale ; pour pouvoir estimer s'ils sont atteints ou non, il faut identifier quels indices factuels de qualité peuvent être repérés dans les documents fournis. Ceux-ci sont nécessairement développés ad hoc ;
- Enfin, il faut s'assurer que toutes les évaluations du corpus seront analysées exactement de la même façon, ce qui passe notamment par des procédures qualité strictes.

Par ailleurs, il fallait compter avec les principes de la MAP en termes d'évaluation, en particulier dans les cas où des consignes précises avaient été diffusées auprès des évaluateurs. Néanmoins, ces principes étant généralement en accord avec les standards, cela ne représentait pas un réel problème dans le cas d'espèce.

CONSTRUCTION D'UNE GRILLE D'ANALYSE

Nous avons appliqué les principes suivants à la construction de la grille d'analyse :

- Les registres et les critères d'évaluation ont été choisis à partir d'un état de l'art sur les standards de qualité en évaluation (voir ci-dessous) ;
- Les items ont été choisis et précisément formulés lors d'ateliers réunissant le prestataire et les chefs de projet évaluation du SGMAP ; ils ne contiennent en eux-mêmes aucun jugement sur les évaluations et visent seulement à rapporter des éléments factuels utiles à l'analyse ;
- Ils ont ensuite été testés sur 6 cas tests, et amendés de façon à réduire les ambiguïtés et assurer la bonne cohérence des appréciations réalisées.

CHOIX DES REGISTRES, CRITÈRES ET INDICES DE QUALITÉ

S'il n'y a pas de standards opposables en évaluation, il y a en revanche un certain nombre de chartes ou standards qualité reconnus à l'échelle internationale. Nous avons recherché des standards :

- ne portant pas uniquement sur le rapport d'évaluation, mais sur le processus dans son ensemble ;
- et intégrant à la fois le point de vue du producteur de l'évaluation (l'évaluation pouvait-elle effectivement être réalisée et atteindre ses résultats attendus ?) comme du point de vue de son utilisateur (l'évaluation est-elle utile, ses résultats sont-ils suffisamment solides ?).

Les principaux standards et approches qualité qui ont été utilisés ici sont ceux de *l'American Evaluation Association* (qui ont inspiré la plupart des standards nationaux, dont les standards allemands et suisses), de la Commission européenne, de l'OCDE (CAD) et de la Banque mondiale. Nous nous sommes également appuyés sur les réflexions du DfID britannique et de la société suisse d'évaluation (SEVAL) qui ont récemment révisé leurs propres standards.

Ces critères comprennent à la fois des éléments propres à l'évaluation de politiques publiques (qu'est-ce qui rend l'évaluation utile ? faisable ?) et des critères méthodologiques traditionnels des sciences sociales (L'analyse est-elle robuste ? traçable ?).

Pour hiérarchiser et choisir parmi les critères proposés, nous avons utilisé un travail récent de synthèse de différentes métaévaluations qui identifie des facteurs clés de qualité, publié par le Centre for Development Impact (CDI). L'avantage de l'approche du CDI est qu'elle ne s'appuie pas tant sur des principes que sur des éléments de preuve issus de la pratique évaluative.

Pour identifier les indices factuels permettant de caractériser l'atteinte ou non des critères qualité, nous nous sommes appuyés sur l'expertise d'Euréval/Quadrant Conseil, en particulier une métaévaluation réalisée en 2001 pour la Région Limousin et trois métaévaluations réalisées pour la Commission européenne entre 2007 et 2011.

MISE EN ŒUVRE ET EXPLOITATION

Les 75 évaluations retenues sont appréciées au regard de la grille d'analyse. Chaque grille d'appréciation remplie fait l'objet d'un contrôle qualité interne portant sur une sélection d'items sensibles en termes d'homogénéité du remplissage².

Un contrôle qualité externe indépendant est mené par Jacques Toulemonde sur 15 évaluations qu'il choisit sur un critère d'opportunité. Le contrôle qualité reprend intégralement l'appréciation sur les grilles choisies, propose des améliorations appliquées par les analystes.

L'exploitation des grilles donne lieu à une analyse agrégée des différents critères et indicateurs, en particulier au regard de son évolution dans le temps. Les données obtenues sont commentées en prenant en compte les situations rencontrées lors de l'analyse. À noter que certaines évaluations peuvent être données comme exemple de l'atteinte de tel ou tel critère, mais que les analyses individuelles des évaluations resteront confidentielles.

Le rapport donnera lieu à un certain nombre de préconisations en regard des principaux écarts aux standards d'évaluation, qui pourront être discutés dans les phases ultérieures de la mission.

BIBLIOGRAPHIE INDICATIVE

Métaévaluations / Meta-analyses effectuées par Euréval / Quadrant Conseil :

- Meta-study on lessons from existing evaluations as an input to the Review of EU spending (DG BUDG, 2008)
- Evaluation of the EU decentralized agencies (DG BUDG, 2009).
- Overview of the evaluations of DG INFSO's acquis (DG INFSO, 2011)
- Setting a Quality assurance system for the EIB Evaluation Unit (EIB, 2016)

Cooksy, L. J., & Caracelli, V. J. (2008). Metaevaluation in practice: Selection and application of criteria. *Journal of MultiDisciplinary Evaluation*, 6(11), 1-15.

Daigneault, P. M. (2010). L'examen de la qualité des évaluations fédérales : une méta-évaluation réussie ? *The Canadian Journal of Program Evaluation*, 23(2), 191-224.

Delahais T. (2014), Ex post evaluation of regulation and regulatory policies - The case of EU regulation, "Regulatory Governance between Global and Local », ECPR Congress, Barcelona

² *Items 2, 9, 18, 20 à 22.*

European Commission (2004). Evaluation Guidelines. Annex A: Evaluation Standards

Lloyd, R., & Schatz, F. (2015). Improving Quality: Current Evidence on What Affects the Quality of Commissioned Evaluations. Centre for Development Impact (CDI) Working Paper.

Loud, M. M. L. (2004). Setting Standards and Providing Guidelines. Evaluation, 10, 2.

OCDE (2010). Normes de qualité pour l'évaluation du développement. Lignes directrices et ouvrages de références du CAD.

Stufflebeam, D. L. (1974). Toward a Technology for Evaluating Evaluation.

Toulemonde, J. (2007). L'appropriation des résultats de l'évaluation. Leçons de la pratique en Région Limousin. S. Jacob, F. Varob & J.-L. Genard (Éds.), L'évaluation des politiques au niveau régional, 131-142.

