ACADÉMIE HUMANITAIRE · HUMANITAIRES JOBS

Mini MBA Programmes et MEAL : concevoir, piloter et prouver l'impact

Support de cours · Semaine 9 : Évaluer : critères, termes de référence, matrice et utilisation

Programme de 3 mois exactement (12 semaines, 144 heures) pour concevoir, piloter et évaluer des programmes humanitaires : évaluation des besoins et sévérité, chaîne de résultats et indicateurs, ciblage et transferts monétaires, planification et budget, plan MEAL et qualité des données, échantillonnage, analyse, redevabilité, évaluation, mesure d'impact, optimisation des ressources et gestion adaptative. Exercices chiffrés corrigés, trois examens mensuels avec études de cas, et un diplôme vérifiable de l'Académie Humanitaire.

https://humanitairejobs.org/fr/academie/mini-mba-programmes-meal/semaine-9/

Document gratuit de l'Académie Humanitaire. Il peut être imprimé et partagé librement à des fins de formation, sans modification.

-

Semaine 9 : Évaluer : critères, termes de référence, matrice et utilisation

Objectifs

- Choisir le type d'évaluation et les critères adaptés à la décision qu'elle doit éclairer.
- Rédiger des termes de référence précis et construire une matrice d'évaluation rigoureuse.
- Gérer une évaluation de façon indépendante et transformer ses recommandations en décisions suivies.

Pourquoi et quand évaluer

Une **évaluation** est une appréciation systématique et objective d'une intervention, de sa conception, de sa mise en œuvre et de ses résultats. Elle diffère du suivi : le suivi est continu et interne, il demande « faisons-nous ce que nous avions prévu ? » ; l'évaluation est ponctuelle, souvent externe, et elle demande « faisons-nous la bonne chose, de la bonne manière, et qu'est-ce que cela change ? ».

Elle a deux fonctions, parfois en tension : la **redevabilité** (rendre compte aux personnes, aux bailleurs, au conseil d'administration) et l'**apprentissage** (améliorer ce programme et les suivants). Une évaluation conçue uniquement pour la redevabilité produit des notes ; une évaluation conçue pour l'apprentissage produit des décisions.

Type	Moment	Question principale
Évaluation en temps réel	Premières semaines d'une urgence	Faut-il corriger la réponse maintenant ?
Évaluation à mi-parcours	Milieu du projet	Que faut-il ajuster pour la seconde moitié ?
Évaluation finale	Fin du projet	Quels résultats, quelles leçons pour la suite ?
Évaluation d'impact	Pendant et après	Quel changement est attribuable à l'intervention ?
Évaluation ex post	1 à 3 ans après la fin	Les effets ont-ils duré ?

Avant de commander une évaluation, écrivez qui l'utilisera et pour quelle décision. Si personne ne peut répondre, l'évaluation sera un rapport de plus sur une étagère.

Les critères du CAD de l'OCDE

Les critères d'évaluation du Comité d'aide au développement de l'OCDE, révisés en 2019, sont au nombre de six :

- **Pertinence** : l'intervention répond-elle aux besoins et priorités des personnes, et reste-t-elle pertinente si le contexte change ?
- **Cohérence** (ajoutée en 2019) : est-elle compatible avec les autres interventions dans le même contexte, internes et externes à l'organisation ?
- **Efficacité** : atteint-elle ses objectifs, y compris pour les différents groupes ?
- **Efficience** : produit-elle ses résultats de façon économique et à temps ?
- **Impact** : quels effets de plus long terme, positifs ou négatifs, voulus ou non ?
- **Durabilité** : les bénéfices continueront-ils après la fin du financement ?

En humanitaire, on utilise aussi les critères adaptés par l'ALNAP : **couverture** (a-t-on atteint toutes les personnes dans le besoin, sans exclusion ?), **connectivité** (lien entre l'urgence et le plus long terme) et **coordination**.

Les critères ne s'appliquent pas tous avec le même poids. Une évaluation en temps réel d'une réponse de trois mois se concentre sur la pertinence, la couverture et l'efficacité ; parler de durabilité à ce stade a peu de sens. Choisir trois ou quatre critères et les approfondir vaut mieux qu'en survoler six.

Des termes de référence qui obtiennent une bonne évaluation

Les **termes de référence** (TdR) sont le contrat intellectuel de l'évaluation. Des TdR vagues produisent des évaluations vagues. Ils contiennent :

- **Contexte** et description de l'intervention (avec la théorie du changement et le cadre logique en annexe).
- **Objet et utilisateurs** : pourquoi maintenant, pour qui, pour quelles décisions.
- **Portée** : période, zones, composantes, ce qui est exclu.
- **Questions d'évaluation** : cinq à huit questions principales, pas trente. Chaque question est évaluable avec les moyens et le temps disponibles.
- **Approche et méthodes** attendues, en laissant l'équipe proposer, et les données existantes disponibles.
- **Éthique** : consentement, protection des données, protocoles pour les sujets sensibles, participation des personnes.
- **Livrables et calendrier** : rapport initial (inception), restitution, rapport provisoire, rapport final, résumé.
- **Profil de l'équipe**, budget, gouvernance (groupe de référence) et règles d'indépendance.

Ordre de grandeur. Une évaluation finale externe de qualité d'un projet multisectoriel coûte souvent entre 25 000 et 80 000 USD selon le terrain et les méthodes. Un budget trop faible conduit à quelques jours de terrain et à un rapport sans preuves ; il vaut mieux moins d'évaluations, mieux faites.

La matrice d'évaluation et le rapport initial

La **matrice d'évaluation** est l'outil central du rapport initial : pour chaque question, elle précise les sous-questions, les indicateurs ou critères de jugement, les sources de données, les méthodes de collecte et d'analyse, et les limites.

Exemple. Question : « Dans quelle mesure les transferts monétaires ont-ils amélioré la sécurité alimentaire des ménages les plus vulnérables ? » Critères de jugement : évolution du score de consommation et du rCSI entre la référence et la fin, comparée à une zone de comparaison ; différences selon le sexe du chef de ménage et la présence d'une personne handicapée. Sources : enquêtes de référence et finale, suivi post-distribution, discussions de groupe. Méthodes : différence de différences, analyse qualitative par groupe. Limites : zone de comparaison partiellement aidée par d'autres acteurs à partir du mois 6.

Le **rapport initial** est le moment de corriger le tir : l'équipe confirme ce qui est faisable, signale les questions impossibles à traiter, propose l'échantillonnage. Le gestionnaire le valide avec le groupe de référence avant le terrain. C'est le meilleur investissement de temps de toute l'évaluation.

La **qualité** d'une évaluation se juge sur la clarté du lien entre **constats** (appuyés par des données), **conclusions** (jugements fondés sur les constats) et **recommandations** (actions qui découlent des conclusions). Une recommandation sans constat correspondant est une opinion de l'évaluateur.

Indépendance, éthique et gestion de l'évaluation

L'**indépendance** protège la crédibilité : l'équipe d'évaluation n'a pas participé à la conception ni à la mise en œuvre ; le gestionnaire de l'évaluation n'est pas le chef du projet évalué ; l'équipe a accès librement aux données et aux personnes ; le programme peut commenter les faits du rapport provisoire, mais pas imposer de changer les conclusions.

Les normes et règles d'évaluation du groupe des Nations unies pour l'évaluation (UNEG) et les principes de l'ALNAP insistent sur l'**éthique** : ne pas nuire aux personnes interrogées, consentement éclairé, confidentialité, attention aux personnes en détresse et aux risques de représailles, restitution aux communautés.

Le **gestionnaire de l'évaluation** a un rôle actif : il prépare les données, organise l'accès au terrain, réunit le groupe de référence, contrôle la qualité aux étapes clés, protège l'indépendance de l'équipe, et organise la restitution. Les principaux risques sont connus : retards de démarrage qui réduisent le terrain, données de référence inexistantes, accès limité, pression pour adoucir les constats.

> Une évaluation qui ne trouve que des points positifs n'est pas crédible. Un programme qui accepte des constats négatifs bien étayés montre qu'il est sérieux.

La réponse de la direction et l'utilisation

La **réponse de la direction** (management response) est le document qui transforme une évaluation en décisions. Pour chaque recommandation : acceptée, partiellement acceptée ou

rejetée, avec la justification ; actions prévues ; responsable ; échéance. Elle est rédigée dans le mois qui suit le rapport final et publiée avec lui.

Le **suivi** des actions se fait lors des revues trimestrielles de direction : où en est chaque action ? Les organisations matures publient chaque année un taux de mise en œuvre des recommandations acceptées.

Une recommandation peut être légitimement rejetée : trop coûteuse, hors du mandat, déjà en cours. Mais la justification doit être écrite. Ce qui n'est pas acceptable est l'absence de réponse.

Enfin, la **diffusion** adapte le format au public : résumé de deux pages pour la direction et les bailleurs, présentation pour les équipes, version simple et orale pour les communautés, et publication du rapport complet (par exemple dans la bibliothèque de l'ALNAP) pour le secteur. La publication des évaluations, y compris critiques, est un engagement du Grand Bargain en matière de transparence.

Exercice d'application : écrire la question et la matrice d'une évaluation

Situation. Un projet de 18 mois (eau, hygiène et protection) dans trois camps de déplacés arrive à mi-parcours. La direction pays doit décider dans quatre mois s'il faut étendre le projet à un quatrième camp et comment. Le bailleur a demandé une évaluation mais n'a pas fixé de questions. Budget : 30 000 USD. Les données de référence existent pour l'eau, pas pour la protection.

Travail demandé. 1) Quel type d'évaluation et quels critères principaux ? 2) Écrivez trois questions d'évaluation. 3) Pour l'une d'elles, remplissez une ligne de matrice (critères de jugement, sources, méthodes, limites).

Corrigé. Évaluation à mi-parcours, formative, avec pour utilisateurs la direction pays et le bailleur ; critères principaux : pertinence, efficacité, couverture, efficience (la durabilité est secondaire à ce stade). Questions possibles : 1) Dans quelle mesure les services d'eau et d'hygiène atteignent-ils les standards prévus pour tous les groupes, y compris les personnes handicapées ? 2) Les activités de protection répondent-elles aux risques prioritaires identifiés par les femmes et les enfants ? 3) Quelles conditions faut-il réunir pour étendre le projet à un quatrième camp avec la même qualité et un coût par personne comparable ? Ligne de matrice pour 2) : critères : correspondance entre risques prioritaires déclarés et activités, utilisation des services, perception de sécurité ; sources : discussions de groupe séparées par sexe et âge, entretiens avec les acteurs de protection, données des services (anonymisées) ; méthodes : analyse qualitative codée, cartographie participative des risques ; limites : absence de référence en protection, d'où une analyse de contribution plutôt qu'une mesure d'évolution.

À retenir

- Le suivi demande « faisons-nous ce que nous avions prévu ? » ; l'évaluation demande « est-ce la bonne chose, bien faite, et que change-t-elle ? ».
- Six critères du CAD (2019) : pertinence, cohérence, efficacité, efficience, impact, durabilité ; plus couverture, connectivité et coordination en humanitaire.
- Des TdR précis : utilisateurs, décisions, cinq à huit questions évaluables, éthique, livrables.
- La matrice d'évaluation relie chaque question à des critères, des sources, des méthodes et des limites.
- Réponse de la direction : chaque recommandation acceptée ou non, avec action, responsable et échéance.

Quiz

1. Quel critère a été ajouté par le CAD de l'OCDE en 2019 ?
   A. Pertinence
   B. Cohérence
   C. Efficience
   D. Impact

2. Une évaluation qui demande si toutes les personnes dans le besoin ont été atteintes, sans exclusion, porte sur le critère :
   A. Durabilité
   B. Couverture
   C. Efficience
   D. Connectivité

3. Le moment le plus utile pour corriger une évaluation mal orientée est :
   A. La lecture du rapport final
   B. La validation du rapport initial et de la matrice d'évaluation
   C. La publication
   D. La réponse de la direction

4. Une recommandation qui ne s'appuie sur aucun constat du rapport est :
   A. Une bonne pratique
   B. Une opinion de l'évaluateur, à remettre en question
   C. Obligatoire
   D. Une conclusion

5. Qui ne devrait pas gérer l'évaluation d'un projet ?
   A. Un responsable MEAL pays
   B. Le chef du projet évalué
   C. Un conseiller du siège
   D. Un comité de pilotage

Corrigé du quiz : 1. B · 2. B · 3. B · 4. B · 5. B