

EVALUATION DE L'ACTION HUMANITAIRE

GUIDE PILOTE



EVALUATION DE L'ACTION HUMANITAIRE

A propos d'ALNAP

ALNAP est un réseau d'apprentissage composé d'adhérents issus de l'ensemble du secteur qui vise à améliorer la qualité et la redevabilité de l'action humanitaire.

Les membres d'ALNAP travaillent à l'identification d'approches communes dans le but d'améliorer la performance, d'explorer de nouvelles façons de renforcer l'apprentissage et la redevabilité au moyen d'évaluations et d'autres mécanismes, et de partager des idées, des leçons et des innovations. Son objectif général est de fournir une contribution active pour relever les défis auxquels le secteur est confronté. Ses membres sont des organisations et des experts issus de l'ensemble du secteur humanitaire : bailleurs, ONG, Mouvement international de la Croix-Rouge et du Croissant-Rouge, agences des Nations unies, consultants indépendants et institutions académiques.

www.alnap.org/EHA



Conçu par Deborah Durojaiye
Revu par Amanda A. Morgan

Traduction en français : Pierre-Julien
Brunet et Caroline Gassiat (Groupe URD)

A propos des auteurs

Margie Buchanan-Smith est chercheuse associée (senior) à l'Overseas Development Institute (ODI). Basée à Londres, elle a été impliquée dans le secteur de l'aide humanitaire depuis plus de 25 ans en tant qu'évaluatrice, conseillère et chercheuse sur les questions de politiques, responsable de programmes humanitaires, formatrice et coach. Elle est aussi actuellement chargée de cours invitée au Feinstein International Center de la Tufts University.

John Cosgrave est un évaluateur indépendant basé en Irlande. Il a plus de 30 ans d'expérience dans l'action humanitaire et le développement, et a travaillé dans plus de 60 pays. Conseiller à l'évaluation et coordinateur pour la Tsunami Evaluation Coalition, il a co-écrit le rapport de synthèse. John fut également l'auteur principal du guide pilote ALNAP sur les évaluations en temps réel dans le cadre de l'action humanitaire publié en 2009.

Margie et John co-animent des sessions de formation annuelles sur l'"Évaluation de l'action humanitaire" pour Channel Research en Europe et pour le programme IPDET financé par la Banque Mondiale en Amérique du Nord.

AVANT-PROPOS



Parce que la pratique de l'évaluation humanitaire devient de plus en plus cruciale dans un système toujours plus protéiforme et complexe, ce guide fournit une aide utile aux spécialistes et non-spécialistes de l'évaluation. La pratique de l'évaluation a considérablement évolué et de nombreux conseils techniques existent déjà. Toutefois, ce guide pilote ALNAP consolide de façon unique la plupart des connaissances actuelles sur l'ensemble des phases d'une évaluation humanitaire : depuis la décision initiale jusqu'à la diffusion finale.

Le format de ce guide est un PDF interactif facile d'utilisation et accessible au plus grand nombre. Il contient des exemples tirés d'expériences réelles, des astuces pratiques, des définitions et des conseils étape par étape sur des éléments spécifiques des évaluations à différentes phases du processus.

Nous voulons nous assurer que ce guide est aussi pratique que possible aussi bien pour les spécialistes de l'évaluation que pour le grand public. Pour cela, nous sommes engagés dans une minutieuse démarche de test. Comme toujours, nous profiterons du potentiel du réseau ALNAP pour nous aider à piloter le Guide durant ce travail d'évaluation. De cette manière, nous savons que nous pouvons améliorer sa qualité, sa pertinence et son utilité pour aider au final à fournir de meilleurs résultats aux populations affectées par les crises.

John Mitchell

COMMENT UTILISER CE GUIDE INTERACTIF

Ce guide utilise des icônes du début à la fin. Survolez les icônes ci-dessous pour découvrir leur signification.

Vous trouverez au bas de chaque page de ce PDF une série d'icônes.



Revenir à la page précédente



Aller à la page "Comment utiliser"



Aller à la page suivante



Revenir à la vue précédente

Ce guide interactif s'affiche de façon optimale avec Adobe Reader. Vous pouvez télécharger la dernière version gratuite sur <http://get.adobe.com/uk/reader/>.

Ce guide peut être vu sur tablettes et smartphones mais certaines fonctions de navigation peuvent être désactivées.

Travail en mode déconnecté – Aucune connexion Internet n'est nécessaire pour utiliser ce guide : nous vous encourageons à télécharger ce PDF interactif et à l'ouvrir dans Adobe Reader. En revanche, les commentaires nécessitent une connexion Internet.

Le Guide EAH est aussi disponible en anglais et en espagnol. Vous pouvez télécharger ceux-ci ici: www.alnap.org/EHA

LES SIX RAISONS D'ETRE DE CE GUIDE

Voici les six raisons pour lesquelles nous pensons qu'il est temps de disposer d'un guide complet relatif à l'EAH :

- 1. L'aide officielle des bailleurs à destination de l'action humanitaire a augmenté** par six environ en valeur réelle de 1990 à 2011.
- 2. Plus d'intérêt et d'investissement au profit de l'évaluation** car l'efficacité de l'aide au développement et de l'action humanitaire suscite des inquiétudes.
- 3. Il existe à présent une masse critique de savoirs collectifs sur laquelle il est possible de s'appuyer** : la base de données des évaluations d'ALNAP contient à elle seule plus de 500 exemples couvrant la dernière décennie.
- 4. Les commandes d'évaluations se sont déplacées des sièges des organisations vers le personnel basé sur le terrain** à mesure que les organisations se décentralisent. Pourtant, les responsables basés sur le terrain ont souvent peu d'expérience dans la planification et la gestion des évaluations, en particulier les EAH.
- 5. Il existe peu de preuves que les résultats de l'évaluation mènent à un changement ou à une réflexion sur les politiques et les pratiques** : des évaluations mieux conçues pourraient fournir des éléments de preuves plus forts en faveur d'un changement de politiques et promouvoir leur utilisation.
- 6. La demande de conseils sur l'EAH est croissante** : un sondage des membres du Humanitarian Practice Network réalisé en 2009 a montré que l'EAH était le premier type d'outil demandé.

Ce guide ALNAP fournit des conseils pratiques et détaillés ainsi que des exemples concrets à ceux qui planifient, conçoivent, réalisent et utilisent des évaluations d'action humanitaire. L'accent est mis sur l'utilisation pour vous encourager à prendre en compte la manière dont on peut s'assurer depuis le début qu'une évaluation sera utilisée.

Ce guide vise à aider des évaluations de haute qualité qui contribuent à améliorer les performances en fournissant le plus d'éléments de preuves possibles sur ce qui fonctionne bien, ce qui ne fonctionne pas et pourquoi. Son but ultime consiste à mieux satisfaire les besoins des populations affectées par les crises humanitaires qui, tout au long de ce guide, seront désignées comme les populations affectées.



Populations affectées

Toutes les populations affectées (en général négativement) par une crise humanitaire.

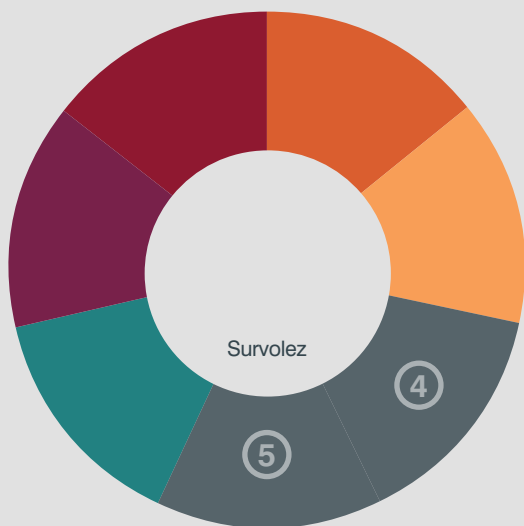
Ce terme est plus neutre que celui de bénéficiaire, lequel implique que la population cible a tiré un bénéfice de l'intervention humanitaire.

Malheureusement, cela n'est pas toujours le cas.

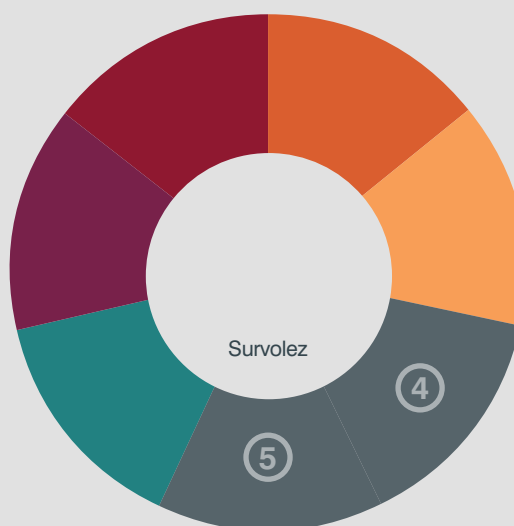
A QUI S'ADRESSE CE GUIDE ?

Ce guide a été conçu pour toutes les personnes directement impliquées dans la réalisation d'EAH.

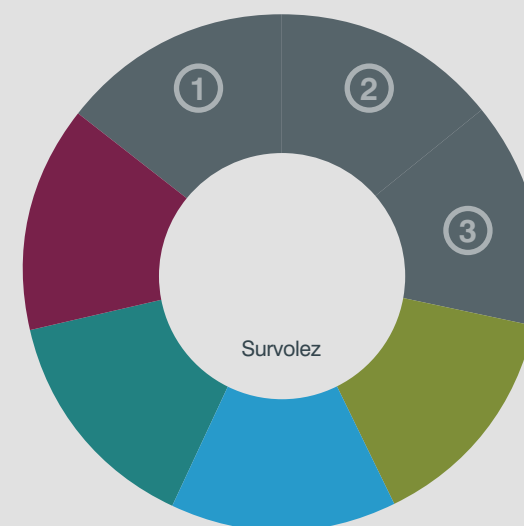
Survolez les numéros et les images en-dessous pour découvrir dans quelle mesure un chapitre vous concerne.



Responsable de l'évaluation



Responsable du programme

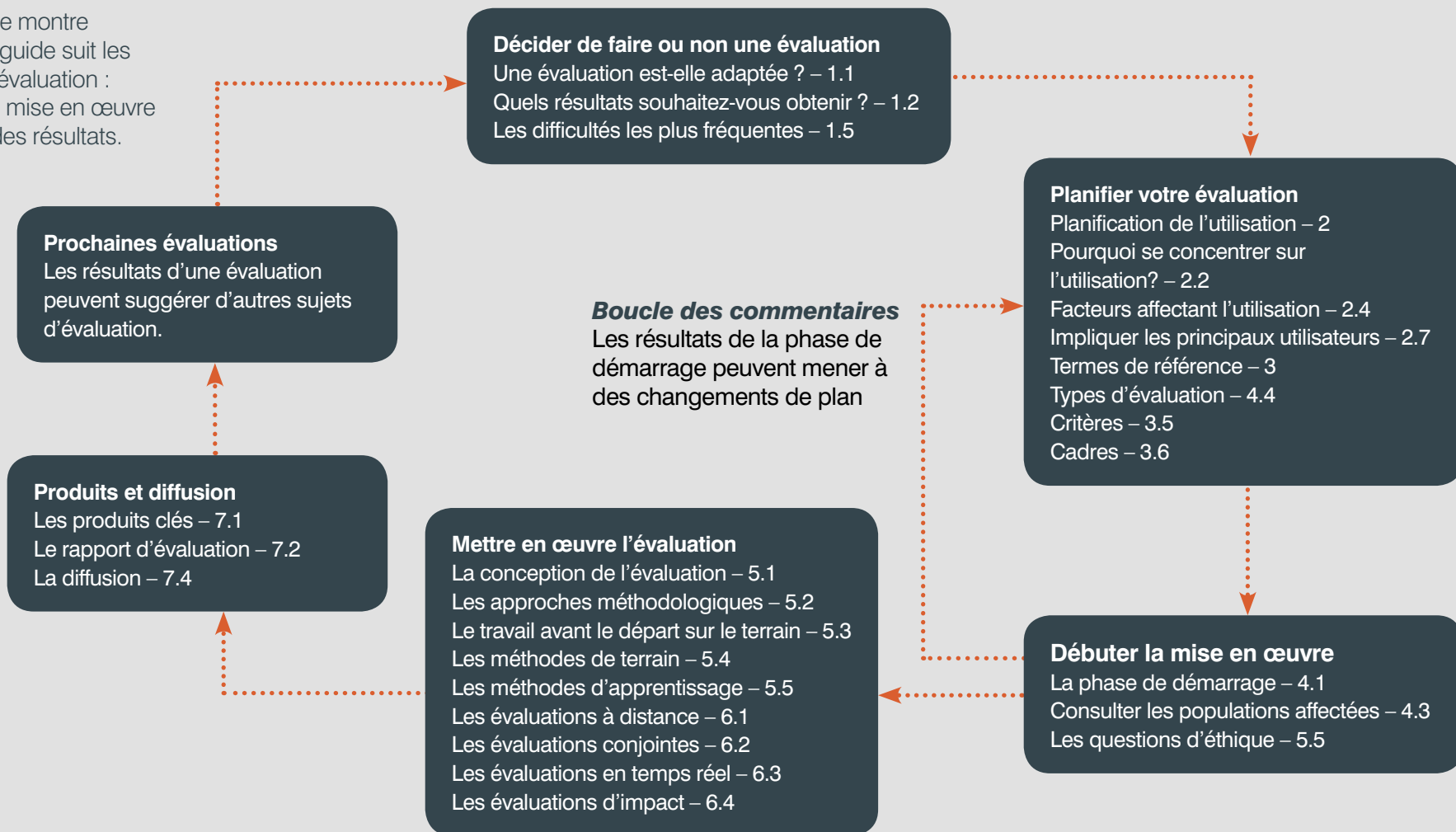


Consultant

QUAND UTILISER CE GUIDE ?

Schéma 1: Le cercle de projet de l'évaluation

Ce graphique montre comment le guide suit les phases de l'évaluation : planification, mise en œuvre et diffusion des résultats.



LE PROCESSUS DE PILOTAGE

Pour renforcer encore les dimensions d'utilisation et d'attention à l'utilisateur de ce guide sur l'EAH, nous nous engageons en 2013 et 2014 dans un processus de pilotage actif de cette version provisoire actuelle.

Nous travaillons avec un noyau dur d'organisations pilotes issues de tout le secteur humanitaire afin d'utiliser ce guide dans diverses activités liées à l'EAH : depuis la gestion et la conduite de l'ensemble des exercices d'évaluation – ou une seule étape au sein d'une plus grande initiative EAH – jusqu'à l'animation de formations et d'activités de renforcement des capacités liées à l'EAH.

Si vous n'êtes pas l'un de ces testeurs clés, nous vous invitons à nous faire connaître vos impressions relatives à l'utilisation de ce guide. Aucun enregistrement n'est demandé pour participer, nous demandons simplement que toute personne qui télécharge cette version pilote partage ses commentaires quant à l'utilisation d'un chapitre ou d'une partie de chapitre dans son travail. Cela peut être fait en cliquant sur les boutons de commentaires confidentiels situés en bas des pages du guide.

Si vous souhaitez vous engager davantage en tant qu'organisation pilote – ce qui pourrait impliquer des sondages en ligne, des conversations téléphoniques et un atelier de discussion –, la manière la plus simple de nous contacter est de nous écrire par e-mail à eha@alnap.org

www.alnap.org/EHA

Durant la phase pilote, ALNAP continuera d'alimenter ce guide avec de nouveaux regards d'évaluateurs et d'autres exemples de bonnes pratiques en lien avec l'évaluation des résultats de protection humanitaire. L'objectif est de contribuer à développer des conseils pour renforcer la qualité des évaluations humanitaires focalisées sur la protection. Si vous souhaitez vous impliquer dans cette démarche, ou si vous avez une certaine expérience à partager en matière d'évaluation, en particulier sur la manière d'appliquer une approche protection à l'évaluation humanitaire, veuillez nous contacter à l'adresse suivante : eha@alnap.org

Cet ouvrage est la traduction française de la version pilote du Guide EAH. Les traducteurs ont fait de leur mieux pour préserver le contenu original de la version en test depuis juin 2013 et garantir la cohérence du texte entre toutes les versions produites durant cette phase pilote. Nous vous demandons ainsi de bien vouloir nous aider à identifier les meilleurs termes liés à l'EAH pour établir la version française finale de ce guide. Pour cela, si vous remarquez que certains passages peuvent être améliorés, merci de nous le faire savoir en utilisant le bouton Commentaires.

Remerciements :

Les auteurs remercient les commentaires très utiles fournis par les membres d'un groupe consultatif composé d'acteurs de l'évaluation du secteur humanitaire. Il s'agit de Jock Baker, Tony Beck, Hana Crowe, Stefan Dahlgren, Wendy Fenton, Enrique Garcia, Josse Gillijns, Saul Guerrero, Babar Kabir, Peter Klansoe, Caroline Loftus, Rob McCouch, Joakim Molander, Jonathan Patrick, Nicoletta Pergolizzi, Riccardo Polastro, Hannah Reichardt, Peta Sandison, Vivien Walden.

L'élaboration de ce guide a été facilitée par le secrétariat d'ALNAP : John Mitchell, Paul Knox-Clarke, Francesca Bonino et, précédemment, Yuka Hasegawa. L'appui en matière de communication a été fourni par Patricia Curmi et Alexandra Warner.

TABLE DES MATIERES

COMMENT UTILISER CE GUIDE

INTERACTIF ? 4

A QUI S'ADRESSE CE GUIDE ? 7

QUAND UTILISER CE GUIDE ? 8

LE PROCESSUS DE PILOTAGE 9

① DECIDER DE FAIRE OU NON UNE EVALUATION 14

1.1 L'action humanitaire 15

1.2 L'évaluation de l'action humanitaire . 16

1.3 Peser le pour et le contre d'une
évaluation et d'autres options 20

1.3.1 Les évaluations tournées vers la
redevabilité. 20

1.3.2 Les évaluations tournées vers
l'apprentissage. 22

1.4 L'évaluation dans le cycle de
la réponse 24

1.5 Les difficultés les plus fréquentes. . . 26

② PROMOUVOIR L'UTILISATION ET ASSURER LA QUALITE 34

2.1 Se concentrer dès le départ sur la
qualité et l'utilisation. 35

2.2 Ce que signifie être focalisé sur
l'utilisation 35

2.3 Comment les évaluations de l'action
humanitaire sont utilisées. 36

2.4 Les facteurs affectant l'utilisation . . . 38

2.5 Les parties prenantes. 41

2.6 Se focaliser sur ce que les principales
parties prenantes doivent savoir . . . 42

2.7 Impliquer les utilisateurs prévus
de l'évaluation 43

2.8 Les facteurs affectant la qualité
de l'évaluation 44

③ PLANIFICATION ET CONCEPTION 48

3.1 Les évaluations tournées vers
l'apprentissage et la redevabilité. . . . 49

3.2 Les types d'évaluation 51

3.2.1 Les évaluations présentant des
périmètres différents. 52

3.2.2 Les évaluations d'impacts ou
de processus 53

3.2.3 Les évaluations à des timings
différents 54

3.2.4 Les évaluations de et par
différents acteurs 56

3.2.5 Les évaluations à base
théorique 57

3.2.6 Les autres types d'évaluation . . . 58

3.3 Déterminer les informations à connaître 60

3.4 Les critères d'évaluation. 62

3.5 Les cadres d'évaluation 67

3.6 Les méthodes d'évaluation. 74

3.6.1 Les approches quantitatives
et qualitatives 74

3.6.2 Consulter les populations
affectées 75

3.6.3 La recherche bibliographique . . . 71

3.7 Gestion et gouvernance 71

3.7.1 Les groupes consultatifs 71
 3.7.2 Gérer les conflits. 73
3.8 Choisir une équipe d'évaluation 75
3.9 Identifier des consultants externes . . 84
3.10 Budgéter une évaluation. 87
3.11 Négocier les Termes de référence . . 90

④ REALISER UNE EVALUATION D'ACTION HUMANITAIRE 98

4.1 La phase de démarrage 99
 4.1.1 Préparer le rapport de démarrage 99
 4.1.2 La matrice d'évaluation. 103
 4.1.3 Utiliser un modèle logique. 105
 4.1.4 L'étude documentaire. 110
 4.1.5 Evaluer le rapport de démarrage. 110

4.2 Les défis liés au leadership et au travail en équipe. 114

4.3 Impliquer les populations affectées 117

4.4 Les questions d'éthique 119

⑤ CONCEPTION ET METHODES D'EVALUATION 126

5.1 La conception de l'évaluation 127
 5.1.1 Les conceptions expérimentales. 128
 5.1.2 Les conceptions statistiques . . . 131
 5.1.3 Les conceptions à base théorique 132
 5.1.4 Les conceptions fondées sur des cas pratiques 133
 5.1.5 Les conceptions participatives. . 133
 5.1.6 Les conceptions ponctuelles vs. les conceptions longitudinales . 134
 5.1.7 Choisir une conception. 134

5.2 Les méthodes de recherche 134
 5.2.1 Les méthodes 134
 5.2.2 L'échantillonnage 137
 5.2.3 Contrôler la qualité de la recherche 142
 5.2.4 Constituer un corpus de données probantes 146
 5.2.5 Utiliser la triangulation. 147

5.2.6 De la collecte à l'analyse des données. 148
 5.2.7 Les hypothèses contrefactuelles 150

5.3 Le travail avant le départ sur le terrain 150
 5.3.1 La recherche de documents-clés 150
 5.3.2 Utiliser un grand nombre de documents. 152
 5.3.3 Réaliser une chronologie 154
 5.3.4 L'analyse du contenu et des mots-clés. 157
 5.3.5 L'analyse de données chiffrées 160
 5.3.6 Le rapport coût-efficacité et efficacité 161
 5.3.7 Evaluer l'attention aux questions de genre dans la revue documentaire. 163
 5.3.8 Les sondages en ligne 165

5.4 Les méthodes de terrain. 166
 5.4.1 Les entretiens avec des informateurs-clés 166

5.4.2 Les entretiens avec les populations affectées	167	6.2.3 Quelles approches et méthodes utiliser ?	200	7.4 Approuver et finaliser l'évaluation . .	231
5.4.3 L'observation	176	6.2.4 Comment assurer le suivi ?	201	7.5 La diffusion	231
5.4.4 Les méthodes d'évaluation participative rapide	176	6.3 Les évaluations en temps réel. . . .	202	7.6 Faciliter l'adoption des recommandations d'une évaluation.	234
5.4.5 Les enquêtes sur le terrain	181	6.3.1 Avantages	202	7.7 Synthèses d'évaluation, revues thématiques et méta-analyses	235
5.5 Les approches tournées vers l'apprentissage	186	6.3.2 Caractéristiques	203		
5.5.1 Les revues après action	187	6.3.3 Méthodes.	207		
5.5.2 Le « storytelling »	188	6.4 Les évaluations d'impact	209		
5.5.3 L'enquête d'encouragement (« appreciative enquiry »)	191	6.4.1 Les évaluations d'impact conjointes	210		
		6.4.2 Comment faire une évaluation d'impact ?	210		
⑥ EVALUATIONS A DISTANCE, CONJOINTES, EN TEMPS REEL ET D'IMPACT	195	⑦ PRODUITS, DIFFUSION ET ADOPTION	215		
6.1 L'évaluation lorsque l'accès est limité	196	7.1 Les produits clés	216		
6.2 Les évaluations conjointes	198	7.2 Le rapport d'évaluation.	220		
6.2.1 Les bénéfices d'une évaluation conjointe	198	7.2.1 Les pages d'introduction.	220		
6.2.2 Comment planifier et gérer une évaluation conjointe	199	7.2.3 Les recommandations	226		
		7.2.4 Les annexes.	227		
		7.3 Faire circuler et commenter le rapport provisoire.	229		

BIBLIOGRAPHIE ANNOTÉE



DECIDER DE FAIRE OU NON UNE EVALUATION

Contenu de ce chapitre

1.1	L'action humanitaire	15
1.2	L'évaluation de l'action humanitaire	16
1.3	Peser le pour et le contre d'une évaluation et d'autres options	20
1.4	L'évaluation dans le cycle de la réponse	24
1.5	Les défis communs	26

Comment utiliser ce chapitre ?

Ce chapitre est particulièrement utile pour les responsables d'évaluation et les responsables de programmes qui décident ou non de lancer une évaluation. Les chapitres **1.1** et **1.2** expliquent les concepts clés de l'évaluation en général, et de l'évaluation de l'action humanitaire en particulier. Le chapitre **1.3** passe en revue les types d'objectifs pour lesquels une évaluation est (ou non) la meilleure solution. Le chapitre **1.4** explique comment l'évaluation est liée au suivi et comment elle s'inscrit dans le cycle de la réponse d'urgence. Le chapitre **1.5** présente la manière dont l'EAH diffère des autres types d'évaluation et souligne les difficultés communes ainsi que les solutions possibles.

Tableaux, schémas et encadrés

Encadré 1 :	Les concepts clés de l'évaluation	17
Tableau 1 :	L'évaluation et les autres processus de redevabilité	20
Tableau 2 :	L'évaluation et les trois piliers d'une organisation apprenante	22
Schéma 2 :	Suivi et évaluation dans le cycle de réponse d'urgence	24
Schéma 3 :	Exemple de tableau de gestion des recommandations	30

1.1 L'action humanitaire

Il existe plusieurs définitions de l'action humanitaire. L'une des plus complètes et des plus largement utilisées est celle proposée par la Good Humanitarian Donorship Initiative (2003):



Action humanitaire

Action entreprise dans le but de sauver des vies, d'atténuer les souffrances et de préserver la dignité humaine pendant et après des crises provoquées par l'homme ou des catastrophes naturelles, ainsi que de prévenir de tels événements et d'améliorer la préparation à leur survenue.

Les paramètres de ce qui est appelé action humanitaire se sont progressivement élargis. Le terme désignait autrefois seulement le fait de sauver des vies ; aujourd'hui, il recouvre aussi l'importance de sauver des moyens d'existence. Notre définition de travail renvoie au fait de préserver la dignité humaine et la capacité de subvenir aux besoins des personnes en est une part importante.

La protection des civils est désormais considérée comme une responsabilité de l'ensemble de la communauté humanitaire et pas seulement de la Croix-Rouge et des Nations unies. Il s'agit de fournir plus que des moyens de secours comme l'alimentation, l'eau et l'assainissement, les abris et les services de santé.

L'action humanitaire ne comprend pas seulement la réponse à une crise mais aussi le soutien à la préparation et à la réduction des risques de catastrophes avant une crise, et à la réhabilitation et au relèvement après celle-ci. Ces deux derniers points sont une zone grise qui apparaît souvent entre l'aide humanitaire et l'aide au développement. On constate actuellement une prise de conscience croissante quant à l'importance de la réponse aux besoins de relèvement après une catastrophe naturelle. Lors des conflits et d'autres crises prolongées, le moment où l'urgence se termine et où le relèvement commence n'est pas toujours clair ; dans la pratique, les deux types de soutien sont souvent nécessaires et fournis de façon simultanée.

L'action humanitaire doit être guidée par les principes d'humanité, d'impartialité, de neutralité et d'indépendance^[1]. Elle est à la fois *compliquée* (avec de nombreux aspects différents inter-reliés) et *complexe* (non-linéaire et se produisant dans un environnement dynamique avec des tensions importantes et une exigence de jugement) (Glouberman, 2002). Ce guide utilise le terme "complexité" pour renvoyer à ces deux notions.

1.2 Evaluation de l'action humanitaire

L'EAH peut être définie comme « un examen systématique et objectif de l'action humanitaire dans le but de tirer les leçons permettant d'améliorer la politique et la pratique, ainsi que d'augmenter la redevabilité ». Une observation attentive de certains des termes clés de cette définition fait ressortir les éléments suivants :

- *Systématique implique une approche planifiée et cohérente.*
- *Objectif implique de prendre du recul par rapport à l'immédiateté de l'action*

¹ Une présentation de l'approche Croix-Rouge sur les principes humanitaires peut être trouvée dans [Wortel \(2009\)](#); [Leader \(2000\)](#).

humanitaire et de réfléchir sur elle.

- *Examen implique la recherche ou l'analyse en vue de déterminer la valeur ou l'importance de l'action.*
- *Tirer les leçons permettant d'améliorer la politique et la pratique, ainsi que d'augmenter la "redevabilité" est la raison d'être d'une évaluation, mais aussi l'utilisation qui en sera faite.*

La mesure dans laquelle une évaluation est réellement indépendante dépend vraiment de son objectif (voir Chapitre 4.2). Ce point est particulièrement critique pour les évaluations tournées vers la redevabilité. Cela peut être moins important pour les évaluations tournées vers l'apprentissage si ceux qui s'occupent de l'apprentissage sont impliqués dans l'évaluation et ont été responsables de la mise en œuvre de l'action humanitaire en cours d'évaluation. Malgré cela, même dans de tels cas, un certain niveau d'objectivité doit être injecté dans le processus, par exemple en ayant recours à un facilitateur externe ou à des personnes ressources expérimentées et indépendantes comme membre ou chef de l'équipe. Réduire les biais doit être un but dans toutes les évaluations. Deux types clés d'évaluation, qui sont souvent

opposées, sont les évaluations pour la redevabilité et pour l'apprentissage.

**Redevabilité**

Processus qui consiste à se considérer comme responsable d'une intervention et à se montrer redevable envers les différentes parties prenantes, y compris les personnes affectées par l'intervention, celles qui la financent et les autres acteurs humanitaires.

**Apprentissage**

Processus via lequel l'expérience et la réflexion mènent à des changements de comportement ou à l'acquisition de nouvelles capacités.

Encadré 1 : Les concepts clés de l'évaluation

**Produits**

Livrables définis produits par l'acteur ou les acteurs évalué(s).

Les produits sont le résultat des moyens reçus et des activités développées par l'acteur. Un produit doit être pleinement attribuable à un acteur ou un groupe d'acteurs (par exemple, les points d'eau fournis dans un camp de personnes déplacées internes).

**Résultats**

Utilisations faites des produits par la population affectée.

Un résultat est seulement attribuable en partie à un acteur (par exemple, l'usage fait de l'eau à partir de points d'eau nouvellement installés : pour la consommation domestique, la consommation animale ou d'autres activités liées aux moyens de subsistance comme la fabrication de briques).

**Impact**

Effet social, économique, technique ou environnemental d'un projet sur les individus, les questions de genre et de groupes d'âge, les communautés et les institutions.

Les impacts peuvent être positifs ou négatifs, macro (niveau sectoriel) ou micro (niveau des ménages), intentionnels ou non. Par exemple, la mise en œuvre d'une aide alimentaire peut prévenir la malnutrition (soit un impact positif intentionnel), mais dans certaines circonstances, cela peut également gêner la production alimentaire locale (soit un impact non intentionnel).

**Impact sur la chaîne de résultats**

Effets à long terme de l'utilisation des produits par les bénéficiaires.

La définition de ce terme est plus étroite que celle d'impact. Par exemple, l'impact d'une eau potable librement disponible peut être la réduction des maladies au sein des populations déplacées internes ou le fait que cela attire d'autres groupes de populations vers le camp. Les acteurs qui débutent une intervention contribuent à l'impact sur la chaîne de résultats mais cela doit également être attribué à différents facteurs sur lesquels les acteurs n'ont pas le contrôle.

**Attribution**

Hypothèse qu'un moyen particulier ait causé un impact spécifique.

Lors d'interventions humanitaires complexes et compliquées, il est rarement possible d'attribuer un résultat à une seule cause en particulier. Une organisation spécialisée dans l'aide alimentaire peut attribuer une réduction de la malnutrition à des distributions alimentaires mais celle-ci peut aussi être due à une amélioration de la qualité de l'eau, des pratiques de soins infantiles, de l'hygiène, des soins de santé, de l'assainissement, du contrôle des vecteurs ou même des variations saisonnières normales.

**Contribution**

Mesure dans laquelle un moyen donné est supposé avoir aidé l'atteinte d'un impact particulier.

Dans le domaine de l'évaluation de l'action humanitaire, il est en général beaucoup plus facile d'évaluer la contribution que l'attribution.



Réfléchissez attentivement au périmètre de votre évaluation. Sur quoi voulez-vous vous concentrer ? Etes-vous plus intéressé par la première phase de l'urgence ou par une phase ultérieure ? Souhaitez-vous également observer le volet de la préparation ? Ainsi que le soutien au relèvement et à la réhabilitation ? Plus vous ciblez votre évaluation et plus vos résultats seront susceptibles d'être utilisables.

La redevabilité et l'apprentissage sont difficiles à séparer dans la vie réelle. Les audits, qui se concentrent très fortement sur la redevabilité, peuvent aussi mener à de l'apprentissage. De même, une amélioration du comportement au moyen de l'apprentissage est à même d'améliorer la redevabilité.

1.3 Peser le pour et le contre d'une évaluation et d'autres options

L'objectif principal d'une évaluation, que ce soit en général la redevabilité ou l'apprentissage, aidera à déterminer le choix de la conception de l'évaluation (Chapitre 4.2).

1.3.1 Les évaluations tournées vers la redevabilité

L'évaluation est l'un des nombreux processus qui puisse permettre à une organisation de s'acquitter de sa responsabilité de rendre des comptes aux autres, que ce soit à son conseil d'administration, à ses bailleurs ou aux populations affectées. Vous devez décider si c'est la solution la plus appropriée et la plus rentable.



Evaluation tournée vers la redevabilité

Une évaluation portant sur la qualité de l'utilisation des ressources (également appelée une évaluation sommative).

Tableau 1 présente sous forme synthétique les différents types de redevabilité et suggère d'autres solutions en plus de l'évaluation. En matière d'EAH, il est difficile d'équilibrer la redevabilité entre plusieurs acteurs présentant des intérêts et des objectifs différents ainsi que des niveaux de pouvoir variables. Par exemple, la redevabilité envers les populations affectées est vue comme extrêmement importante ; pourtant, dans de nombreux environnements de conflits, les populations affectées sont privées de pouvoir et n'ont qu'une faible voix. Dans la pratique, les organisations de bailleurs sont parmi les acteurs les plus puissants.

Tableau 1 : Evaluation et autres processus de redevabilité

Type de redevabilité	L'évaluation est-elle appropriée?	Autres processus de redevabilité
Redevabilité stratégique (par exemple envers le mandat et les objectifs de l'organisation)	Oui	Analyse stratégique
Redevabilité managériale (par exemple, pour l'utilisation des ressources au sein d'une organisation)	Oui	Gestion de la performance et autres outils de gestion
Redevabilité financière (par exemple envers les bailleurs)	Oui, en particulier si la rentabilité et l'efficacité sont les principales préoccupations	Audit (peut être plus approprié si le contrôle financier et la conformité sont les principales préoccupations)
Redevabilité contractuelle (par exemple, pour tenir ses engagements)	Oui, en particulier s'il existe une obligation contractuelle à faire une évaluation	Audit et autres processus spécifiés dans le contrat
Redevabilité relationnelle (par exemple, envers d'autres organisations impliquées dans une opération)	Oui, si incluse dans les Termes de référence de l'évaluation	Examen institutionnel
Redevabilité juridique (par exemple, envers les droits locaux ou internationaux)	Non	Examen du respect des conditions légales (par exemple, code du travail)
Redevabilité envers les populations affectées	Oui, si l'évaluation est conçue de façon appropriée	

La redevabilité présente souvent des aspects externes et internes. La direction d'une organisation peut être redevable de l'utilisation des ressources non seulement envers les donateurs mais aussi envers son conseil d'administration puisqu'une mauvaise utilisation des ressources peut menacer la survie de l'organisation. De la même manière, la redevabilité envers la population affectée peut refléter des obligations auxquelles la direction et le conseil d'administration se sont engagés (voir [HAP, 2010](#)). Si la redevabilité est le principal objectif de votre évaluation, demandez-vous :

- *Quel type de redevabilité vous intéresse le plus ?*
- *Une évaluation est-elle le meilleur moyen de procéder ?*

1.3.2 Les évaluations tournées vers l'apprentissage



Evaluation tournée vers l'apprentissage

Evaluation conçue pour faciliter l'apprentissage individuel, de groupe ou organisationnel (également appelée évaluation formative).

Les évaluations tournées vers l'apprentissage peuvent être extrêmement efficaces en ce qui concerne l'analyse de ce qui a fonctionné ou non et la manière dont la performance peut être améliorée. Ce n'est pourtant pas le seul processus d'apprentissage disponible ou le plus rentable. Plusieurs sont examinés au Chapitre 5.5 ; ils peuvent être entrepris seuls ou intégrés dans une évaluation tournée vers l'apprentissage.

Une organisation apprenante ("**Tableau 2**") est plus susceptible de tirer profit d'une évaluation.

Tableau 2 : L'évaluation et les trois piliers d'une organisation apprenante

Pilier	Caractéristiques distinctives	Implications pour l'évaluation
Un environnement propice à l'apprentissage	Les membres de l'équipe se sentent libres d'être en désaccord entre eux, de poser des questions candides, d'admettre leurs erreurs et de représenter les points de vue des minorités. Ils reconnaissent la valeur des idées opposées et sont prêts à prendre des risques ainsi qu'à explorer des domaines inconnus.	Utiliser les méthodes d'évaluation qui encouragent l'ouverture d'esprit comme l'enquête appréciative, la technique dite du « changement le plus significatif » et la revue après action.
Processus d'apprentissage concrets	Des processus formels existent pour générer, collecter, interpréter et diffuser de l'information.	Demander éventuellement au personnel de faire partie de l'équipe d'évaluation. Définir un processus formel pour impliquer le personnel dans l'analyse des résultats et les leçons tirées.
Une direction stratégique qui renforce l'apprentissage	Les dirigeants de l'organisation démontrent leur volonté d'alimenter des points de vue alternatifs ; pointent l'importance de consacrer du temps à l'identification du problème, au transfert de connaissances et à la réflexion ; s'engagent dans un questionnement et une écoute actifs.	Permet que l'évaluation soit soutenue au niveau de la gestion. Permet que l'évaluation ait une durée adaptée. Crée un espace pour les retours et la discussion des résultats avec la direction générale.

Source: [Garvin et al., \(2008\)](#).

1.4 L'évaluation dans le cycle de la réponse

L'EAH est habituellement une activité unique réalisée à un moment clé du cycle de la réponse d'urgence humanitaire dans le but de fournir des informations à ce cycle ainsi que de nourrir les réponses futures^[2]. Dans certains cas, une série d'évaluations peut être planifiée à différentes étapes du cycle de la réponse, comme pour l'évaluation du Darfour ([Broughton et al., 2006](#)). Par ailleurs, un suivi doit être réalisé tout au long d'une action humanitaire. Le suivi et l'évaluation sont des outils complémentaires pour aider à déterminer dans quelle mesure une intervention fonctionne bien.



Suivi

Processus continu de collecte des données : aide à répondre aux questions quantitatives, se concentre sur les moyens et les produits.

Le suivi est en général réalisé par les personnes qui mettent en œuvre un programme pour s'assurer

² Chapitre développé d'après [Imas Morra and Rist \(2009\)](#) et [Hallam \(2011\)](#).

qu'il reste sur la bonne voie^[3]. Par exemple, le suivi d'un programme d'aide alimentaire d'urgence se concentrera probablement sur les moyens tels que les quantités d'aide alimentaire distribuée, et sur les produits, comme le nombre de personnes recevant l'aide alimentaire. Il s'intéresse aussi parfois aux résultats – par exemple, les taux de malnutrition et les prix du marché pour les produits alimentaires. Le suivi est souvent faible dans l'action humanitaire et le manque de données de suivi fiables crée souvent des problèmes lors de l'évaluation.



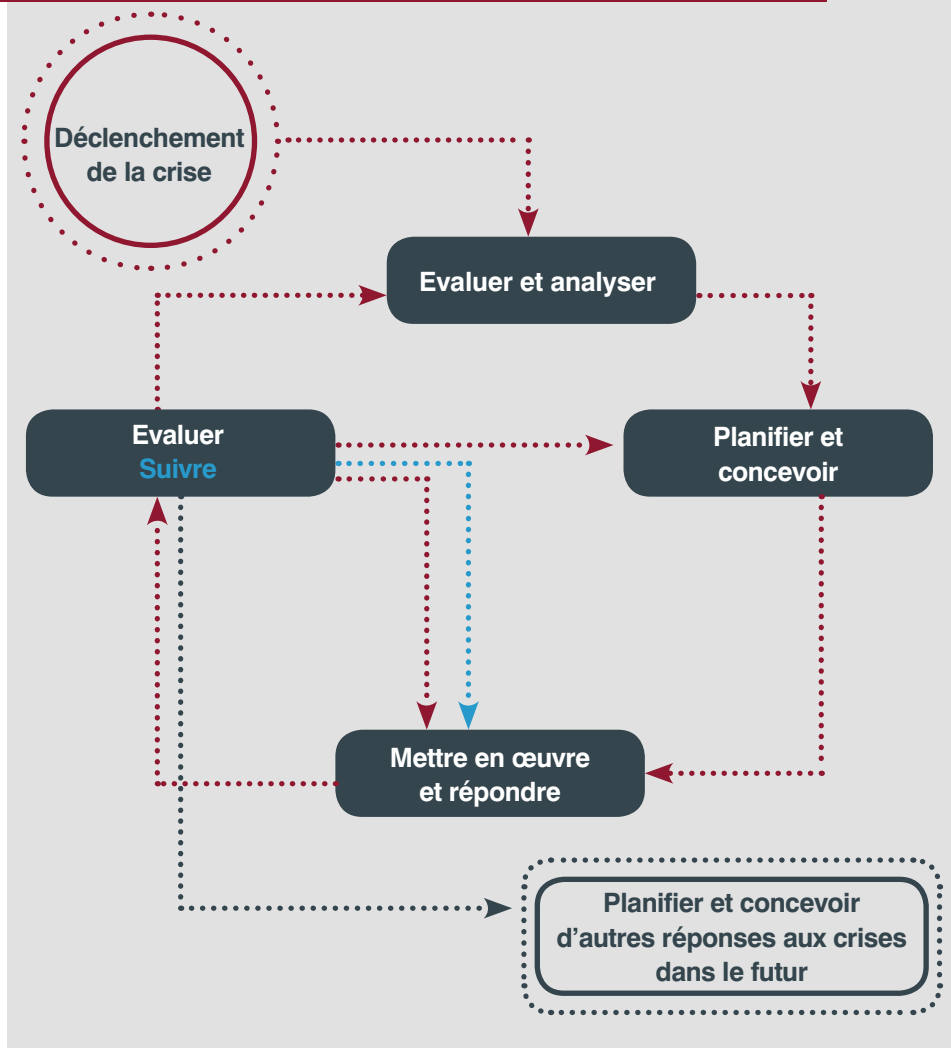
Evaluation

Examen ponctuel de la performance d'un programme ou d'un projet : aide à répondre aux questions quoi et pourquoi, se concentre sur les résultats et l'impact.

Un exemple d'impact serait par exemple l'effet d'un programme d'aide alimentaire d'urgence sur les marchés alimentaires locaux. L'évaluation se concentre sur les résultats atteints ou non atteints

³ Il existe des formes extérieures de suivi, comme les rapports d'évaluation citoyenne (Citizen Report Cards de la [World Bank, 2004](#)), mais on les trouve généralement plutôt dans les contextes de développement qu'humanitaires.

Schéma 2 : Suivi et évaluation dans le cycle de la réponse d'urgence



avec les moyens et les produits, sur la manière dont les organisations atteignent ou non leurs objectifs principaux et sur ce qui devrait changer pour qu'elles les atteignent. Elle peut être réalisée par du personnel interne ou externe, mais est souvent conduite par des personnes non impliquées dans la mise en œuvre du programme afin de bénéficier d'une meilleure crédibilité vis-à-vis des parties prenantes.

Pour s'assurer qu'une évaluation est utile et utilisée, elle doit être planifiée au moment où ses résultats peuvent au mieux contribuer à des instants clés de prise de décision au sein du cycle de réponse (**Schéma 2**).

Il n'est pas approprié de commanditer des évaluations pour chaque projet ou programme si cela est susceptible de surcharger l'organisation et si des systèmes de suivi efficaces sont en place. Les évaluations stratégiques sont plus à même de satisfaire des besoins d'informations clés de manière opportune.

Les évaluations choisies de façon stratégique sont fondamentalement différentes de l'approche cyclique décrite dans le **Schéma 2**.



Les questions clés à prendre en compte sont :

- comment le processus d'évaluation présente une valeur ajoutée pour l'organisation dans son ensemble ?
- combien d'évaluations stratégiques l'organisation a-t-elle la capacité d'absorber ?

L'unité d'évaluation de l'Agence suédoise de coopération internationale au développement a expérimenté avec succès cette approche.



Exemple
de bonne
pratique 1

Sélection stratégique des évaluations

Le cycle de planification des évaluations de l'Agence suédoise de coopération internationale au développement commence avec les échanges que l'unité d'évaluation mène auprès de toutes les unités opérationnelles afin de déterminer ce qu'elles aimeraient savoir et comment une évaluation pourrait les aider en cela. Une liste d'environ 100 idées d'évaluations est tout d'abord compilée, liste à partir de laquelle le département de l'évaluation en choisit 15. Celles-ci sont à ajouter aux 80 évaluations réalisées chaque année au niveau opérationnel (SIDA, 2011).

Il peut être particulièrement approprié de réaliser une évaluation :

- Pour un programme dont les résultats ne

sont pas connus ou sujets à débats

- Pour des interventions de grande taille et coûteuses
- Pour des initiatives pilotes
- Là où l'organisation a des intérêts stratégiques
- Là où les parties prenantes souhaitent une évaluation

Il n'est pas approprié de réaliser une évaluation :

- Quand il est peu probable qu'elle apporte de nouvelles informations
- Quand des questions de sécurité ou de manque de données pourraient réduire sa crédibilité

1.5 Les difficultés les plus fréquentes

Ce sont les mêmes circonstances de crise qui nécessitent une action humanitaire et expliquent que l'EAH est différente d'autres formes d'évaluation, en particulier l'évaluation de projets de développement, et souvent plus difficile^[4]. Les évaluateurs doivent clarifier les contraintes qu'ils rencontrent depuis le début de l'évaluation et la manière dont ils essaient de faire avec, du point de vue de la gestion des

⁴ Chapitre adapté de Beck (2006).

risques. Le rapport final doit indiquer comment ces contraintes ont affecté le processus et les résultats de l'évaluation. La liste qui suit décrit des difficultés communes à toutes les EAH et leurs solutions éventuelles.

- **L'urgence et le chaos des secours humanitaires** – Une réponse humanitaire rapide est, par définition, planifiée rapidement et souvent in extremis. Les documents de planification et de suivi peuvent être rares, les objectifs imprécis et les plans de départ peuvent rapidement devenir obsolètes à mesure que le contexte change ou est clarifié.

Solution éventuelle : penser à construire une chronologie de la crise et de l'action humanitaire à évaluer. Utiliser des entretiens avec le personnel de l'organisation pour identifier des objectifs réels ou implicites et la manière dont ils auraient pu changer au fil du temps.

- **Manque de données de référence** – Les données peuvent avoir été détruites lors de la catastrophe, elles peuvent aussi ne plus être pertinentes (par exemple, quand une grande proportion de la population a été déplacée).

Solution éventuelle : réaliser des entretiens pour demander aux bénéficiaires et aux informateurs locaux dans quelle mesure les conditions ont changé et pourquoi.

Inclure des questions similaires dans les enquêtes. Voir Chapitre 5.4 et Encadré 7.

- **Renouvellement du personnel** – Le fort renouvellement du personnel qui travaille dans l'action humanitaire, en particulier le personnel expatrié, complique le travail des évaluateurs pour ce qui est de trouver et interroger les informateurs clés.

Solution éventuelle : inviter les informateurs clés qui font ou faisaient partie du personnel à participer à un entretien téléphonique ou à une enquête en ligne (Chapitre 5.3.8).

- **Points de vue divisés** – Les conflits renforcent souvent les différences de points de vue : les événements sont sujets à des interprétations très diverses qui rendent difficile une évaluation « objective ».

Solution éventuelle : rassembler autant de points de vue différents que possible. S'assurer que différents avis sont représentés dans le rapport d'évaluation final bien que l'équipe d'évaluation puisse avoir comme consigne de donner son jugement sur certains aspects de la réponse humanitaire en fonction de l'information rassemblée. Voir Chapitre 5.2.5.

- **Perte de confiance en raison de la politisation ou d'un traumatisme** – Lors de conflits et d'autres crises humanitaires

qui se sont politisées et où l'on retrouve de manière généralisée des maltraitances, des traumatismes et de la peur, la confiance peut être perdue au sein de la population affectée, et il peut être difficile d'obtenir des réponses approfondies et précises aux questions.

Solution éventuelle : concevoir des méthodes de collecte de données et des manières de s'engager avec la population affectée qui soient sensibles aux traumatismes et aux peurs. Penser à associer intervieweur et interviewé selon leur sexe. Voir Chapitres 4.3 et 4.4.

- **Insécurité et manque d'accès** – Des questions de sécurité peuvent compliquer ou rendre impossible, en particulier dans des environnements de conflit, l'atteinte de la population affectée par les évaluateurs.

Solution éventuelle : explorer des manières innovantes de réaliser l'évaluation à distance. Voir Chapitre 6.1.

- **Manque de temps du personnel pour parler aux évaluateurs** – Les organisations humanitaires et leur personnel peuvent rencontrer des difficultés pour mettre en œuvre leurs programmes dans des environnements à fort stress. Prendre du temps pour le passer avec des évaluateurs peut alors être une priorité faible. La culture opérationnelle de l'action

humanitaire peut aussi constituer un obstacle à la réflexion.

Solution éventuelle : certains ateliers de réflexion et d'apprentissage assez courts peuvent apporter de bons éclairages s'ils sont bien animés. Voir Chapitre 5.5.1. Demander au personnel ce qu'il veut savoir à propos de ses opérations et dans quels domaines il est désireux de changer la manière dont l'organisation aborde une tâche.

- **Manque de lignes de responsabilité claires** – Le manque de responsabilités clairement définies entre acteurs humanitaires peut gêner la redevabilité et l'attribution de l'impact, en particulier lors d'une crise impliquant de nombreux acteurs humanitaires.

Solution éventuelle : se concentrer sur la contribution plutôt que sur l'attribution. Penser à réaliser une évaluation conjointe pour explorer l'impact combiné d'un certain nombre d'organisations ou de la réponse humanitaire internationale dans son ensemble, plutôt qu'à essayer d'isoler artificiellement l'impact d'une seule agence. Voir Chapitre 5.1.

- **Charges de travail non réalistes** – Les Termes de référence peuvent indiquer une activité de 24 personnes-mois avec un budget prévu de 4 personnes-mois.

Solution éventuelle : *utiliser un rapport de démarrage pour préciser le travail d'évaluation de manière à ce qu'il puisse être terminé dans le temps alloué. C'est une manière utile de gérer les attentes en amont. Voir Chapitre 4.1.1.*



Démarrage

Phase initiale d'une évaluation durant laquelle l'équipe d'évaluation essaie de développer une compréhension globale du travail d'évaluation et prépare un rapport détaillé décrivant le plan de l'évaluation.

- **Terrains isolés et infrastructures endommagés** – *Ces difficultés peuvent rendre l'accès difficile.*

Solution éventuelle : planifier avec soin le travail de terrain avec une personne qui connaît bien la situation actuelle, y compris les probables temps de transport et contraintes d'accès. De même, penser à confier au partenaire le calendrier des visites de terrain en fonction des critères définis par l'équipe d'évaluation.

- **Pression horaire sur la population affectée** – *Les personnes peuvent avoir peu de temps à consacrer à une évaluation à cause d'activités urgentes nécessaires à leur propre survie.*

Solution éventuelle : *éviter les méthodes chronophages comme les focus groups, et utiliser plutôt à la place des évaluations de groupe. Consulter les bénéficiaires quand ils attendent des distributions ou un transport. Utiliser l'observation directe du comportement des bénéficiaires au sein des populations affectées. Voir Chapitre 5.4.3.*

- **Contraintes éthiques concernant les approches expérimentales** – *L'hypothèse selon laquelle "si les personnes affectées par la sécheresse n'avaient pas reçu d'aide alimentaire, beaucoup d'entre elles seraient mortes" est souvent appelée contrefactuelle car elle va à l'encontre des faits (puisque l'on part du principe qu'elles ont reçu de la nourriture). Des considérations éthiques interdiraient bien sûr de tester cette hypothèse qui consisterait à priver certaines personnes de l'assistance.*

Solution éventuelle : *comparer les différentes méthodes d'assistance (par exemple, les dons en espèces, les bons d'achats, et la distribution alimentaire directe). Profiter de toutes les « expériences naturelles » susceptibles de se produire, comme un groupe qui n'aurait pas reçu d'assistance en raison de son isolement.*

- **Manque d'apprentissage dans la communauté humanitaire** – Les analyses d'ALNAP au sujet de l'action humanitaire révèlent que les mêmes problèmes refont surface année après année bien qu'ils aient été identifiés lors d'évaluations.

Solution éventuelle : se concentrer sur l'utilisation dès le début d'une évaluation. Consulter les probables utilisateurs et utiliser des méthodes qui leur semblent crédibles. Voir Chapitre 2. Encourager la conception d'un tableau de gestion des recommandations (**Schéma 3**) accompagnée d'un suivi régulier de la part des directeurs.



Tableau de gestion des recommandations

Enregistrement de la réponse associée à chaque recommandation de l'évaluation ainsi que des étapes que les dirigeants prévoient de suivre pour y répondre, comprenant une date butoir et la partie responsable de chacune de ces étapes.

Schéma 3 : Exemple de tableau de gestion des recommandations

	Nécessité d'un financement supplémentaire ?		Réponse de la direction			Commentaire	Action à mener	Timing	Unité responsable
	Oui	Non	Acceptée	Partiellement acceptée	Rejetée				
Recommandation 1									
Recommandation 2									

Source: FAO.

Certaines difficultés de l'évaluation peuvent trouver une réponse dans le fait de mener au préalable un examen de l'évaluabilité afin de déterminer si l'évaluation est faisable et la manière dont elle pourrait être conduite de façon optimale.



Il est bien plus facile d'organiser la gestion des recommandations lorsque celles-ci sont ciblées par rapport à des dirigeants ou des niveaux de gestion bien particuliers.



Etude de faisabilité d'une évaluation

Analyse de la mesure dans laquelle une activité ou un programme peut être évalué(e) de façon fiable et crédible.

NOTES

DECIDER DE FAIRE OU NON UNE EVALUATION

Partagez vos commentaires:

Aidez-nous à rendre le Guide plus pratique et focalisé sur l'utilisateur en partageant vos commentaires sur son contenu et la navigation. Les commentaires peuvent inclure des questions, des suggestions, des ressources utiles ainsi que des conseils pratiques. Vous pouvez partager vos idées en cliquant sur le bouton de commentaires en bas de chaque page. Ceci générera un courriel confidentiel à l'équipe EAH. Pour plus d'informations sur la façon dont vous, votre équipe ou votre organisation peut s'engager plus activement dans le pilotage du Guide contactez-nous directement à eha@alnap.org

PROMOUVOIR L'UTILISATION ET ASSURER LA QUALITÉ

Contenu de cette section

2.1	Se concentrer dès le départ sur la qualité et l'utilisation	35
2.2	Ce que signifie être focalisé sur l'utilisation	35
2.3	Comment les évaluations de l'action humanitaire sont utilisées	36
2.4	Les facteurs affectant l'utilisation	38
2.5	Les parties prenantes	41
2.6	Se focaliser sur ce que les principales parties prenantes doivent savoir	42
2.7	Impliquer les utilisateurs ciblés de l'évaluation	43
2.8	Les facteurs affectant la qualité de l'évaluation	43

Comment utiliser ce chapitre

Ce chapitre explicite ce que signifie être focalisé sur l'utilisation ainsi que les différentes manières d'utiliser l'évaluation (Chapitres 2.1 à 2.3). Les chapitres 2.4 à 2.8 proposent des façons de promouvoir l'utilisation en identifiant les principales parties prenantes en les faisant participer au processus d'évaluation et en prenant des mesures pour optimiser les chances que l'évaluation soit utilisée. Ces considérations sont importantes pour les responsables qui commanditent et gèrent les évaluations, et doivent être prises en compte au début du processus d'évaluation afin de réussir^[5].

⁵Chapitre inspiré de [Patton \(2008\)](#), [Sandison \(2007\)](#), et [Hallam \(2011\)](#).

Tableaux, schémas et encadrés

Tableau 3:	Les différentes manières d'utiliser les évaluations de l'action humanitaire	34
Schéma 4:	Les facteurs affectant l'utilisation d'une évaluation de l'action humanitaire	36
Encadré 2:	Cartographie des parties prenantes	38

2.1 Se concentrer dès le départ sur la qualité et l'utilisation

Même si de nombreuses évaluations de l'action humanitaire ont mis en lumière les mérites et l'importance de différents programmes et projets, et ont pu identifier des leçons clés pour favoriser l'apprentissage, leur capacité à susciter un changement est bien plus faible. Elles ont eu tendance par le passé à être quelque peu déconnectées de la culture et de l'état d'esprit des organisations humanitaires, ce qui risque d'endommager leur crédibilité. Par conséquent, une bien plus grande attention doit être accordée à la manière dont les évaluations sont utilisées si l'on veut qu'elles atteignent leur potentiel.

Une planification minutieuse, des temps de travail suffisants et des mécanismes adéquats pour un suivi convenable au service des utilisateurs caractérisent une réelle intention d'évaluation. (Sandison, 2007)

Au-delà de la qualité du seul rapport final, la qualité du processus d'évaluation dans son ensemble est un élément critique de son utilité et de sa potentielle utilisation (voir Chapitre 2.8). D'autres facteurs sont également essentiels pour assurer son utilisation

(voir Chapitre 2.4), mais le facteur sur lequel l'équipe d'évaluation a le plus de contrôle reste la qualité.

2.2 Ce que signifie être focalisé sur l'utilisation

Une évaluation focalisée sur l'utilisation s'effectue en ayant à l'esprit les principaux utilisateurs pour servir des usages bien particuliers, déclarés et pratiques. Lors de la planification d'une évaluation, il est essentiel d'identifier dès le début les utilisateurs visés et de les aider à décider ce qu'ils veulent accomplir avec une évaluation. Si cela n'a pas été fait au moment où une équipe d'évaluation est désignée, comme c'est souvent le cas, alors c'est à l'équipe de le faire. Afin d'être plus efficace, ce processus doit inclure des interactions fréquentes entre les utilisateurs et l'équipe. Les objectifs de l'utilisateur doivent guider le choix de l'approche et des méthodes qui devrait également prendre en compte des contraintes comme les limitations de temps et de ressources. L'accent mis sur l'utilisation et la collaboration avec les principaux utilisateurs devrait continuer de guider le processus d'évaluation, de la planification jusqu'à la mise en œuvre. Cette manière de penser et de travailler demande de l'engagement et du temps.



Exemple
de bonnes
pratiques
n°2

Aider les parties prenantes clés à clarifier leurs objectifs

En 2006, une évaluation orientée vers l'apprentissage a été réalisée pour l'organisation religieuse allemande Diakonie (**Welsh et Otto, 2006**). Les évaluateurs ont entièrement impliqué les partenaires de mise en œuvre de Diakonie sur le terrain. Ils ont encouragé l'organisation à impliquer ses partenaires locaux dans la préparation d'un rapport de démarrage pour identifier les objectifs et les méthodes de l'évaluation. Cette approche partagée de l'apprentissage a été nourrie par des recherches sur le terrain et les éclairages communs ont ensuite été partagés à des niveaux supérieurs. Des ateliers participatifs se sont déroulés lors des étapes préliminaires (orientation commune) et finales (interprétation commune) de chaque phase.

2.3 Comment les évaluations de l'action humanitaire sont utilisées

Les EAH sont utilisées de deux façons : (1) dans une optique de redevabilité (ou sommative), et (2) pour l'apprentissage, en général focalisé sur un projet spécifique (formatif), mais parfois focalisé sur un large système de connaissances (de développement). Une évaluation contient habituellement un mélange de différents éléments, même si l'un peut prévaloir sur les autres (**Tableau 3**).

Tableau 3 : Les différentes façons d'utiliser les évaluations de l'action humanitaire

Utilisation	Questions	Exemples
Sommative – juger le mérite ou la valeur d'un programme (par exemple, pour remplir la redevabilité vis-à-vis des acteurs ou les informer des décisions de financement)	Le programme répond-t-il aux besoins ? Quels sont les résultats ?	La publication du département britannique en charge du développement international (DFID) <i>Multilateral Aid Review (2011)</i> examine la manière dont le travail de différentes organisations a pu concilier les priorités des bailleurs.
Formative – pour renforcer l'apprentissage (améliorer un programme, par exemple)	Qu'est-ce qui fonctionne et qu'est-ce qui ne fonctionne pas ? Quelles sont les forces et les faiblesses actuelles?	<i>The Organisational Learning Review of Caritas Internationalis' Response to the Tsunami Emergency (Otto et al., 2006)</i> a facilité le processus d'apprentissage, en soulignant l'ouverture et la participation des parties prenantes clés.
De développement – pour renforcer l'apprentissage (par exemple, en introduisant de nouvelles idées au sein d'une organisation ou d'un secteur dans son ensemble)	Le programme prend-il en compte l'actualité et les limitations ? Quels sont les modèles généraux que l'on retrouve parmi tous les programmes ?	L'évaluation conjointe menée au Rwanda (<i>Borton et al., 1996</i>) a introduit de nouvelles idées concernant la redevabilité des organisations humanitaires, et a précipité l'arrivée de nombreuses innovations majeures en matière de politiques, comme les standards Sphere.

Note: Tableau réalisé d'après [Patton \(2008\)](#).

Les évaluations ne sont pas toujours utilisées de manière efficace. Ce phénomène peut revêtir différentes formes :

- **Utilisation ritualiste** – l'évaluation n'a de fonction que symbolique, par exemple pour remplir une obligation contractuelle vis-à-vis du bailleur de fond, mais avec peu ou pas du tout d'engagement de la part de l'organisation qui l'utilise.
- **Utilisation impropre** – des résultats sont supprimés, mal interprétés, ou faussés afin de servir à des fins personnelles ou politiques.
- **Non-utilisation** – les résultats sont ignorés car les utilisateurs y trouvent une valeur faible ou inexistante, ou alors ils ne sont pas conscients de sa valeur, ou encore le contexte a radicalement changé. La non-utilisation peut être une conséquence d'un manque d'engagement de la part de la direction, d'une mauvaise conception de l'évaluation, ou d'une incapacité de l'évaluation à répondre à ses propres questions ou à fournir des preuves irréfutables de ses résultats.

2.4 Les facteurs affectant l'utilisation

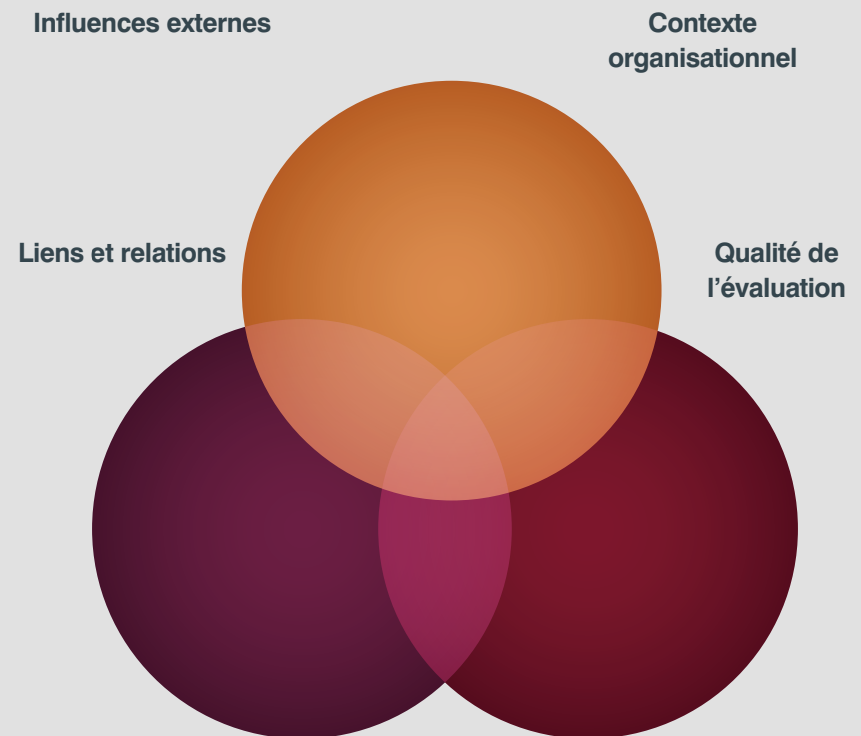
Les facteurs qui affectent l'utilisation sont classés en quatre grandes catégories :

1. **Qualité** – par exemple, la qualité de la période de démarrage, de la collecte d'informations, de l'analyse ou du produit final. Il s'agit du facteur le plus important qui affecte la crédibilité.
2. **Contexte organisationnel** – par exemple, la culture organisationnelle de l'apprentissage et de la gestion du savoir, et la structure organisationnelle en termes de proximité des unités d'évaluation vis-à-vis des décideurs. Les évaluations sont bien plus susceptibles d'être utilisées au sein d'une organisation qui recherche activement des informations sur la performance afin d'améliorer la gestion et la mise en œuvre du programme. Ce processus peut être fortement influencé, soit quand les principaux dirigeants soutiennent l'apprentissage, soit quand ils attachent peu d'importance à une prise de décision basée sur des données factuelles, et sont donc enclins à être sur la défensive vis-à-vis de l'évaluation.

3. **Relations** – par exemple, entre les évaluateurs et les utilisateurs de l'évaluation. Le facteur personnel – par exemple, la présence de poids lourds de l'évaluation parmi les acteurs clés avec lesquels il est difficile pour les évaluateurs d'établir et de maintenir une relation – peut être crucial pour déterminer si une évaluation est utilisée ou non. La position formelle et l'autorité de tels poids lourds peuvent avoir moins d'importance que leur enthousiasme et leur intérêt.
4. **Influences externes** – le public et les médias par exemple, ou des parties prenantes extérieures dont les actions peuvent influencer l'utilisation d'une évaluation.

Ces catégories sont représentées sous la forme d'un modèle à trois cercles dans le **Schéma 4**, les circonstances extérieures étant représentées par la zone ombrée en bleu autour des cercles.

Schéma 4 : Les facteurs affectant l'utilisation d'une évaluation de l'action humanitaire



Ce modèle est basé sur le cadre RAPID d'évaluation des liens entre recherche et politiques. Voir : <http://www.odi.org.uk/rapid/tools/Documents/Framework.pdf>.



Exemple
de bonnes
pratiques
n°3

Éléments pour une évaluation extrêmement influente

L'évaluation conjointe de l'aide d'urgence au Rwanda (Borton et al. 1996) est l'une des évaluations les plus influentes de l'action humanitaire jamais entreprises. Son succès peut être expliqué à l'aide du modèle à trois cercles (Schéma 4) :

- Cette évaluation était de grande qualité, en raison de la minutie et de la rigueur du travail, des compétences de l'équipe, et de la manière dont elle a su faire des recommandations. Cela lui a valu une grande crédibilité auprès des utilisateurs potentiels.
- En termes de structure et de culture organisationnelle, le système humanitaire international était ouvert au changement à ce moment-là, en partie à cause du malaise répandu concernant les performances très variables au sein des organisations humanitaires. Un grand nombre d'individus haut-placés ont défendu l'utilisation de l'évaluation et ont habilement su créer et exploiter des alliances ou des réseaux en son nom.
- Pour ce qui est des relations, les évaluateurs avaient des liens très forts avec les décideurs politiques du secteur.
- Les influences externes étaient toutes aussi importantes. Après le choc de la crise au Rwanda et de son intense couverture médiatique, les décideurs politiques étaient plus ouverts au changement

Durant le processus de planification d'une évaluation tournée vers l'utilisation, il est utile de se poser les questions suivantes :

- *Comment s'assurer que l'évaluation soulève des questions dont les réponses entraîneront une action ?*
- *Comment s'assurer que l'évaluation est d'assez bonne qualité pour lui garantir une crédibilité auprès de ses utilisateurs ?*
- *Comment le contexte organisationnel peut-il favoriser ou entraver l'adoption des résultats de l'évaluation ? Que pouvons-nous y faire ?*
- *Comment pouvons-nous développer les liens relationnels, par exemple entre les évaluateurs et les utilisateurs, afin d'assurer que les résultats et les recommandations de l'évaluation seront pris en compte ?*
- *Quelles influences externes peuvent favoriser ou entraver l'adoption des résultats de l'évaluation.*

Dans l'idéal, chacun de ces trois cercles devrait être puissant et chevaucher les autres. Le contexte externe devrait être propice aux évaluations. S'il n'est pas possible de créer ces conditions, il faut se demander si le moment est vraiment choisi pour conduire une évaluation, ou si une évaluation est le

meilleur outil pour atteindre vos objectifs.

2.5 Les parties prenantes

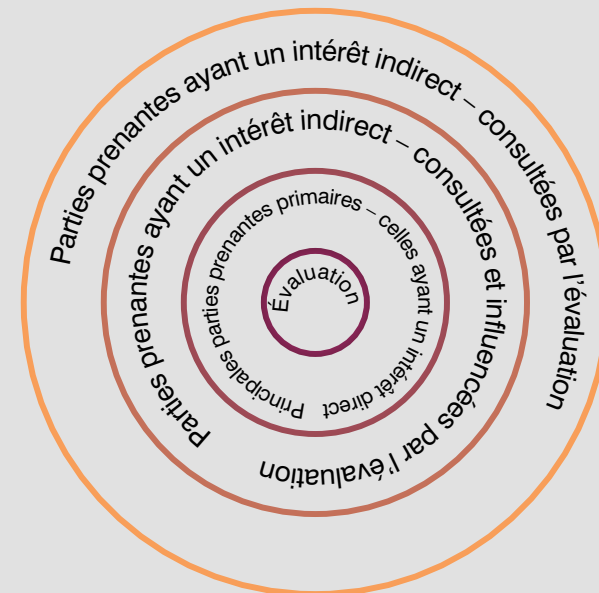
Identifier les parties prenantes clés est une étape cruciale dans la planification d'une EAH qui devrait être effectuée tôt. Les parties prenantes sont des individus qui seront affectés par les résultats de l'évaluation. Certains peuvent y trouver un intérêt direct – par exemple les organismes de financement, le personnel de mise en œuvre et les populations affectées. D'autres peuvent y trouver un intérêt indirect – par exemple, le personnel travaillant pour d'autres organisations humanitaires, les fonctionnaires et les contribuables.

Cartographier les parties prenantes est une façon utile de catégoriser les intérêts et les besoins de différents groupes d'acteurs. Si ce processus est fait visuellement, il peut également être participatif et impliquer de nombreuses personnes. Voir **Encadré 2**.

Encadré 2 : Cartographie des parties prenantes

Différentes parties prenantes peuvent être représentées en une série de cercles concentriques, en distinguant les principales parties prenantes qui sont censées être pleinement engagées dans l'évaluation, et celles ayant des intérêts indirects, qui peuvent jouer un rôle différent. Celles qui ont un intérêt indirect sont également divisées entre celles qui devraient être consultées et influencées par l'évaluation, et celles qui devraient à peine être consultées.

Une manière de réaliser ce type de carte est de noter toutes les différentes parties prenantes sur des post-it, de dessiner des cercles sur un tableau à feuilles, et de placer chaque post-it sur le cercle approprié selon les intérêts des différentes parties prenantes et leur degré d'engagement annoncé dans l'évaluation.



Les questions utiles à se poser à propos des parties prenantes sont entre autres :

- *Quelles sont les parties prenantes ayant un intérêt direct ? Parmi elles, qui sont les principaux utilisateurs attendus de l'évaluation ?*
- *Comment impliquer les parties prenantes primaires, encourager leur appropriation de l'évaluation et s'assurer que cette dernière réponde à leurs besoins de manière pertinente ?*
- *Quelle partie prenante primaire devrait constituer la première priorité afin de s'assurer que l'évaluation demeure pratique et ciblée ?*

Cette dernière question est particulièrement importante s'il existe un grand nombre de parties prenantes primaires avec des intérêts divergents et en compétition – par exemple, des organismes de financement souhaitant connaître l'efficacité et la rentabilité du programme par souci de redevabilité ou encore le personnel du programme souhaitant se renseigner sur ce qui a fonctionné et ce qui n'a pas fonctionné ainsi que sur la manière de s'améliorer. À moins de clairement identifier et prioriser les parties prenantes primaires et les utilisateurs attendus, les buts concurrentiels de l'évaluation peuvent devenir ingérables, et vous risquez d'en atteindre aucun.

2.6 Se focaliser sur ce que les principales parties prenantes doivent savoir

Une fois que vous avez identifié les parties prenantes primaires et les utilisateurs attendus, demandez-leur : qu'avez-vous besoin de savoir pour mieux décider quoi faire et comment ?

Cette question encourage une focalisation sur l'utilisation et minimise le risque de recevoir une liste de questions non ciblées. Il est probable que des problèmes difficiles surviennent, et les responsables de programmes doivent en prendre connaissance afin d'améliorer les programmes. Sondez les parties prenantes pour savoir ce qu'ils feront avec les informations qu'ils recevront de l'évaluation.



Cartographiez les parties prenantes primaires et regroupez-les par aires présumées d'intérêts communs pour simplifier le processus de questionnement sur ce qu'ils feront avec les résultats de l'évaluation.

Une fois obtenue la réponse à cette question cruciale, un suivi doit être assuré en demandant :

- *Qui a besoin de quoi, et pourquoi ?*
- *Quel est le but de cette évaluation ?*
- *Qu'évalue-t-on exactement ?*

2.7 Impliquer les utilisateurs attendus de l'évaluation

Les gens auront davantage tendance à utiliser les évaluations s'ils comprennent et se sentent faire partie du processus d'évaluation et des résultats. Ils sont aussi plus susceptibles de comprendre et de s'approprier l'évaluation s'ils ont été activement impliqués tout au long du processus. Pour cela, il faut notamment :

- *Découvrir ce que veulent les gens et pourquoi lors de la première étape de la conception de l'évaluation (Chapitre 3.3).*
- *Impliquer les utilisateurs potentiels dans la compilation des Termes de référence (Chapitre 3.11).*
- *Envisager la possibilité de permettre à quelques utilisateurs de rejoindre l'équipe d'évaluation (Chapitre 3.9).*

- *Constituer un groupe de référence pour l'évaluation en incluant les utilisateurs clés (Chapitre 3.7.1).*
- *Organiser un atelier où les utilisateurs et l'équipe de l'évaluation réalisent ensemble la conception de l'évaluation.*
- *S'assurer que l'équipe d'évaluation communique de manière régulière avec les utilisateurs durant le processus d'évaluation, et que les utilisateurs sont impliqués dans les décisions clés, par exemple à propos d'un recentrage de l'évaluation et des compromis.*
- *Organiser un atelier pour présenter les résultats de l'évaluation. Envisager de demander aux utilisateurs de participer à la rédaction des recommandations.*
- *Demander aux utilisateurs de concevoir une stratégie de diffusion (Chapitre 7.5).*

Les évaluations peuvent être une source d'inquiétude et de résistance chez ceux dont le travail est évalué. Ce phénomène peut être exacerbé dans les crises humanitaires à grande visibilité qui ont attiré l'attention des médias internationaux. Dans ces cas-là, il est d'autant plus important de construire un sentiment d'appropriation. Une communication fréquente tout au long de l'évaluation peut également aider à cela. Favoriser

l'appropriation parmi les évalués doit parfois être équilibré avec le besoin de préserver une certaine objectivité. Engendrer un engagement vis-à-vis de l'évaluation va souvent de pair avec une attitude d'ouverture au changement.

2.8 Les facteurs affectant la qualité de l'évaluation

La qualité d'une évaluation dépend de six facteurs clés : la conception, la participation, la planification, les produits finaux, les mécanismes de suivi et la crédibilité de l'évaluateur^[6]. Si vous pouvez répondre « oui » à toutes les questions ci-dessous, vous êtes sur la bonne voie pour mener à bien et produire une évaluation de grande qualité.

- **Conception** – *Le but de l'évaluation est-il clair, et a-t-il été décidé en accord avec toutes les parties prenantes clés ? L'approche a-t-elle été conçue en fonction de l'objectif mais aussi des besoins et des intérêts des utilisateurs ? (par exemple, pour une évaluation orientée vers l'apprentissage, l'équipe sur le terrain est-elle impliquée ?) La conception et les méthodes fourniront-elles des preuves suffisantes pour répondre aux*

questions soulevées par l'évaluation ? Des mécanismes appropriés sont-ils mis en place pour s'assurer que les menaces pesant sur l'indépendance de l'évaluation sont gérées ?

- **Participation et appropriation** – *Les parties prenantes clés seront-elles impliquées tout au long du processus d'évaluation ?*
- **Planification** – *L'équipe dispose-t-elle de suffisamment de temps pour s'impliquer adéquatement dans le processus d'évaluation ? Le moment de l'évaluation est-il bien choisi pour influencer correctement la prise de décision ?*
- **Produits** – *Les rapports et autres produits sont-ils accessibles, faciles à lire et à comprendre ? Les éléments de preuve sont-ils crédibles (bien documentés, objectifs) ? Les méthodes sont-elles clairement décrites et les limitations inhérentes identifiées ? Y a-t-il un cheminement logique de la preuve aux résultats, aux conclusions et enfin aux recommandations ? Les recommandations sont-elles claires, spécifiques, priorisées, constructives, mesurables, pertinentes, définies dans le temps et ciblées ?*

⁶Cette section se base sur [Sandison \(2007\)](#).

- **Mécanismes de suivi** – *Des plans de suivi spécifique ont-ils été établis dès le début ?*
- **Crédibilité de l'évaluateur** – *Les compétences et la réputation des évaluateurs sont-elles crédibles ? L'équipe d'évaluation possède-t-elle à la fois les compétences et l'expérience nécessaires pour cette évaluation ? Les évaluateurs sont-ils assez indépendants (et capables de maintenir cette indépendance lors de l'évaluation) ? Possèdent-ils les qualités nécessaires en matière de relations interpersonnelles ? Sont-ils capables de rester équilibrés, impartiaux, objectifs et constructifs ?*

NOTES

PROMOUVOIR L'UTILISATION ET ASSURER LA QUALITÉ

Partagez vos commentaires:

Aidez-nous à rendre le Guide plus pratique et focalisé sur l'utilisateur en partageant vos commentaires sur son contenu et la navigation. Les commentaires peuvent inclure des questions, des suggestions, des ressources utiles ainsi que des conseils pratiques. Vous pouvez partager vos idées en cliquant sur le bouton de commentaires en bas de chaque page. Ceci générera un courriel confidentiel à l'équipe EAH. Pour plus d'informations sur la façon dont vous, votre équipe ou votre organisation peut s'engager plus activement dans le pilotage du Guide contactez-nous directement à eha@alnap.org

PLANIFICATION ET CONCEPTION

Contenu de cette section

3.1	Les évaluations tournées vers l'apprentissage et la redevabilité	49
3.2	Les types d'évaluation	51
3.3	Déterminer les besoins en connaissances	60
3.4	Les critères d'évaluation	62
3.5	Les cadres d'évaluation	67
3.6	Les méthodes d'évaluation	74
3.7	Gestion et gouvernance	76
3.8	Choisir une équipe d'évaluation	80
3.9	Identifier les consultants externes	84
3.10	Budgéter une évaluation	87
3.11	Négocier les Termes de référence	90

Comment utiliser ce chapitre ?

Le chapitre **3.1** apporte des conseils pour identifier l'objectif d'une évaluation alors que le chapitre **3.2** se concentre sur le choix du type d'évaluation. Les chapitres **3.3** et **3.4** identifient des objectifs de connaissances et des critères d'évaluation alors que le chapitre **3.6** aide à décider le degré de spécificité à adopter pour déterminer le cadre de l'évaluation ainsi que les méthodes. Les chapitres **3.7** à **3.10** explorent la gestion, les contrats passés et l'élaboration du budget pour une évaluation. Le chapitre **3.11** se focalise sur l'élaboration des Termes de référence.

Tableaux, schémas et encadrés

Tableau 4 :	Les évaluations tournées vers l'apprentissage et la redevabilité	50
Encadré 3 :	Questions de premier niveau	61
Tableau 5 :	Les cadres d'évaluation	69
Schéma 5 :	Cadre conceptuel de l'UNICEF pour la malnutrition	72
Encadré 4 :	Évaluer la composante « ressources humaines » de l'AH	73
Tableau 6 :	Avantages et inconvénients relatifs à l'utilisation d'évaluateurs internes et externes	81
Tableau 7 :	Avantages et inconvénients à engager des consultants indépendants ou à signer un contrat avec un bureau d'études	85
Tableau 8 :	Éléments de coûts pour une évaluation	89
Tableau 9 :	Éléments des Termes de référence	91

3.1 Les évaluations tournées vers l'apprentissage et la redevabilité

Comme mentionné dans le chapitre 1.3, deux raisons principales poussent à entreprendre une évaluation : la redevabilité et l'apprentissage. La plupart des évaluations pratiques abordent ces deux aspects dans une certaine mesure, mais en théorie, l'un des deux devrait prévaloir. L'un des problèmes fréquents avec les EAH est qu'elles choisissent mal leur cible principale et accordent la même importance à ces deux objectifs. Il est important de clarifier l'objectif principal puisque cela peut influencer le style de l'évaluation et le choix des méthodes.

À titre d'exemple, une évaluation tournée vers la redevabilité est davantage susceptible d'adopter un style plus proche de l'investigation, cherchant à attribuer la responsabilité à la fois pour les succès et les échecs. Mais cela n'est pas forcément bénéfique à l'apprentissage si cela suscite une attitude défensive de la part de ceux qui ont besoin d'apprendre. L'apprentissage a besoin d'un environnement psychologiquement sécurisé (Edmondson, 2004) où il est possible de reconnaître les difficultés et d'admettre les erreurs.

Les évaluations participatives qui comprennent des ateliers sont une approche plus appropriée lorsque l'apprentissage est l'objectif prioritaire, car ceux qui veulent apprendre ou ont besoin d'apprendre sont impliqués de manière plus centrale dans le processus de réflexion. Le Tableau 4 montre à quel point les deux approches peuvent être différentes.

Tableau 4 : Les évaluations tournées vers l'apprentissage et la redevabilité

Élément de l'évaluation	Évaluation tournée vers la redevabilité	Évaluation tournée vers l'apprentissage
Termes de référence	Basée sur les contributions des parties prenantes externes ainsi que des responsables du programme	Devrait être mise en place par les personnes directement impliquées dans le programme qui souhaitent ou ont besoin d'apprendre
Composition de l'équipe	Personnel extérieur indépendant	Personnel interne, avec éventuellement un facilitateur externe ou une combinaison de personnels interne et externe
Approche privilégiée	Méthodologie de collecte et d'analyse des données (plus objective)	Processus de réflexion et d'élaboration de résultats (plus participatif)
Méthodes	Un mélange de méthodes quantitatives et qualitatives qui fourniront des preuves solides	Méthodes participatives impliquant les personnes qui doivent apprendre
Direction	Les personnes en charge de la redevabilité	Les personnes responsables de la gestion des connaissances et de l'apprentissage
Mode de direction	Plus directif	Plus axé sur la facilitation
Circulation du rapport	Dans le domaine public	Peut être limitée à l'organisation pour encourager une participation ouverte et franche



Lors de la planification et de la conception de l'évaluation, voici les questions clés à se poser :

- Quel objectif est le plus important – la redevabilité ou l'apprentissage ? Où, dans le continuum entre la redevabilité et l'apprentissage, placeriez-vous l'objectif ? Si vous ne voulez pas le perdre de vue, il ne doit pas être au milieu.
- Si la redevabilité est l'objectif principal : redevabilité envers qui et pour quoi ?
- Si l'apprentissage est l'objectif principal : par qui ?

Une fois cet objectif principal clarifié, il est possible de concevoir un processus et des méthodes d'évaluation pour l'atteindre.

3.2 Les types d'évaluation

Une fois que vous avez décidé qui sont les principales parties prenantes et les utilisateurs attendus, ce qu'ils veulent et ce qu'ils doivent savoir, ainsi que l'objectif global de l'évaluation, il vous faut ensuite décider quel sera le type d'évaluation. Voici les considérations à prendre en compte :

- *La portée de l'évaluation - par exemple, se concentre-t-elle sur un projet, un programme ou un secteur ?*
- *Faut-il privilégier les processus ou les résultats ?*
- *Faut-il privilégier la redevabilité ou l'apprentissage ?*
- *Le timing de l'évaluation – temps réel, à mi-parcours, ou après la fin de l'intervention.*
- *Combien d'acteurs seront évalués ?*
- *L'évaluation sera-t-elle basée sur la théorie du programme ? (Voir Chapitre 3.2.5)*
- *Tout autre aspect remarquable de l'évaluation.*

En général, les évaluations appartiennent simultanément à plusieurs catégories. Le type d'évaluation, ou une combinaison de plusieurs types, détermine également la conception de

l'évaluation (voir Chapitre 5.1).

3.2.1 Les évaluations présentant des périmètres différents

La plupart des évaluations se concentrent sur le niveau projet, programme, cluster ou sectoriel.



Évaluation au niveau projet

Évaluation d'une intervention humanitaire unique avec des objectifs, des ressources et un calendrier de mise en œuvre spécifiques, qui se déroule souvent dans le cadre d'un programme plus vaste.

Un exemple illustrant ce type est l'évaluation des projets de construction d'abris Tearfund et Tear Netherlands après le tremblement de terre de 2009 de Padang, à Sumatra (Goyder, 2010).



Évaluation au niveau programme

Évaluation d'un ensemble d'interventions présentant un objectif humanitaire global.

Un exemple illustrant ce type est l'évaluation de l'aide de Danida aux personnes déplacées internes en Angola (Cosgrave, 2004). Des programmes émergent parfois d'une série d'interventions avec des objectifs cohérents (comme dans cet exemple) plutôt qu'ils ne sont conçus dès le départ comme un programme cohérent, comme pour l'évaluation du programme de l'UNICEF au Timor-Leste (Tolani-Brown et al., 2010).



Évaluation au niveau cluster

Évaluation de multiples projets au sein d'un plus grand programme.

Deux exemples d'évaluations au niveau cluster sont l'évaluation réalisée par AusAid d'un cluster d'ONG après le tremblement de terre au Pakistan (Crawford et al., 2006) et du travail des ONG dans l'océan Pacifique (Crawford et Eagles, 2008).

Le terme d'évaluation au niveau cluster fait parfois référence à l'évaluation du système de Coordination par Cluster des Nations Unies, comme l'évaluation initiale réalisée par Stoddard et al., 2007, une approche de l'évaluation développée par la Fondation Kellogg dans les années 80 dans laquelle

l'évaluateur accompagne des projets disparates au début d'un programme lorsque les financements sont accordés et promeut l'apprentissage dans ces projets ([Barley et Jenness, 1993](#) ; [Sanders, 1998](#) ; [WK Kellogg Foundation, 1998, p. 17](#)).



Évaluation au niveau sectoriel

Évaluation d'un groupe d'interventions dans un secteur, toutes contribuant à l'accomplissement d'un objectif humanitaire spécifique. L'évaluation peut couvrir une partie d'un pays, un pays entier, ou plusieurs pays.

Un exemple illustrant ce type est l'évaluation réalisée sur le secteur Eau, Assainissement et Hygiène (WASH en anglais) en République Démocratique du Congo ([van der Wijk et al., 2010](#)).

3.2.2 Les évaluations d'impact ou de processus



Évaluation d'impact

Évaluation qui se concentre sur les effets à plus long terme plutôt que sur l'acheminement de l'aide (en général, mais pas toujours, effectuée quelques temps après la réalisation du projet).

L'évaluation du Bureau de la population, des réfugiés et des migrations du Département d'État américain de l'aide aux réfugiés du Burundi ([Telyukov et al., 2009](#)) constitue un bon exemple d'évaluation d'impact.

L'évaluation de la reconstruction du comté de Lofa sur une initiative communautaire est un autre excellent exemple. Toutes les évaluations qui étudient les résultats d'une intervention sont des formes d'évaluation d'impact.



Évaluation de processus

Évaluation qui se concentre sur les processus par lesquels les moyens sont convertis en produits. Peut aussi examiner l'intervention dans son ensemble.

Les évaluations de processus ont l'avantage de pouvoir se dérouler longtemps avant que les résultats ne soient clairement apparents. De nombreuses évaluations en temps réel contiennent une part importante d'évaluation de processus. L'évaluation conjointe de la réponse humanitaire à la crise au Rwanda ([Borton et al., 1996](#)) était à l'origine une évaluation de processus, même si elle en étudiait aussi les résultats. La plupart des évaluations humanitaires sont un mélange d'évaluation de processus et d'évaluation d'impact. L'évaluation normative est un type particulier d'évaluation de processus.



Évaluation normative

Évaluation qui compare ce qui a été mis en œuvre avec ce qui était planifié ou avec des normes spécifiques.

Les évaluations normatives sont un sous-ensemble des évaluations de processus. Elles peuvent évaluer la mise en place du projet en la comparant avec les plans du projet mais dans les EAH, elle est plus susceptible d'être évaluée par rapport à des standards. C'était le cas de l'évaluation de la réponse au tremblement de terre du Gujarat réalisée par le Disasters Emergency Committee qui s'est appuyée sur le Code de conduite des ONG ([SCRH et FICR, 1994](#)) pour réaliser des comparaisons. ([Humanitarian Initiatives UK et al., 2001](#)).

3.2.3 Les évaluations à des timings différents

Les évaluations s'inscrivent dans un continuum qui s'étend du plus rétrospectif au plus prospectif. Le timing de l'évaluation détermine à quel point ces deux éléments peuvent avoir de l'importance.

Les études de faisabilité de l'évaluation sont utiles s'il n'est pas clair qu'une évaluation crédible sera possible – comme, par exemple, l'examen de l'évaluabilité en Somalie (Cosgrave, 2010). Ces exercices sont parfois combinés avec les examens initiaux des programmes, comme ce fut le cas avec l'examen du programme et l'étude de l'évaluabilité du programme de l'UNICEF en faveur de l'éducation dans les situations d'urgence et de transition post-crisis (Barakat et al., 2010) ainsi que pour l'examen préparatoire du Pôle Mondial d'éducation (Reid et al., 2010)



Évaluation en temps-réel

Evaluation d'une opération humanitaire en cours à mesure qu'elle se déroule.

On trouve comme exemples d'évaluations en temps réel celle de la crise dans la vallée du Swat au Pakistan (Cosgrave et al., 2010) et celle des inondations de 2010 au Pakistan (Polastro et al., 2011b).



Évaluation à mi-parcours

Evaluation réalisée à la moitié d'une intervention.

L'évaluation à mi-parcours de la décision régionale « Sécheresse » de la DG ECHO en faveur de la Corne de l'Afrique (Wilding et al., 2009) est un bon exemple.



Évaluation ex-post

Evaluation réalisée à la fin d'une intervention.

L'évaluation qu'a réalisée CARE à propos de son intervention d'urgence post-tsunami dans deux régions au Sri Lanka (Bhattacharjee et al., 2007) est un bon exemple.



Évaluation itérative

Série d'évaluations conçue pour se dérouler tout au long d'une intervention.

Les évaluations de l'intervention humanitaire en Haïti menées par le IASC 3 et 20 mois après le tremblement de terre (Grünewald et al., 2010 ; Hidalgo et Théodate, 2012) constituent un exemple d'évaluation itérative. On peut également mentionner une évaluation de l'UNHCR (Haut-Commissariat des Nations Unies pour les réfugiés) qui s'est prolongée six ans après la fin de l'intervention (Skran, 2012).



Évaluation ex-ante

Evaluation conduite avant qu'une intervention ne commence.

De telles évaluations, rares dans le secteur humanitaire, sont basées sur les leçons apprises lors d'opérations précédentes similaires. C'est par exemple le cas de la mise en place d'un Corps volontaire européen d'aide humanitaire (Bruaene et al., 2010).

3.2.4 Les évaluations de et par différents acteurs



Auto-évaluation

Evaluation par ceux qui conçoivent et mettent en œuvre une intervention.

La plupart de ces évaluations sont des documents internes ; l'un des exemples publiés est l'évaluation de la réponse de Medair après le tsunami de 2004 au Sri Lanka, réalisée par un membre de l'équipe initiale au Sri Lanka qui est retourné par la suite sur le terrain pour réaliser l'évaluation (Lee, 2005).



Évaluation participative

Evaluation durant laquelle les parties prenantes, y compris les bénéficiaires, travaillent ensemble pour concevoir, mener et interpréter une évaluation.

Les évaluations participatives sont relativement rares dans le secteur humanitaire ; l'étude pays Colombie (Mendoza et Thomas, 2009) pour l'évaluation de la prise en compte de l'âge, du genre et de la diversité (Thomas et Beck, 2010) représente un exemple d'évaluation participative



Évaluation menée par une seule organisation

Evaluation menée par l'organisation qui a mis en place l'intervention.

On peut citer plusieurs exemples de ce type d'évaluation humanitaire : l'évaluation réalisée par CARE de son intervention après le tremblement de terre au Pakistan ([Kirkby et al., 2006](#)) et l'évaluation des programmes de transferts monétaires au Soudan réalisée par Oxfam ([Bush and Ati, 2007](#)).

L'un des exemples est l'évaluation par la Coalition pour l'évaluation des opérations de secours et de relèvement après le tsunami de 2004 dans l'Océan Indien ([Telford et al., 2006](#)) ; un autre est celui de l'évaluation de l'aide humanitaire dans le Centre-Sud de la Somalie ([Polastro et al., 2011a](#)).



Évaluation conjointe

Évaluation menée par deux organisations ou plus, évaluant le travail de deux organisations ou plus.

Ce type d'évaluation peut impliquer un bailleur et une organisation bénéficiaire, plusieurs organisations avec des mandats similaires, ou différents acteurs travaillant dans le même secteur. L'un des exemples est l'évaluation conjointe d'ONG partenaires dans un projet d'urgence de renforcement des capacités après le tremblement de terre de Yogyakarta ([Wilson et al., 2007](#)).



Évaluation de l'ensemble du système

Évaluation de la réponse du système humanitaire dans son ensemble, ouverte à tous les acteurs du système



Évaluation à base théorique

Évaluation axée sur la théorie sous-jacente du programme.



Théorie d'un programme

Théorie ou modèle présentant la manière dont une intervention peut contribuer à des résultats spécifiques à travers une série de résultats intermédiaires.

3.2.5 Les évaluations à base théorique

Les évaluations à base théorique ont surtout été utilisées dans le secteur du développement, en particulier pour des projets d'éducation et de santé publique. Les évaluations à base théorique explicite sont rares dans le secteur de l'humanitaire. ([Boller et al. 2011](#)) adoptent une approche à base théorique dans leur évaluation des programmes

de l'UNICEF de soutien au développement du jeune enfant. La théorie d'un programme apparaît parfois sous d'autres appellations, y compris celle de « logique de programme » (Funnell et Rogers, 2011). Identifier la théorie du programme est une étape cruciale pour réaliser une évaluation à base théorique.



Cadre logique

Matrice identifiant les moyens, les produits, les résultats et impacts, leurs relations causales, les hypothèses sous-jacentes, les indicateurs et les niveaux à atteindre.



Évaluation réaliste

Type d'évaluation à base théorique qui se concentre attentivement sur le contexte de l'intervention et envisage la cible de l'intervention non pas comme un groupe monolithique, mais plutôt comme un ensemble de sous-groupes distincts.

L'évaluation utilise le cadre logique du programme comme cadre. Voir le Chapitre 4.1.3 pour des détails

supplémentaires sur l'évaluation à base théorique et les cadres logiques.

Les évaluations réalistes ne posent pas la question « Qu'est-ce qui fonctionne ? » ou « Est-ce que ce programme fonctionne ? », mais « Qu'est-ce qui fonctionne, pour qui, dans quelles circonstances, dans quelle mesure, et comment ? » (Pawson et Tilley, 2004). Ainsi, une évaluation réaliste ne va pas examiner la manière dont différents sous-groupes au sein d'une population affectée vont bénéficier d'une intervention. Par exemple, le Rapport sur les liens entre Urgence, Réhabilitation et Développement dans la réponse au Tsunami (Christoplos, 2006) souligne que seuls les groupes avec des moyens de subsistance basés sur la pêche ont reçu une aide leur permettant de restaurer leurs moyens de subsistance.

3.2.6 Les autres types d'évaluation



Évaluation des mécanismes de financement

Évaluation portant sur une série d'interventions financées par un même mécanisme.

L'un des exemples est l'évaluation des projets de la FAO financés par le Fonds central d'intervention d'urgence de l'ONU (CERF en anglais) ([Cossée et al., 2010](#)).



Évaluation technologique

Evaluation d'une technique ou d'une technologie spécifique.

Ces évaluations ne sont pas fréquentes ; un exemple à citer est l'examen de l'utilisation du Module d'Assainissement à grande échelle de la Croix-Rouge britannique après le tremblement de terre de 2010 en Haïti ([Fortune et Rasal, 2010](#)). De telles évaluations peuvent aussi porter sur des approches innovantes comme l'utilisation de l'approche communautaire pour gérer la malnutrition aiguë au Népal ([Guerrero, 2010](#)).



Évaluation institutionnelle

Evaluation des dynamiques internes des organisations chargées de mettre en œuvre l'aide, de leurs instruments politiques, de leurs mécanismes d'acheminement des services et de leurs pratiques de gestion ainsi que des liens entre ces éléments.

On peut citer comme exemple l'évaluation du programme de l'UNICEF pour le renforcement de capacités humanitaires, financé par le Département pour le développement international du Royaume-Uni (DFID) ([Bhattacharjee et al., 2010](#) ; [Brown et al., 2005](#)) et l'examen indépendant de la réponse opérationnelle de l'UNICEF au tremblement de terre de 2010 en Haïti ([Bhattacharjee et al., 2011](#)).



Évaluation des politiques

Evaluation qui examine les compréhensions, croyances et hypothèses qui rendent les projets individuels possibles et souhaitables. Elle peut évaluer à la fois l'efficacité d'une politique et la manière dont la politique a été mise en œuvre.

Ces évaluations sont souvent de grande ampleur et ont tendance à être davantage des examens que des évaluations. Les exemples suivants illustrent ce type d'évaluation : l'examen de la réponse humanitaire des Nations Unies ([Adinolfi et al., 2005](#)), par DFID ([Ashdown, 2011](#)) et le document du bureau des Nations Unies pour la coordination des Affaires humanitaires (OCHA) : *To Stay and*

Deliver: Good Practice for Humanitarians in Complex Security Environments (Egeland et al., 2011).



Méta-évaluation

Il s'agit soit d'une évaluation conçue afin d'agrèger des résultats issus de plusieurs évaluations, soit d'une évaluation analysant la qualité d'une ou plusieurs évaluation(s).

La méta-évaluation d'évaluations conjointes réalisée par ALNAP (Beck et Buchanan-Smith, 2008) en est un exemple. Scriven (2011) fournit une liste de contrôle pour les méta-évaluations du second type.



Évaluation thématique

Évaluation portant sur une sélection d'interventions qui répondent à une priorité humanitaire spécifique commune à plusieurs pays, régions, et éventuellement organisations et secteurs.

L'un des exemples illustrant ce type d'évaluation est celle réalisée sur le rôle de l'évaluation des besoins après le tremblement de terre et le tsunami de décembre 2004 en Asie (de Ville de Goyet et Morinière, 2006). Ce fut l'une des cinq études thématiques réalisées pour la Tsunami Evaluation Coalition. L'examen récent du HCR sur l'éducation des réfugiés constitue un exemple d'évaluation thématique portant sur un seul secteur (Dryden-Peterson, 2011).

3.3 Déterminer les informations à connaître

Les dépenses engagées lors d'une évaluation ne sont justifiables que si celle-ci fournit des informations utiles et exploitables, et pas seulement ce que vous aimeriez savoir, mais aussi ce que vous avez besoin de savoir pour améliorer les programmes, c'est-à-dire des informations qui vont améliorer votre travail (voir Chapitre 2.6).

Deux niveaux de questions sont présents dans une évaluation : les questions de premier niveau qui apparaissent dans les Termes de référence (TdR) et auxquelles on demande aux évaluateurs de répondre, et les questions que les évaluateurs posent ensuite aux autres pour recueillir les données nécessaires permettant de répondre aux questions de premier niveau. Le Comité International de la Croix-Rouge limite ses évaluations à trois questions de premier niveau. Cela permet aux questions les plus essentielles de bénéficier d'une attention adaptée.

La Croix-Rouge avait l'habitude de demander à toutes les parties intéressées d'identifier la gamme de questions et de sujets qu'ils voudraient voir inclus dans une évaluation. Le département de l'évaluation réorganisait ensuite les réponses en questions « évaluables ». Cependant, il a été prouvé que le périmètre de l'évaluation augmentait toujours jusqu'à ce qu'il devienne difficile de gérer le processus. Afin d'atténuer cela, le département de l'évaluation essaye à présent de se concentrer sur trois questions clés seulement pour chaque évaluation (Hallam, 2011, p. 11).

L'un des plus grands problèmes avec de nombreux Termes de référence actuels dans l'EAH est que les questions sont présentées de manière trop détaillée et empiètent ainsi sur le deuxième niveau de questions.

Encadré 3 : Questions de premier niveau

Les questions de premier niveau doivent se concentrer sur les informations nécessaires pour prendre des décisions politiques et opérationnelles majeures. On peut retrouver par exemple les questions suivantes :

- L'intervention a-t-elle atteint son objectif, et si non, pourquoi ?
- Devons-nous poursuivre cette approche, et pourquoi ?
- Quels changements devons-nous effectuer afin de pouvoir mieux répondre à une crise similaire dans le futur ?

Lors de la phase de démarrage, les évaluateurs compétents doivent être en mesure d'identifier les questions clés de « second niveau » qu'ils devront poser afin de répondre aux questions de haut niveau.



Lorsque vous faites circuler des TdR provisoires, ne demandez pas que l'on vous suggère des questions. Vous vous retrouveriez invariablement face à des douzaines de questions très difficiles à rationaliser puisque chaque question est la préférée de quelqu'un. À la place, demandez aux parties prenantes quelles informations elles souhaitent obtenir et à quelles fins elles les utiliseraient. Nombre de ces demandes d'informations peuvent être incorporées dans des questions de haut niveau plus larges.

3.4 Les critères d'évaluation

Nous sommes engagés dans des processus d'évaluation au quotidien. Que ce soit en achetant des pommes au supermarché ou en choisissant un sandwich à la cafétéria, nous émettons un jugement sur les options qui s'offrent à nous en utilisant des critères de qualité et de valeur. Une EAH formelle utilise une approche bien plus structurée, mais se concentre globalement sur des critères similaires, même si la notion de « rapport coût-efficacité » (en comparant différentes approches de la même tâche) est souvent favorisée au détriment de la valeur (qui pourrait sous-entendre que la vie humaine et la souffrance ont une valeur monétaire).

Ces deux critères fondamentaux doivent être subdivisés en une série de critères qui peuvent être évalués de manière systématique. En 1991, le Comité d'aide au développement de l'Organisation de coopération et de développement économique (OCDE/CAD), se basant sur les problèmes les plus communs rencontrés dans les projets de développement, a proposé quatre critères de qualité - pertinence, efficacité, durabilité et impact - et le critère de valeur de l'efficacité (OCDE/CAD, 1991). Quelques années plus tard, l'OCDE/CAD a

adapté ces critères pour l'EAH dans des situations d'urgence complexes (OCDE/CAD, 1999), ajoutant les critères de couverture et de cohérence, mais aussi suggérant que le caractère approprié (« appropriateness ») constitue une alternative à la pertinence et que l'interdépendance constitue une alternative à la durabilité. L'organisation propose également deux nouveaux critères : la coordination et la protection. Ces critères reflétaient les plus grands problèmes rencontrés par l'action humanitaire dans les années 90.

Même si les critères de l'OECD/DAC ne sont pas parfaits, il existe de bonnes raisons pour les utiliser plutôt que d'inventer de nouveaux critères lors de chaque évaluation.

- *Utiliser les critères standards rend beaucoup plus facile le processus de méta-évaluation (tirer des leçons d'un large éventail d'évaluations).*
- *Les critères standards sont susceptibles de capturer les points faibles récurrents dans l'action humanitaire puisqu'ils sont nés de l'expérience et de la recherche.*
- *Il est plus facile pour les responsables d'évaluations et les évaluateurs de réaliser des évaluations qui utilisent les critères standards.*



Pour les nouveaux évaluateurs, le plus grand défi est bien souvent de choisir les meilleurs critères. Telle question relève-t-elle de l'efficacité ou de la pertinence ? Le choix de la réponse dépend fréquemment d'une question de préférence personnelle ou de convention. Il existe plusieurs façons valables de subdiviser les critères de qualité et de valeur, et les sous-catégories se chevauchent souvent.

Lors de la planification d'une évaluation, identifiez tout d'abord ce qu'il vous faut savoir, puis classez ces points au sein de critères – et non le contraire. Les critères sont des outils de réflexion qui peuvent faire naître de nouvelles questions pertinentes. Posez seulement des questions d'évaluation sur lesquelles vous êtes prêts à agir en vous basant sur les réponses. N'utilisez que les critères qui concernent les questions auxquelles vous voulez répondre. Les questions sont les plus importantes, et non les critères.

Sur la page suivante, nous résumons les critères de base de l'évaluation sous la forme d'une question tirée de la « vraie vie », suivie par un critère et sa définition, eux-mêmes suivis par une action qui peut être entreprise en se basant sur la réponse à la question. Ces critères prennent en compte les impacts, qu'ils soient positifs ou négatifs, intentionnels ou non intentionnels.

**CRITÈRES POUR
L'ÉVALUATION
DE L'ACTION
HUMANITAIRE**

Question : Dans quelle mesure la ration alimentaire était-elle adaptée aux besoins locaux ?

Critère : Caractère approprié



Caractère approprié

Dans quelle mesure les activités humanitaires sont adaptées aux besoins locaux.

Le caractère approprié augmente l'appropriation, la redevabilité et le rapport coût-efficacité. Il s'agit d'une variante du traditionnel critère OCDE/CAD, il mesure dans quelle proportion un projet correspond aux besoins locaux, aux priorités et aux politiques des bailleurs.

Actions requises : Il est possible que la direction change les procédures de planification et que les opérations transforment la conception des futurs paniers d'aide.

Question : Dans quelle mesure les distributions alimentaires ont-elles permis de réduire la malnutrition ?

Critère : Efficacité



Efficacité

Degré selon lequel une activité a atteint son objectif ou doit le faire en se basant sur les produits existants.

Actions requises : La direction peut affiner ses politiques pour de telles interventions ; les opérations peuvent changer la conception des programmes actuels et concevoir différemment leurs futurs programmes.

Question : En termes de coûts, a-t-il été plus efficace d'utiliser des partenaires de distribution plutôt que des prestataires ?

Critère : Efficience



Efficience

Mesure à la fois qualitative et quantitative des produits considérés comme un résultat des moyens.

Évaluer l'efficience requiert en général de comparer des approches alternatives permettant d'atteindre un produit, pour voir si l'approche la plus efficiente a bien été utilisée.

CRITÈRES POUR L'ÉVALUATION DE L'ACTION HUMANITAIRE

Actions requises : La direction peut changer ou conserver ses politiques pour ce qui est de donner la priorité aux partenaires ; les opérations peuvent négocier des tarifs plus justes avec leurs partenaires.

Question : Quel impact a eu la décision de distribuer de la nourriture aux femmes âgées plutôt qu'aux chefs de famille masculins en termes de dynamiques liées aux foyers ?

Critère : Impact

La notion d'impact a précédemment été définie comme tout effet sur un individu, un groupe, une communauté ou une institution. Un impact peut être positif ou négatif, intentionnel ou non intentionnel.

Actions requises : La direction peut changer ses politiques en ce qui concerne le mode de distribution ; les opérations peuvent affiner les pratiques de distribution, par exemple en prenant mieux en compte les dynamiques de genre.

Question : De quelle manière la distribution de l'aide alimentaire a-t-elle affecté le niveau de plantation pour la prochaine récolte ?

Critère : Interdépendance



Interdépendance

Mesure dans laquelle les opérations d'urgence à court terme prennent en compte les problèmes à plus long terme et interdépendants.

L'interdépendance est liée au concept de durabilité, à l'idée selon laquelle les interventions doivent soutenir des objectifs à plus long terme et, au final, être gérées sans apport des bailleurs.

Actions requises : Les opérations peuvent améliorer la conception des futurs programmes.

Question : Dans quelle mesure la distribution alimentaire ciblait-elle les plus vulnérables ?

Critère : Couverture



Couverture

Mesure dans laquelle l'aide atteint les principaux groupes de population affectés par la crise.

CRITÈRES POUR L'ÉVALUATION DE L'ACTION HUMANITAIRE

Actions requises : Les opérations peuvent améliorer la conception des futurs programmes.

Question : Dans quelle mesure la décision d'importer du sucre plutôt que d'en acheter au moulin public était-elle conforme aux politiques des bailleurs visant à encourager les achats locaux ?

Critère : Cohérence



Cohérence

Mesure dans laquelle il existe une cohérence entre les politiques de sécurité, de développement, de commerce, militaires et humanitaires, et dans laquelle toutes ces politiques prennent en compte les considérations humanitaires et liées aux droits de l'homme.

La cohérence est une question politique qui peut ne pas être pertinente dans le cadre d'évaluations d'une seule organisation ou d'un seul projet.

Actions requises : La direction peut chercher une meilleure façon de trouver un équilibre entre le rapport coût-efficacité et l'adhésion à des politiques globales.

Question : Existait-il des lacunes ou des doublons dans la distribution alimentaire qu'ont réalisée cette organisation et d'autres organisations.

Critère : Coordination



Coordination

Mesure dans laquelle les interventions de différents acteurs sont harmonisées, promulguent une synergie et évitent les lacunes, les doublons et les conflits liés aux ressources.

La coordination est parfois intégrée dans le critère d'efficacité plutôt que traitée comme un critère indépendant.

Actions requises : La direction et les opérations peuvent rechercher de meilleures façons de collaborer avec d'autres acteurs.

Beck (2006) donne des détails supplémentaires concernant l'utilisation des critères d'évaluation standards de l'OCDE/CAD, y compris à propos des débats autour des définitions, de l'utilisation des critères et des exemples de questions.

3.5 Les cadres d'évaluation

Les évaluations représentent un défi, il existe toujours le danger qu'un aspect essentiel soit négligé, que l'évaluation manque de rigueur conceptuelle ou qu'elle ne respecte pas les standards reconnus. Différents cadres peuvent être utiles dans ces cas :

- *Ils apportent une structure qui permet de diviser les activités de l'évaluation en de plus petits éléments plus faciles à gérer.*
- *Ils réduisent le risque que des éléments clés soient négligés, en dirigeant systématiquement l'attention de l'équipe vers tous les éléments du sujet de l'évaluation.*
- *Ils confèrent aux évaluations une structure qui sera reconnue par les parties prenantes. Celles-ci peuvent avoir été très peu habituées aux critères standards de l'évaluation, mais elles sont souvent à l'aise avec la structure des cadres conceptuels ou des standards utilisés couramment dans leur secteur.*
- *Ils peuvent constituer une référence en matière de bonnes pratiques généralement reconnues qui peut être testée dans le cadre du projet évalué.*

Les critères OCDE/CAD décrits dans le chapitre 3.4 fournissent un cadre d'évaluation possible, mais les utiliser en tant que cadre empêche la focalisation sur les questions d'évaluation.

D'autres cadres peuvent ajouter des détails et spécificités à l'évaluation et la rendre ainsi plus accessible à ses utilisateurs. De tels cadres peuvent être utilisés en complément des critères OCDE/CAD, ou même se substituer à eux le cas échéant. On demande souvent aux évaluateurs d'utiliser au moins l'un de ces types de cadres :

- *De larges cadres normatifs reflètent les normes qui définissent l'engagement humanitaire. Ils comprennent le droit international humanitaire, les principes humanitaires et diverses conventions.*
- *Les cadres conceptuels sont de larges modèles qui reflètent la manière dont un système est censé fonctionner. Ils sont souvent bien testés et validés. Les sources de ces cadres conceptuels comprennent les récents travaux de recherche académique, les principales évaluations et les documents de conception de projets.*

- *Des standards et des guides peuvent être utilisés à la fois comme des standards pour l'évaluation et comme une manière de subdiviser l'action humanitaire en de petites composantes plus faciles à examiner. Ces éléments comprennent certains guides sectoriels ainsi que des manuels opérationnels et des manuels d'urgence des organisations.*

Les cadres et leurs utilisations sont résumés dans le **Tableau 5.**

Tableau 5 : Les cadres d'évaluation

Cadres et exemples	Utilisation possible
<p>Les grands cadres normatifs :</p> <ul style="list-style-type: none"> • Droit humanitaire international comme le Protocole sur la protection des victimes de conflits armés non internationaux (CICR, 1977) • Principes humanitaires (Wortel, 2009) • Convention relative aux droits de l'enfant (Assemblée General de la ONU, 1989) 	<p>Les cadres normatifs peuvent être utilisés durant la phase de démarrage pour voir quels aspects de l'intervention pourraient susciter des inquiétudes et devraient être examinés de plus près. Certains cadres fournissent également des listes de vérification ou des standards avec lesquels examiner les politiques et la performance.</p>
<p>Les grands cadres conceptuels :</p> <ul style="list-style-type: none"> • Les catégories de consolidation de la paix d'Utstein (Smith, 2004) • Le cadre stratégique pour les moyens de subsistance dans les situations humanitaires complexes (Lautze et Raven-Roberts, 2003) • Le cadre conceptuel de l'UNICEF pour la malnutrition (Black et al., 2008) 	<p>Les cadres conceptuels peuvent être utilisés pour évaluer le succès de chaque élément d'une intervention et la mesure dans laquelle tous les éléments fonctionnent entre eux. Ils peuvent conférer une cohérence intellectuelle à une évaluation et assurer que tous les éléments sont abordés de façon appropriée.</p>
<p>Standards et lignes directrices concernant l'ensemble du système :</p> <ul style="list-style-type: none"> • Principes et bonnes pratiques pour l'aide humanitaire (Good Humanitarian Donorship, 2003) • Code de conduite des ONG et du Mouvement international Croix-Rouge/ Croissant-Rouge (Borton, 1994) • Code de People in Aid (People In Aid, 2003) • Standards pour un Partenariat en faveur de la Redevabilité Humanitaire (HAP, 2010) • Standards SPHERE (SPHERE Project, 2011) • Principes directeurs relatifs au déplacement de personnes à l'intérieur de leur propre pays (OCHA, 2004) • Principes directeurs relatifs aux États fragiles (OECD/DAC, 2007) 	<p>Ces normes peuvent à la fois servir de liste de vérification ou de référence pour évaluer la performance, de base pour subdiviser l'évaluation en plusieurs tâches gérables, et de structure pour le rapport. Elles sont plus efficaces lorsque qu'une organisation s'est formellement engagée à y adhérer ; si ce n'est pas le cas, leur utilisation peut être difficile à justifier. La phase de démarrage est le meilleur moment pour proposer l'utilisation d'un standard particulier.</p>

Cadres et exemples	Utilisation possible
<p>Standards sectoriels</p> <ul style="list-style-type: none"> • Éléments sectoriels des standards SPHERE (Sphere Project, 2011) • Standards minimum de l'INEE pour l'éducation en situations d'urgence (INEE, 2006) 	<p>Les standards sectoriels représentent une bonne base pour organiser des évaluations sectorielles. Dans certains cas, elles sont fondées sur des standards généraux ; par exemple, le standard SPHERE sur la consultation des populations affectées est censé s'appliquer à tous les secteurs.</p>
<p>Standards et guides des organisations</p> <ul style="list-style-type: none"> • Le manuel du Haut Commissariat des Nations Unies pour les réfugiés (UNHCR, 2007) • Le manuel du Programme alimentaire mondial relatif aux opérations d'urgence sur le terrain (PAM, 2002) • Le manuel de l'UNICEF relatif aux situations d'urgence (UNICEF, 2005) 	<p>Ces documents peuvent fournir une bonne base pour vérifier la conformité (redevabilité) et également pour subdiviser et organiser le travail d'évaluation.</p>

Dans l'idéal, les Termes de référence doivent identifier les cadres conceptuels et les standards que les évaluateurs sont censés utiliser. Par ailleurs, l'équipe d'évaluation peut vouloir proposer un cadre conceptuel et des standards pertinents lors de la phase de démarrage.



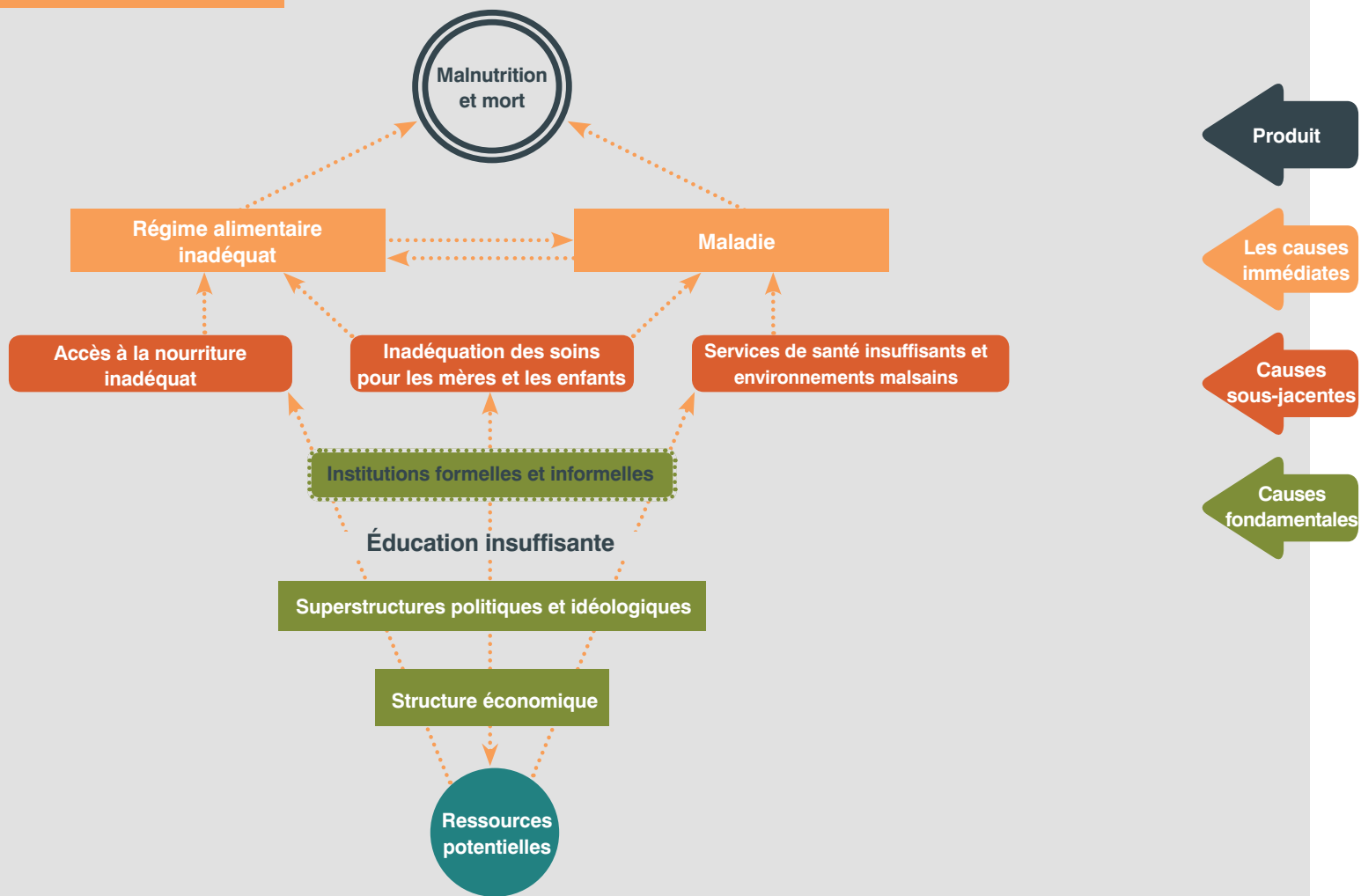
Parmi les questions clés à prendre en compte :

- Les documents de planification du projet font-ils référence à des cadres conceptuels ou à des normes particulières ?
- Est-il approprié d'utiliser l'un d'entre eux comme cadre pour l'évaluation ?



Si l'évaluation utilise un cadre qui s'avère central dans le travail des évaluateurs, cela peut être – comme les critères d'évaluation - une bonne façon de structurer le rapport. Cependant, si l'évaluation utilise des standards multiples, il peut être conseillé d'en choisir un comme cadre principal et d'aborder rapidement les autres dans les annexes du rapport principal.

Schéma 5 : Cadre conceptuel de l'UNICEF pour la malnutrition



Un exemple de cadre conceptuel est le cadre conceptuel de l'UNICEF pour la malnutrition présenté dans le **Schéma 5**. Les modèles conceptuels sont très souvent présentés sous la forme de diagrammes. Ce diagramme présente trois niveaux de causes : basiques, sous-jacentes et immédiates. Si vous n'évaluez pas un projet destiné à réduire la malnutrition et les morts qui lui sont associées, vous pouvez examiner dans quelle mesure le projet a exploré et traité au moins les causes immédiates et sous-jacentes. L'un des exemples de l'utilisation de ce cadre est le projet de recherche du SCF sur les causes de la malnutrition chez les enfants de moins de trois ans en Éthiopie (**Meyer, 2007**).

L'**encadré 4** présente un exemple d'utilisation pratique du cadre où les questions de l'évaluation de second niveau s'inspirent d'un cadre fourni par le Code de bonnes pratiques de People in Aid (**People in Aid, 2003**).

Encadré 4 : Évaluer la composante « ressources humaines » de l'AH

Le personnel des organisations humanitaires, recruté à la fois au niveau national et international, est un facteur clé dans le succès ou l'échec d'une intervention humanitaire. Pourtant, la composante « ressources humaines » de l'action humanitaire est souvent négligée dans l'EAH, sauf lorsqu'il y a de réels problèmes. Les questions clés de l'évaluation pour ce qui est de la gestion des ressources humaines sont les suivantes :

- Dans quelle mesure la structure et la capacité des effectifs sont-elles adaptées à une mise en place efficace des opérations ?
- Dans quelle mesure les capacités de gestion sont-elles adaptées à une mise en place efficace des opérations ?
- Dans quelle mesure les systèmes, pratiques et politiques liés aux ressources humaines soutiennent-ils ou entravent-ils les opérations ?
- Comment la direction a-t-elle soutenu ou entravé la performance globale ?

Cet encadré a été rédigé à partir de contributions et de conseils fournis par People in Aid.

3.6 Les méthodes d'évaluation

La sélection de méthodes d'évaluation détaillées (voir Chapitre 5) fait normalement partie du travail de l'équipe d'évaluation. Cependant, le responsable d'évaluation doit prendre en compte certaines questions méthodologiques car elles affectent la durée et le budget nécessaires à l'évaluation.

Il est du rôle d'un responsable d'évaluation de spécifier que l'équipe d'évaluation est censée utiliser à la fois des méthodes quantitatives et qualitatives de façon rigoureuse. Il peut également y avoir des cas où le responsable d'évaluation doit spécifier qu'une méthode en particulier doit être utilisée – par exemple, pour des raisons institutionnelles ou de compatibilité avec d'autres évaluations.

En général, le rapport de démarrage de l'équipe doit proposer, à l'intérieur du cadre fixé par le responsable d'évaluation les méthodes spécifiques qu'elle va utiliser, dans l'attente de l'approbation par le responsable d'évaluation. Pour des évaluations plus importantes, un groupe de revue par les pairs apporte souvent des commentaires sur la méthodologie proposée.



Un responsable d'évaluation qui spécifie une méthode particulière prend sa part de responsabilité quant aux résultats si cette méthode n'est pas adaptée pour répondre à la question posée par l'évaluation dans ce contexte précis. En général, il vaut mieux spécifier soit les méthodes, soit les résultats attendus, mais pas les deux.

3.6.1 Les approches quantitatives et qualitatives

Les méthodes d'évaluation se divisent en deux grandes catégories : quantitatives et qualitatives. Les responsables d'évaluation sont parfois contraints d'utiliser l'un ou l'autre.



Méthode quantitative

Méthode utilisée pour collecter des données numériques.

Si une étude quantitative est bien conçue, ses résultats peuvent être généralisés pour une population plus importante en utilisant des méthodes statistiques (voir Chapitre 5.2). En raison de leur origine (les sciences naturelles), les méthodes quantitatives sont parfois considérées comme plus scientifiques que les méthodes qualitatives. Les méthodes quantitatives

prédominant dans les sciences naturelles qui sont la branche la plus ancienne des sciences. Cependant, leurs fondements – les observations sont indépendantes de l’observateur, les faits sont séparés de la théorie – s’avèrent problématiques dans le cadre de recherches sociales. Les méthodes qualitatives ne reposent pas sur de tels fondements et se nourrissent d’une riche littérature abordant ces interactions.



Méthode qualitative

Méthode utilisée pour récolter des données non-numériques, en général des descriptions de significations ou d’expériences.

La plupart des EAH reposent sur un mélange de ces méthodes, avec cependant une prédominance de méthodes qualitatives. De telles méthodes sont en effet plus à même de répondre à des questions d’évaluation types ; de plus, l’utilisation rigoureuse des méthodes quantitatives dans la complexité d’un contexte humanitaire type pose des difficultés techniques (voir Chapitre 5.4.5).

3.6.2 Consulter les populations affectées

Il existe de nombreuses options pour consulter les populations affectées (voir Chapitre 5.4). Par exemple, vous pouvez mener un sondage quantitatif, ou utiliser un sondage qualitatif détaillé déjà réalisé par un chercheur en sciences sociales. Vous pouvez aussi mener des entretiens ou utiliser des techniques d’évaluations participatives rapides.

Des données détaillées sur le suivi de la performance peuvent être disponibles pour certains projets. Allez-vous directement consulter les populations affectées ou le faire à travers les structures administratives ? Êtes-vous dans une simple optique de consultation, ou voulez-vous que les populations participent directement à l’évaluation dans laquelle elles sont engagées depuis le début ?

Si vous souhaitez que l’équipe d’évaluation mène des consultations détaillées avec la communauté plutôt que des visites rapides comme c’est actuellement la norme, accordez du temps supplémentaire pour réaliser un travail de terrain. La consultation des populations affectées a longtemps constitué la faiblesse de l’EAH (voir Chapitre 4.3).

3.6.3 Recherche bibliographique

Les considérations à ne pas négliger lors d'une recherche bibliographique sont les suivantes : (1) quels sont les documents pertinents et comment les rendre accessibles ? (2) combien de temps doit être alloué à cette recherche bibliographique ? Il s'agit d'un élément essentiel de l'évaluation pour lequel on n'a généralement pas assez de temps. Le chapitre 5.3 couvre les méthodes d'études de cas, se concentrant sur des tâches spécifiques comme l'analyse du contenu et le travail sur des grands ensembles de documents.

3.7 Gestion et gouvernance

Le travail du responsable d'évaluation démarre lorsque la décision a été prise (ou s'apprête à être prise) de mener une évaluation. Dans certains cas, la politique ou le plan d'évaluation d'une organisation peut rendre cette décision routinière ; dans d'autres cas, le responsable d'évaluation peut avoir besoin de faire pression sur les parties prenantes pour obtenir l'appui nécessaire afin de réaliser une évaluation.



Une visite sur le terrain de la part du responsable d'évaluation en amont de l'évaluation peut s'avérer utile pour affiner les Termes de Référence et obtenir du soutien de la part des parties prenantes sur le terrain.

3.7.1 Les groupes consultatifs

Les dispositifs qui permettent au mieux de gérer une évaluation font en sorte que les principales parties prenantes demeurent engagées dans leurs décisions. Les groupes consultatifs sont souvent formés avec cet objectif. Les types les plus communs sont présentés ci-dessous.

GRUPE DE PILOTAGE



Groupe de pilotage

Groupe établi pour piloter une évaluation à travers des étapes clés comme l'établissement des Termes de Référence, ou encore la rédaction du rapport initial et du rapport final.

En général, un groupe de pilotage est composé de responsables d'évaluation provenant de plusieurs organisations.

Un leadership solide est nécessaire pour permettre au groupe de travailler de manière efficace et d'exploiter pleinement son potentiel, en particulier si les membres proviennent d'horizons divers.

GROUPE DE GESTION



Groupe de gestion

Groupe qui gère l'évaluation au jour le jour et dont les tâches comprennent la rédaction des Termes de Référence, le recrutement et la gestion de l'équipe d'évaluation, la gestion de la relecture et la finalisation du rapport d'évaluation.

Il s'agit en général d'un sous-groupe du groupe de pilotage ; ce groupe peut être remplacé par un seul gestionnaire d'évaluation. Pour un grand groupe de gestion, il est conseillé d'établir un plus petit groupe de pilotage qui peut rapidement prendre des décisions si nécessaire sans avoir besoin d'appeler à la tenue d'une réunion du groupe de gestion.

GROUPE DE RÉFÉRENCE



Groupe de référence

Groupe constitué des principales parties prenantes connaissant bien l'environnement local et pouvant apporter des conseils sur des questions pratiques en lien avec l'évaluation et sur la faisabilité des recommandations qui en découlent.

Établir un groupe de référence est une bonne façon d'impliquer les principales parties prenantes quand une évaluation a été commanditée par le siège.

GROUPE DE REVUE PAR LES PAIRS



Groupe de revue par les pairs

Groupe qui conseille sur les questions de qualité ; souvent composé d'évaluateurs et d'autres spécialistes choisis pour leurs connaissances de l'évaluation, de la région ou du type d'intervention évalué.

Un groupe de revue par les pairs peut aussi apporter des conseils sur l'analyse contextuelle et la méthodologie, mais aussi émettre des commentaires sur la rédaction du rapport. Il n'est pas nécessaire que ce groupe comporte beaucoup de personnes. Il peut s'avérer particulièrement utile pour des évaluations qui sont gérées par un directeur général plutôt que par un évaluateur et quand l'équipe d'évaluation ne connaît pas bien la région.



Tenez séparés le groupe de revue par les pairs et le groupe de référence afin que chacun puisse se concentrer sur la tâche qui lui a été assignée et pour éviter qu'ils ne soient distraits. Même si leurs responsabilités semblent se recouper, ils fonctionnent mieux de manière séparée.

Les groupes consultatifs ont parfois des appellations différentes selon les organisations. Les vastes évaluations conjointes tireront profit de tous ces groupes, mais des dispositifs plus simples sont plus appropriés pour de plus petites évaluations. Même la plus petite évaluation peut tirer partie d'un groupe de référence et d'un groupe de revue par les pairs ou même d'un seul conseiller.

3.7.2 Gérer les conflits

Les disparités entre l'échelle du travail d'évaluation et les ressources disponibles peuvent mener à un conflit entre le responsable d'évaluation et l'équipe d'évaluation. Un grand nombre de ces conflits tirent leur origine de différences de compréhension. Le rapport de démarrage peut aider à réduire le risque d'un tel conflit durant les premières phases, mais un conflit peut aussi émerger lors de la rédaction du rapport, entraînant de multiples révisions et une frustration accrue des deux côtés.



L'une des façons de minimiser les conflits dans le cadre d'évaluations litigieuses est de faire appel à un groupe de pilotage pour se faire conseiller sur la recevabilité des rapports (même si la décision finale appartient toujours au responsable d'évaluation). Les groupes de pilotage peuvent aussi contribuer à assurer la qualité de l'ensemble du processus d'évaluation.

Les conflits de personnalité sont une autre source potentielle de désaccord, l'évaluation étant un domaine qui attire les fortes personnalités. Il peut être utile d'adopter une politique formelle de résolution des conflits spécifiant quelles sont les différentes étapes en cas de conflit et identifiant une personne neutre qui pourrait servir d'arbitre entre les parties.



Exemple
de bonne
pratique
n°4

Établir une procédure de résolution des différends

La Tsunami Evaluation Coalition a adopté une politique sur la résolution des différends entre les différentes équipes d'évaluation, entre une équipe et le comité de pilotage, au sein de l'équipe de synthèse et entre l'équipe de synthèse et le groupe de direction. Grâce à cette politique, l'auteur principal était chargé de gérer les relations au sein de l'équipe.

Lorsqu'il y avait un conflit important entre les membres de l'équipe sur une question précise, ou entre un membre et le chef de l'équipe, que l'auteur principal n'était pas capable de régler, le groupe de direction demandait au Secrétariat d'ALNAP de fournir un rapport présentant les deux aspects du problème. Le chef du secrétariat d'ALNAP pouvait soit directement préparer un rapport, soit embaucher un évaluateur expérimenté pour le faire. Le groupe de direction pouvait ensuite décider de demander à l'auteur principal de mener l'une des actions suivantes :

- Ne présenter qu'une seule des interprétations
- Inclure les deux interprétations dans le rapport
- N'inclure qu'une seule interprétation, mais mentionner que cette interprétation n'était pas unanime.

Dans le cas où les membres de l'équipe de synthèse n'étaient pas satisfaits de la résolution du problème, ils avaient le droit d'exiger que leur nom soit supprimé du rapport.



Les procédures formelles de résolution des conflits sont appropriées pour toute évaluation nécessitant de telles procédures, pas seulement pour de grandes évaluations.

3.8 Choisir une équipe d'évaluation

La décision de faire appel à une équipe d'évaluation extérieure, interne ou mixte dépend de l'objectif de l'évaluation. Si l'évaluation vise principalement la redevabilité, les évaluateurs doivent être externes. En fonction du contexte organisationnel, il peut être préférable d'inclure un ou deux employés dans l'équipe – provenant éventuellement du département d'évaluation de l'organisation, s'il existe, ou, si cela est pertinent, de l'un des départements opérationnels. Ces membres ne doivent jamais être majoritaires ou assumer un rôle de leadership, l'inclusion de tout autre personnel extérieur dans l'équipe pouvant mettre en danger l'indépendance perçue de l'évaluation.

Si l'objectif principal de l'évaluation est l'apprentissage, l'équipe d'évaluation doit soit inclure une majorité de personnel interne (qui est supposé bénéficier des processus d'apprentissage) ou engager une équipe d'évaluateurs externes dont

le rôle principal sera de faciliter l'apprentissage de l'équipe.

Les avantages et inconvénients relatifs à l'utilisation d'évaluateurs externes et internes sont présentés dans le **Tableau 6**.

Tableau 6 : Avantages et inconvénients relatifs à l'utilisation d'évaluateurs internes et externes

Évaluateurs internes		Évaluateurs externes	
+	Les bénéfices que les évaluateurs de l'apprentissage et de la réflexion retirent durant le processus d'évaluation demeurent au sein de l'organisation.	+	Ils sont souvent plus objectifs.
+	Ils connaissent l'organisation et sa culture.	+	Ils risquent moins d'être influencés par l'organisation.
+	Ils sont connus du personnel.	+	Ils apportent un regard neuf.
+	Ils peuvent être moins menaçants et bénéficier d'une plus grande confiance.	+	Ils peuvent avoir une plus grande expérience sur laquelle s'appuyer.
+	Les résultats et recommandations peuvent être plus adaptés à l'organisation.	+	Ils peuvent consacrer plus de temps à l'évaluation.
+	Les recommandations ont souvent beaucoup plus de chance d'être adoptées.	+	Ils peuvent servir d'experts ou de facilitateurs extérieurs.
+	Ils coûtent moins cher.	+	Ils ne font pas partie de la structure de pouvoir de l'organisation.
+	Cette option renforce la capacité d'évaluation interne et contribue, de manière générale, au renforcement de la capacité du programme.	+	Ils peuvent apporter des ressources supplémentaires.
-	Leur objectivité peut être questionnée.	+	Ils sont plus souvent formés et expérimentés en matière d'évaluation.

Évaluateurs internes	Évaluateurs externes
<ul style="list-style-type: none"> — La structure organisationnelle peut limiter leur participation. 	<ul style="list-style-type: none"> + Ils sont considérés comme des experts.
<ul style="list-style-type: none"> — Leurs engagements professionnels peuvent limiter leur participation. 	<ul style="list-style-type: none"> — Ils peuvent ne pas connaître l'organisation.
<ul style="list-style-type: none"> — Leurs motivations peuvent être questionnées. 	<ul style="list-style-type: none"> — Ils peuvent ne pas être au courant des contraintes qui vont affecter l'adoption des recommandations.
<ul style="list-style-type: none"> — Ils peuvent éventuellement accepter trop facilement les hypothèses de l'organisation. 	<ul style="list-style-type: none"> — Les bénéfices qui découlent des réflexions et de l'apprentissage effectués durant le processus d'évaluation ne demeurent pas au sein de l'organisation.
<ul style="list-style-type: none"> — Ils peuvent ne pas être formés aux méthodes d'évaluation. 	<ul style="list-style-type: none"> — Ils peuvent être perçus comme des adversaires.
<ul style="list-style-type: none"> — Faire appel à des évaluateurs internes peut réduire la crédibilité de l'évaluation en dehors de l'organisation. 	<ul style="list-style-type: none"> — Ils coûtent plus cher.
<ul style="list-style-type: none"> — Ils peuvent avoir des difficultés à éviter certains biais. 	<ul style="list-style-type: none"> — Les embaucher peut demander de longues négociations de contrat
<ul style="list-style-type: none"> — Ils peuvent manquer d'expertise technique. 	<ul style="list-style-type: none"> — Utiliser des évaluateurs externes peut compliquer le suivi des recommandations.

Évaluateurs internes	Évaluateurs externes
	<ul style="list-style-type: none"> — Ils peuvent ne pas connaître l'environnement.
	<ul style="list-style-type: none"> — Ils peuvent être influencés par leur désir d'être rappelé pour un nouveau travail et donc être moins indépendants qu'ils ne le disent.

Les biais peuvent également s'avérer un vrai enjeu dans les évaluations.



Biais

Soit les préjugés de l'évaluateur, soit des erreurs systématiques qui surviennent en raison de la conception de l'évaluation.

Le premier type de biais se rencontre par exemple lorsqu'un évaluateur n'aime pas une organisation en particulier ou une approche adoptée par un programme, ou quand il est tenté de ne pas divulguer des résultats problématiques par peur de ne plus se voir confier à l'avenir de travaux d'évaluation par son client.

Le second type de biais peut être illustré par le fait de peu couvrir des endroits moins accessibles pour avoir fait le choix de ne visiter que les endroits les plus faciles d'accès.

3.9 Identifier les consultants externes

Obtenir le(s) consultant(s) le(s) plus approprié(s) pour votre évaluation est une étape importante qui prend du temps. Plus vous recrutez en amont, plus vous aurez de choix.

Le type de consultant varie selon la tâche d'évaluation. Lorsque l'on traite des questions sensibles de partenariats, vous avez besoin de consultants ayant une compréhension des relations institutionnelles en général et des partenariats en particulier. Pour les évaluations en temps réel, vous aurez besoin de consultants ayant suffisamment d'expérience opérationnelle pour être crédibles face au personnel sur le terrain. Avant de choisir des consultants, plusieurs considérations sont à prendre en compte :

- *À quel point le consultant connaît-il l'organisation ? Un consultant déjà familier de l'organisation aura besoin de passer moins de temps à apprendre à la connaître, mais sera peut être moins à même de remettre en cause ses idées préconçues. À quel point le consultant connaît le contexte. Cet aspect est particulièrement*

important pour des crises et des conflits complexes. En fonction de l'évaluation, il peut être plus important de comprendre le contexte ou le type d'activité à évaluer.

- *Le consultant possède-t-il l'expérience et les compétences prérequisées pour l'approche d'évaluation proposée ?*
- *La composition globale de l'équipe. Les plus grandes équipes peuvent offrir davantage de compétences, mais sont aussi plus coûteuses et requièrent plus de support logistique sur le terrain. De plus, elles impliquent un plus grand risque de conflits internes et prennent plus de temps pour écrire un rapport.*

Un autre problème est celui du recrutement des consultants de manière individuelle ou en tant qu'équipe via une société de conseil. Les avantages et inconvénients de chaque choix sont résumés dans le **Tableau 7**.

Tableau 7 : Avantages et inconvénients à engager des consultants indépendants ou à signer un contrat avec un bureau d'études

Consultant indépendant		Bureau d'études	
+	Vous sélectionnez exactement les personnes et les compétences que vous souhaitez.	+	Il n'y a qu'un seul contrat à gérer.
+	Vous pouvez vous assurer que tous les membres de l'équipe ont quelque chose à offrir.	+	Vous avez plus d'influence sur les questions de qualité.
+	La contractualisation peut être moins complexe qu'avec un évaluateur externe.	+	La résolution des conflits au sein de l'équipe est la responsabilité du contractant.
+	Vous bénéficiez d'une relation directe avec chaque membre de l'équipe.	+	Vous ne devez gérer qu'un petit nombre de paiements.
+	Cette option est souvent moins coûteuse.	+	Le contractant assume une part des risques.
-	Les limitations financières peuvent vous empêcher d'embaucher les consultants les plus qualifiés.	-	Cette option est généralement plus coûteuse.
-	Il est de votre responsabilité de résoudre les conflits au sein de l'équipe.	-	L'équipe peut parfois être composée d'un seul expert confirmé entouré d'une équipe de « juniors ».
-	Il vous faut fournir des ressources humaines pour soutenir l'équipe.	-	Certains des membres de l'équipe peuvent ne jamais avoir travaillé ensemble auparavant.
-	Vous êtes en charge d'organiser la logistique sur le terrain.	-	Il vous faudra parfois gérer la logistique sur le terrain et d'autres tâches si le contractant n'est pas compétent.

Consultant indépendant	Bureau d'études
<ul style="list-style-type: none"> Vous devez gérer de nombreux paiements différents. 	<ul style="list-style-type: none"> Les limitations financières peuvent vous empêcher d'embaucher une équipe dans le temps opportun.
<ul style="list-style-type: none"> Vous assumez tous les risques encourus. 	

Vous pouvez parfois sélectionner des consultants sur la base de vos expériences conjointes passées, mais vous ne disposez pas toujours de ce luxe. Bien sûr, vous pouvez demander de des précisions sur l'approche qu'ils souhaitent adopter, mais la meilleure option est de sélectionner les consultants sur la base de leurs précédents travaux. Qui a écrit les rapports d'évaluation que vous considérez de bonne qualité ?

La relation contractuelle dépendra de la politique de l'organisation. Une pratique commune est de lier les paiements aux différents stades de l'évaluation, comme la signature du contrat, la validation du rapport de démarrage, la fin du travail sur le terrain et l'acceptation des diverses versions et du rapport final.



Ne vous basez pas simplement sur le potentiel d'un auteur apparaissant dans une évaluation passée, contactez le responsable d'évaluation et renseignez-vous sur les antécédents professionnels expériences professionnelles du consultant. Il est possible que le rapport final soit en réalité la 19e version et ait été complètement réécrit par un autre membre de l'équipe.

3.10 Budgéter une évaluation

L'un des principaux coûts d'une évaluation est le nombre total de jours de consultance externe. En général, il s'agit du produit de la taille de l'équipe et du nombre de jours de travail sur le terrain, multiplié par les taux de rémunération des consultants et les indemnités journalières de subsistance (perdiems). Un budget peut être divisé en plusieurs phases : préparatoire, travail sur le terrain et travail de retour du terrain.

- **La phase préparatoire** peut comprendre une première réunion pour discuter de l'évaluation, des lectures de fond et du travail de bureau, un entretien préliminaire au siège et la rédaction du rapport de démarrage. La réunion *initiale* peut prendre 1 ou 2 jours, en comptant les déplacements, mais tous les membres de l'équipe ne sont pas obligés d'y participer. Les lectures de fond, les entretiens préliminaires et le rapport initial peuvent prendre de 1 à 5 jours pour une petite évaluation, et un mois ou plus pour une évaluation plus vaste et plus complexe. Une petite évaluation (d'une seule activité sur un seul site par une seule organisation) prendra en général de 3 à 7 jours par consultant pour le travail préparatoire.
- **Travail sur le terrain** : les coûts sont fonction de la durée du travail sur le terrain et du prix des consultants. En règle générale, les méthodes quantitatives demandent davantage de travail de terrain et de plus larges équipes ; les méthodes qualitatives demandent davantage de planification et de capacité d'analyse. Une petite évaluation prend généralement une semaine d'observation et d'entretiens sur les lieux choisis pour l'observation, comprise entre une demi-semaine pour le premier briefing et une autre demi-semaine pour les entretiens de suivi et le débriefing. Plus de temps est requis s'il y a plusieurs lieux à visiter sur le terrain, ou si l'équipe est censée produire un rapport avant de quitter le pays. Une petite évaluation comprend généralement entre 12 et 14 jours de travail sur le terrain, une grande évaluation peut prendre des mois de travail de terrain. En règle générale, on compte une semaine de travail pour chaque pays visité, et une semaine de plus pour chaque site majeur (une province ou un district) à visiter. Si l'équipe s'est engagée dans un processus détaillé de consultation des bénéficiaires, il faudra compter 2 ou 3 semaines par site. Les méthodes d'enquêtes quantitatives sont plus coûteuses et devraient être budgétées selon un échantillon et selon le temps qu'il faudra pour traiter et analyser les informations.

- **Le travail de retour du terrain** comprend avant tout l'assemblage et l'analyse des données collectées puis l'écriture des rapports. Cela peut inclure les débriefings, les cycles d'examen et la diffusion. Le fait d'assembler des données et de rédiger des rapports nécessite toujours plus de temps que la période allouée. Le temps requis varie selon la complexité du rapport et la quantité d'analyse requise. Les débriefings peuvent prendre de 1 à 2 jours, rédiger le rapport entre 1 et 20 jours, et chaque cycle de relecture de 2 à 10 jours. Lorsqu'une équipe est composée de beaucoup de membres, une bonne règle consiste à ajouter 5 jours à la préparation du rapport pour permettre au responsable d'évaluation d'incorporer leur contribution. Une petite évaluation consacre généralement entre 7 et 12 jours aux activités de retour du terrain.

Les éléments budgétaires de base sont énoncés dans le **Tableau 8**.

Tableau 8 : Éléments de coûts pour une évaluation

Ligne budgétaire	Éléments possibles
Personnel	Rémunérations et indemnités du personnel, indemnités du personnel des partenaires et autre personnel.
Consultants	Chef d'équipe, consultants internationaux, consultants nationaux, autres consultants.
Personnel de soutien	Rémunérations et indemnités de l'administration, des chercheurs, des interprètes, des conducteurs, du personnel de sécurité, et autres.
Déplacements	Visas, vols pour l'équipe d'évaluation et le personnel accompagnateur, vols pour les réunions, déplacements internes de l'équipe et du personnel accompagnateur.
Indemnités journalières	Logement et perdiem des consultants.
Saisie de données	Saisie et nettoyage de données pour supprimer les réponses qui n'ont pas de sens, comme quelqu'un enregistré comme étant à la fois un homme et enceinte.
Réunions et ateliers	Location du lieu, repas, et indemnités.
Production du rapport	Relecture, traduction, illustrations, maquette, mise en page, impression, développement de supports électroniques, distribution.
Autres	Communication, e-mails et courriers postaux, téléconférences, licences et frais juridiques, sécurité.
Coûts administratifs	

3.11 Négocier les Termes de référence

La conception de l'évaluation est clarifiée dans les Termes de Référence, un outil clé de l'évaluation qui fournit aux parties prenantes une compréhension commune de ses objectifs et finalités, des questions clés et d'autres questions pratiques et stratégiques.

Les Termes de référence devraient être négociés avec les principales parties prenantes de l'évaluation. Ils permettent au responsable d'évaluation de définir les tâches de l'évaluation et d'exposer aux membres de l'équipe d'évaluation ce que l'on attend d'eux. Cette compréhension commune de la tâche est essentielle pour réussir une évaluation (**Tableau 9**).



Élaborer les Termes de Référence en organisant des visites de cadrage sur le terrain peut aider à s'assurer que l'équipe sur le terrain et le bureau d'évaluation se comprennent bien.

Tableau 9 : Éléments des Termes de référence

Élément	Commentaires
Contexte	Cela ne nécessite qu'un paragraphe ou deux ; la plus grande partie des connaissances contextuelles de l'équipe d'évaluation viendra de lectures préliminaires. Les introductions de propositions de projet ou les documents d'appel sont de bonnes sources.
Objectifs et finalités de l'évaluation	<p>L'objectif principal de l'évaluation est-il l'apprentissage ou la redevabilité (Chapitre 3.1) ? S'il s'agit des deux, quel objectif a la priorité ? Par exemple, une évaluation peut servir à informer les bailleurs de l'efficacité avec laquelle leur argent a été utilisé (redevabilité) mais peut également examiner les facteurs dans la conception et la mise en œuvre du projet qui ont conduit à des projets plus efficaces (apprentissage). Toutes les évaluations, qu'elles soient essentiellement pour la redevabilité ou l'apprentissage, fournissent des opportunités d'apprentissage.</p> <p>Il est utile de se demander si une évaluation représente la meilleure façon d'atteindre les objectifs indiqués. Un audit permettrait-il de mieux satisfaire les exigences de redevabilité ? D'autres processus (revue après action, ateliers) satisferaient-ils mieux les exigences d'apprentissage ?</p>
Périmètre de l'évaluation	<p>Quels secteurs, géographie, phase d'intervention, et laps de temps l'évaluation va-t-elle couvrir ? Se concentrera-t-elle sur le niveau politique ? Le niveau du programme ? Les processus ou opérations spécifiques ? S'agit-il d'une évaluation en temps réel, d'une évaluation à mi-parcours ou d'une évaluation ex-post ? S'agit-il d'une évaluation d'une seule organisation, d'une auto-évaluation ou d'une évaluation conjointe ?</p> <p>La direction ou la politique de l'organisation définit parfois le périmètre de l'évaluation. Dans d'autres cas, vous aurez peut être à développer le périmètre en discutant avec des collègues. De bons Termes de référence spécifient non seulement ce qui doit être évalué mais aussi ce qui ne doit pas être évalué, et pourquoi. Par exemple, si une organisation a récemment évalué son travail sur l'eau et l'assainissement d'urgence, il serait logique d'exclure ce secteur d'une autre évaluation en cours.</p>
Utilisation attendue de l'évaluation	Comment l'évaluation va-t-elle être utilisée, et par qui ? Clarifier ces questions va aider l'évaluation à présenter un bon rapport coût/efficacité. Voir Chapitre 2 pour des astuces sur la manière de s'assurer que votre évaluation est focalisée sur l'utilisation.

Élément	Commentaires
Pourquoi l'évaluation est-elle entreprise, et pourquoi maintenant ?	Y a-t-il des délais internes ou externes ? L'évaluation est-elle dépendante de décisions dans le cadre d'un cycle de financement ou de décisions concernant la stratégie future de l'organisation ? De tels délais et liens doivent être clarifiés si vous avez impliqué les principales parties prenantes depuis le début, et doivent être explicités dans les Termes de référence.
Modèles conceptuels et standards internationaux pertinents	Existe-t-il un modèle conceptuel que vous souhaitez voir utilisé par les membres de l'équipe d'évaluation lors de l'élaboration de leur méthodologie et leur analyse – par exemple, le cadre des moyens de subsistance relatif aux urgences humanitaires complexes ? Quels sont les standards internationaux pertinents pour cette évaluation – par exemple, les standards SPHERE ou le code de conduite de People in Aid ?
Les principales questions d'évaluation	<p>Quelles sont les questions principales auxquelles vous voulez que l'évaluation réponde ? Quels critères OCDE/CAD voulez-vous utiliser et comment sont-ils reliés ?</p> <p>Conseil: <i>Conservez un faible nombre de questions d'évaluation afin que l'évaluation reste ciblée.</i></p>
Méthodologie	Cette partie ne doit pas être un exposé détaillé mais doit plutôt s'attacher à indiquer les préférences méthodologiques – par exemple, si vous souhaitez que l'équipe consulte des personnes en particulier, comme des représentants du gouvernement ou des bénéficiaires, ou encore utiliser des méthodes spécifiques comme une enquête formelle. En général, l'équipe doit développer la méthodologie détaillée. Si le responsable d'évaluation spécifie à la fois le produit et le processus, cela risque de réduire le sentiment d'appropriation de l'équipe d'évaluation et leur part de responsabilité dans le produit final.
Calendrier	<p>Spécifiez des dates pour les produits livrables clés et toute autre date butoir. Demandez aux parties prenantes si ces dates posent problème – par exemple, si elles tombent pendant la saison des pluies, la saison de récolte, des élections, ou des périodes administrativement chargées.</p> <p>Conseil : <i>Allouez suffisamment de temps afin de faire circuler un rapport provisoire auprès des relecteurs pour commentaires (généralement 2 ou 3 semaines) puis pour l'intégration des révisions (2 ou 3 semaines de plus).</i></p>

Élément	Commentaires
Rôles et responsabilités	Précisez qui est la personne responsable du transport sur le terrain, qui est responsable de prendre les rendez-vous et d'autres tâches. Vérifiez avec vos collègues sur le terrain qu'ils peuvent correctement soutenir l'équipe d'évaluation.
Dispositifs de gestion	Spécifiez s'il y a aura des comités consultatifs et quels seront leurs rôles.
Produits	<p>Les produits comprennent en général le rapport de démarrage, les comptes-rendus de visites sur le terrain (sous forme de notes ou de présentation), un rapport principal, le résumé de l'évaluation et les comptes-rendus des réunions (voir Chapitre 7.1). Précisez la longueur et le format si vous avez une préférence. Il est également utile de mettre en place des paiements par phase pour chaque produit spécifique afin d'encourager l'équipe d'évaluation à maintenir une approche focalisée sur son but.</p> <p>Si le rapport de l'évaluation doit adhérer à un certain « style » propre à votre organisation, ou s'il existe de fortes préférences pour le format et le style, spécifiez-le afin d'éviter d'avoir à reformater le rapport plus tard. Précisez la longueur du rapport (le nombre de mots est une mesure plus précise et moins susceptible de faire l'objet de malentendus que le nombre de pages) et quelles sont les annexes souhaitées. Vous pouvez également préciser la structure globale du rapport même s'il est préférable de donner à l'équipe d'évaluation une certaine flexibilité à ce sujet.</p>
Gestion des risques	Décrivez les risques et les défis susceptibles d'émaner de l'évaluation, et la manière dont le responsable d'évaluation propose de les gérer.
Budget	Donnez un budget indicatif pour l'évaluation.

Élément	Commentaires
Évaluation des offres	S'il n'y a pas un document d'appel d'offre formel, vous pouvez inclure les critères sur lesquels vous vous basez pour évaluer les différentes offres (quel pourcentage de la note ira au prix, à la composition de l'équipe, à la méthodologie et à d'autres critères) et le délai de réception des offres.
Références clés	Vous pouvez attacher une liste des références clés pour l'évaluation. La plupart sont des documents internes, mais vérifiez auprès des principales parties prenantes si d'autres documents doivent être ajoutés aux lectures de démarrage.



Les questions clés à se poser durant la planification de l'évaluation sont les suivantes :

- Les parties prenantes clés sont-elles prêtes « à intégrer le navire » de l'évaluation et ont-elles été impliquées de façon adéquate dans la rédaction et la validation des Termes de Référence ?
- Les objectifs et finalités sont-ils clairement définis ?
- A-t-on bien défini le périmètre de ce qui va être (et de ce qui ne va pas être) évalué ?
- L'utilisation prévue et toutes les contraintes temporelles sont-elles clairement énoncées ?
- Le nombre des questions d'évaluation est-il raisonnable et gérable ?
- Existe-t-il des exigences méthodologiques clairement définies ?
- Les rôles et responsabilités, les dispositifs de gestion et les produits désirés sont-ils clairement définis ?
- Les risques et les réponses prévues pour y faire face sont-ils identifiés ?
- Le budget est-il clair ?

NOTES

PLANIFICATION ET CONCEPTION

Partagez vos commentaires:

Aidez-nous à rendre le Guide plus pratique et focalisé sur l'utilisateur en partageant vos commentaires sur son contenu et la navigation. Les commentaires peuvent inclure des questions, des suggestions, des ressources utiles ainsi que des conseils pratiques. Vous pouvez partager vos idées en cliquant sur le bouton de commentaires en bas de chaque page. Ceci générera un courriel confidentiel à l'équipe EAH. Pour plus d'informations sur la façon dont vous, votre équipe ou votre organisation peut s'engager plus activement dans le pilotage du Guide contactez-nous directement à eha@alnap.org

REALISER UNE EVALUATION D'ACTION HUMANITAIRE

Contenu de ce chapitre

4.1	La phase de démarrage	99
4.2	Les défis liés au <i>leadership</i> et au travail en équipe	114
4.3	Impliquer les populations affectées	117
4.4	Les questions d'éthique	119

Comment utiliser ce chapitre ?

Ce chapitre est destiné aux évaluateurs et aux responsables d'évaluation. Le chapitre 4.1 couvre la phase de démarrage durant laquelle les évaluateurs explicitent le sens du travail de l'évaluation, et prévoient leur approche et leur méthodologie. Le chapitre 4.2 apporte des conseils sur le travail en équipe et le leadership afin de surmonter des écueils fréquents dans la gestion de l'EAH. Le chapitre 4.3 donne des conseils pour consulter les populations affectées et le chapitre 4.4 répond à certaines questions éthiques liées à la réalisation d'EAH.

Tableaux, schémas et encadrés

Tableau 10 :	Éléments du rapport de démarrage	94
Tableau 11 :	Exemple d'une matrice d'évaluation	96
Tableau 12 :	Exemple de cadre logique	99
Schéma 6 :	Exemple de chaîne des résultats	100
Schéma 7 :	Modèle logique pour une évaluation humanitaire	102
Tableau 13 :	Liste de vérification du rapport de démarrage	104
Schéma 8 :	Les quatre éléments d'une évaluation de l'action humanitaire réussie	107
Encadré 5 :	Prendre en compte l'âge et le genre dans la consultation des populations affectées	111
Schéma 9 :	Fiche-conseil sur les droits des personnes interrogées	114
Encadré 6 :	Le paiement des personnes interrogées	115

4.1 La phase de démarrage

La phase de démarrage a lieu avant le travail de terrain : son produit est le rapport de démarrage. Ce dernier peut être relativement court, comme dans le cas de l'évaluation du financement de la réponse à l'urgence en Haïti (Moriniere, 2011a), si on le compare au rapport final (Moriniere, 2011b).

La plupart des Termes de référence (TdR) des évaluations humanitaires sont fixés plutôt que négociés entre le responsable de l'évaluation et l'équipe d'évaluation. Le rapport de démarrage permet à l'équipe d'évaluation de définir un plan spécifique et de trouver un accord sur lui avec le responsable de l'évaluation ainsi que de soulever des points problématiques et de répondre à toute éventuelle ambiguïté des TdR. Le travail sous-entendu par les TdR prend souvent beaucoup plus de temps que ce qui est budgété. Le rapport de démarrage constitue donc une opportunité pour que l'équipe d'évaluation clarifie ce qu'elle peut ou non couvrir.



Pour l'équipe d'évaluation, il est presque toujours intéressant de produire un rapport de démarrage et ce, même s'il n'est pas demandé, car cela aide l'équipe à planifier son travail et à éliminer de possibles malentendus entre l'équipe et l'organisation commanditaire.

4.1.1 Préparer le rapport de démarrage

Le rapport de démarrage offre les opportunités suivantes :

- *Pour le responsable de l'évaluation : évaluer comment l'équipe comprend l'évaluation et se prépare à l'approcher.*
- *Pour l'équipe d'évaluation : convertir les TdR en un plan réalisable validé par le responsable de l'évaluation.*
- *Pour l'équipe d'évaluation : clarifier les TdR et souligner les tensions qui doivent être résolues par le commanditaire (par exemple, les attentes contradictoires au sein de l'organisation vis-à-vis de l'évaluation).*
- *Pour l'équipe d'évaluation : planifier son travail de façon cohérente.*
- *Pour les autres parties prenantes : recevoir une déclaration d'intention claire de la part de l'équipe d'évaluation de manière à ce qu'elles puissent rapidement repérer de potentielles*

difficultés liées à la méthodologie proposée.

Le rapport de démarrage doit présenter un plan de travail et un calendrier clairs et réalistes qui prennent en compte les contraintes existantes. Le plan de travail doit inclure l'allocation des rôles et des responsabilités au sein de l'équipe, ainsi que les dates butoir pour la remise des rapports internes et les plans de voyage détaillés. Le contenu du rapport de démarrage varie en fonction du contexte (**Tableau 10**). Si l'opportunité/rapidité (« timeliness ») est un élément à évaluer, une ébauche de chronologie devra faire partie du rapport de démarrage. Sur le terrain, l'équipe peut ajouter d'autres données à la chronologie.

Le rapport de démarrage doit présenter les méthodologies proposées, notamment un plan initial des entretiens prioritaires pour les entretiens à venir. Il doit remercier tout groupe consultatif et, là où cela est possible, spécifier leurs rôles. En annexe, il doit inclure un guide d'entretien et une liste des sujets à aborder en focus groups si ces activités sont prévues.

4.1.2 La matrice d'évaluation

La matrice d'évaluation est l'une des parties clés de la plupart des rapports de démarrage : l'équipe y définit un plan pour répondre à chacune des questions d'évaluation (un extrait de matrice pour un seul élément d'évaluation est donné dans le **Tableau 11**). La matrice d'évaluation fournit trois éléments d'informations :

Tableau 10 : Eléments du rapport de démarrage

Elément	Commentaires
Contexte	<p>Ce chapitre doit résumer le contexte dans lequel l'évaluation prend place.</p> <p><i>Astuce : une chronologie peut être assez utile pour fixer le contexte et peut être poursuivie dans le rapport principal avec les données rassemblées durant le travail de terrain (voir Chapitre 5.3.3).</i></p>
Action à évaluer	<p>Ce chapitre doit montrer que les membres de l'équipe comprennent en quoi consiste l'action à évaluer. Il peut comprendre une description de l'intervention avec les données de base collectées par l'équipe durant le travail de bureau et inclure un modèle logique (voir Chapitre 4.1.3). Les tableaux de données peuvent être mis en annexe.</p>
Objectif de l'évaluation	<p>Ce chapitre doit résumer la compréhension que les membres de l'équipe se font de l'objectif général et des objectifs spécifiques de l'évaluation ainsi que de l'utilisation qui en sera faite. Il doit aussi expliquer comment cela a influencé leur choix méthodologique.</p>
Méthodologie	<p>Ce chapitre doit fixer les méthodes que l'équipe propose d'utiliser pour rassembler et analyser les données permettant de répondre aux questions d'évaluation. Il doit également indiquer (potentiellement en annexe) tous les guides d'entretien, les formulaires d'enquête ou autre instrument de collecte de données qui seront utilisés. Il doit clairement définir les limites des méthodes de collecte de données proposées, dont la stratégie d'échantillonnage, et toute autre limite due aux ressources disponibles.</p>

Élément	Commentaires
Questions d'évaluation	Ce chapitre est seulement nécessaire si les questions d'évaluation sont retravaillées par l'équipe d'évaluation : par exemple, pour synthétiser certaines questions de façon à réduire leur total jusqu'à un nombre gérable.
Matrice d'évaluation	Cet élément montre que les évaluateurs prévoient de répondre à chacune des questions d'évaluation (voir Chapitre 4.1.2). La matrice doit refléter les méthodes définies dans le chapitre Méthodologie. Elle peut être présentée en annexe.
Plan de travail détaillé	Cet élément précise où les membres de l'équipe prévoient d'aller et quand, ainsi que les jours proposés pour les visites des sièges. Il indique également quel membre de l'équipe sera responsable de quelle tâche. Astuce : dans des environnements plus dangereux, il peut être déconseillé d'énoncer en détail le plan de visite dans le rapport de démarrage car une flexibilité de dernière minute peut s'avérer nécessaire.
Plan général du rapport	Il prend en général la forme d'une table des matières.
Personnes à cibler dans les entretiens	Cet élément fournit une liste préliminaire des personnes que l'équipe souhaite interroger ou, au moins, les types de personnes à interroger.
Questions et sujets en suspens	Opportunité que l'équipe d'évaluation peut saisir pour souligner des ambiguïtés, des sujets de préoccupations ou des contradictions auxquels elle aimerait que le commanditaire réponde en les clarifiant avant la prochaine étape de l'évaluation.

1. *Questions à laquelle il faut répondre : celles-ci sont reprises des TdR (de façon explicite ou implicite).*
2. *Comment le jugement est étayé : les critères ou indicateurs sur lesquels les réponses se baseront doivent être précisés.*
3. *Les sources et méthodes d'information attendues : vous devez normalement avoir plusieurs sources et méthodes pour chacune des questions d'évaluation (Chapitre 5.2.5).*

Une matrice d'évaluation montre comment l'équipe prévoit de répondre à chaque question. En la parcourant, le responsable de l'évaluation pourra voir si l'équipe a négligé certaines sources majeures.

Tableau 11: Extrait d'une matrice d'évaluation

Élément	Question d'évaluation	Comment le jugement est étayé	Sources et méthodes probables
Caractère inclusif du processus de priorisation au niveau pays	Dans quelle mesure la préparation des documents de projet est-elle inclusive et transparente sur le terrain ?	Preuve de processus inclusifs et transparents sur le terrain.	Entretiens avec des informateurs clés auprès des Coordinateurs humanitaires, des ONG et des organisations aux sièges et sur le terrain.
	Dans quelle mesure les gouvernements hôtes sont-ils impliqués dans la préparation des documents de projet ? Ce niveau d'implication est-il adapté et suffisant ? Quel est le point de vue des gouvernements sur ce processus ?	Preuve de l'implication Points de vue exprimés par les représentants des gouvernements et d'autres acteurs humanitaires.	Analyse des mots-clés des documents de projet. Entretiens avec des informateurs clés auprès des gouvernements, des Coordinateurs humanitaires et les représentants d'organisations. Enquête sur la manière dont est perçu le processus.

Élément	Question d'évaluation	Comment le jugement est étayé	Sources et méthodes probables
Caractère inclusif du processus de priorisation au niveau pays	Dans quelle mesure les principaux besoins sont-ils priorisés et sur quelles données les décisions de priorisation sont-elles basées ?	Preuve d'une priorisation au niveau pays Proportion dans laquelle les projets priorisés au niveau terrain sont financés	Entretiens avec des informateurs clés auprès des Coordinateurs humanitaires, d'OCHA, des équipes pays et des équipes de coordination humanitaire Comparaison des demandes financées avec les demandes rejetées par les bailleurs ou retirées par les soumissionnaires.

Source: adapté d'une évaluation de 5 ans du CERF.



Les responsables d'évaluation peuvent trouver utile de préparer leur propre matrice d'évaluation pour vérifier l'exhaustivité de chaque méthodologie proposée par les équipes d'évaluation. Cependant, cette matrice ne doit pas être partagée avec l'équipe qui remporte l'appel d'offre tant que celle-ci n'a pas préparé sa propre version. Le processus de planification est toujours plus important que prévu et l'équipe doit avoir l'opportunité de se confronter en détail aux questions d'évaluation selon sa propre manière de faire.

4.1.3 Utiliser un modèle logique



Modèle logique

Diagramme présentant la théorie du programme (modèle des résultats intermédiaires et finaux attendus) en vue d'une intervention.

Les TdR comprennent souvent un modèle logique de l'opération (par exemple, l'évaluation de 5 ans du CERF) ou demandent aux évaluateurs d'en développer un. La popularité croissante de l'évaluation à base théorique et l'attention croissante aux théories du changement comme cadres d'intervention sont susceptibles d'augmenter la demande de modèles logiques. La théorie du programme et l'évaluation à base théorique sont analysés plus en détail dans le Chapitre 3.2.5.

Les cadres logiques sont l'une des formes les plus simples de modèle logique avec ses activités en lien avec les résultats. Le Tableau 12 présente un extrait de cadre logique adapté d'une approche originellement défendue par NORAD (1999).

Tableau 12: Extrait de cadre logique

Hiérarchie des résultats	Indicateurs	Hypothèses
Objectif général – l'objectif de niveau supérieur auquel le projet est censé contribuer	Indicateurs mesurant l'atteinte de l'objectif général	Hypothèses concernant la durabilité de l'objectif général
Objectif spécifique – l'effet qu'il est prévu d'atteindre comme résultat du projet	Indicateurs mesurant l'atteinte de l'objectif spécifique	Hypothèses concernant l'atteinte de l'objectif général une fois l'objectif spécifique atteint
Résultats – les résultats que les responsables du projet devraient être capables de garantir	Indicateurs mesurant la proportion dans laquelle les résultats sont produits	Hypothèses concernant l'atteinte de l'objectif spécifique une fois les résultats en place
Activités – les actions entreprises pour produire les résultats	Moyens : les biens et services nécessaires pour entreprendre les activités	Hypothèses en lien avec la production des résultats

Source: adapté de [NORAD \(1999\)](#).

L'approche cadre logique a été critiquée comme excessivement rigide, en particulier dans les contextes humanitaires qui changent rapidement. Un autre problème réside dans le fait que la hiérarchie des résultats peut être source de confusion dans la mesure où l'objectif général d'un projet peut être le même que l'objectif spécifique d'un programme contenant ce projet.

NORAD s'est finalement tourné vers une approche plus flexible en utilisant une chaîne des résultats comme indiqué dans le **Schéma 6** (les noms et les définitions précises des éléments d'une chaîne de résultats peuvent varier selon les organisations).

Schéma 6 : Exemple de chaîne des résultats



On demande parfois aux évaluateurs de développer un cadre logique pour un projet qui n'en avait pas réalisé durant la phase de conception de projet. Dans de telles circonstances, il vaut mieux encourager le client à utiliser un modèle logique moins rigide comme la chaîne de résultats. L'approche de la théorie du changement semble en train de remplacer le cadre logique pour certains bailleurs.

Parfois, le changement souhaité peut être exprimé comme le fait d'éviter des morts et des souffrances ou encore la prévention ou la réduction de facteurs qui y contribuent comme la malnutrition aiguë.



Théorie du changement

Description du mécanisme central par lequel le changement se produit pour les individus, les groupes et les communautés.

Les théories du changement peuvent être soutenues par des théories de l'action définissant quelles actions sont menées pour atteindre le changement. Par exemple, si la théorie du changement consiste dans le fait que l'amélioration de la sécurité alimentaire des ménages réduira la malnutrition au sein d'une population affectée par la crise, la théorie de l'action définira comment cela se réalisera (par exemple, au moyen de transferts monétaires, de distribution alimentaire et de mécanismes de marché). Le récent guide OCDE/CAD sur l'évaluation des interventions de construction de la paix comprend une annexe sur la compréhension et l'évaluation des théories du changement ([OECD/DAC, 2012, p. 80](#)).



Théorie de l'action

Description de la manière dont les interventions sont construites de façon à activer le changement décrit dans la théorie du changement.

La théorie du programme peut être explicite (là où la théorie est présentée comme un modèle logique), semi-explicite (là où elle peut être construite à partir du cadre logique) ou implicite (là où il n'existe pas

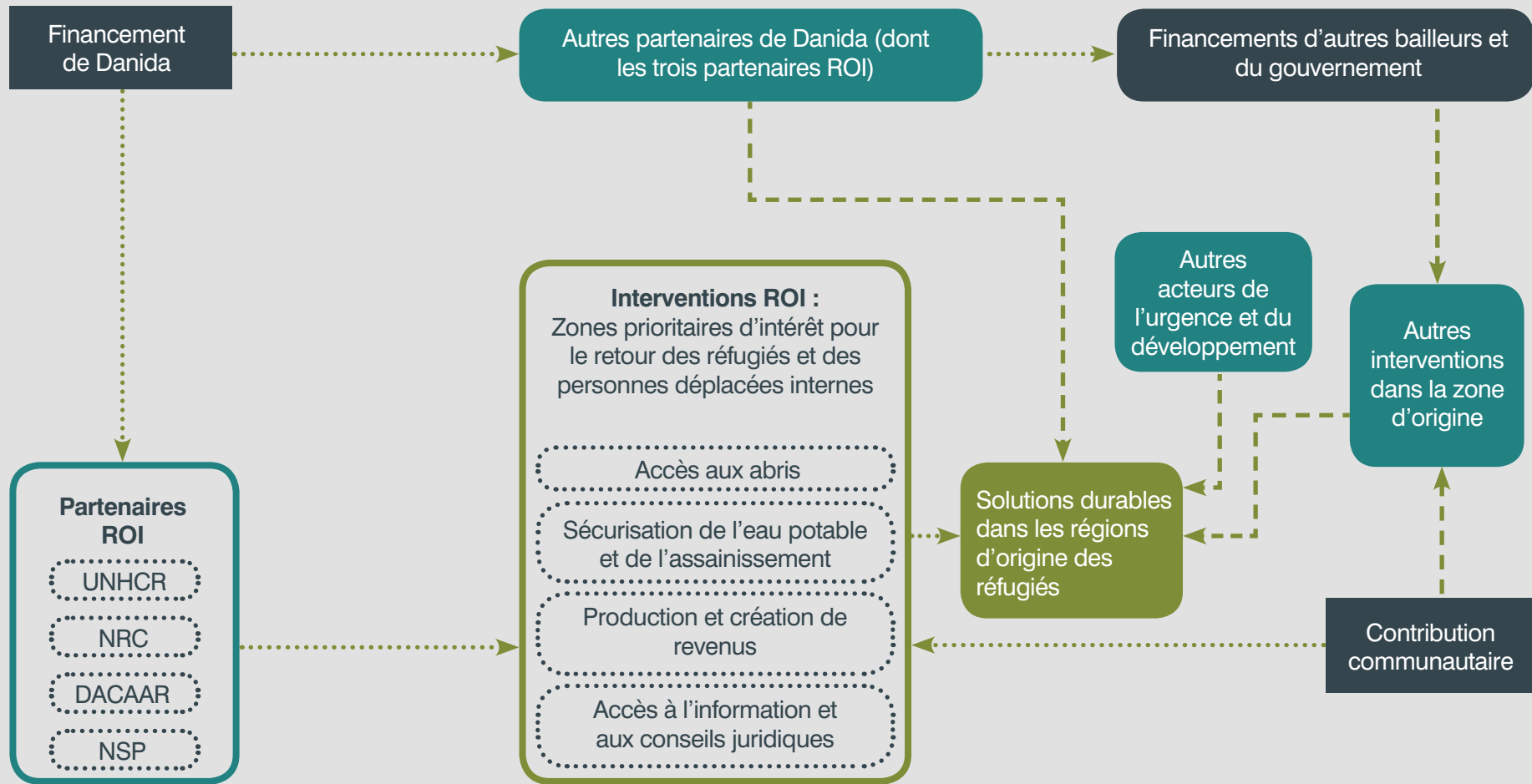
de cadre logique mais où un modèle logique peut être construit à partir de la manière dont le projet a opéré et à partir d'entretiens avec du personnel clé du projet).

Des changements dans l'efficacité de l'intervention parmi les éléments d'un modèle logique peuvent souligner l'exactitude des hypothèses faites à propos des liens entre les différents éléments.

S'il existe un cadre conceptuel plus large pour le secteur, il peut être utile de le comparer avec le modèle logique afin de voir si ce dernier comporte des manques importants.

Les modèles logiques peuvent aussi être développés à partir de liens implicites entre des documents du programme. Le modèle logique présenté dans le **Schéma 7** a été développé pour l'évaluation du ROI (Regions of Origin) du Danemark en Afghanistan. Il a été conçu à partir de la logique de projet qui était implicitement suggérée dans divers documents de projet.

Schéma 7: Modèle logique pour une évaluation humanitaire



La Fondation Kellogg a publié un guide sur le développement des modèles logiques (**WK Kellogg Foundation, 2004**) qui est surtout axé sur des chaînes de résultats assez simples.



Les modèles sont des simplifications commodes d'une réalité complexe. Cette simplification est leur plus grande force car elle permet la manipulation de concepts et de situations complexes mais c'est aussi leur plus grande faiblesse car cela peut vous amener à survoler un facteur clé. Si vous utilisez un modèle lors d'une validation, examinez dans quelle mesure le modèle correspond à la situation réelle que vous examinez et modifiez-le en fonction.

Il existe un certain nombre de logiciels comme DoView qui peuvent faciliter le dessin d'un simple pipeline ou de modèles logiques de chaînes de résultats.

4.1.4 L'étude documentaire

Les rapports de démarrage comprennent souvent une étude documentaire préliminaire qui peut notamment permettre de réaliser les points suivants :

- *Vous aider à choisir un cadre conceptuel approprié si celui-ci n'a pas déjà été mis en place.*

- *Vous aider dans le développement d'une théorie du changement si elle n'a pas déjà été identifiée.*
- *Souligner les sujets clés à approfondir durant le travail de terrain.*
- *Vous permettre de cartographier l'intervention par secteur, pays/région, période ou toute autre dimension.*
- *Identifier les personnes clés à interroger.*
- *Identifier les cibles potentielles pour un échantillonnage raisonné.*

Les ressources consultées doivent idéalement être présentées dans une bibliographie annotée qui démontre leur pertinence dans le cadre de cette évaluation. Le Chapitre 5.3 analyse les méthodes de revue documentaire plus en détail.

4.1.5 Evaluer le rapport de démarrage

Le rapport de démarrage permet au responsable de l'évaluation de voir comment l'équipe comprend le travail de l'évaluation et prévoit de l'aborder. Il permet également aux parties prenantes sur le terrain de voir comment l'équipe d'évaluation prévoit de travailler de manière à ce qu'elles puissent identifier tous les sujets ou les défis qu'elles doivent avoir à l'esprit.

Un rapport de démarrage donne au responsable de l'évaluation l'occasion de répondre aux problèmes grâce à la compréhension et à l'approche qu'en a l'équipe et ce, avant qu'ils ne deviennent des sujets majeurs. Il doit démontrer que l'équipe

comprend le contexte de la crise humanitaire, le contexte de la réponse et des actions à évaluer, l'objectif et l'intention de l'évaluation ainsi que les préoccupations des parties prenantes (**Tableau 13**).

Tableau 13: Liste de vérification du rapport de démarrage

<input type="checkbox"/>	Le rapport démontre-t-il une compréhension précise du contexte ?
<input type="checkbox"/>	Le rapport démontre-t-il une compréhension précise de l'intervention ?
<input type="checkbox"/>	Le rapport démontre-t-il une compréhension précise des objectifs et de l'intention de l'évaluation ?
<input type="checkbox"/>	Existe-t-il un plan de travail précis ?
<input type="checkbox"/>	Est-il réaliste ?
<input type="checkbox"/>	Le plan prévu évite-t-il de rentrer en conflit avec toutes les activités planifiées du programme ?

<input type="checkbox"/>	Le plan identifie-t-il les dates clés pour toutes les activités et tous les produits ?
<input type="checkbox"/>	Les rôles au sein de l'équipe sont-ils clairement décrits et les responsabilités clairement assignées ?
<input type="checkbox"/>	Le plan démontre-t-il une progression logique claire ?
<input type="checkbox"/>	Le plan identifie-t-il clairement les probables contraintes et y répond-il ?
<input type="checkbox"/>	Tous les itinéraires de voyages et les dates prévus sont-ils indiqués ?
<input type="checkbox"/>	La méthodologie proposée est-elle claire ?
<input type="checkbox"/>	Le rapport montre-t-il une compréhension précise des risques et des limites inhérents aux méthodes proposées ?
<input type="checkbox"/>	La matrice d'évaluation propose-t-elle des méthodes et des critères de jugement (ou des indicateurs) appropriés pour répondre à chaque question d'évaluation ?
<input type="checkbox"/>	La méthodologie ainsi que la quantité et le type de travail de terrain proposés permettront-ils à l'équipe de répondre aux questions d'évaluation telles que définies dans la matrice d'évaluation ?



L'équipe a-t-elle accordé une attention suffisante au rôle de chaque groupe de pairs, de référence ou consultatif ?



Les groupes consultatifs pertinents ont-ils accepté le rapport ?



Le rapport démontre-t-il que l'équipe a conscience des préoccupations potentielles des parties prenantes ?

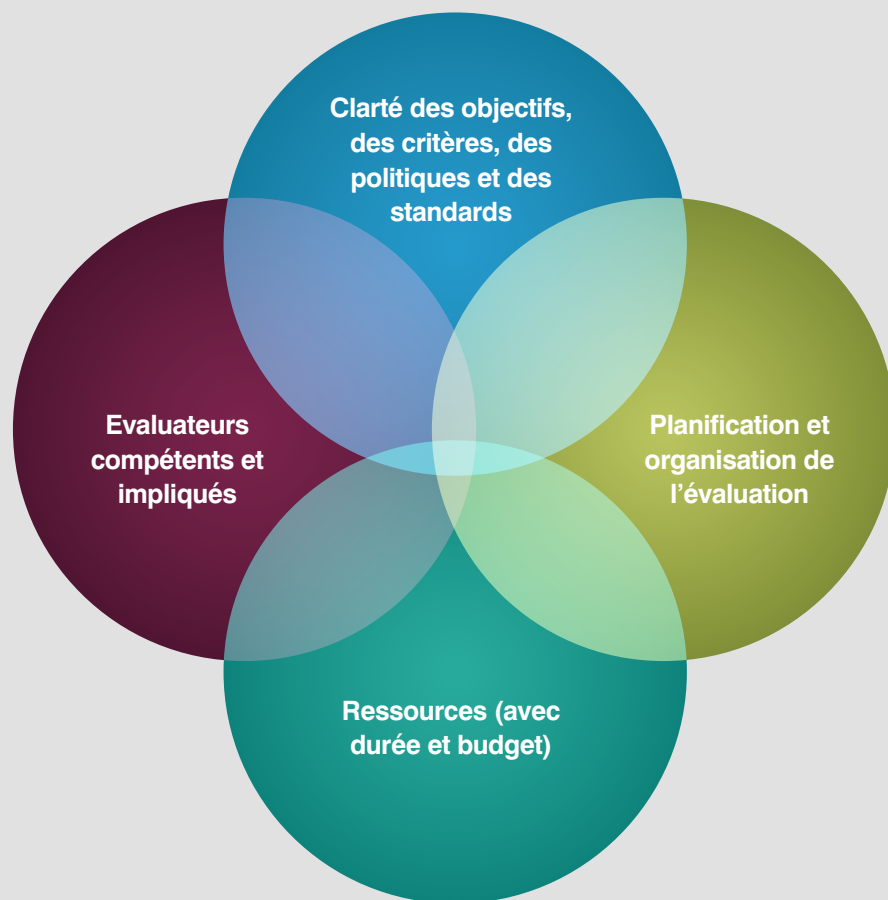


Le guide d'entretien et les autres outils de collecte des données sont-ils présentés en annexe ?



Le rapport indique-t-il toutes les zones d'ambiguïté ou les contradictions à propos desquelles l'équipe a besoin de clarification avant de passer à la phase suivante ?

Schéma 8 : Les quatre éléments d'une évaluation d'action humanitaire réussie



4.2 Les défis liés au *leadership* et au travail en équipe

Une EAH est comme tout autre projet : son succès dépend de sa bonne gestion. Il existe potentiellement trois niveaux de gestion : la gestion par le département de l'évaluation du commanditaire (le responsable de l'évaluation), la gestion par le responsable du contrat lorsqu'une société de conseil est utilisée et la gestion du processus et de l'équipe par le chef d'équipe. Chacun de ces trois niveaux a des préoccupations différentes. Ce chapitre se concentre sur la gestion du processus de l'EAH et de l'équipe d'évaluation par le chef d'équipe. Le **Schéma 8** résume les quatre éléments clés pour gérer une évaluation réussie.

Même lorsque les évaluateurs sont compétents, que les ressources sont adaptées et que les objectifs sont clairs, une bonne planification est néanmoins essentielle pour assurer le succès d'une évaluation. Presque toutes les EAH sont réalisées en devant faire face à des contraintes de temps. Se mettre d'accord sur les TdR, recruter une équipe, écrire le rapport provisoire et intégrer les commentaires des relecteurs prennent presque toujours plus longtemps que prévu. Le travail de terrain est

celui qui souffre le plus des retards et autres dépassements des coûts. Une planification très soignée peut aider à éviter cela.



Clarifier depuis le commencement quel appui fournira le client à l'équipe d'évaluation. L'équipe aura-t-elle accès aux dossiers, droit à un espace de travail au siège et sur le terrain, bénéficiera-t-elle d'un soutien pour les rendez-vous au siège et sur le terrain, de services de sécurité, de transports sur le terrain et d'assistance dans la réservation des logements sur le terrain ?

L'ampleur du travail d'évaluation détermine la taille de l'équipe. Cependant, comme indiqué plus haut, les grandes équipes s'accompagnent de problèmes et de risques qui peuvent dépasser les bénéfices liés à leur plus large palette d'expertise. Les grandes équipes signifient plus de travail pour le chef d'équipe, par exemple dans la gestion des activités et dans la centralisation des données. Elles posent également un problème dans les environnements dangereux. Il peut être difficile de trouver un hébergement sûr dans certains contextes d'EAH et il est possible que certaines grandes équipes doivent se scinder, ce qui peut rendre difficile la tenue de discussions le soir.



Distribuez des modèles de documents pour les listes de personnes rencontrées, les bibliographies et les itinéraires de manière à obtenir l'information telle que vous la souhaitez. Obtenir les données sous un format standard réduit en effet le travail de conversion et de collecte.



Si vous travaillez avec certains évaluateurs pour la première fois, et que vous ne les connaissez pas de réputation, organisez le travail de terrain de manière à passer une journée ensemble au début de l'évaluation et à pouvoir vous faire une idée de leurs forces et faiblesses, dont les biais potentiels.

Les équipes d'EAH sont en général constituées en contractant une société de conseil ou par une embauche directe. Dans les deux cas, le chef d'équipe peut ne pas avoir travaillé par le passé avec l'ensemble des membres de l'équipe. Il se peut ou non que le chef d'équipe connaisse l'une des personnes de réputation.

Même les meilleurs évaluateurs ont des jours « sans » et une bonne performance précédente n'est pas une garantie de bonne performance pour l'évaluation en cours. Les membres de l'équipe

peuvent être affectés par des préoccupations personnelles, la maladie d'un proche ou d'autres problèmes. Le chef d'équipe doit définir des dates butoir pour les différentes tâches et s'assurer immédiatement si ces dates sont tenues ou non. Cela peut aider à identifier rapidement tous les problèmes de performance et permettre des actions correctives.

Parfois, des problèmes se produisent au sein de l'équipe en raison de conflits de personnalités, de problèmes de performance ou encore de valeurs divergentes. Certains environnements d'EAH peuvent être stressants. Parmi ces problèmes, les questions de performance sont les plus faciles à traiter. Toutefois, le temps de s'assurer qu'il existe bien un problème de performance, il peut être trop tard pour écarter la personne sans mettre en péril l'évaluation.



Si un expert insatisfaisant ne peut être abandonné par l'équipe, mettez-le en binôme avec un membre de l'équipe plus compétent pour aider à minimiser les problèmes de qualité.

Les grandes équipes impliquent des risques de conflit accrus par rapport aux résultats, cela en

raison de différences de points de vue ou de différences entre les zones que les différents membres de l'équipe ont observées. Voici quelques conseils testés pour gérer les grandes équipes d'évaluation :

- *Utilisez un tableau de synthèse des observations (voir Chapitre 5.2.4) pour garder une trace des données probantes et des premiers résultats. Le fait de le partager peut permettre de maintenir l'équipe informée et d'aider à construire un point de vue commun.*
- *Si vous avez une grande équipe pour un seul pays, consacrez un peu de temps à une réflexion sur le programme de travail de manière à ce que l'équipe puisse discuter des premiers résultats.*
- *Planifiez des réunions d'équipe après le travail de terrain et avant la rédaction de la première version du rapport dans le but d'échanger sur les données probantes, les résultats, les conclusions et les recommandations.*

Parfois, un membre de l'équipe peut être très attentif à un sujet en particulier qui, bien que pertinent, soit d'une certaine manière périphérique par rapport à l'objet de l'évaluation, ou trop limité, ou encore trop complexe pour être inclus dans le rapport global.

Dans ce cas, vous pouvez demander à ce membre de l'équipe d'écrire une annexe sur le sujet.

4.3 Impliquer les populations affectées

La consultation des populations affectées a été pendant de longues années le maillon faible de l'EAH malgré de nombreux efforts pour augmenter la redevabilité envers ce groupe critique de parties prenantes puisqu'il s'agit des personnes au nom de qui l'action humanitaire est entreprise⁷. Si nous voulons véritablement écouter les populations affectées, il est essentiel de les consulter durant une évaluation.

Pour que cela réussisse, les populations affectées doivent être impliquées depuis le début dans la définition du besoin et de l'objectif de l'évaluation ainsi que dans la contribution à la formulation des questions d'évaluation. Cela est rarement le cas et il peut exister d'autres manières plus appropriées

⁷ Près de trois quarts des EAH évaluées par ALNAP entre 2001 et 2004 ont été jugées insatisfaisantes ou faibles en termes de consultation et d'encouragement de la participation par les principales parties prenantes, en particulier les membres des populations affectées ([Beck and Buchanan-Smith, 2008](#)).

de s'assurer que l'action humanitaire est redevable envers les populations affectées tout au long d'un programme ou d'un projet (voir par exemple [HAP, 2010](#)). Quoi qu'il en soit, les populations affectées doivent toujours faire entendre leur voix durant l'évaluation de manière à ce qu'elles puissent donner leur sentiment sur la qualité et l'efficacité de l'action humanitaire à moins qu'il n'y ait une bonne raison de ne pas le faire (par exemple, dans le cadre d'une évaluation entièrement centrée sur des processus institutionnels plutôt que sur l'efficacité d'un programme en matière de réduction des souffrances).

Etre attentif aux questions de genre et d'âge est crucial en matière de consultation des populations affectées car ce sont des déterminants clés de la vulnérabilité, des différences d'impact de la crise et des différences de besoins (voir [Encadré 5](#)). Alors que le genre et l'âge sont toujours importants, d'autres facteurs comme le handicap, la caste ou l'ethnicité peuvent également l'être dans d'autres contextes. Le Groupe de travail des Nations unies en matière d'évaluation recommande actuellement dans ses orientations de prendre en compte les aspects de droits de l'Homme et d'égalité de genre dans toutes les évaluations ([UNEG, 2011](#)).

L'UNICEF a également publié des conseils en matière de conception et de conduite d'évaluations focalisées sur l'équité (**Bamberger et Segone, 2011**). Voir Chapitre 5.4.2 pour des méthodes de consultation des populations affectées.

Encadré 5 : Prendre en compte l'âge et le genre dans la consultation des populations affectées

Les catastrophes naturelles et les conflits armés n'affectent pas chaque individu de la même manière. Le genre et l'âge influent fortement sur la manière dont les personnes sont affectées durant une crise et après. Lors du tsunami de 2006 en Asie du Sud-est, les taux de mortalité ont été par exemple bien supérieurs pour les femmes, les enfants et les personnes âgées que pour les hommes. En règle générale, les urgences liées à un conflit tuent plus d'hommes et les catastrophes naturelles plus de femmes. Pour les catastrophes naturelles, un fossé plus important entre le statut socio-économique des hommes et des femmes est associé à une plus grande différence entre les taux de mortalité des hommes et des femmes (Neumayer et Plümper (2007).

Le genre et l'âge peuvent être des déterminants clés

des dynamiques de pouvoir et de la vulnérabilité en matière de violence liée au genre. Les opinions des femmes et des jeunes peuvent être passées sous silence par les hommes au sein d'une communauté avant, durant et après une crise. Les projets d'urgence humanitaire doivent prendre en compte les manières dont le genre et l'âge affectent les besoins et les vulnérabilités. Par exemple, dans les camps de réfugiés et de personnes déplacées, les latrines doivent être placées dans des lieux où les femmes peuvent les utiliser de façon sûre sans risque d'attaques ; les hommes qui ont perdu leur femme peuvent avoir des besoins particuliers en termes de soins aux enfants.

Les droits légaux des hommes et des femmes, par exemple à détenir des terres, peuvent être différents

dans certaines sociétés. Cela a été parfois une question centrale pour les foyers dont le chef était désormais une femme à cause de la crise, comme par exemple après le génocide du Rwanda.

Pour toutes ces raisons, le sexe et l'âge doivent être considérés avec attention au moment de collecter

de l'information auprès des populations affectées, et sur elles, durant une évaluation. La définition des personnes âgées et des enfants doit être adaptée au contexte lors de la consultation avec la population locale.

Source: [Mazurana et al. \(2011\)](#).

4.4 Les questions d'éthique

Chacun de nous occupe des rôles différents avec des responsabilités différentes et bien faire dans un rôle peut entrer en conflit avec le fait de bien faire dans un autre. Cela s'appelle un dilemme éthique. Par exemple, en tant qu'employé, il est de notre responsabilité de fournir des informations à ceux qui nous encadrent lorsque cela est nécessaire mais cela peut aussi entrer en conflit avec notre obligation de confidentialité envers les personnes interrogées.

Diverses structures d'évaluation ont produit des guides d'éthique. Les principes directeurs à destination des évaluateurs développés par l'Association américaine de l'évaluation (2004)

insistent sur le questionnement permanent, la compétence, l'intégrité et l'honnêteté, le respect des personnes et la responsabilité vis-à-vis du bien-être public et général. Les standards d'évaluation des programmes développés par le Joint Committee (1994) se concentrent sur l'utilité, la faisabilité, la propriété et la précision. Les Lignes directrices éthiques du Groupe d'évaluation des Nations unies (2008) se focalisent quant à elles sur l'intentionnalité de l'évaluation, l'obligation des évaluateurs, les obligations envers les participants, le processus et le produit de l'évaluation.



Familiarisez-vous avec toute politique éthique développée par votre organisation et les organisations professionnelles dont vous êtes membre ou que vous vous êtes engagé à respecter.

La principale question éthique en matière d'EAH réside dans la capacité des évaluations à avoir un impact négatif sur les informateurs et autres parties prenantes. Mettre en danger des informateurs est un sujet critique, en particulier lors d'urgences complexes où le simple fait de parler à l'équipe d'évaluation peut placer une personne interrogée en situation de risque de détention, voire plus grave encore. Là où l'identification des personnes interrogées peut les mettre en danger, des termes comme « réfugié 1, réfugié 2 » peuvent être utilisés. Voir Chapitre 5.4.2 pour une analyse plus approfondie des questions éthiques lors d'entretiens auprès des populations affectées. Une autre manière dont les évaluations peuvent avoir un impact négatif sur les parties prenantes est leur emploi inapproprié pour justifier de façon impropre des décisions prises par avance mais qui ne sont pas soutenues par des données probantes.

Les enfants sont souvent oubliés lors des consultations. Ils ont pourtant un point de vue qui leur est propre et il est important de les prendre en compte quand on évalue des programmes qui essaient de satisfaire leurs besoins. Cependant, réaliser des travaux de recherche avec des enfants soulèvent des préoccupations éthiques bien particulières. L'UNICEF a publié une note technique sur la manière de répondre à ces préoccupations (**Bureau de l'évaluation de l'UNICEF, 2002**). Les évaluations sont une forme de recherche sociale qui est habituellement conduite sur la base d'un consentement éclairé. Les personnes interrogées doivent être informées des règles de base de l'entretien, ce qui est important étant donné les très grandes différences de pouvoir qui peuvent exister entre celui qui interroge et celui qui est interrogé. Cela peut être réalisé au moyen d'une simple fiche (**Schéma 9**). La plupart des évaluations humanitaires sont conduites sur la base d'une variante de la règle de Chatham House (**Chatham House, 2007**) selon laquelle les commentaires des personnes interrogées ne leur sont attribués d'aucune façon, ni directe ni indirecte.

Schéma 9 : Fiche-conseil sur les droits des personnes interrogées

Vos droits en tant que personne interrogée : Vous avez le droit de clôturer l'entretien à tout moment. Vous avez le droit de ne pas répondre aux questions de votre choix.

Rien de ce que vous dites ne vous sera attribué de façon directe ou indirecte sans votre permission expresse. Les notes tirées de cet entretien ne seront pas partagées en dehors de l'équipe d'évaluation. Si le client est d'accord et que vous possédez une adresse de courrier électronique, **nous nous ferons parvenir le rapport pour que vous le commandiez.**

De même, si les résultats d'une évaluation indiquent que des interventions en particulier sont inefficaces ou inadaptées, cela peut causer des pertes d'emploi ou des pertes de bénéfiques pour ceux qui les réalisent ou reçoivent l'aide.

Parfois pourtant, de tels effets négatifs sur des individus peuvent être tout à fait appropriés. Même si les observations de l'évaluation envers des actions humanitaires de faible qualité peuvent être désagréables pour les personnes visées, il ne faut pas oublier que ces actions consomment des ressources qui pourraient autrement mieux prévenir des décès et des souffrances.

Encadré 6 : Le paiement des personnes interrogées

Les personnes interrogées qui participent à des évaluations humanitaires ne sont pas payées en général. Ce sujet pose la question de savoir si la relation relève de l'exploitation car l'évaluateur reçoit un bénéfice direct alors que le bénéfice immédiat de la personne interrogée n'est pas clair. Cette question apparaît le plus souvent dans le cadre des discussions de type *focus groups* qui requièrent deux heures ou plus de présence de la part des participants. Dans les pays industrialisés, les participants de *focus groups* sont habituellement rémunérés.

La plupart des débats sur l'éthique de recherche dans les pays en voie de développement se concentrent sur la recherche médicale (Molyneux et al., 2012), en particulier la recherche clinique (Benatar, 2002). Pour l'EAH, le risque consiste dans le fait qu'une évaluation payée limite la participation à d'autres évaluations. Nous vous proposons donc les conseils généraux suivants :

- Les évaluateurs ne doivent pas payer les personnes interrogées en argent ou en nature. Cependant, il est le bienvenu de proposer des boissons fraîches et des collations légères. On peut également donner aux participants des sessions de focus groups un T-shirt de l'organisation ou tout autre cadeau promotionnel en guise de remerciement pour le temps qu'ils vous ont consacré. Toutefois, aucun cadeau ne doit être fait dans des situations dangereuses où ils pourraient encourager la participation mais aussi le ciblage des bénéficiaires.
- Les évaluateurs doivent traiter les personnes interrogées avec respect, arriver à l'heure aux entretiens planifiés, retranscrire aussi précisément que possible les opinions des personnes interrogées dans leurs rapports et remercier l'assistance.

Les évaluateurs ne doivent pas critiquer des individus ou des organisations de façon nominative pour des exemples précis de mauvaises pratiques à moins que l'évaluation n'ait été capable d'examiner l'ensemble du travail de la personne ou de l'organisation et puisse replacer cet exemple en contexte.

Les individus et les organisations ne devraient pas être critiqués à moins qu'ils/elles n'aient eu l'occasion de se défendre de leurs actions. Si les évaluateurs sont confrontés à des actions clairement ou potentiellement criminelles, il est alors plus approprié de les notifier directement au responsable de l'évaluation et à votre supérieur hiérarchique plutôt que de les inclure dans le rapport d'évaluation. Les évaluateurs ne sont pas des détectives et doivent laisser l'enquête approfondie relative aux actes potentiellement criminels aux personnes les mieux équipées pour les réaliser.



En général, le devoir d'un évaluateur est de signaler une irrégularité au responsable de l'évaluation mais il peut exister des situations où cela n'est pas suffisant et où l'évaluateur se trouve juridiquement et éthiquement obligé d'en informer les autorités locales. Dans de telles circonstances, l'évaluateur doit rechercher des conseils juridiques avant de signaler l'irrégularité à quiconque d'extérieur à l'organisation.

NOTES

REALISER UNE EVALUATION D'ACTION HUMANITAIRE

Partagez vos commentaires:

Aidez-nous à rendre le Guide plus pratique et focalisé sur l'utilisateur en partageant vos commentaires sur son contenu et la navigation. Les commentaires peuvent inclure des questions, des suggestions, des ressources utiles ainsi que des conseils pratiques. Vous pouvez partager vos idées en cliquant sur le bouton de commentaires en bas de chaque page. Ceci générera un courriel confidentiel à l'équipe EAH. Pour plus d'informations sur la façon dont vous, votre équipe ou votre organisation peut s'engager plus activement dans le pilotage du Guide contactez-nous directement à eha@alnap.org

CONCEPTION ET METHODES D'ÉVALUATION

Contenu de ce chapitre

5.1	La conception de l'évaluation	127
5.2	Les méthodes de recherche	134
5.3	Le travail avant le départ sur le terrain	150
5.4	Les méthodes de terrain	166
5.5	Les approches tournées vers l'apprentissage	186

Comment utiliser ce chapitre ?

Ce chapitre peut aider les évaluateurs et les responsables d'évaluation à comparer les avantages et les inconvénients des différentes approches et méthodes d'évaluation. Le chapitre 5.1 donne un aperçu en matière de conception d'évaluation. Les chapitres 5.2 à 5.4 décrivent les différentes méthodes de collecte de données qui conviennent à l'EAH, à la fois au bureau et sur le terrain, avec des exemples d'utilisation dans différents contextes d'évaluation et des astuces relatives au processus d'analyse. Le chapitre 5.5 présente certains processus participatifs qui sont particulièrement bien adaptés aux évaluations tournées vers l'apprentissage.

Tableaux, schémas et encadrés

Schéma 10 :	Une théorie de programme implicite pour l'action humanitaire	132
Schéma 11 :	Les facteurs qui affectent la crédibilité et l'utilisation	135
Schéma 12 :	Fiabilité et validité de la recherche quantitative	142
Tableau 14 :	Les questions à (se) poser en matière de recherche qualitative	144
Schéma 13 :	Chaîne de preuves idéalisée	146
Tableau 15 :	Extrait d'un tableau de synthèse des observations	149
Schéma 14 :	Nombre de documents postés sur ReliefWeb durant une récente réponse humanitaire	153
Schéma 15 :	Événements clés dans le déclenchement d'une épidémie de fièvre de la dengue en 2009 au Cap Vert	156
Tableau 16 :	L'outil d'analyse du genre du Comité permanent inter-agences (IASC)	164
Encadré 7 :	Utiliser les souvenirs	168
Tableau 17 :	Les défis et l'éthique de la consultation des populations affectées lors d'une crise humanitaire	171
Tableau 18 :	Les manières de consulter la population affectée	174
Tableau 19 :	Les méthodes d'évaluation participative rapide	178
Encadré 8 :	Prendre en compte le sexe et l'âge durant la collecte des données sur le terrain	184

5.1 La conception de l'évaluation

Avant de choisir des méthodes, l'évaluateur doit réfléchir à la conception de l'évaluation. De nombreuses méthodes peuvent être utilisées avec différentes conceptions d'évaluation, mais certaines méthodes conviennent mieux que d'autres à certaines conceptions.



Conception de l'évaluation

Approche de recherche à utiliser lors d'une évaluation.

La conception de l'évaluation est parfois divisée en trois grandes catégories : expérimentale, quasi-expérimentale et non-expérimentale ([Bamberger et al., 2011a](#)). Cependant, cette catégorisation est d'une certaine manière limitante pour une évaluation humanitaire et un récent rapport de DFID ([Stern et al., 2012](#)) a identifié cinq types de conception pour les évaluations d'impact :

1. *Expérimentale*
2. *Statistique*
3. *A base théorique*

4. *Basée sur un cas*

5. *Participative*

Les évaluations de processus peuvent utiliser n'importe laquelle de ces conceptions mais tendent à se concentrer sur les approches participatives et celles basées sur un cas.

Les différentes conceptions utilisent des bases différentes pour déduire les causes des résultats. Les conceptions expérimentales impliquent une comparaison avec une hypothèse contrefactuelle.



Hypothèse contrefactuelle

Situation qui aurait prévalu si une action humanitaire donnée n'avait pas eu lieu.

Les conceptions statistiques impliquent des corrélations et des procédures différentes pour contrôler les variables confusionnelles. Les approches à base théorique établissent des parallèles entre des cas ou des groupes de cas. Les conceptions participatives impliquent une validation des causes des résultats par les participants.

[Bamberger et al. \(2011a, p. 31\)](#) adoptent une approche différente en catégorisant les évaluations d'impact d'après le moment où le travail de terrain a lieu (avant, pendant ou après une intervention) et selon s'il existe ou non un groupe de comparaison.

La majorité des évaluations humanitaires se déroulent durant ou après une intervention, et l'utilisation de groupes de comparaison ou de contrôle est rare. Les évaluations de processus ont lieu durant la mise en œuvre.

5.1.1 Les conceptions expérimentales

Cette catégorie comprend les essais contrôlés randomisés (RCT en anglais), les conceptions quasi-expérimentales et les expérimentations naturelles.

LES ESSAIS CONTROLES RANDOMISES



L'essai contrôlé randomisé

Expérimentation où les sujets sont désignés de façon aléatoire pour des groupes de traitement et de contrôle.

Toute véritable conception expérimentale prend la forme de RCT. Avec une sélection aléatoire, tout

individu a une chance égale d'être sélectionné pour le groupe de traitement ou de contrôle. Pour les évaluations humanitaires, des familles ou des communautés peuvent être choisies plutôt que des individus. Le RCT est la conception d'évaluation la plus efficace et la plus robuste pour éliminer les biais de sélection.



Biais de sélection

Risque que ceux qui participent à une évaluation ou bénéficient d'une intervention humanitaire diffèrent d'une manière difficilement observable des personnes dont ce n'est pas le cas.

La forme la plus robuste de RCT est l'essai à double insu où ni l'individu ni le chercheur ne savent si l'individu fait partie ou non du groupe de traitement. Un exemple d'utilisation de l'essai à double insu dans le secteur humanitaire a été le test d'efficacité des compléments probiotiques et prébiotiques auprès d'enfants sévèrement malnutris au Malawi ([Kerac et al., 2009](#)).

Les RCT sont confrontés à une série de problèmes potentiels en contexte humanitaire. Tout d'abord se

pose la question éthique d'attribuer une assistance à une population affectée de façon aléatoire plutôt que seulement sur la base des besoins. Ce point trouve normalement une réponse dans le fait de fournir une assistance différente mais de même type aux groupes de traitement et de contrôle.

Dans le cas de l'essai au Malawi ([Kerac et al., 2009](#)), tous les enfants malnutris ont reçu des montants équivalents d'aliments thérapeutiques avec l'élément de test inclus dans les aliments pour le seul groupe de traitement. Dans le cadre d'évaluations de programmes de transferts monétaires contre nourriture, le Programme alimentaire mondial a fourni l'un ou l'autre à des groupes de traitement et de contrôle sélectionnés de façon aléatoire au Sri Lanka ([Sandström and Tchatchua, 2010](#)) et au Malawi ([Audsley et al., 2010](#)).



Exemple
de bonne
pratique 5

Construire l'expérimentation au sein de l'intervention depuis le départ

L'Institut International de Recherche sur les Politiques Alimentaires a conduit une évaluation qui comparait la distribution alimentaire et la distribution d'argent par le PAM au Sri Lanka après le tremblement de terre et le tsunami de 2004 dans l'océan Indien. L'équipe a distribué de façon aléatoire de l'aide alimentaire ou de l'argent à chaque bénéficiaire. Parce que les deux groupes ont reçu des types d'assistance attractifs, le projet a évité les problèmes éthiques et pratiques qui surviennent quand un groupe est assisté lors d'une crise humanitaire et qu'un autre ne l'est pas (Sandström et Tchatchua, 2010).

Les problèmes techniques liés aux RCT sont notamment les suivants : le besoin de fournir le traitement de façon standardisée, les difficultés à préserver une composition stable des groupes de traitement et de contrôle ainsi que les dépenses et les efforts nécessaires. Ces contraintes ont amené à un examen des évaluations de programmes de promotion de la santé par l'Organisation Mondiale de la Santé en Europe afin de découvrir que l'utilisation des RCT en la matière était « dans la plupart des cas, inappropriée, trompeuse et inutilement onéreuse » ([WHO Regional Office for Europe, 1998, p. 5](#)). Les RCT requièrent aussi que

l'évaluation soit conçue depuis le début du projet avec une sélection aléatoire des groupes de contrôle et de traitement. Pour toutes ces raisons, les RCT sont rares dans le cadre d'évaluations humanitaires.



Exemple
de bonne
pratique 6

Faire une sélection à l'aveugle pour réduire le risque de biais

Lors de l'évaluation IRC de la reconstruction sur initiative communautaire de la Province de Lofa County, les communautés ont été désignées pour des groupes de traitement et de contrôle dès le début du projet au moyen d'un tirage au sort (Fearon et al., 2008). Une partie de la recherche était conduite par une ONG qui n'avait pas été impliquée dans le programme. L'ONG a réalisé les tests de la fonction communautaire sans savoir dans quel groupe se trouvait chaque communauté.

CONCEPTIONS QUASI-EXPERIMENTALES



Conception quasi-expérimentale

Conception d'évaluation dans laquelle un groupe de comparaison est choisi en dehors du groupe de traitement.

Les conceptions quasi-expérimentales dépendent de la capacité d'un certain nombre de techniques à faire correspondre le groupe de comparaison avec

le groupe de traitement. Elles sont susceptibles de biais de sélection car le groupe de comparaison peut être très différent du groupe de contrôle. Par ailleurs, une excellente évaluation des transferts monétaires de l'UNHCR au Burundi a cherché à comparer la situation de rapatriés récents qui avaient reçu de l'argent avec des rapatriés plus anciens qui n'en avaient pas reçu mais n'ont pas pu arriver à leur fin en raison des autres différences qui existaient entre les groupes (Haver et al., 2009).

EXPERIMENTATIONS NATURELLES



Experimentation naturelle

Situation dans laquelle la composition d'un groupe de traitement ou de contrôle est déterminée par des facteurs situés en dehors du contrôle du projet.

Les expérimentations naturelles peuvent se produire lorsqu'un groupe est affecté ou assisté et qu'un autre groupe similaire ne l'est pas ou que les deux groupes sont affectés ou assistés de manière différente. Par exemple, un examen de l'impact de l'ouragan Mitch sur la survie des enfants a comparé des enfants de zones frappées par l'ouragan

avec d'autres habitants des zones qui n'avaient pas été frappées ([Baez and Santos, 2007](#)). Les expérimentations naturelles peuvent également se produire lorsque les conditions varient avec le temps comme dans le cas d'arrivée en masses de réfugiés à l'ouest de la Tanzanie ([Maystadt, 2011](#)).

Les expérimentations naturelles peuvent être très simples, comme celle au Niger qui comparait les salaires journaliers avec les tarifs de rémunération à la pièce ([Kevlihan, 2010](#)). Dans un autre exemple, certaines juridictions du gouvernement local étaient plus préparées pour le tremblement de terre de Yogyakarta en 2006 car elles étaient déjà en alerte pour une éruption du Mont Merapi. L'Institut Fritz a noté dans ces zones des niveaux de satisfaction élevés vis-à-vis de la réponse d'urgence ([Bliss and Campbell, 2007](#)).

[Brancati \(2007\)](#) a utilisé une conception expérimentale naturelle dans son article sur l'impact des tremblements de terre dans les conflits inter-étatiques. Lors des évaluations conjointes, il est possible de comparer des groupes qui ont été assistés par différentes organisations en utilisant différentes approches.



Les évaluateurs doivent toujours être à la recherche de groupes qui n'ont pas reçu d'assistance ou qui ont reçu différents types d'assistance durant une réponse humanitaire. Parfois, de tels groupes (dans la mesure où ils sont identiques en tout autre point) peuvent donner naissance à une expérimentation naturelle. Les enquêtes auprès des populations affectées peuvent révéler de telles expérimentations naturelles..

5.1.2 Conceptions statistiques

Les conceptions statistiques reposent sur des outils statistiques de comparaison des différents groupes. Les conceptions statistiques, en particulier économétriques, sont largement utilisées en matière de recherche sur les conflits ([Do and Lyer, 2006](#)) ; mais elles ont été peu utilisées dans les EAH. Elles sont assez exigeantes sur le plan des connaissances statistiques nécessaires à la bonne conduite d'études. Les études économétriques conduites par la Banque mondiale en sont un exemple. D'autres exemples sont l'utilisation d'une analyse de régression linéaire pour déterminer l'importance de différents facteurs de résilience après le séisme d'Haïti ([Tulane University's Disaster Resilience Leadership Academy and University of Haiti, 2012](#)) et l'utilisation d'un modèle économétrique dans le cadre de l'évaluation des rapatriés vers le

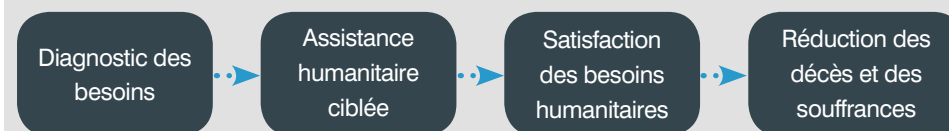
Burundi qu'a réalisée le Département d'Etat des Etats-Unis dans le but de comparer les rapatriés avec les non-rapatriés ([Telyukov et al., 2009](#)).

Les séries chronologiques interrompues sont souvent utilisées pour de la recherche en lien avec les catastrophes naturelles, comme les études d'impact du tremblement de terre de Taïwan en 1999 sur le suicide ([Yang et al., 2005](#)) et du Guatemala en 1976 sur la division du travail.

5.1.3 Les conceptions à base théorique

Comme mentionné dans le chapitre 3.2.5, les évaluations à base théorique sont basées sur une théorie de programme. Actuellement, peu d'évaluations humanitaires ont explicitement une base théorique bien qu'un très grand nombre soit basé sur une théorie implicite de l'action humanitaire tel qu'indiqué dans le **Schéma 10**, ou sur le cadre logique de l'intervention.

Schéma 10 : Une théorie de programme implicite pour l'action humanitaire



Parmi les évaluations récentes qui se considèrent comme étant à base théorique, on trouve une évaluation multi-bailleurs de la réponse humanitaire au Sud Soudan ([Bennett et al., 2010](#)), une évaluation de la résilience en Haïti ([Tulane University's Disaster Resilience Leadership Academy and University of Haiti, 2012](#)), et une évaluation du programme de l'UNICEF sur le développement de la petite enfance ([Boller et al., 2011](#)). Bien que des évaluations humanitaires puissent être partiellement à base théorique, cela n'est pas souvent souligné ou même clairement identifié dans les rapports d'évaluation.

La popularité croissante de l'approche de la théorie du changement et l'insistance de plus en plus forte sur la théorie du programme signifient que les conceptions à base explicitement théorique sont susceptibles de devenir de plus en plus populaires.

5.1.4 *Les conceptions basées sur un cas*

Une étude de DFID ([Stern et al., 2012](#)) utilise le terme « basé sur un cas » pour faire référence à une gamme beaucoup plus large que celle habituellement comprise sous la dénomination « étude de cas ».

Le terme basé sur un cas comprend toutes les évaluations qui se concentrent sur des cas spécifiques plutôt que sur le fait de comparer et d'opposer un certain nombre de cas différents comme les études de cas le font traditionnellement. Beaucoup d'évaluations humanitaires en cours appartiennent à cette catégorie. Les études de cas qui comparent les interventions à travers plusieurs pays sont courantes parmi les grandes évaluations humanitaires. Par exemple, l'évaluation du CERF ([Channel Research, 2011](#)) a examiné l'utilisation des financements à travers 16 pays (6 au moyen de visites de terrain et 10 au moyen d'une revue documentaire).

5.1.5 *Les conceptions participatives*

Un aspect de la participation est la participation des parties prenantes de l'évaluation. Certaines évaluations participatives n'incluent pas les populations affectées comme de simples sources d'informations mais aussi comme des éléments

actifs de l'évaluation. Ainsi, les participants ne sont pas vus comme des bénéficiaires passifs mais comme des acteurs de la réponse. Il est toutefois rare de voir cela dans les évaluations humanitaires qui, parfois, n'utilisent même pas les populations affectées comme source d'informations. La consultation des populations affectées continue d'être un problème dans l'action humanitaire ([Cosgrave et al., 2010, p. 4](#)), y compris dans les EAH. La [ALNAP de 2002 \(pp. 178–179\)](#) a constaté que seulement 27% des évaluations examinées avaient consulté les populations affectées et en avaient rendu compte de façon satisfaisante.

Un autre aspect est la participation des personnes évaluées. L'évaluation en temps réel des inondations au Pakistan a organisé des ateliers de parties prenantes pour valider les résultats et les conclusions mais aussi prioriser les recommandations de l'équipe ([Polastro et al., 2011b](#)). L'évaluation conjointe des ONG en 2007 suite à la réponse au séisme de Yogyakarta ([Wilson et al., 2007](#)) est allée encore plus loin dans l'organisation d'événements avec les parties prenantes afin de réexaminer et de modifier les conclusions et d'élaborer les recommandations.

5.1.6 *Les conceptions ponctuelles vs. les conceptions longitudinales*

La plupart des évaluations humanitaires comprennent seulement une période de travail sur le terrain. Les évaluations qui revisitent le même sujet au fil du temps, appelées évaluations longitudinales, sont courantes dans les contextes de développement mais rares dans l'évaluation humanitaire. Elles sont particulièrement efficaces pour identifier des changements survenus au fil du temps. Un exemple d'étude longitudinale en contexte humanitaire est la seconde évaluation des liens entre urgence, réhabilitation et développement suite au tsunami ([Brusset et al., 2009](#)). On retrouve également l'évaluation de World Vision sur les communautés affectées par le tsunami en Inde qui a rendu visite aux communautés deux ans plus tard, et une analyse des moyens de subsistance urbains en Afghanistan durant laquelle l'équipe a rendu visite aux foyers étudiés une fois par mois pendant un an ([Beall and Schutte, 2006](#)).

Bien que les évaluations longitudinales soient rares, un bon suivi est longitudinal et les différences entre les rapports de suivi du début et de la fin peuvent fournir des données probantes de changement longitudinal.

5.1.7 *Choisir une conception*

Le choix de la conception d'évaluation sera déterminé par le contexte et vos objectifs. Chaque conception a différents avantages et inconvénients. Les options peuvent être réduites par le temps dont vous disposez, votre budget et les compétences disponibles.

5.2 Les méthodes de recherche

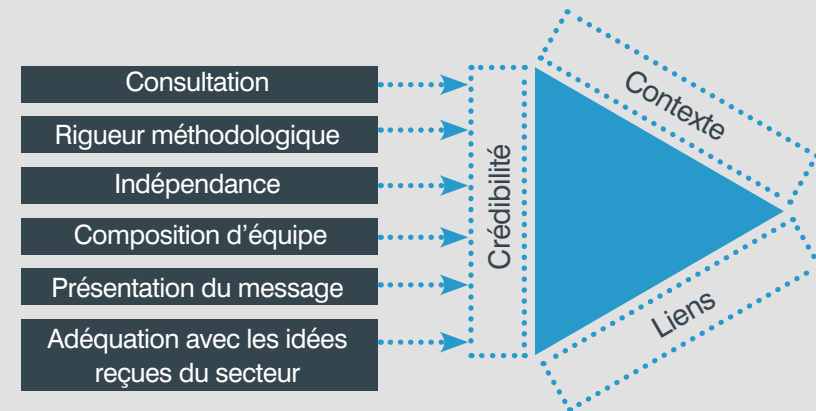
5.2.1 *Les méthodes*

L'EAH utilise à la fois des méthodes quantitatives, centrées sur des données chiffrées, et des méthodes qualitatives, centrées sur des données non chiffrées. Par exemple, des enquêtes structurées sont une méthode quantitative dans laquelle les questions non modifiables et standardisées sont posées à un échantillon de la population sélectionné de façon aléatoire. Parmi les autres types d'enquêtes, on compte les questions ouvertes qui ne sont pas chiffrées par l'intervieweur et qui sont en général qualitatives. Les entretiens semi-structurés d'informateurs clés sont un exemple de méthode qualitative. Pourtant, les données collectées à partir d'une recherche qualitative peuvent être enregistrées sous forme chiffrée.

L'évaluation de développement a été l'objet d'une redite des « guerres de paradigmes » des années 70 et 80 (Denzin, 2010), avec certains experts avançant que les méthodes quantitatives sont la seule approche de recherche capable de produire des données probantes objectives et d'autres s'opposant fortement à ce point de vue (Picciotto, 2012). De nombreuses critiques des méthodes quantitatives d'évaluation ne sont pas contre la catégorie en général mais uniquement contre l'utilisation d'enquêtes (en particulier les face-à-face).

L'important est d'utiliser des méthodes réalistes par rapport au contexte, bien adaptées à la question de recherche et perçues comme crédibles par les parties prenantes. La crédibilité ou la perception de la qualité influencent fortement l'utilisation (Court and Young, 2003); les facteurs qui affectent la crédibilité comprennent la rigueur, l'indépendance et la consultation (Schéma 11).

Schéma 11: Les facteurs qui affectent la crédibilité et l'utilisation



Les évaluateurs doivent clairement annoncer quelles méthodes ils ont utilisées. De plus, une autre bonne pratique consiste à ce qu'ils indiquent aussi sur quel standard ou quel guide ils se basent pour les appliquer de façon spécifique. Cela ne fournit pas seulement aux lecteurs plus d'informations avec lesquelles estimer la qualité de l'évaluation, cela aide aussi les autres personnes éventuellement désireuses d'utiliser la même méthode dans leur propre travail.

Les conseils actuels en matière d'évaluation dans les contextes d'urgence et de développement encouragent l'utilisation de méthodes mixtes dans un grand nombre de conceptions ([Bamberger et al., 2010](#)), y compris pour l'évaluation d'impact ([Leeuw and Vaessen, 2009](#); [Stern et al., 2012](#)), de manière à aboutir à des évaluations plus cohérentes, fiables et utiles ([Bamberger, 2012](#); [Adato, 2011](#)). Les méthodes mixtes sont utilisées de façon optimale lorsqu'elles se renforcent mutuellement. Cela peut se produire, par exemple, lorsque la recherche quantitative (une enquête) est non seulement basée sur le niveau de référence initial mais aussi sur une première série de recherches qualitatives qui a identifié les variables et hypothèses clés. Une fois l'enquête réalisée, une

recherche qualitative complémentaire peut explorer les questions soulevées par l'enquête et utiliser l'enquête pour aider à un échantillonnage raisonné de façon à ce que les catégories clés soient couvertes ([Adato, 2011](#)).

Une approche alternative, avec un travail de terrain qualitatif encadré par deux périodes de travail quantitatif, est défendue par [Bolton et al. \(2007\)](#). Ces auteurs recommandent des entretiens qualitatifs pour établir des listes libres d'impacts entre les enquêtes quantitatives pré et post-intervention, et donnent des exemples d'endroits où un tel travail de terrain quantitatif a identifié des impacts inattendus.

Malheureusement, pour la plupart des évaluations humanitaires, des pressions de temps font que les recherches qualitatives et quantitatives ont lieu en parallèle, ce qui réduit grandement les opportunités de synergie entre les deux méthodes. Un article d'InterAction sur les méthodes combinées lors d'évaluations d'impact ([Bamberger, 2012](#)) identifie un certain nombre d'autres séquences possibles mais aucune n'offre toutes les possibilités de synergie qu'offre la séquence qualitative-quantitative-qualitative.

5.2.2 L'échantillonnage

Il est normalement impossible d'interroger toutes les personnes au sein des populations affectées ou d'évaluer tous les éléments d'un programme. Cela peut être compensé en travaillant avec un échantillonnage représentatif de la population.



Echantillonnage

Sélection de personnes ou de cas à partir d'une population plus large en vue d'une recherche approfondie.

En matière d'échantillonnage, une approche consiste à définir un échantillon aléatoire au sein de la population. L'échantillonnage aléatoire n'est pas réservé aux méthodes quantitatives mais il est essentiel si vous voulez généraliser les résultats à partir de l'échantillon pris au sein de l'ensemble de la population.



Echantillon aléatoire

Echantillon défini à partir d'une population dans laquelle chaque membre a une chance égale d'être inclus dans l'échantillon.

Les évaluateurs ne disposent normalement pas de tous les détails sur la population et utilisent à la place certains proxy pour établir l'échantillon. Les listes de bénéficiaires peuvent être utilisées comme un cadre d'échantillonnage dans les enquêtes humanitaires mais de telles listes sont souvent obsolètes, « gonflées » ou incorrectes. Dans un camp de réfugiés bien organisé, il peut être possible d'utiliser une liste principale des abris comme cadre d'échantillonnage.

Un véritable échantillonnage aléatoire est souvent très difficile dans le cadre du bouleversement caractéristique des contextes humanitaires. Par conséquent, des formes d'échantillonnage pseudo-aléatoire peuvent être utilisées. Un type courant d'échantillonnage peut consister dans le fait de sélectionner une maison toutes les cinq, en faisant utiliser à la personne interrogée une grille préconçue. Les ordinateurs peuvent être utilisés pour générer des numéros aléatoires (techniquement, ils sont pseudo-aléatoires puisqu'ils sont générés par une règle mathématique) et permettre de faire une sélection dans l'assistance.



Exemple
de bonne
pratique 7

Utiliser les listes des sites affectés comme cadre d'échantillonnage

Les enquêtes réalisées en 2006 lors de la Tsunami Evaluation Coalition ont utilisé des listes de villages affectés et de camps de personnes déplacées internes comme le cadre initial de l'échantillonnage. Des sites ont alors été sélectionnés au sein de ce cadre d'échantillonnage et des bénéficiaires ont été sélectionnés pour des entretiens sur ces sites (**Brusset et al., 2006**).



Si vous devez réaliser un échantillonnage aléatoire à partir d'une liste de projet, un tableur peut générer des numéros aléatoires que vous pouvez utiliser pour sélectionner des projets. Vous pouvez souvent éprouver le besoin de stratifier l'échantillon de manière à assurer qu'une section transversale des types de projet est représentée.



Exemple
de bonne
pratique 8

Utiliser des numéros aléatoires pour générer un échantillon aléatoire stratifié

Lors d'une évaluation du CERF en 2011, des financements furent échantillonnés à partir de 16 pays sur une période de 5 ans et deux différentes fenêtres de financement. L'équipe voulait un échantillon aléatoire mais aussi s'assurer que chaque année, chaque organisation et chaque fenêtre de financement était bien représentée dans l'échantillon. A chaque projet était assigné un numéro généré de façon pseudo-aléatoire. Ensuite, les projets étaient classés par pays, année, fenêtre de financement et organisation bénéficiaire. Chaque pays avait un taux d'échantillonnage nominal compris entre 20 et 30 projets par pays. Le taux d'échantillonnage nominal était alors appliqué pour calculer combien de projets devaient être sélectionnés pour chaque organisation (pour chaque pays, chaque année et chaque fenêtre de financement), avec un minimum de 1. Ce nombre de projet était alors sélectionné pour chaque organisation (pour chaque pays, chaque année et chaque fenêtre de financement avec le projet portant le numéro aléatoire le plus élevé qui était tiré au sort en premier.

ECHANTILLONS PAR GRAPPES

Certaines enquêtes cherchent à éviter un cadre d'échantillonnage en utilisant une approche d'échantillonnage par grappes.



Echantillon par grappes

Echantillon déterminé en divisant tout d'abord la population dans son ensemble par grappes (en général des zones géographiques) puis en sélectionnant à l'intérieur de ces grappes.

Les grappes peuvent ne pas être des zones géographiques mais des écoles ou des sections de camps. En général, la sélection au sein des grappes est réalisée en utilisant certaines règles pseudo-aléatoires systématiques à l'intérieur de chaque grappe (par exemple, la maison la plus proche du point sélectionné puis chaque Xème maison dans la même direction). Une recherche récente a également suggéré que les sélections aléatoires peuvent être faites au moyen de l'utilisation de photos aériennes ou satellite ([Shannon et al., 2012](#)).

Les échantillons par grappes sont couramment utilisés pour des enquêtes nutritionnelles avec un échantillon de 30 tiré de chacune des 30 grappes. De toutes les techniques d'échantillonnage aléatoire, l'échantillonnage par grappes présente le plus grand risque d'erreurs d'échantillonnage (c'est-à-dire le risque le plus élevé de ne pas être représentatif de l'ensemble de la population). Les tailles d'échantillons doivent d'être plus grandes que les échantillons aléatoires tirés de l'ensemble de la population de manière à atteindre la même précision.

TAILLE DE L'ECHANTILLON

La taille de l'échantillon nécessaire pour être en mesure de généraliser les résultats à partir de l'échantillon relatif à l'ensemble de la population de façon fiable du point de vue statistique dépend des éléments suivants :

- *Plus proche est cette proportion de la population qui a le facteur qui vous intéresse : plus proche de 50 % est la proportion, plus grand est l'échantillon nécessaire pour l'estimer avec précision.*
- *La variabilité de ce facteur dans la population (en général, exprimé comme*

l'écart standard, selon les estimations tirées de l'expérience ou d'enquêtes passées) : plus grande est la variabilité, plus grande est la taille nécessaire de l'échantillon.

- *La marge d'erreur ou de précision nécessaire : une précision plus fine exige une plus grande taille d'échantillon.*
- *La taille de la population : elle est importante seulement si la taille de l'échantillon est supérieure de 5 % à la taille de la population.*

Les enquêtes nutritionnelles de grande échelle utilisent en général des échantillons par grappes avec 30 grappes de 30 échantillons (30 x 30) et les échantillons de l'enquête de la Tsunami Evaluation Coalition en ont utilisé environ 1.200. Des variantes de type 30 x 7, 33 x 6 et 67 x 3 ont été suggérées comme des alternatives là où de hauts niveaux de prévalence sont attendus, comme pour la couverture vaccinale et la distribution de kits d'urgence ([Bilukha, 2008](#)).



Calculer la taille de l'échantillon est techniquement complexe et il est préférable de rechercher de l'aide auprès d'un statisticien spécialisé dans les enquêtes pour s'assurer que votre enquête produira les informations dont vous avez besoin.



Au moment d'établir le budget, si vous n'avez pas de meilleures informations, il est probablement prudent de partir du principe que vous aurez besoin d'un échantillon de 1.000 à 1.200 personnes.

Les échantillons destinés à des méthodes qualitatives sont en général relativement peu nombreux et raisonnés plutôt qu'aléatoires. L'avantage d'un échantillon raisonné réside dans le fait que vous pouvez vous concentrer sur ces cas présentant un très fort intérêt.



Echantillon raisonné

Echantillon sélectionné de façon dirigée.

En général, les échantillons raisonnés sont sélectionnés pour inclure des personnes que l'on pense les plus capables de fournir les données recherchées (par exemple, des informateurs clés). La stratégie inhérente aux échantillons raisonnés les distingue de simples échantillons de commodité.



Echantillon de commodité

Echantillon choisi par commodité plutôt que d'après une stratégie sous-jacente.

L'échantillonnage de commodité est la stratégie d'échantillonnage la plus faible puisque ce qui est commode du point de vue de l'échantillonnage peut être inadapté à la population affectée. Par exemple, les personnes les plus proches de la route sont bien plus faciles à échantillonner mais elles reçoivent aussi plus d'assistance que celles qui sont plus difficiles d'accès comme l'a montré l'évaluation sur l'éducation au Timor oriental de l'UNICEF ([Tolani-Brown et al., 2010, p. 51](#)).

Une technique d'échantillonnage raisonné largement utilisée en recherche qualitative est l'échantillonnage en boule de neige.



Echantillonnage en boule de neige

Echantillonnage élargi en demandant aux premières personnes interrogées d'en recommander d'autres qu'elles considèrent comme d'intéressantes sources d'informations

L'échantillonnage en boule de neige est souvent utilisé pour définir un échantillon lors d'entretiens semi-structurés avec des informateurs clés. Il peut également être utilisé pour les populations qui seraient difficiles à échantillonner d'une autre manière.



Exemple
de bonne
pratique 9

L'échantillonnage en boule de neige lors d'enquêtes téléphoniques

L'évaluation de 2009 qui portait sur l'utilisation des distributions d'argent par l'UNHCR en Afghanistan a utilisé l'échantillonnage en boule de neige lors d'entretiens téléphoniques en Iran. Il était demandé à chaque ménage interrogé d'identifier d'autres personnes qui pourraient fournir des informations utiles. Sans cela, il aurait été très difficile d'échantillonner ce groupe de façon efficace ([Davin et al., 2009, p. 13](#)).



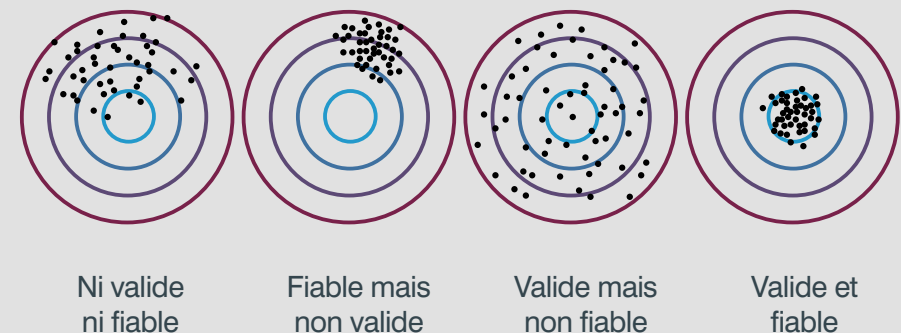
Que vous utilisiez des méthodes quantitatives ou qualitatives ou une combinaison des deux, énoncez clairement quelle stratégie d'échantillonnage vous utilisez et quelles implications cela a sur votre recherche. Le fait d'être transparent par rapport à toutes les contraintes rencontrées augmente votre crédibilité.

5.2.3 Contrôler la qualité de la recherche

Le premier test concernant la qualité d'une évaluation est le test de valeur; et celui-ci se trouve au cœur de toute approche centrée sur l'utilisation probable de l'évaluation (voir Chapitre 2.4). Les méthodes d'évaluation doivent fournir des valeurs fiables et valides. [Hammersley \(1992, p. 69\)](#) définit une valeur comme valide ou vraie si elle « représente précisément les éléments des phénomènes qu'elle cherche à décrire, expliquer ou théoriser ». L'EAH utilise principalement les méthodes qualitatives et, dans une moindre mesure, les méthodes quantitatives. L'une des plus grandes différences entre les deux est la manière dont elles déterminent la proportion dans laquelle elles se révèlent une authentique réflexion sur la réalité sous-jacente.

Des méthodes quantitatives essaient d'être à la fois valides (c'est-à-dire de refléter l'authentique situation sous-jacente) et fiables (c'est-à-dire de reproduire le même résultat lors d'essais répétés de la même situation) [Schéma 12](#)). Des approches de mesures et des séries de questions standardisées sont utilisées pour assurer la fiabilité des approches quantitatives. Par conséquent, on recommande aux enquêteurs d'utiliser des formulations spécifiques et de ne pas proposer d'explications supplémentaires.

Schéma 12 : Fiabilité et validité de la recherche quantitative



Les schémas reflètent dans quelle mesure les résultats sont conformes à la réalité sous-jacente.

Les cercles bleu clair du centre sont les réalités et les points représentent des données.

Les méthodes quantitatives dépendent des statistiques pour déterminer la validité des affirmations. Si l'échantillon est sélectionné de façon aléatoire et passe certains tests statistiques, alors les résultats de l'échantillon peuvent être généralisés à l'ensemble de la population. Les données quantitatives sont en général perçues comme faisant autorité et scientifique. Les limites de la recherche qualitative sont souvent bien plus visibles aux parties prenantes que celles de la recherche

quantitative. Cette perception provient d'un manque de compréhension des données quantitatives. Par exemple, dans une enquête formelle, les réponses que les enquêteurs attendent des personnes interrogées dépendent de qui sont ces personnes et de comment elles sont perçues par l'intervieweur. De tels effets sont très difficiles à contrôler et des pré-tests sont rarement réalisés à une échelle suffisante pour les mettre en lumière.

L'enquêteur est l'une des principales sources d'erreurs lors d'entretiens en face-à-face (O'Muircheartaigh et Campanelli, 1998, p. 63). Il existe une riche littérature sur la manière dont les réponses peuvent être influencées par les caractéristiques de l'enquêteur, y compris le genre, ([Flores-Macias and Lawson, 2008](#); [Kane and Macaulay, 1993](#); [Webster, 1996](#)), le statut ([Riessman, 1977](#)), la race ([Hill, 2002](#)), l'origine ethnique ([Webster, 1996](#)), la distance sociale générale ([Dohrenwend et al., 1968](#)), et même le fait que l'enquêteur soit mince ou non ([Eisinga et al., 2011](#)).



Observez toujours avec attention les résultats de la recherche quantitative. Accordez de l'importance à l'intervalle de confiance. Par exemple, un titre de journal peut dire que le soutien à un parti politique a baissé de 4 %. Cependant, de tels sondages ont en général une limite de confiance de 95 % à plus ou moins 3 %, il est ainsi possible que le soutien à un parti de 20 % soit indiqué à 18 % dans une enquête et 22 % dans une autre sans avoir changé. Dans les rapports d'EAH, observez la manière dont les données ont été collectées et comment l'enquête a traité les contraintes courantes des crises humanitaires comme la réticence des intervieweurs à révéler des informations personnelles.



A quel point l'échantillon était-il représentatif ? Dans votre propre compte-rendu, assurez-vous que vous indiquez l'intervalle de confiance et, si approprié, la signification statistique de vos résultats.

Les méthodes qualitatives doivent être évaluées d'une manière différente; [Mays and Pope \(2000\)](#) ont suggéré les questions résumées dans le [Tableau 14](#).

Tableau 14 : Les questions à (se) poser en matière de recherche qualitative

Question	Comment traduire cela dans le cadre de l'EAH ?
<p>Valeur ou pertinence : Cela valait-il la peine de réaliser ce travail ? A-t-il apporté une connaissance utile ?</p>	<p>L'utilisation de l'évaluation établira sa valeur.</p>
<p>Clarté de la question de recherche : La question de recherche était-elle claire ? Si ce n'est pas le cas au début, au moins à la fin du processus de recherche ? Le chercheur a-t-il été capable de mettre de côté ses idées préconçues ?</p>	<p>Le rapport de démarrage est l'outil clé pour clarifier les questions d'évaluation.</p>
<p>Caractère approprié de la conception : Une méthode différente aurait-elle été plus appropriée ? Par exemple, si une hypothèse causale a été testée, une approche qualitative était-elle appropriée ?</p>	<p>Le chapitre sur la méthodologie doit indiquer pourquoi une combinaison bien particulière de méthodes a été choisie.</p>
<p>Contexte : Le contexte ou le cadre sont-ils décrits de façon adéquate de manière à ce que les lecteurs puissent associer les résultats à d'autres situations ?</p>	<p>Le rapport d'évaluation doit inclure un chapitre qui définit le contexte.</p>
<p>Echantillonnage : L'échantillon incluait-il toute la gamme de cas ou de contextes de manière à pouvoir faire des généralisations conceptuelles plutôt que statistiques (autrement dit, était-ce plus qu'un simple échantillonnage de commodité ?). Le cas échéant, des efforts ont-ils été faits pour obtenir des données qui puissent contredire ou modifier l'analyse en élargissant l'échantillon (par exemple, à un autre type de zone) ?</p>	<p>Le chapitre sur la méthodologie doit décrire l'approche d'échantillonnage utilisée et son effet attendu sur les données ainsi que les contraintes pratiques rencontrées par l'équipe d'évaluation comme le manque d'accès à certaines zones ou certains groupes de population.</p>

Question

Collecte et analyse des données : Les procédures étaient-elles systématiques ? Un fichier d'enregistrement des procédures a-t-il été élaboré afin que quelqu'un d'autre puisse répéter chaque étape, y compris l'analyse ? L'analyse a-t-elle bien intégré toutes les observations ? Dans quelle mesure l'analyse a-t-elle développé des concepts et des catégories capables d'expliquer des processus clés ou de précieux compte-rendus et observations des personnes interrogées ? L'interprétation des données et la théorie sur laquelle elle est basée étaient-elles faciles à comprendre ? Le chercheur a-t-il recherché des éléments contredisant ses résultats ?

Réflexivité : Le chercheur a-t-il attentivement évalué l'impact probable des méthodes utilisées et sa propre influence sur les données obtenues ? Les données incluses dans le rapport ont-elles fourni des données probantes suffisantes pour que les lecteurs puissent évaluer si les critères d'analyse ont été atteints ?

Comment traduire cela dans le cadre de l'EAH ?

Le chapitre méthodologique doit contenir des exemples d'instruments de collecte de données utilisés et expliquer le processus d'analyse. Le rapport doit montrer une relation claire entre les données probantes, les résultats, les conclusions et les recommandations. L'échelle de temps des actions humanitaires demande souvent aux évaluateurs d'utiliser des raccourcis (comme le fait de travailler directement à partir de notes manuscrites plutôt que d'utiliser des procédures d'analyse recourant à un logiciel de codage), mais la méthodologie doit rester aussi rigoureuse que possible et être clairement décrite dans le rapport.

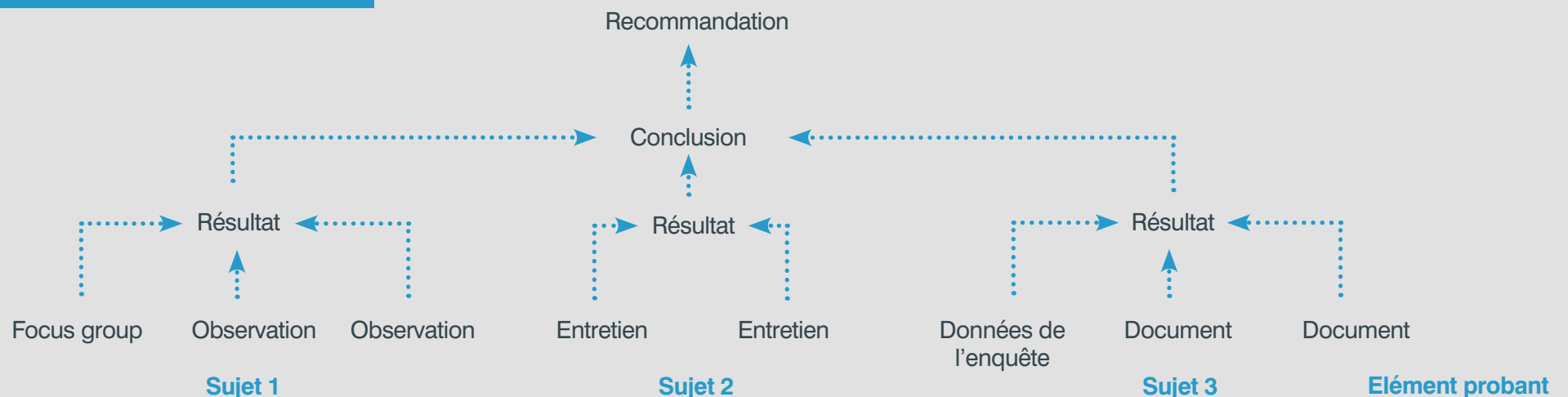
L'évaluation doit inclure de courtes biographies des évaluateurs pour permettre aux lecteurs de juger du poids à accorder à l'analyse et aux biais potentiels. Le chapitre sur la méthodologie doit détailler toutes les limites de la méthodologie.

Source: [Mays and Pope \(2000\)](#), ([ALNAP, 2005](#)).

En résumé, les évaluations doivent justifier les méthodes qu'elles ont choisies et décrire en détail la stratégie d'échantillonnage, les contraintes, les limites imposées par l'absence de réponse et toute autre source potentielle de biais.

5.2.4 Constituer un corpus de données probantes
 La critique la plus courante des évaluations est le fait que les recommandations ne soient pas basées sur les conclusions, que les conclusions ne soient pas basées sur les résultats et que les résultats ne soient pas basées sur les données probantes (bien sûr, la critique peut aussi provenir d'un désaccord avec les conclusions et les recommandations). Le **Schéma 13** montre comment une recommandation doit être basée sur des données probantes qui, dans l'idéal, sont de multiples éléments de preuves collectés en utilisant différentes méthodes d'évaluation.

Schéma 13 : Chaîne de preuves idéalisée



Un élément de preuve est une donnée collectée à partir d'une observation, d'un entretien, d'un document ou d'une autre source. Tous les éléments de preuve ne sont pas d'égale importance ; gardez cela à l'esprit s'il y a des contradictions dans les preuves. Les résultats sont les déductions tirées des preuves. Les conclusions sont des déductions d'ordre supérieur tirées des résultats après réflexion sur leur signification. Les recommandations sont des suggestions d'action pour répondre aux manques qui émergent dans les conclusions et pour améliorer la performance.

Vous pouvez renforcer votre chaîne de preuves en utilisant un simple tableau dans le but d'enregistrer chaque élément de preuve pour chaque sujet d'évaluation. Cela pourrait être un tableau avec des colonnes qui listent les sujets, la preuve, ses sources (par exemple, les entretiens, les documents, les observations et les enquêtes), la date d'enregistrement (pour faciliter la collecte) et les initiales du membre de l'équipe qui a entré l'élément de preuve.

Utiliser un tableau de preuves vous permet de voir pour quels sujets vous manquez de preuves pour obtenir un résultat. Parfois, il peut être possible de

rediriger l'effort de collecte de données pour se concentrer sur ces sujets.



Structurez votre tableau de preuves (et tous vos instruments de collecte des données) autour de la structure attendue pour le rapport d'évaluation. Le fait de relier immédiatement les données collectées à leur utilisation finale évite de perdre du temps à une phase ultérieure en réarrangeant votre matériau de travail.



Dans une étude multi-pays ou multi-sites, prévoyez assez de temps après le premier pays ou la première visite de site pour affiner vos instruments de collecte de données.

5.2.5 Utiliser la triangulation



Triangulation

Fait de comparer des données de différentes sources pour voir si elles tendent vers le même résultat.

La triangulation est une technique clé pour assurer la précision et la fiabilité des recherches qualitatives et par méthodes combinées.

- **La triangulation de méthodes** compare les données générées par différentes méthodes de recherche – par exemple, les observations et les entretiens de groupe.
- **La triangulation de sources** compare les informations obtenues à partir de différentes sources – par exemple, les informateurs clés au siège et sur le terrain.
- **La triangulation d'évaluateurs** compare les informations collectées par différents évaluateurs – par exemple, les différents membres de l'équipe.
- **La triangulation analytique** compare les résultats de différentes techniques analytiques – par exemple, une étude focalisée sur le nombre de références à un sujet en particulier dans les documents officiels (analyse de contenu) et une autre focalisée sur le niveau de financement accordé à ce même sujet (analyse de données chiffrées).

5.2.6 De la collecte à l'analyse des données

Le fait d'utiliser un tableau de preuves facilite également l'analyse car cela organise tous les éléments probants relatifs à un sujet en particulier (**Tableau 15**).



Au moment de concevoir votre méthodologie, prévoyez la triangulation dès le début en utilisant plus d'une méthode pour observer chaque question.

Tableau 15 : Extrait d'un tableau de synthèse des observations

Sujet	Élément probant	Résultat	Conclusion	Recommandation
Niveau de qualification des volontaires	<ol style="list-style-type: none"> 1. <i>Lors de l'exercice de simulation n°5, le formateur n'a pas poussé la tête assez en arrière pour libérer la langue de la voie aérienne en montrant la réanimation par bouche-à-bouche.</i> 2. <i>Un tiers des nœuds de jonction de cordes montrés en exercice de simulation étaient inappropriés ou dangereux.</i> 3. <i>La méthode de réanimation montrée le plus fréquemment n'est plus considérée comme très efficace.</i> 	Les qualifications des volontaires et du formateur demandent à être affinées.	La formation demande un mécanisme de contrôle qualité.	L'organisation doit mettre en place un mécanisme de contrôle qualité dans les 6 mois pour assurer que la formation est de bonne qualité et que les qualifications sont maintenues à jour.

Quand tous les éléments probants relatifs à un sujet particulier sont listés ensemble, il est plus facile pour l'évaluateur de les prendre en compte avec précision avant d'écrire sur le sujet. Il est aussi plus facile pour les multiples membres d'une équipe de participer à l'analyse.



Lorsqu'un tableau de preuves contient des informations issues d'entretiens, il ne peut être partagé en dehors de l'équipe sans enfreindre la confidentialité.

5.2.7 Les hypothèses contrefactuelles

Les évaluations quantitatives recourent en général à des comparaisons avec une hypothèse contrefactuelle valide du point de vue statistique pour déterminer quel effet a eu une intervention. C'est particulièrement le cas avec les évaluations d'impact (voir Chapitre 6.4).

Les RCT, dans lesquelles on affecte de façon aléatoire les personnes à des groupes de traitement et de contrôle, sont la conception la plus robuste du point de vue statistique en matière d'évaluation quantitative (voir Chapitre 5.1.1), mais les approches quasi-expérimentales sont plus courantes dans les évaluations quantitatives que les véritables RCT. Du point de vue statistique, les conceptions d'évaluation robustes sont rares et difficiles à atteindre dans les contextes de développement ([Bamberger et al., 2011b, p. 20](#)). et encore plus rares dans les contextes humanitaires où les contraintes sont proportionnellement plus importantes.

Bien que le terme "contrefactuel" soit traditionnellement utilisé pour des hypothèses valides du point de vue statistique, il est aussi possible d'examiner l'hypothèse contrefactuelle en utilisant des méthodes quantitatives et qualitatives non-statistiques. Des approches aussi simples

que demander aux populations affectées quels changements ont été dus à une intervention peuvent aider à déterminer l'hypothèse contrefactuelle, et la triangulation peut être utilisée pour soutenir cette approche. De telles questions doivent être formulées avec soin pour éviter de provoquer une réponse positive par politesse (voir [Encadré 7](#) pour des précisions sur l'utilisation des souvenirs).

Un article issu d'une réunion de l'Association américaine de l'évaluation ([Bamberger et al., 2009](#)) liste une série de méthodes non statistiques pour déterminer l'hypothèse contrefactuelle qui n'est pas en général : « Que ce serait-il produit si l'intervention n'avait pas eu lieu » mais « Qu'aurait-il prévalu si une action humanitaire donnée n'avait pas eu lieu ». Cette liste comprend toute une gamme de méthodes qualitatives et quantitatives.

5.3 Le travail avant le départ sur le terrain

5.3.1 La recherche de documents clés

Une recherche documentaire initiale ou étude de bureau est la première étape de toute évaluation ; c'est une manière économique et efficace d'apprendre sur le contexte, la crise humanitaire et la réponse avant de se lancer dans le travail de terrain.

Les documents clés peuvent souvent être trouvés en recourant aux méthodes suivantes :

- *Demandez des suggestions aux parties prenantes clés et au groupe de référence.*
- *Consultez ReliefWeb (<http://reliefweb.int/>) et d'autres sites humanitaires clés comme AlertNet et Humanitarian Practice Network.*
- *Recherchez des bibliographies sur le sujet.*
- *Consultez la base de données des rapports d'évaluation d'ALNAP pour rechercher des évaluations passées.*
- *Faites des recherches sur Internet, y compris sur Google Scholar.*
- *Lorsque vous trouvez un document utile, vérifiez une liste de référence ou d'autres pistes pour des travaux utiles supplémentaires..*
- *Faites circuler une liste de documents provisoires à l'équipe et au groupe de référence pour qu'ils vous suggèrent des sources complémentaires.*

Une fois que vous avez identifié les documents, les membres de l'équipe doivent les lire. Ils peuvent soit noter les points importants dans un tableau de preuves (voir Chapitre 5.2.4), soit préparer un bref résumé qui souligne les points clés. Parce que tous les documents ne sont pas aussi pertinents les uns

que les autres, il est intéressant de disposer d'un système d'enregistrement où les auteurs des notes de lecture évaluent chaque document.



Pour éviter d'être surchargés de textes à lire, concentrez-vous sur les documents qui vous apportent des éléments nouveaux sur un point clé, vous fournissent une analyse utile, ou que vous allez probablement citer dans le rapport d'évaluation. Il vous faudra également peut-être inclure des documents que votre utilisateur final attend de consulter même si vous ne pensez pas que ces documents soient utiles.

L'utilisation d'un logiciel de gestion bibliographique comme EndNote ou Zotero pour enregistrer les documents utilisés dans le cadre de votre étude peut vous faire gagner beaucoup de temps au moment d'organiser votre bibliographie. Ajoutez un résumé à chaque entrée que vous créez. Cela facilite la production d'une bibliographie annotée qui peut s'avérer très utile. Des commentaires spécifiques, par exemple sur la pertinence d'une source pour une évaluation particulière ou votre opinion sur sa qualité, doivent être enregistrés dans un champ de notes distinct plutôt que comme une partie du résumé.

Si vous réalisez une importante revue documentaire, pensez à faire une bibliographie annotée disponible publiquement pour aider d'autres recherches sur les questions humanitaires.



Utiliser DropBox, Google Drive ou d'autres logiciels de collaboration à distance pour partager le même ensemble de documents. Cela permet une mise à jour automatique de la copie de l'ordinateur local de chaque membre de l'équipe lorsque l'un d'eux ajoute un nouveau fichier ou en modifie un existant.

Tout bon guide de méthodologie de recherche couvre la recherche documentaire bien que le traitement puisse être bref. Voir par exemple [Thomas and Mohnan \(2007\)](#). Les textes qui se concentrent sur la recherche de documents clés, comme [Hart \(1998\)](#), tendent à se focaliser sur la recherche de littérature utile à une thèse de doctorat.



Exemple
de bonne
pratique 10

Publier une bibliographie des sources utilisées pour une évaluation

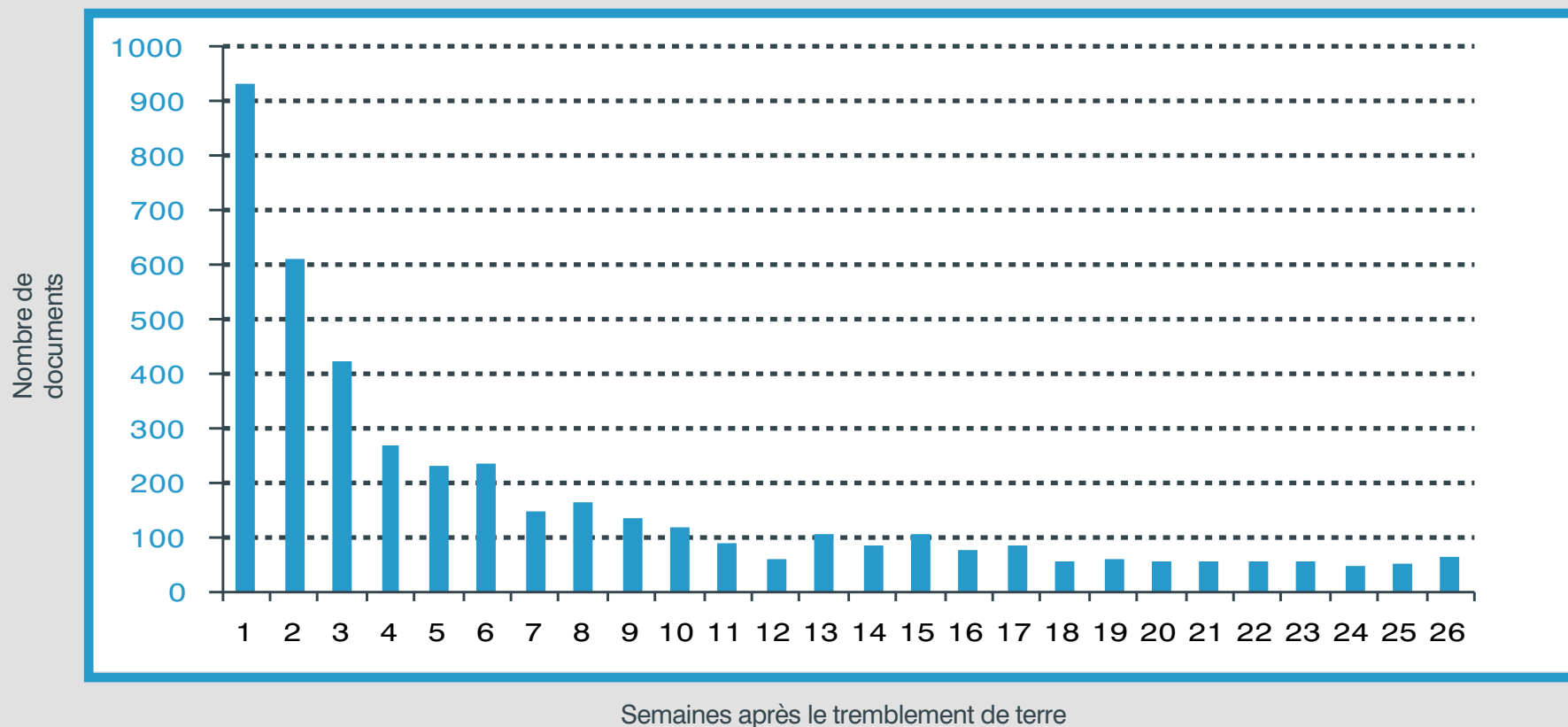
Le travail de suivi de la Tsunami Evaluation Coalition réalisé en 2009 a publié une bibliographie annotée dans un volume séparé qui accompagnait le rapport principal. La bibliographie contenait 610 entrées et a été élaborée en partie sur la base de la bibliographie développée pour le travail de la Coalition en 2005 et 2006 ([Cosgrave et al., 2009a](#)).

5.3.2 Utiliser un grand nombre de documents

Après de récentes urgences humanitaires, les organisations ont généré un grand nombre de documents décrivant leurs actions et relatant leurs avancées. Une organisation peut à elle seule produire de nombreux rapports de situation et autres matériaux écrits en peu de temps. De nombreuses organisations publient leurs rapports sur le site de ReliefWeb (<http://reliefweb.int/>), géré par le Bureau de la Coordination des Affaires humanitaires [Schema 14](#)).

Schéma 14 : Nombre de documents postés sur ReliefWeb durant une récente réponse humanitaire

Nombre de documents postés par semaine sur ReliefWeb à propos du tremblement de terre au Pakistan (4213 pendant les 26 premières semaines)



Une équipe d'évaluation ne peut évidemment pas lire des centaines de documents. Beaucoup seront d'un intérêt secondaire mais certains contiendront des informations essentielles. Vous pouvez séparer le bon grain de l'ivraie en utilisant un logiciel pour classer les documents d'une manière qui vous permette de faire des recherches complexes et de mesurer les occurrences de mots clés.

Utilisez une approche structurée, comme une liste de vérification ou un outil de notation, pour résumer l'information de manière rigoureuse. Lorsque vous disposez de trop de documents pour utiliser cette approche, vous pouvez utiliser un échantillon aléatoire.

5.3.3 Réaliser une chronologie

Une chronologie peut être utile pour comprendre comment le contexte et la réponse ont changé au fil du temps ainsi que pour souligner les liens qui méritent une étude plus approfondie.



Impliquez toute l'équipe dans la réalisation de la chronologie car cela peut servir à familiariser chacun de ses membres avec le contexte



Exemple
de bonne
pratique 11

Préparer une chronologie détaillée

L'évaluation finale du projet Emergency Capacity Building Project a présenté une chronologie détaillée pour expliquer l'évolution du projet au fil du temps (Morris et Shaughnessy, 2007). L'extrait suivant de la chronologie montre les premiers pas du projet.

Date	Événement	Description
Mars 2003	Création d'un groupe de travail inter-agences (IWG)	L'objectif général du groupe était de faciliter une mise en œuvre plus efficace de l'assistance humanitaire par les organisations non gouvernementales (ONG) au moyen d'une stratégie conjointe de construction des capacités. Les objectifs spécifiques étaient de créer de nouveaux mécanismes durables pour partager les bonnes pratiques et pour s'inspirer de l'analyse d'autres organisations spécialisées, de tester collectivement des pratiques innovantes dans le cadre des opérations de terrain des ONG ainsi que de produire et partager des définitions et des standards qui peuvent être appliqués collectivement à l'ensemble du secteur humanitaire.
Juillet 2004	Rapport sur Emergency Capacity	L'analyse a montré que, pendant que des membres de l'IWG avaient progressé en matière de construction des compétences d'urgence, plus particulièrement dans les systèmes de financement rapide et de réponses, des manques majeurs en termes de capacités demeuraient, gênant à la fois la vitesse et l'efficacité de l'action humanitaire mais aussi la coordination avec les autres.
Août 2004	Dossier de demande de subvention pour le projet ECB Project	Le dossier de demande de subvention ECB, dont CARE était le coordinateur pour les 7 organisations, a été soumis à la Fondation Bill et Miranda Gates pour une aide au développement des initiatives 1, 2 et 3 du projet ECB.

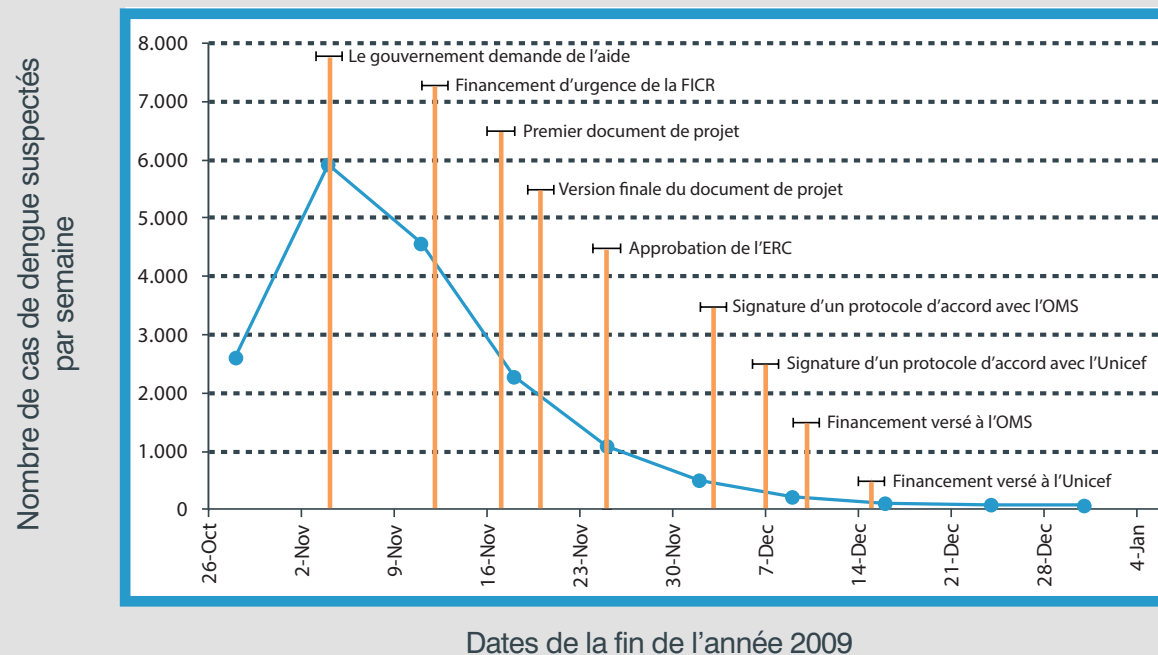
Les chronologies sont en général comme des tables de dates ou d'événements. Parfois, les événements qui se produisent au sein d'une organisation sont placés dans une colonne distincte des événements qui ont lieu en dehors.



Pensez à souligner les points clés d'une chronologie sous la forme d'une ligne de temps visuelle (**Schema 15**).

Schéma 15 : Evénements clés dans le déclenchement d'une épidémie de fièvre de la dengue en 2009 au Cap Vert

Ligne de temps pour le déclenchement d'une épidémie de fièvre de la dengue au Cap Vert et le processus d'application du CERF



5.3.4 L'analyse du contenu et des mots-clés

Comme indiqué ci-dessus, les réponses humanitaires peuvent générer un très grand nombre de documents. Plus de 100 documents par jour ont été postés sur ReliefWeb durant la première semaine qui a suivi le tremblement de terre et le tsunami de décembre 2004 en Asie, et près de 2000 documents ont été postés le mois suivant. Il n'est pas possible qu'une équipe d'évaluation les lisent tous mais ils peuvent toujours constituer une source d'informations utile tout au long de l'analyse du contenu.



Analyse du contenu

Analyse de l'information textuelle d'une manière standardisée permettant aux évaluateurs de procéder à des déductions à propos de l'information.

L'analyse du contenu prend en général la forme d'un codage des documents dans le but de réduire le texte à une série de catégories. Une forme d'analyse de contenu utilisée dans un certain nombre d'évaluations, y compris dans la récente évaluation du CERF ([Channel Research, 2011](#)), est l'analyse des mots-clés.



Analyse des mots-clés

Forme d'analyse de contenu qui examine la fréquence de l'occurrence des mots-clés pour mettre en évidence les tendances au fil du temps dans une série de documents ou pour comparer deux documents.

Une liste de mots-clés peut présenter jusqu'à 500 termes. Certaines catégories seront représentées par de multiples mots-clés. La catégorie Eau, Assainissement et Hygiène, comprend les mots-clés suivants :

- *puits de forage*
- *chloration*
- *matières fécales*
- *latrine*
- *assainissement*
- *source*
- *WASH*
- *chlore*
- *défécation*
- *hygiène*
- *eaux usées*
- *savon*
- *toilettes*
- *eau*

- *approvisionnement en eau*
- *puits*
- *WATSAN*

Deux guides très utiles en matière d'analyse textuelle sont ceux de [Krippendorff \(2004\)](#) et [Benini \(2009\)](#), a short guide to using three textual analysis tools for humanitarian and development workers.



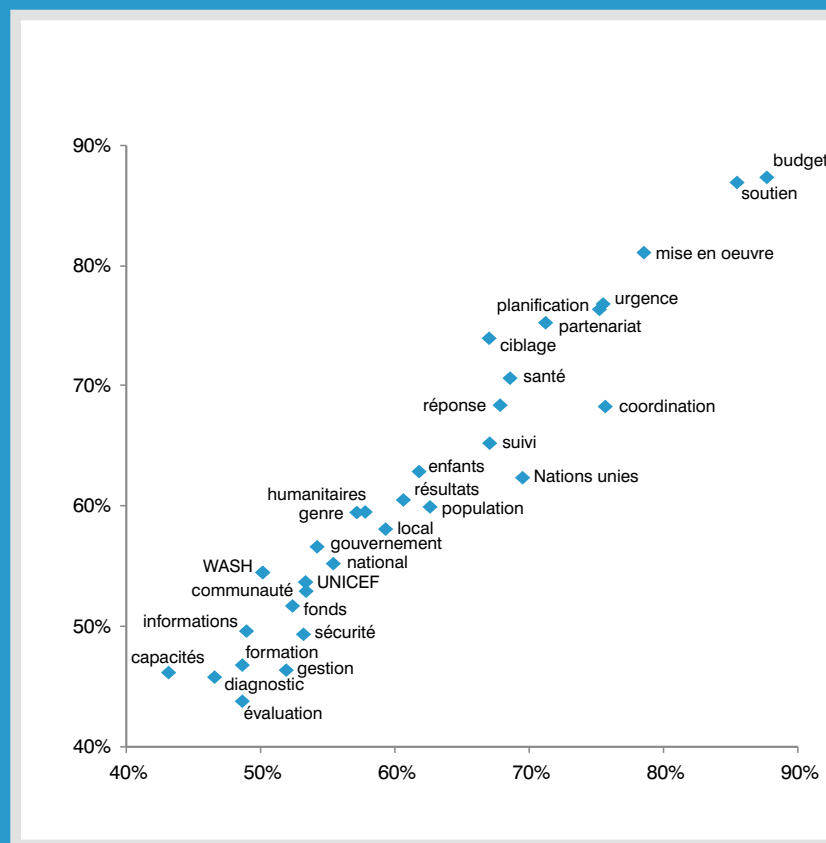
Exemple
de bonne
pratique 12

Utiliser l'analyse de mots-clés pour vérifier qu'un échantillon est représentatif

Dans le cadre de l'évaluation des cinq ans du CERF, des pays ont été sélectionnés par le groupe de pilotage pour des études plus détaillées basées sur un échantillon raisonné plutôt que sur un échantillon aléatoire. L'équipe d'évaluation voulait s'assurer qu'il n'y ait pas de différences fondamentales entre les documents de projet des 16 pays de l'étude et ceux des 63 pays non sélectionnés. La prévalence et la fréquence de 405 mots-clés représentant 327 catégories dans 7.797 documents ont été calculées en utilisant le logiciel dtSearch. La fréquence de l'occurrence a été comparée parmi les 2.897 documents issus des pays de l'étude et 4.900 documents issus d'autres pays pour s'assurer que les projets de l'échantillon étaient représentatifs de l'ensemble de la population

Comparaison entre deux séries de documents de l'occurrence d'un terme donné

% de documents de projet comportant le terme parmi les autres pays bénéficiaires du CERF



% de documents de projet comportant le terme au sein des pays ciblés par l'évaluation

5.3.5 L'analyse de données chiffrées

Les opérations humanitaires peuvent générer de grandes quantités de données chiffrées, par exemple en ce qui concerne les financements, les distributions ou le nombre de personnes recevant une assistance. Analyser ces données peut mettre en valeur des sujets particuliers et déterminer les éléments suivants :

- *Moyennes* et standards
- *Données aberrantes* (cas inhabituels)
- *Cohérence* des modèles d'activités avec les objectifs déclarés d'un projet.

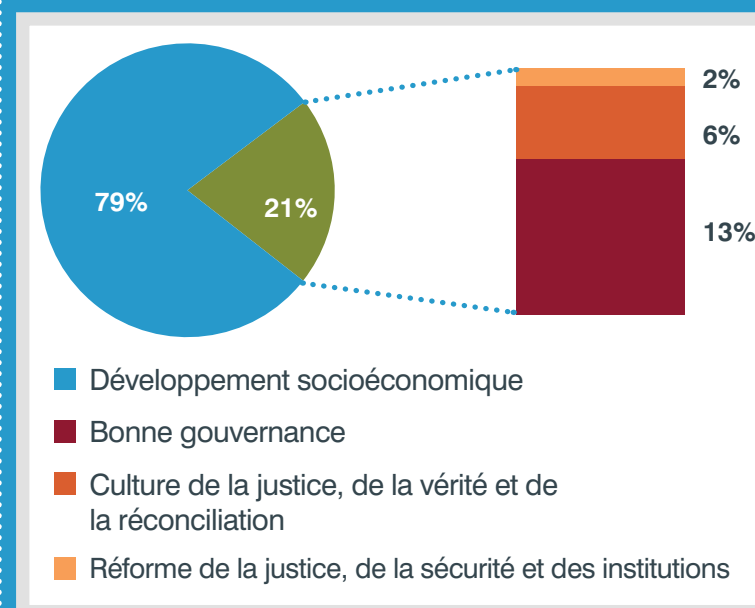


Exemple
de bonne
pratique 13

Analyser les dossiers des bailleurs pour voir comment les financements ont été utilisés

Une évaluation conjointe multi-bailleurs de l'appui à la prévention des conflits et à la consolidation de la paix au Sud Soudan comprenait une analyse des dossiers qui étudiait les dépenses de 2.724 projets et les classait par catégorie de consolidation de la paix. Comme indiqué dans le schéma ci-dessous, l'analyse a montré que près de 80 % de l'appui des bailleurs allaient au développement socio-économique (Bennett et al., 2010).

Répartition du budget de 4,05 milliards de dollars par catégorie de consolidation de la paix





Un tableur permet une analyse plus flexible qu'une base de données. Apprenez comment utiliser des tableaux dynamiques croisés et des fonctions de base de données dans votre tableur pour réaliser l'analyse.

5.3.6 Rapport coût-efficacité et efficience

Le rapport coût-efficacité constitue un aspect complexe des évaluations humanitaires. Les bailleurs lui ont accordé une attention accrue de manière à ce que leurs financements obtiennent le plus fort impact possible. L'EAH compare en général différentes approches parce qu'il est impossible de fixer une valeur monétaire à la vie humaine et à la souffrance. Malgré cela, même la comparaison des coûts de différentes approches en vue d'un même effort peut s'avérer difficile pour deux raisons :

1. // n'existe pas assez de données pour permettre des comparaisons valides.
2. // est difficile d'évaluer le compromis entre vitesse et coût.

S'il est un domaine où les données ne sont pas fiables, c'est l'estimation des coûts unitaires. Dans le cas le plus fréquent, le reporting d'une organisation vis-à-vis d'un de ses bailleurs comprend les

moyens et les coûts. Cependant, il n'existe pas une terminologie standardisée pour les moyens et, de ce fait, des moyens portant les mêmes noms au sein de deux organisations peuvent ne pas être comparables.

Par exemple, deux organisations peuvent installer des latrines à des coûts unitaires très différents dans le même camp de personnes déplacées internes. La première peut installer seulement des dalles de latrines pour des familles en zones rapidement accessibles alors que la seconde peut former des groupes communautaires à construire les dalles au lieu de leur acheter de manière à ce qu'ils puissent installer des latrines quand ils rentreront dans leurs villages, fournit des matériaux pour la construction de la superstructure des latrines, travaille dans des zones relativement inaccessibles et met également en œuvre un programme complet d'éducation à l'hygiène.

Par ailleurs, des données financières ne sont pas toujours utilisables pour réaliser des comparaisons. Le rapport financier remis à un bailleur ne comprend pas les contributions reçues d'autres bailleurs et les rapports financiers regroupent souvent un ensemble d'activités sous une seule entrée budgétaire. Pour

prolonger l'exemple des latrines, une organisation peut utiliser des transports gratuits fournis par les militaires alors qu'une autre paie ses transports de même que l'une peut avoir un seul bailleur pour son programme de latrine alors que l'autre en a plusieurs.

L'assistance mise en œuvre peut aussi être difficile à déterminer. En 2011, une évaluation conjointe de l'assistance dans le centre et le sud de la Somalie n'a pas été en mesure de définir avec précision dans quelle proportion l'assistance avait atteint la population ciblée, les estimations fournies par les personnes interrogées allant de 30 à 70 % ([Polastro et al., 2011a, p. 48](#)).



Lorsque vous interrogez les plus hauts dirigeants d'une organisation, demandez-leur les coûts unitaires des différents éléments de leur opération. Les dirigeants qui connaissent leurs coûts unitaires et inducteurs de coûts sont plus susceptibles de rechercher activement à maintenir leurs coûts bas pour maximiser l'assistance qu'ils mettent en œuvre.

Malgré ces difficultés, il peut toujours être utile de calculer les coûts unitaires pour trois raisons :

1. Cela met en valeur les différences qualitatives entre programmes qui peuvent être suivies durant le travail de terrain.
2. Cela pose la question de la nécessité de tous les services mis en œuvre par les organisations à des coûts supérieurs ou de la possibilité que la communauté affectée effectue facilement elle-même certains de ces services.
3. Cela identifie l'approche qui fournisse une assistance au plus faible coût.

Au moment d'évaluer le rapport coût-efficacité, recherchez des activités qui ont été menées par plus d'une organisation et pour lesquelles le nombre d'intrants (latrines, consultations médicales, personnes desservies) est clair. Essayez d'évaluer leur coût unitaire et, si nécessaire, demandez aux organisations plus d'éléments détaillés sur les coûts.



Lorsqu'il existe d'énormes différences dans les coûts unitaires, pensez à inclure des exemples des différentes approches des projets à échantillonner sur le terrain afin de voir si les différents niveaux de coût ont amené des résultats et des impacts différents.

Il est plus difficile de comparer les cas où une approche à coût élevé fournit une assistance plus rapidement que celle à bas coût. Des réponses plus rapides peuvent sauver des vies et réduire les souffrances, mais cela est difficile à quantifier. L'exemple classique est celui de l'approvisionnement des secours par voie aérienne plutôt que par la route. Les programmes nutritionnels mis en œuvre par des ONG peuvent avoir un coût unitaire plus faible que les programmes gouvernementaux mais ces derniers sont bien plus durables.

Dans de tels cas, les évaluateurs doivent étudier l'ensemble de la chaîne, pas seulement l'élément qui a été géré rapidement, pour voir si l'option la plus onéreuse a bien réellement fourni l'assistance la plus rapide. Par exemple, dans certains cas des opérations d'urgence du Rwanda en 1994, des opérations aériennes ont continué après que des mécanismes d'approvisionnement par la route ont été mis en place, et des biens qui étaient facilement disponibles localement ont été acheminés par les airs ([Borton et al., 1996, chapitre 5](#)).

Il se peut qu'un transport aérien prenne tellement de temps à être organisé que les camions arrivent avant le premier avion ou que les biens acheminés par

voie aérienne aient passé tellement de temps aux douanes ou à l'aéroport qu'il n'y a pas de gain de vitesse au final.



Il peut être très instructif d'examiner les différents coûts unitaires pour les opérations qui sont directement mises en œuvre et celles qui le sont via des partenaires. Dans certains cas, une organisation peut opérer des deux manières.

5.3.7 *Evaluer l'attention aux questions de genre dans les études sans terrain*

La question du genre est clé dans l'action humanitaire et constitue un important facteur de vulnérabilité (voir [Encadré 5](#)). Evaluer la sensibilité au genre est particulièrement complexe dans le cadre d'études sans terrain. Une approche consiste à définir dans quelle mesure un projet se concentre sur le genre et à calculer quel pourcentage des projets répondent à la question à des degrés différents de concentration. Le Comité permanent inter-agences (IASC) a défini 4 niveaux d'attention au genre ([Tableau 16](#)).

Tableau 16 : L'outil d'analyse du genre du Comité permanent inter-agences (IASC)

Niveau	Description
0	Le genre n'est traité dans aucun composant du projet.
1	Le projet est conçu pour contribuer de façon limitée à l'égalité de genre. Les dimensions relatives au genre sont comprises de façon significative dans un ou deux des trois principaux composants : évaluation des besoins, activités et résultats.
2a	Intégration du genre : le projet est conçu pour contribuer de façon significative à l'égalité de genre. Les différents besoins des femmes/filles et des hommes/garçons ont été analysés et pris en compte avec soin dans les trois principaux composants : évaluation des besoins, activités et résultats.
2b	Actions ciblées : le principal objectif du projet est de faire avancer l'égalité de genre. L'ensemble du projet cible (a) les femmes ou les hommes, les filles ou les garçons qui présentent des besoins particuliers ou souffrent de discrimination ou (b) concentre toutes les activités qui mettent en place des services propres au genre ou des relations plus équitables entre les femmes et les hommes.

Source: (IASC, 2010a, 2010b).

La même approche peut également être appliquée à des sujets transversaux comme la vulnérabilité, l'âge, le VIH/SIDA et la viabilité écologique afin de classer les projets comme : non concentrés sur le sujet, concentrés sur le sujet de façon limitée, répondant au sujet de façon significative (mais pas comme priorité principale) ou le traitant comme priorité principale. [HelpAge International \(2010\)](#) a utilisé une classification similaire dans une étude récente sur les financements humanitaires pour les personnes âgées.

Si votre étude comprend des entretiens avec des informateurs clés, vous devez également poser des questions sur le sujet durant les entretiens afin de trianguler les informations issues de l'analyse du dossier.



S'il y a trop de projets dans le dossier à étudier pour pouvoir tous les analyser, vous pouvez sélectionner un échantillon. Toutefois, un tel échantillon doit être un échantillon aléatoire si vous voulez à partir de lui généraliser à l'ensemble du dossier.

5.3.8 Enquêtes en ligne

Celles-ci sont beaucoup plus faciles à organiser que les enquêtes réalisées par des personnes en face-à-face. Elles peuvent être un complément utile à d'autres méthodes de collecte de données. Cependant, elles souffrent du potentiel biais induit par les non-réponses. Elles sont utiles pour identifier les sujets à étudier de façon plus approfondie auprès du personnel d'une organisation ou de ses organisations partenaires. L'ouvrage de [Dillman et al. \(2009\)](#) est la principale référence en la matière. Pour une analyse des enquêtes de terrain, voir Chapitre 5.4.5.



Exemple
de bonne
pratique 14

Utiliser des enquêtes en ligne pour rassembler des opinions provenant d'un groupe éparé

L'étude Stay and Deliver portant sur les bonnes pratiques en matière de mise en œuvre de l'assistance humanitaire dans des environnements différents au niveau de la sécurité et des risques a étudié plusieurs stratégies que les acteurs humanitaires ont utilisées pour maintenir une présence opérationnelle et continuer leur travail. L'équipe a effectué un travail de terrain dans six pays et réalisé des revues documentaires dans six autres. De manière à enregistrer les opinions du personnel national, l'équipe a mis en place une enquête en ligne multilingue (anglais, français, espagnol et arabe) pour collecter les opinions de 1148 acteurs locaux (Egeland et al., 2011).



Demandez toujours aux personnes interrogées si elles souhaiteraient répondre à d'autres questions ou apporter des clarifications par téléphone ou e-mail.



Pensez à traduire l'enquête dans d'autres langues pour la rendre accessible à tous ceux que vous voulez interroger.

5.4 Les méthodes de terrain

5.4.1 *Les entretiens avec des informateurs-clés*

Les entretiens avec des informateurs-clés sont la colonne vertébrale des évaluations humanitaires. Souvent, des aspects d'une réponse humanitaire ne sont pas bien documentés et interroger les personnes impliquées dans cette réponse est la meilleure façon de déterminer ce qui s'est passé et pourquoi. Les informateurs clés sont les personnes dont on peut penser, d'après leur rôle (par exemple de responsable d'organisation ou de chef de communauté), qu'ils connaissent bien l'intervention en cours d'évaluation. Les informateurs clés qui ont réfléchi sur l'opération et en ont tiré des conclusions générales sont des sources particulièrement utiles.

Les entretiens avec des informateurs clés sont une manière efficace de réunir de l'information quand le temps et les ressources sont limités. Les entretiens durent en général de 45 minutes à 1 heure et peuvent réunir une ou plusieurs personnes. Les entretiens sont semi-structurés et utilisent un guide d'entretien ou une liste de vérification. Par exemple, les TdR peuvent inclure la question « Dans quelle mesure le projet a-t-il pris en compte les questions de genre dans sa conception ? » mais cette formulation peut simplement évoquer une répétition de la politique relative au genre actuellement en place au sein de l'organisation. Si vous demandez plutôt « Qui a le plus bénéficié de ce projet – les hommes, les femmes, les garçons ou les filles ? » et que vous approfondissez la réponse à travers d'autres questions, vous obtiendrez plus facilement une indication concernant la proportion dans laquelle les responsables de projet sont conscients de l'importance des données désagrégées ou ont réalisé une analyse de genre de l'intervention. Rappelez-vous que c'est le travail de l'évaluateur et non de la personne interrogée de répondre aux questions des TdR.

Au moment de planifier votre guide d'entretien, prenez en compte les éléments suivants :

- *Limitez vos questions à 20 ou 25, soit le maximum que vous puissiez parcourir en une heure d'entretien.*
- *Débutez chaque entretien avec une explication de l'objectif de l'entretien et les droits de la personne interrogée, y compris les questions de confidentialité.*
- *Séquencez vos questions de manière à ce que les premières créent un climat de confiance et que les questions plus délicates soient abordées à la fin. Lors d'un entretien semi-structuré, vous pouvez modifier l'ordre des questions pour ne pas rompre le fil de la conversation. Assurez-vous de donner suite à toute information inattendue mais pertinente.*
- *Incluez toujours une question finale sur ce que la personne interrogée a appris ou ferait différemment la prochaine fois.*
- *Demandez à la personne interrogée si elle s'attendait à d'autres questions car cela peut vous aider à faire ressortir des points survolés, à résoudre des incompréhensions sur l'objectif de l'évaluation et parfois à découvrir de nouvelles pistes.*
- *Demandez toujours aux personnes interrogées*

de suggérer d'autres personnes avec qui parler.

En fonction de la culture, certaines questions peuvent être considérées comme sensibles de même que les réponses peuvent être dictées par des normes sociales. Par exemple, si vous demandez à des bénéficiaires « Dans quelle mesure l'assistance que vous avez reçue a-t-elle été utile ? », ils peuvent se sentir obligés de faire l'éloge de l'assistance par politesse. En revanche, si vous leur demandez « Comment pensez-vous que vos voisins ont perçu l'assistance qu'ils ont reçue ? », ils peuvent être moins contraints par les bonnes manières.

Encadré 7 : Utiliser les souvenirs

Il existe parfois peu de données historiques ou de niveaux de référence pour indiquer à quoi ressemblait la situation avant la crise. Vous pouvez compenser cela en demandant aux personnes interrogées comment est à présent la situation par rapport au passé. Les personnes se font parfois du passé une idée plus favorable qu'il n'était en réalité et, pour cela, il est important de procéder à des vérifications en posant des questions complémentaires. Par exemple, si l'on vous dit durant un entretien que chaque famille avait un char à bœufs avant la grande sécheresse d'il y a 10 ans, demandez quelle sorte de char avait la personne interrogée ou les membres de sa famille. Le fait de poser des questions à propos des coûts et des frais de transport peut aider à déterminer le niveau de précision.

La mémoire ne remplace pas une étude sur les niveaux de référence. Elle devient moins précise au fil du temps, en particulier lorsque le souvenir évoqué n'est pas d'une importance centrale pour la personne interrogée ([de Nicola and Giné, 2011](#)). La précision d'un souvenir varie aussi par rapport à la situation sociale de la personne interrogée ([Manesh et al., 2008](#)).



Si vous utilisez les souvenirs, limitez-les aux événements importants plutôt qu'aux données d'ordre général et essayez d'introduire un élément qui puisse être vérifié à partir d'autres sources

Les critiques concernant les entretiens semi-structurés avancent qu'elles collectent des perceptions mais pas des faits et que les personnes

interrogées peuvent donner des réponses biaisées ou des réponses qui arrangent leurs propres intérêts. Ces critiques sont vraies dans une certaine mesure. Cependant, aucune méthode de recherche n'est totalement dépourvue de biais. Même dans les sciences physiques, l'observateur influence l'observation, d'où l'utilisation de tests en « double-aveugle ». Même pour des méthodes quantitatives comme l'enquête, la réponse à une question dépend de la personne qui la pose ([Davis et al., 2010](#)).

Las percepciones son en sí mismas Las perceptions sont des informations valables en elles-mêmes ; par exemple, elles sont utilisées comme base de certains indicateurs économiques clés comme l'Indice de confiance des entreprises et l'Indice des directeurs d'achats. Comme ces indices, les entretiens semi-structurés enregistrent les connaissances et les perceptions des personnes présentant une expérience intéressante.

Il est évident que toutes les personnes interrogées ont des intérêts propres à défendre. Certaines veulent que leur travail soit largement reconnu, d'autres veulent en cacher les défauts et d'autres encore veulent que l'assistance se poursuive. L'évaluateur doit rester conscient de l'influence de

tels biais et également du fait que les entretiens sont des processus structurés. Lorsque vous utilisez des données d'entretien, pensez au poids à accorder aux différents points de vue.

- *Donnez un poids plus important aux opinions des personnes les plus affectées. Par exemple, les points de vue des bénéficiaires pour savoir s'ils ont été consultés ou pas a beaucoup plus de poids que ceux des organisations sur le même sujet.*
- *Le point de vue d'une personne ayant une expérience significative dans un secteur peut avoir plus de poids que celui d'une personne ayant peu d'expérience. Cependant, ce n'est pas toujours le cas : une personne avec beaucoup d'expérience peut être enfermée dans une approche ou un état d'esprit bien particulier.*
- *Pensez à donner un poids plus important aux points de vue qui contredisent l'apparent intérêt personnel de la personne interrogée. Par exemple, accordez une attention particulière aux bénéficiaires qui demandent une diminution de l'assistance ou à un représentant du gouvernement qui demande d'acheminer les ressources en passant par les ONG.*

Idéalement, les évaluateurs devraient être capables de parler la langue des informateurs clés qu'ils

interrogent. Lorsque cela n'est pas possible, ils doivent travailler avec des interprètes, ce qui augmente le temps de travail nécessaire pour passer en revue toutes les questions. Les évaluateurs doivent faire tout leur possible pour engager des interprètes de bon niveau. Les interprètes professionnels valent bien leur prix car ils sont moins susceptibles de modifier la question ou les réponses. Durant une évaluation récente en Afghanistan, il a été demandé aux anciens de la communauté de s'exprimer sur l'impact des Talibans au niveau des services gouvernementaux qu'ils avaient en charge. Le membre du personnel de l'organisation qui servait d'interprète a dit à l'équipe que les anciens se plaignaient de ce que les actions des Talibans avaient réduit leur accès aux services. Plus tard, l'assistant de l'équipe d'évaluation leur a rapporté que les anciens avaient en réalité dit qu'ils étaient les Talibans et que le gouvernement refusait de leur fournir les services à moins qu'ils ne rendent leurs armes.



Si vous devez utiliser un membre de l'organisation comme interprète et que vous n'avez aucune notion de la langue, il est bon d'avoir un assistant qui comprenne la langue et puisse prévenir l'équipe s'il constate des imprécisions majeures. Vous pouvez aussi enregistrer la conversation à la vue de tous afin de l'analyser plus tard et d'encourager l'interprète à être précis.

5.4.2 Les entretiens avec les populations affectées

Les entretiens avec la population sont une source importante de données pour les évaluations humanitaires ainsi qu'une étape essentielle pour satisfaire les besoins de la population affectée et les consulter.

Ils fournissent souvent de très bonnes informations sur l'assistance mise en œuvre, en particulier sur le fait qu'elle soit adaptée ou non.

CONCEVOIR VOTRE APPROCHE

Les entretiens avec les populations affectées sont particulièrement complexes (**Tableau 17**). Les connaissances locales sont essentielles pour être conscient de ces difficultés et y répondre de façon éthique.

Tableau 17 : Les défis et l'éthique de la consultation des populations affectées lors d'une crise humanitaire

Défi	Réponse éthique
Des groupes de population courent des risques de maltraitance ou de violence ciblée.	Si le fait de parler à une équipe d'évaluation peut augmenter les risques de maltraitance ou de violence qu'encourent les populations, ne les interrogez pas.
Des groupes puissants politiquement dominent le processus de consultation.	Soyez vigilant par rapport aux dynamiques de pouvoir et cherchez des moyens pour atteindre les populations marginalisées et moins puissantes.
Des membres de la population affectée sont traumatisés par leur expérience de la crise.	Faites preuve de sensibilité avec les personnes interrogées afin d'éviter de créer des tensions, par exemple en évitant de leur faire revivre une expérience traumatique.
La crise a eu pour résultat une rupture de la confiance et des relations au sein de la population affectée.	Demandez à des membres respectés de la communauté de présenter le processus d'évaluation et de consultation. Assurez une transparence complète en expliquant l'objectif, les contraintes et l'utilisation prévue de l'évaluation ainsi que la manière dont les données seront utilisées et conservées.



Dans la plupart des environnements, à moins que l'entretien ne se déroule dans un espace privé comme le foyer, un entretien en tête-à-tête tournera rapidement à l'entretien de groupe au fur et à mesure que les amis et les voisins se joignent à vous.



Les entretiens individuels avec un bénéficiaire sont plus susceptibles de fournir une image précise de l'aide qui a été reçue alors que les entretiens de groupe aborderont plus les aspects « politiques » car les personnes interrogées jouent un rôle devant l'assistance en même temps qu'elles répondent à votre question.

Les étapes à suivre pour concevoir un processus de consultation avec les populations affectées sont les suivantes :

- *Identifier quels groupes de population vous voulez interroger durant l'évaluation et pourquoi. Souvenez-vous que les connaissances locales sont fondamentales pour prendre les bonnes décisions.*
- *Devez-vous seulement inclure les bénéficiaires directs de l'action humanitaire ou aussi ceux qui n'en ont pas bénéficié mais ont été affectés par la crise ?*
- *Comment devez-vous désagréger les groupes de population à interroger ? (par exemple, par genre, ethnicité, état socio-économique, et selon le statut : personne déplacée, réfugié ou résident)*
- *Choisissez l'approche de votre entretien (voir [Tableau 18](#)).*
- *Préparez-vous à être flexible sur vos méthodes de consultation au cas où un manque de temps ou d'accès ou encore des contraintes logistiques vous obligerait à changer vos plans.*
- *Soyez sensible aux défis et à l'éthique de la consultation des populations affectées lors d'une crise humanitaire (voir [Tableau 17](#)).*

- *Décrivez l'ensemble du processus des entretiens dans le chapitre *Méthodologie de votre rapport d'évaluation*.*



Exemple
de bonne
pratique 15

Consultation des populations affectées

Dans une évaluation de l'UNHCR sur l'intégration de l'âge, du genre et de la diversité ([Thomas and Beck, 2010](#)), l'étude pays de la Colombie a utilisé une approche extrêmement participative. Quatre communautés de différentes parties du pays ont été sélectionnées pour participer à l'évaluation tout au long des étapes suivantes ([Mendoza and Thomas, 2009](#)):

- Un atelier avec des membres de la communauté pour élaborer une chronologie des événements ayant mené au diagnostic participatif de l'UNHCR (un outil clé dans le processus d'intégration) et le plan d'action qui en a résulté pour évaluer la compréhension du plan d'action et juger si le diagnostic participatif avait généré des changements dans la communauté.
- Des réunions avec des sous-groupes de femmes, d'hommes, d'adolescent(e)s, d'enfants, de personnes âgées et de personnes handicapées pour analyser les différentes perceptions qu'avaient les sous-groupes des résultats du diagnostic participatif et du plan d'action

RESUME DES DIFFERENTES METHODES

La plupart des entretiens avec les populations affectées prennent la forme d'entretiens de groupes. Ceux-ci sont souvent désignés de façon incorrecte comme des entretiens en « *focus groups* ». Les « *focus groups* » sont des conversations intensives facilitées par un animateur réunissant six à huit personnes issues d'un groupe spécifique de la population partageant une même expérience (voir [Krueger and Casey, 2009](#)). Dans la pratique, il est rarement possible de contrôler suffisamment l'environnement pour conserver un « *focus group* » car la plupart évolue vers un entretien de groupe moins contrôlé. De façon idéale, un entretien de groupe ne doit pas comprendre plus de 10 à 12 personnes mais ce chiffre peut rapidement augmenter jusqu'à 100 ou plus. Le [Tableau 18](#) résume les différentes manières de consulter la population affectée.

Tableau 18 : Les manières de consulter la population affectée

Méthode	Avantages	Inconvénients
<p><i>Focus groups</i> et entretiens de groupes (6 à 10 personnes)</p>	<p>C'est une façon efficace de rechercher les points de vue d'un groupe homogène (par exemple, des personnes déplacées internes ou un groupe particulier de ménages qui peut encore être désagrégé selon les catégories hommes-femmes).</p> <p>Elle se prête à une discussion ouverte qui peut être guidée par une liste de vérification dans un souci de cohérence et donne aux intervieweurs la flexibilité d'approfondir certaines informations intéressantes (par exemple, en demandant « Est-ce valable pour tous ? » ou « Est-ce que les femmes sont d'accord avec cela ? ») et de poser directement les questions aux individus.</p> <p>Les intervieweurs peuvent utiliser des méthodes d'évaluation participative rapide avec le groupe (voir Chapitre 5.4.4).</p>	<p>Il peut être difficile de contrôler la taille et la composition du groupe. Si d'autres personnes se joignent spontanément, cela peut affecter la dynamique du groupe.</p> <p>Cette approche requiert un facilitateur qualifié et un preneur de notes.</p> <p>Il peut être difficile de poser des questions sensibles à un groupe.</p>
<p>Réunions communautaires</p>	<p>Il est possible d'utiliser les structures communautaires existantes pour consulter les populations affectées bien que ce choix demande que les dynamiques de pouvoir au niveau local soient prises en compte.</p> <p>Il est possible d'atteindre un grand nombre de personnes, voire tous les membres de la communauté</p>	<p>Les réunions peuvent être dominées par des chefs communautaires.</p> <p>Les femmes et les jeunes peuvent être réticents à parler de façon ouverte.</p> <p>Il peut ne pas être possible de suffisamment désagréger les réponses pour permettre une analyse utile.</p>

Méthode	Avantages	Inconvénients
Entretiens individuels et entretiens de ménages	<p>Il peut être plus facile de poser des questions sensibles.</p> <p>Les entretiens peuvent produire des informations plus approfondies et plus précises.</p> <p>Les entretiens peuvent être utilisés comme illustration pour des études de cas sur les ménages.</p> <p>Cette approche peut compléter les entretiens de groupe de façon très utile.</p>	<p>Cette approche est plus chronophage parce qu'elle nécessite un plus grand nombre d'entretiens.</p> <p>Un échantillon plus réduit peut ne pas permettre de généralisations bien qu'il puisse néanmoins servir d'illustration.</p>
Entretiens semi-structurés avec des informateurs clés	<p>Ils peuvent apporter une bonne vue d'ensemble du programme.</p> <p>Ils peuvent apporter des éclairages qui peuvent être approfondis lors de discussions en focus groups ou d'entretiens individuels.</p> <p>Suivre une liste de vérification assure une cohérence tout en permettant la flexibilité nécessaire pour approfondir une information intéressante.</p>	<p>De bonnes connaissances locales sont nécessaires pour sélectionner des informateurs clés appropriés et être conscients des biais potentiels.</p>
Enquêtes formelles	<p>Elles fournissent des statistiques utiles et des séries de données comparables.</p> <p>Les données qui en résultent sont facilement généralisables du moment où la méthode d'échantillonnage est fiable.</p> <p>Les données quantifiées peuvent être plus convaincantes pour les preneurs de décision.</p>	<p>Elles peuvent être trop chronophages et onéreuses à utiliser lorsque les ressources sont limitées (par exemple, elles demandent une conception minutieuse et des tests avant la mise en œuvre).</p> <p>Elles peuvent ne pas être faisables dans un contexte de conflit ou dans d'autres environnements imprévisibles où l'accès est limité ou encore là où les personnes manquent de confiance et sont réticentes à l'idée de s'exprimer.</p> <p>Elles autorisent peu de marge de manœuvre pour approfondir les réponses intéressantes ou inattendues.</p>

Une approche déjà testée consiste à confier les entretiens approfondis avec les bénéficiaires à un chercheur en sciences sociales. L'équipe peut ensuite identifier à partir de ce travail les sujets clés qu'elle souhaite davantage examiner lors de ses propres entretiens sur le terrain avec les bénéficiaires.

5.4.3 L'observation

L'observation directe est très utile, en particulier pour trianguler des données obtenues à partir d'autres sources. L'observation est également utile pour collecter des données primaires mais elle peut devoir être complétée par d'autres méthodes car la signification de ce que vous voyez n'est pas toujours évidente.



Emportez une petite caméra pour capturer des observations visuelles là où cela est culturellement acceptable. Les photographies sont une manière efficace de montrer au personnel du siège quelles sont les problématiques sur le terrain.



Exemple
de bonne
pratique 16

Utiliser des photographies pour étayer vos messages clés.

Durant la réponse à la crise du Kosovo, une organisation dotée d'une solide expérience dans l'eau et l'assainissement a réalisé de nombreux projets qui se trouvaient en-dessous des standards habituels. Au moment de l'introduction de la réunion de débriefing avec l'organisation, l'idée que celle-ci n'ait pas réussi à atteindre ses propres standards a été fortement rejetée par le personnel de l'organisation. Toutefois, l'atmosphère a changé lorsque l'évaluateur a utilisé une série de photos - dont l'une montrait une poche d'eau de 5 tonnes en équilibre instable et située dans un environnement avec de très nombreux enfants – pour illustrer la négligence envers les règles de sécurité et les politiques de l'organisation (Wiles et al., 2000).





Lorsque cela est possible, regardez autour de vous avant de vous engager dans des entretiens individuels ou de groupes avec la population affectée. L'évaluateur est souvent vu comme faisant partie de la machine de l'aide extérieure et comme ayant la capacité d'influencer le flux de l'aide. Dans certaines cultures, cela peut avoir un effet sur les réponses apportées. Le fait de commencer par regarder autour de vous peut vous donner des informations qui vous permettront de questionner les réponses et de tendre vers un échange plus franc.

Lorsque les évaluateurs s'attendent à voir un certain nombre de cas similaires, il peut être très utile de structurer les observations de manière à faciliter l'identification des similarités et des dissemblances entre les cas. Cela peut prendre la forme d'une simple liste de vérification.

Les évaluateurs doivent toujours être à l'affût de petits signes qui fournissent un indicateur indirect des données clés. Un exemple bien connu est celui de l'usure des tapis dans un musée qui indique quelles sont les expositions les plus visitées ([Webb et al., 1966](#)). Les évaluateurs peuvent utiliser l'absence de fils d'attente aux latrines (ou l'absence de défécations sauvages) comme un indicateur du nombre suffisant de latrines installées. Lors

d'une évaluation au Malawi, l'absence de chants de bienvenue à l'arrivée dans un village (par rapport à l'accueil en musique des autres sites) a conduit l'évaluateur à penser que les villageois n'étaient pas très satisfaits du travail de l'organisation qu'il était en train d'évaluer.

5.4.4 *Les méthodes d'évaluation participative rapide*



L'évaluation participative rapide

Série de méthodes permettant aux populations locales de renforcer, d'analyser et de partager facilement leurs connaissances et leur apprentissage avec les chercheurs venus de l'extérieur.

De nombreuses techniques d'évaluation participative rapide (EPR) sont visuelles et peuvent capturer une grande quantité d'informations dans un diagramme, une carte ou une image. D'autres sont utiles pour des mesures ou des valeurs relatives. La plupart sont accessibles à tout type de participants (y compris analphabètes). Ces techniques permettent par conséquent aux groupes et communautés de réaliser leur propre analyse, avec l'intermédiaire d'un

facilitateur et dans un format que des chercheurs venus de l'extérieur peuvent comprendre. Ces activités peuvent être intégrées de façon utile à

un entretien de groupe. Le **Tableau 19** fournit des exemples de techniques EPR qui peuvent être utilisées dans le cadre d'EAH.

Tableau 19 : Les méthodes d'évaluation participative rapide

Technique	Description	Utilisation potentielle
Calendrier	Le groupe construit un diagramme qui montre les changements au fil du temps (par exemple, dans le travail agricole, les charges de travail liées aux questions de genre ou les maladies).	<p>Un calendrier sur 24 heures pour les femmes peut montrer la quantité de temps passée à accéder aux ressources issues des secours (comme l'aide alimentaire) durant le premier stade d'une crise.</p> <p>Un calendrier saisonnier peut montrer les périodes de plus grande pénurie alimentaire qui pourraient ensuite être comparées durant l'évaluation avec les périodes de distribution alimentaire.</p>
Chronologie	Le groupe construit une chronologie des événements.	Une première chronologie peut enregistrer les événements clés ou les moments d'insécurité durant une crise liée à un conflit, et une seconde peut enregistrer les moments où l'assistance alimentaire a été apportée.
Empilement proportionnel	On donne au groupe une pile de 100 cailloux, haricots ou graines pour représenter un total lié à une catégorie donnée (comme les revenus d'un ménage) et on lui demande de diviser l'ensemble de manière à illustrer l'importance relative de différents éléments au sein de cette catégorie (comme les sources de revenus).	<p>Le fait de connaître l'importance relative de différentes sources de moyens d'existence peut indiquer l'importance relative d'une intervention d'urgence spécifique comme un programme de travail-contre-argent ou d'aide alimentaire.</p> <p>Cette technique peut aussi être utilisée pour identifier les plus pauvres au sein d'une communauté.</p>

Technique	Description	Utilisation potentielle
Le classement	On demande au groupe de classer différents items, soit les uns par rapport aux autres, soit de façon bidimensionnelle en fonction de certains critères.	Le classement peut être utilisé pour comprendre comment différents types d'assistance humanitaire (comme l'aide alimentaire, les biens non alimentaires, les semences et les outils) satisfont les besoins des bénéficiaires. S'il est fait de façon bi-dimensionnelle, il peut capturer les différents besoins d'un ménage (par exemple, des personnes âgées, des femmes, des hommes ou des enfants).
« Transect walk » (marche d'observation)	L'évaluateur marche à travers le village avec un petit groupe ou avec des informateurs clés et leur demande ce qu'il/elle observe (par exemple, qui a accès aux pâturages pour faire brouter son bétail ou qui vit dans telle maison).	Cela peut aider des évaluateurs à comprendre les différents impacts d'une catastrophe naturelle sur différents groupes d'une communauté ou différentes zones et à approfondir la réflexion sur la sensibilité de la réponse humanitaire par rapport à ces différences.
Diagramme de Venn	Les cercles qui représentent différentes catégories se recoupent là où les catégories ont une valeur en commun.	Ces diagrammes sont habituellement utilisés pour montrer différentes institutions (par exemple, pour les organisations humanitaires internationales, nationales et locales, comment elles sont interconnectées et leur importance relative dans la réponse à une crise).

*Vous pouvez trouver davantage d'informations sur l'utilisation des outils EPR dans *Methods for Community Participation* (Kumar, 2002) et en *Participatory Rapid Appraisal for Community Development* (Theis and Grady, 1991). L'approche EPR découle du travail de Robert Chambers Chambers (1994). Des ressources en ligne sur ces outils ont été mises à disposition par la Banque mondiale (Bank, 2007) et la Canadian International Development Agency (CIDA, 2005).*



Les techniques EPR ont l'air simple mais requièrent de bonnes compétences de facilitateur. Elles doivent être utilisées avec un groupe relativement homogène, comme des femmes récemment déplacées ou des habitants restés dans un village après avoir été attaqués. Il peut être utile de réaliser des exercices EPR en séparant les hommes et les femmes.



Exemple
de bonne
pratique 17

Utiliser les outils d'évaluation participative rapide

En 2001, l'évaluation du Disasters Emergency Committee concernant la réponse au tremblement de terre au Gujarat en 2001 a utilisé des outils EPR pour les recherches sur la communauté. Pour commencer, les sujets clés de recherche ont été développés à partir des critères d'évaluation d'un atelier tenu en présence des membres de l'équipe de recherche.

Pour le jeu des pairs, on demandait à chaque groupe de dessiner ou d'écrire sur des cartes les types d'aide que la communauté dans son ensemble a reçus suite au séisme. Les cartes étaient placées en ligne sur le sol. Le groupe préparait alors une seconde ligne de cartes qui montrait qui avait reçu les différentes interventions. On donnait ensuite aux membres du groupe des cartes représentant le timing (trop tard, trop tôt, au bon moment), la quantité (trop peu, trop, ok) et la qualité (trop faible, trop haute, ok) puis on leur demandait de les placer en-dessous de chaque ligne. En ce qui concerne la chronologie, on demandait au groupe d'indiquer les interventions et dans quelle mesure la communauté s'était impliquée à chaque fois, sur la chronologie de la réponse au tremblement de terre et de préciser à quel moment elle aurait aimé être impliquée dans le processus.

Pour l'exercice de classement, on demandait au groupe de lister les capacités locales qui étaient importantes au moment du tremblement de terre, à savoir les organisations, les structures, les contacts extérieurs, les réseaux, les compétences, les ressources, les connaissances et les personnes clés, et de les classer avec des pierres selon la force de chaque capacité avant, pendant et après le séisme. Après chaque exercice, les évaluateurs engageaient une discussion sur les résultats et le rôle des organisations évaluées dans tout changement ([Humanitarian Initiatives UK et al., 2001](#)).

5.4.5 Les enquêtes de terrain

Les enquêtes peuvent être utilisées pour une large gamme d'objectifs, depuis l'évaluation de la malnutrition jusqu'à l'évaluation de l'intention de retour chez les réfugiés ou encore l'efficacité des messages d'éducation à la santé. [Fowler \(2009\)](#) constitue une introduction brève et intéressante aux enquêtes. Ce chapitre traite principalement des enquêtes structurées réalisées par un enquêteur en face-à-face. Le chapitre [5.3.8](#) traite des enquêtes en ligne. Il existe également un certain nombre de techniques d'enquêtes spécialisées (par exemple, pour la nutrition) qui ne sont pas abordées ici. Les enquêtes de type entretien téléphonique assisté par

ordinateur ne sont en général pas utilisées dans le cadre des évaluations humanitaires.



Les enquêtes peuvent fournir des données utiles quant aux perceptions que les bénéficiaires ont de l'aide. Pensez toujours à réaliser une enquête auprès des bénéficiaires si vous voulez savoir quelles sont leurs perceptions générales à propos de l'assistance.

Les enquêtes sont complexes et nécessitent une gestion attentive. Elles sont l'une des principales méthodes quantitatives d'évaluation. Leur force provient du fait que si un échantillon aléatoire est pris au sein d'une population donnée, il est susceptible d'être représentatif de l'ensemble de la population.

Les enquêtes structurées en face-à-face sont rarement utilisées dans l'évaluation humanitaire pour un certain nombre de raisons :

- **Coût** – une enquête gérée de façon professionnelle peut coûter entre 20,000 US\$ et 100.000 US\$. On retrouve le plus souvent de telles enquêtes dans des évaluations humanitaires conjointes disposant d'importantes ressources comme les évaluations de la Tsunami Evaluation Coalition ([Telford and Cosgrave, 2007](#))¹ ou l'évaluation de suivi de 2010 ([Brusset et al., 2009](#)).
- **Temps** – développer et tester un outil d'enquête, former les enquêteurs, réaliser l'enquête et traiter les résultats prend plus de temps que d'autres méthodes d'évaluation^[8].
- **Absence d'un bon cadre d'échantillonnage** (une liste qui donne une estimation de la population réelle) à partir duquel établir un échantillon aléatoire : les enquêtes dans le contexte humanitaire utilisent parfois un échantillonnage pseudo-

aléatoire pour surmonter ce problème.

- **Absence d'enquêtes sur les seuils de référence** – les enquêtes sont plus utiles lorsqu'il existe un seuil de référence à partir duquel comparer les résultats. C'est rarement le cas dans les évaluations humanitaires.
- **Absence d'évaluateurs présentant les qualifications requises** – la plupart des évaluateurs humanitaires ont une formation en sciences sociales qualitatives et n'ont pas beaucoup d'expérience dans le travail d'enquête formelle.

La prochaine étape consiste à concevoir votre outil d'enquête. (Vous devriez normalement commencer par cela mais étant donné les difficultés à trouver des cadres d'échantillonnage convenables dans les contextes humanitaires, il vaut mieux commencer par le cadre d'échantillonnage car cela peut influencer l'approche de l'enquête). Les questions de l'enquête, bien que plus simples par certains aspects que les questions d'entretiens semi-structurés, sont par ailleurs plus complexes puisque l'intervieweur ne peut pas ajuster la question et réaliser des questions de suivi comme cela est possible avec les entretiens semi-structurés. Les enquêtes structurées en face-à-face partent du principe que les questions seront posées de la

⁸ L'enquête est souvent la méthode de collecte de données la plus lente. La première version de l'évaluation du Disasters Emergency Committee concernant sa réponse au tsunami, qui était basée sur des méthodes qualitatives, a dû être révisée lorsque les données issues de l'enquête quantitative des bénéficiaires ont été disponibles (Tony Vaux, communication personnelle).

même manière à chaque personne interrogée. Les questions d'une enquête doivent être conçues avec précaution puis testées avant d'être utilisées car elles ne peuvent pas être ajustées une fois que l'enquête est en cours.

On teste habituellement les questions d'une enquête en réalisant une enquête pilote. La manière dont les questions sont posées peut avoir un impact important sur les réponses et une enquête pilote peut mettre en évidence certains de ces points. La monographie de [Fowler's \(1995\)](#) sur les questions d'enquête est une ressource utile.



Si l'enquête doit être traduite, faites retraduire la traduction dans la langue originale par un autre traducteur.

Le nombre d'enquêteurs requis dépend de la taille et de la répartition de l'échantillon ainsi que de la stratégie d'échantillonnage. La taille de l'échantillon dont vous avez besoin dépend d'un certain nombre de facteurs parmi lesquels :

- *Dans quelle mesure le sujet de votre étude est-il répandu au sein de la population ? (Par exemple, si vous enquêtez sur l'intention de*

retour et que seulement 5 % de la population souhaite rentrer, il vous faudra une taille d'échantillonnage bien plus grande pour tirer des conclusions valides du point de vue statistique que pour une situation dans laquelle 75 % de la population souhaite rentrer.)

- *Quel niveau d'erreur êtes-vous prêt à accepter ? Plus le niveau acceptable est petit, plus la taille de l'échantillon doit être importante. Les conceptions les plus courantes adoptent 5% de possibilité de réponses positives erronées (quand le résultat dit que l'assistance était efficace alors qu'elle ne l'était pas) et 20% de possibilité de réponses négatives erronées (quand le résultat dit que l'assistance n'était pas efficace alors qu'elle l'était en réalité).*

Testez toujours votre outil d'enquête avant de l'utiliser. Pensez à recruter une société d'études de marché locale pour réaliser l'enquête. Certaines organisations peuvent conduire des enquêtes pour leurs propres intérêts susceptibles d'être utiles dans le cadre de votre évaluation.

Encadré 8 : Prendre en compte le sexe et l'âge durant la collecte des données sur le terrain

Les étapes suivantes vous aideront à assurer que votre évaluation est sensible aux questions de genre et d'âge et produit des données qui pourront être désagrégées par sexe et par âge :

- Équilibrez bien vos équipes de travail de terrain au niveau du ratio hommes/femmes. Les intervieweuses et les traductrices ont habituellement un meilleur accès aux femmes et aux jeunes filles au sein des populations affectées.
- Assurez-vous que vos informateurs clés comprennent des femmes qui connaissent bien la communauté et les besoins particuliers des femmes et dans quelles mesures ceux-ci ont été satisfaits, si c'est le cas. Ces informatrices clés peuvent être des enseignantes, des vendeuses qui occupent un rôle important dans le marché local et des infirmières. Assurez-vous que certains informateurs clés connaissent également bien les besoins des enfants, des jeunes et des personnes âgées.
- Assurez-vous qu'au moins la moitié des participants des focus groups et des entretiens de groupe sont des femmes. Dans l'idéal, les femmes doivent être interrogées à l'écart des hommes qui tendent à dominer les discussions mixtes.
- Organisez des focus groups et des entretiens de groupe par tranches d'âge, par exemple avec des enfants, des jeunes et des personnes âgées. Les adultes tendent à dominer les discussions où se mélangent tous les âges.
- Tenez un registre du genre des informateurs clés et des autres personnes interrogées de manière à être conscient de tout biais potentiel lié au genre dans les informations que vous rassemblerez à partir des entretiens.

Source: Basado en parte en [Mazurana et al \(2011\)](#).

Après avoir collecté les données de votre enquête, vous devez les saisir et les nettoyer. Cela peut être onéreux et chronophage, en particulier si l'enquête contient des questions ouvertes.



Utiliser des aides électroniques pour collecter des réponses dans le cadre d'enquêtes peut presque éliminer le coût de saisie et de nettoyage des données car la forme de l'enquête peut être déterminée de manière à rejeter les combinaisons de réponses non valides. De plus, les données collectées peuvent être simplement chargées dans la principale base de données sans aucune autre saisie. Cela peut enfin accélérer le processus d'enquête

Les outils numériques peuvent être utilisés pour des enquêtes de diagnostic aussi bien que pour des enquêtes d'évaluation. Le mécanisme de diagnostic rapide multi-Clusters au Pakistan a utilisé des assistants numériques personnels (PDA en anglais) pour rassembler des données de diagnostic ([McRAM Team, 2009](#)). Shirima et al. fournissent un exemple d'utilisation de PDA dans le cadre d'une enquête auprès de 21.000 ménages en Tanzanie et apportent quelques éclairages sur les questions pratiques que cela implique ([Shirima et al., 2007](#)). Une étude récente suggère que le fait d'utiliser des

PDA pourrait réduire la charge logistique et le taux d'erreurs lors de la collecte des données ([Seebregts et al., 2009](#)).

Une nouvelle évolution est l'utilisation croissante des téléphones portables pour la collecte de données. Ceux-ci présentent l'avantage que les données sont immédiatement enregistrées, prévenant ainsi le risque de perte des données si le téléphone est endommagé ou perdu. Cela permet également un contrôle qualité en temps réel et une supervision des collecteurs de données afin de réduire les risques de fabrication des données ([Tomlinson et al., 2009](#)).

Les téléphones portables offrent des avantages semblables à ceux des PDA en termes de réduction des erreurs d'enregistrement et de saisie ([Zhang et al., 2012](#)). Plusieurs logiciels sont aujourd'hui disponibles pour réaliser des enquêtes à partir de téléphones portables.

5.5 Les approches tournées vers l'apprentissage

Les évaluations tournées vers l'apprentissage sont en général plus participatives et doivent impliquer ceux qui ont besoin ou envie d'apprendre. Les méthodes doivent encourager la réflexion et la reconnaissance des erreurs aussi bien que des réussites (voir Chapitre 3.1). Ce chapitre présente un certain nombre d'approches testées et analysées en matière d'EAH tournées vers l'apprentissage.

Dans la culture militante des opérations humanitaires, on dit souvent ne pas disposer de suffisamment de temps pour les exercices participatifs et réflexifs. Pourtant, seule une faible portion de l'action humanitaire est soumise à une pression de temps telle qu'elle serait pénalisée si elle consacrait du temps à l'évaluation. Si elles sont bien planifiées et animées, de nombreuses activités tournées vers l'apprentissage peuvent être accomplies en quelques heures ou en une journée, et la richesse née de la réflexion et de l'apprentissage collectifs justifie l'intérêt en temps.

Au moment de concevoir une évaluation tournée vers l'apprentissage, identifiez en premier lieu la cible de l'apprentissage (par exemple, des membres

particuliers du personnel, des partenaires ou encore de membres de la population affectée). Concevez ensuite les processus qui permettent à ces groupes de personnes de réfléchir.

Voici une liste de ces processus:

- *Créez un espace sûr pour la réflexion et la discussion où les participants sentent qu'ils peuvent parler librement et reconnaître leurs erreurs.*
- *Poser des questions simples et ouvertes est en général le moyen le plus efficace d'encourager la réflexion (voir 'Poser des questions et écouter' in [Chapman et al., \(2005\)](#)).*
- *Utilisez des processus participatifs et créatifs.*
- *Assurez-vous que les leçons apprises sont bien documentées.*
- *Soyez particulièrement attentif à la qualité de l'animation.*
- *Encouragez ceux qui ont participé à la réflexion et au processus d'apprentissage à s'approprier les résultats de l'évaluation.*
- *Élaborez un processus de suivi.*

5.5.1 Les revues après action



Revue après action

Echange structuré à propos d'une intervention humanitaire qui permet à une équipe de réfléchir à ce qui s'est passé, à pourquoi cela s'est passé et sur la manière de conserver les points forts et d'améliorer les faiblesses.

Une revue après action est habituellement un processus qui demande une animation dans le cadre d'un atelier et implique tous ceux qui ont été identifiés comme la cible de l'apprentissage. Une atmosphère ouverte qui renforce la confiance entre les participants est essentielle pour réussir sa revue après action. Le principe général est le suivant : « pas d'attribution, pas de rétribution ». Un animateur neutre et objectif est essentiel pour assurer que la discussion reste focalisée sur les sujets à traiter, qu'elle reste positive et ne dégénère pas en auto-justification ou en accusation. La revue doit traiter les questions suivantes :

- *Quels étaient les objectifs ou les intentions de l'action ?*
- *Qu'est-ce qui s'est réellement produit ?*

- *Qu'est-ce qui s'est bien passé ?*
- *Qu'est-ce qui aurait pu mieux se passer ?*
- *Que feriez-vous différemment la prochaine fois ?*

Une revue après action comprend les étapes suivantes. A chaque étape, des participants peuvent écrire leurs réponses sur des post-it, l'animateur les regroupant par thèmes pour réfléchir sur certains points spécifiques, ou alors le groupe peut répondre à chaque question de façon collective, l'animateur écrivant les réponses sur un paper-board.

1. Les participants, seuls ou en binômes, mettent par écrit ce qu'ils comprennent de l'objectif ou de l'intention de l'action. (Parce que l'action humanitaire a souvent lieu dans des environnements chaotiques et que les plans peuvent rapidement devenir obsolètes, cette étape est une manière utile de découvrir si l'objectif est clair et partagé, et s'il est nécessaire de le redéfinir).
2. Les participants écrivent ensuite ce qui s'est réellement passé, en travaillant éventuellement en un seul groupe pour construire une chronologie des événements clés et des changements au fil du temps, soit pour la situation, soit pour le programme.

3. Le groupe répond ensuite à trois questions : Qu'est-ce qui s'est bien passé ? Qu'est-ce qui aurait pu mieux se passer ? Qu'est-ce que vous feriez d'une autre manière la prochaine fois ?
4. Cette étape est suivie par une discussion à la fin de laquelle l'animateur résume tous les enseignements qui ont émergé et demande aux participants de voter pour ce qu'ils considèrent comme les trois plus importantes leçons tirées.
5. Les leçons apprises clés et toutes les recommandations exploitables qui émergent sont documentées et communiquées à tous les participants. Il est possible de se mettre d'accord sur un cadre temporel, par exemple de 6 à 12 mois, pour évaluer la progression de la mise en œuvre des recommandations.

Plusieurs guides de conduite des revues après action sont disponibles, notamment le [guide d'USAID](#), et le Food Security Information for Action's After Action Reviews and Retrospects. Le personnel sur le terrain apprécie l'opportunité de refléter et consolider son apprentissage que lui fournit une revue après action.



Le personnel de terrain profite souvent de l'opportunité qu'offre une revue après action pour réfléchir sur l'apprentissage et le consolider.

5.5.2 Le "storytelling"

Raconter des histoires est une manière naturelle de communiquer que la plupart d'entre nous utilise. Se focaliser sur ce processus et l'animer peut être une façon créative de faciliter l'apprentissage, en particulier lorsque cela s'accompagne d'un processus de questionnements et de réflexions pour approfondir l'analyse. Ce chapitre décrit deux manières d'utiliser le *storytelling* dans des objectifs d'apprentissage.

UTILISER DES METAPHORES

Le meilleur moyen d'utiliser des métaphores est lors d'un atelier, par exemple en respectant les étapes suivantes :

Les participants, individuellement ou en groupes, dessinent des images (par exemple d'un arbre ou d'une rivière) pour représenter leur expérience du programme en cours d'évaluation. Les différentes

parties de l'arbre ou de la rivière représentent les différentes parties du programme.

1. Les participants racontent à leurs collègues une histoire à propos de leur dessin, sans être interrompus.
2. Ceux qui écoutent posent des questions sur l'histoire de manière à en approfondir l'analyse et l'apprentissage ainsi que pour proposer leurs réflexions.
3. A la fin de ce processus, le groupe liste les principales réalisations, difficultés et leçons apprises à partir de l'histoire pour leur expérience combinée du programme.
4. Cette étape peut se focaliser sur l'avenir en posant les questions suivantes à la fin :
 - a. Comment peut-on s'appuyer sur ce qui s'est bien passé ?
 - b. Comment peut-on traiter les difficultés ?
 - c. Comment appliquerons-nous ce que nous avons appris ici ?



Exemple
de bonne
pratique 18

Utiliser des métaphores et le “storytelling” dans un rapport

La revisión de Tearfund sobre el fortalecimiento de las Le rapport qu'a réalisé Tearfund en matière de renforcement des capacités liées à la gestion des catastrophes naturelles a utilisé la métaphore de l'arbre comme base d'un *storytelling*. Lors d'un atelier de suivi et d'apprentissage, il a été demandé à chaque partenaire de Tearfund de dessiner un arbre.

- Les racines représentaient les valeurs et principes qui sous-tendaient le programme.
- Le tronc représentait les partenaires (organisations et individus).
- Les branches représentaient les activités.
- Le fruit représentait les réalisations du programme.
- Les feuilles représentaient les leçons apprises.
- Les branches cassées sur le sol représentaient les difficultés internes.
- Les nuages dans le ciel représentaient les difficultés extérieures.
- Les bourgeons représentaient les activités planifiées mais pas encore mises en œuvre

Le *storytelling* a été suivi par une analyse d'apprentissage. Malgré un certain scepticisme initial, les participants ont dit qu'ils avaient trouvé que ce processus participatif à la fois visuel et créatif avait encouragé la réflexion et l'apprentissage ; avait amené à la prise de conscience collective des réalisations, des difficultés et des leçons apprises ; et avait donné un instantané de l'ensemble du programme. Ils ont aussi reconnu la difficulté de traduire l'apprentissage en action

LA TECHNIQUE DU CHANGEMENT LE PLUS IMPORTANT

La technique du changement le plus important est une approche participative qui peut être utilisée avec les membres du personnel ou de la communauté locale. Lorsqu'elle est utilisée avec la population affectée, elle peut être focalisée sur les résultats ou l'impact. Elle s'appuie sur le niveau du terrain de façon ascendante en utilisant des histoires qui enregistrent ce qui s'est passé, quand, pourquoi et pour quelles raisons cela est important. Cette technique a été utilisée lors de l'évaluation de la réponse de la FICR au tremblement de terre de 2007 au Pérou ; on a alors demandé aux participants d'identifier les changements les plus importants qui s'étaient produits comme résultat de l'intervention de

la Croix-Rouge péruvienne ([Martinez, 2009](#)).

Les étapes étaient les suivantes :

1. Les histoires sont collectées au niveau du terrain en réponse à deux questions : durant le dernier mois [ou une autre période de temps], à votre avis, quel fut le changement le plus important qui a eu lieu dans les vies des personnes qui ont participé au projet ? Pourquoi cela a-t-il été le changement le plus important ?
2. Les plus importantes de ces histoires sont sélectionnées par un panel de parties prenantes ou de membres du personnel.
3. Une fois que les changements ont été enregistrés, les histoires sont lues à haute voix, souvent dans le cadre d'un atelier, de manière à ce que les participants discutent et réfléchissent sur la valeur des changements rapportés.

Vous trouverez de nombreux conseils sur cette technique dans [Davies and Dart \(2005\)](#).

5.5.3 L'enquête d'encouragement ("appreciative inquiry")



L'enquête d'encouragement

Méthode qui cherche à renouveler, développer et s'appuyer sur ce qui s'est bien passé lors d'une intervention en partant de l'hypothèse que les questions tendent à focaliser l'attention dans une direction particulière.

L'enquête d'encouragement se concentre sur les solutions plutôt que sur les problèmes. Elle se concentre sur ce qui fonctionne, en partant de l'hypothèse que les questions sur les forces et les faiblesses sont plus motivantes, illustrent des possibilités et peuvent même aboutir à des transformations. En tant que processus participatif, elle engage les parties prenantes clés et peut les aider à stimuler leur réflexion sur l'évaluation ainsi que leur sentiment d'appropriation. Lorsqu'il existe une crainte et un sentiment négatif vis-à-vis des évaluations, cette approche peut se révéler énergisante, construire de la confiance entre les évaluateurs et ceux dont le travail est en cours d'évaluation, et impliquer les participants réticents.

Durant la phase de planification d'une évaluation, l'enquête d'encouragement peut être utilisée pour construire le modèle logique du programme et pour identifier les questions d'évaluation via un processus participatif, en général un atelier avec animateur durant lequel les participants sont guidés pour se poser des questions entre eux et se mettre d'accord en groupe sur les questions clés de l'évaluation en respectant les étapes suivantes :

1. *Interrogez* les participants sur leurs bonnes pratiques, des exemples de réussite et les valeurs qui sous-tendent leur travail et le programme. Cela peut être fait en demandant aux participants de se poser des questions les uns aux autres.
2. *Demandez* aux participants de se forger pour l'avenir une vision du programme qui soit enracinée dans leurs expériences positives passées.
3. *Identifier* des questions d'évaluation en demandant aux participants : A quelles questions l'évaluation doit-elle répondre pour que vous obteniez les informations nécessaires vous permettant d'avancer vers vos objectifs ?

Durant la phase de collecte des données d'une évaluation, une enquête d'encouragement peut être utilisée pour explorer les exemples de réussites et ce qui peut en être appris, de même que des valeurs et des souhaits. Des questions sur la vision peuvent explorer la manière dont le programme pourrait accéder à un plus haut niveau de performance. Cela contribuera grandement à l'apprentissage dans la mesure où cela est réalisé sous la forme d'un entretien de groupe.

Preskill and Catsambas (2006) analysent en détail l'enquête d'encouragement et donnent des exemples de son utilisation dans des évaluations réalisées dans différents secteurs.



Pensez à utiliser l'enquête d'encouragement avec des personnes qui peuvent être sensibles à la critique.

NOTES

Partagez vos commentaires:

Aidez-nous à rendre le Guide plus pratique et focalisé sur l'utilisateur en partageant vos commentaires sur son contenu et la navigation. Les commentaires peuvent inclure des questions, des suggestions, des ressources utiles ainsi que des conseils pratiques. Vous pouvez partager vos idées en cliquant sur le bouton de commentaires en bas de chaque page. Ceci générera un courriel confidentiel à l'équipe EAH. Pour plus d'informations sur la façon dont vous, votre équipe ou votre organisation peut s'engager plus activement dans le pilotage du Guide contactez-nous directement à eha@alnap.org

EVALUATIONS A DISTANCE, CONJOINTES, EN TEMPS REEL ET D'IMPACT

Contenu de ce chapitre

6.1	L'évaluation lorsque l'accès est limité	196
6.2	Les évaluations conjointes	198
6.3	Les évaluations en temps réel	202
6.4	Les évaluations d'impact	209

Comment utiliser ce chapitre

Ce chapitre présente les différents types d'évaluations utilisées dans le cadre de l'EAH : évaluations où l'accès est limité (Chapitre 6.1), évaluations conjointes (Chapitre 6.2), évaluations en temps réel 6.3), évaluations d'impact (Chapitre 6.4). Ces EAH de pointe méritent d'être étudiées plus en détail afin d'établir les meilleures approches et méthodes.

Tableaux, schémas et encadrés

Tableau 20:	Les méthodes alternatives à utiliser dans des situations d'accès limité	186
Tableau 21:	Les caractéristiques d'une évaluation en temps réel	193
Schema 16:	Aspects prospectifs et rétrospectifs des évaluations en temps réel	195
Tableau 22:	Les méthodes utilisées dans le cadre des évaluations en temps réel	197



6.1 L'évaluation lorsque l'accès est limité

La plupart des méthodes de travail sur le terrain des EAH requièrent que les évaluateurs soient présents sur le terrain pour observer et conduire des entretiens et des exercices. Cela peut s'avérer problématique lorsque des questions de sécurité empêchent l'accès des évaluateurs régionaux ou internationaux.



Demandez-vous si certains évaluateurs peuvent encourir moins de risques, en raison par exemple de leur nationalité ou de leur ethnie. Cependant, gardez à l'esprit que vous avez toujours le devoir de garantir le bien-être des évaluateurs.

Les méthodes utilisées par une équipe d'évaluation peut être affectée par des contraintes de sécurité ou d'accès. La liste suivante présente ces méthodes, classées par ordre de difficulté d'accès selon les situations :

1. *Recherche bibliographique*
2. *Analyse de données numériques*
3. *Entretiens avec des informateurs clés (au siège)*

4. *Entretiens avec des informateurs clés (sur le terrain)*
5. *Entretiens avec des bénéficiaires*
6. *Observation*
7. *Groupes de discussion avec les bénéficiaires*
8. *Enquêtes auprès des bénéficiaires*
9. *Évaluation participative rapide.*

Lorsqu'il est dangereux pour les évaluateurs internationaux de travailler, l'une des solutions peut être d'engager des consultants locaux. Cependant, dans le cadre d'urgences chroniques complexes, tout le monde, y compris les consultants locaux, peut être perçu par les populations comme manquant d'objectivité, ou être associé avec tel ou tel groupe. Les consultants internationaux peuvent encourir des risques lors de conflits internationalisés du fait d'être associés à l'une des parties.



Lors de la constitution de l'équipe d'évaluation, veillez à faire attention à la manière dont les consultants nationaux ou internationaux peuvent être perçus.

Le **Tableau 20** liste les alternatives à l'envoi d'évaluateurs sur le terrain.

Tableau 20 : Les méthodes alternatives à utiliser dans des situations d'accès limité

Méthode	Alternatives
Entretiens avec des informateurs clés sur le terrain	Entretiens via téléphone satellite Entretiens dans le cadre d'un événement extérieur auquel les informateurs participent.
Entretiens avec des bénéficiaires et des groupes de discussion	Entretiens et discussions avec des groupes qui se sont déplacés dans des zones accessibles pour des formations ou autres événements, ou qui se rendent régulièrement dans des zones accessibles (comme par exemple les marchés en villes), ou à l'entrée de zones sécurisées.
Observations	Imagerie satellite, vidéos et photographies réalisées grâce à des appareils utilisant un système de positionnement géographique (GPS) intégré.
Enquêtes auprès des bénéficiaires	Enquêtes réalisées en ligne et par SMS (ces derniers ont tendance à être sujets à un biais d'auto-sélection et doivent être interprétés avec précaution)
Outils d'évaluation participative rapide	Utilisation des outils à travers un processus de diffusion communautaire (pour cela, l'échelle de temps excède souvent l'échelle de temps des évaluations humanitaires)

Les questions de sécurité qui limitent l'accès pour les évaluations peuvent également gêner le suivi des projets, ce qui a pour conséquence de laisser l'équipe d'évaluation avec peu de données de suivi fiables. L'évaluation du IASC en Somalie centrale et du sud a identifié 11 leçons pour un suivi de bonne qualité dans de telles circonstances ([Polastro et al., 2011a, p. 28](#)). La triangulation – en particulier de données issues d'une zone inaccessible par rapport à des données extérieures à la zone – est critique.

6.2 Les évaluations conjointes

Les évaluations conjointes sont de plus en plus fréquentes dans le secteur de l'aide humanitaire, ce qui reflète une tendance plus globale vers des approches conjointes⁹. Les évaluations conjointes peuvent se rencontrer au niveau des circuits de financement multi-bailleurs, comme par exemple pour l'évaluation de soutien à la consolidation de la paix au Sud Soudan ([Bennett et al., 2010](#)); au niveau des consortiums d'ONG, comme par exemple l'évaluation multi-ONG de la réponse au tremblement de terre de Yogyakarta ([Wilson et al., 2007](#)) et l'évaluation ACT de la réponse au

tremblement de terre en Haïti ([McGearty et al., 2012](#)); et au niveau du concept de programme transformateur des Nations Unies comme avec les évaluations en temps réel du IASC ([Cosgrave et al., 2010](#); [Grünewald et al., 2010](#); [Polastro et al., 2011a](#));

Il est également possible de les retrouver dans les évaluations de niveau systémique, comme dans le cas du tsunami ([Telford et al., 2006](#)). Une évaluation conjointe peut être entreprise soit dans un but de redevabilité, soit dans un but d'apprentissage.

6.2.1 Les bénéfices d'une évaluation conjointe

Les raisons pour entreprendre une évaluation conjointe sont notamment les suivantes (voir [Breier, 2005](#)):

- *Pour avoir une vision globale et évaluer en se concentrant davantage sur le niveau politique.*
- *Pour éviter les défis que pose l'attribution d'un impact à une organisation en particulier.*
- *Pour des raisons stratégiques (les évaluations conjointes bénéficient en général d'une plus grande crédibilité et peuvent s'avérer utiles pour promouvoir le changement).*
- *Pour aider les organisations participantes à mieux comprendre leurs approches respectives*

⁹ Cette section est en partie basée sur [Beck and Buchanan-Smith \(2008\)](#).

et à échanger des bonnes pratiques.

- Pour des raisons administratives et financières, par exemple en mettant en commun les ressources et capacités d'évaluation
- Pour réduire les coûts de transaction des pays en développement (une évaluation conjointe peut réduire le nombre d'évaluations d'organisations individuelles consultant toutes les mêmes parties prenantes).

Les inconvénients d'une évaluation conjointe sont notamment la complexité du processus, qui peut transformer l'évaluation en un projet chronophage avec des coûts de transaction élevés et des structure managériales complexes.



Les questions clés à se poser lorsque vous décidez de lancer une évaluation conjointe sont les suivantes :

- Quels sont le périmètre et l'objectif de cet exercice ?
- Quelles sont les questions clés à aborder ?
- Une collaboration est-elle l'approche la plus adaptée ?

Même si l'objectif affiché des EAH conjointes relève en théorie à la fois de la redevabilité et de

l'apprentissage, dans la pratique, les organisations participantes (par exemple celles de la Tsunami Evaluation Coalition), ont souvent constaté qu'elles étaient plus utiles à l'apprentissage.

6.2.2 Comment planifier et gérer une évaluation conjointe

Se focaliser sur l'utilisation demande d'impliquer les principaux utilisateurs attendus de l'évaluation dans la phase de planification qui s'achève avec les Termes de Référence. Lors d'une évaluation conjointe, il existe de nombreux utilisateurs de premier niveau et, par conséquent, la négociation des Termes de référence est susceptible de prendre beaucoup plus de temps. Il s'agit d'une étape essentielle qui mérite qu'on la planifie avec du temps supplémentaire plutôt qu'on ne la raccourcisse. Bien entendu, il peut s'agir de l'une des raisons pour lesquelles les Termes de Référence des évaluations conjointes sont en général mieux élaborés que ceux d'évaluations d'organisations individuelles (voir [Beck and Buchanan-Smith, 2008](#)). Une direction forte et des aptitudes en matière de facilitation sont essentielles pour cette phase de planification initiale afin de gérer les négociations entre de multiples partenaires et de s'assurer que les Termes de référence finaux sont clairs et concis

et non énumérés sous la forme d'une longue liste de questions non priorisées.

Accordez une attention toute particulière au système de gestion (voir Chapitre 3.7). Deux options ont été testées pour les évaluations conjointes :

- **Système de gestion à deux niveaux** – *plus grand groupe de pilotage qui rassemble toutes les organisations parties prenantes et un groupe de gestion plus petit pour gérer le processus et l'équipe d'évaluation au jour le jour. Ce dernier peut aussi servir de tampon entre les organisations parties prenantes et l'équipe d'évaluation si les résultats de l'évaluation sont particulièrement controversés (voir [Dabelstein, 1996](#)).*
- **Système de gestion à trois niveaux** – *un groupe de pilotage et un groupe de gestion comme décrit ci-dessus (en général basés au siège) avec un groupe de référence additionnel dans le pays ou un groupe de pilotage chargé de construire l'appropriation dans le pays où se déroule l'évaluation.*



Les structures de gestion complexes et les hauts coûts de transactions liés à une évaluation conjointe peuvent être un obstacle à l'implication des ONG, en particulier des ONG nationales et de petite taille. Le processus d'identification des options d'engagement (qui est moins exigeant que pour les parties prenantes plus importantes) peut aider à garantir leur implication.



Le nombre d'évaluateurs possédant les compétences techniques, politiques et interpersonnelles nécessaires pour être à la tête d'évaluations conjointes demeure faible. Une planification et un processus d'appel d'offre effectués le plus tôt possible permettront de s'assurer que les candidats sont qualifiés.

6.2.3 Quelles approches et méthodes utiliser ?

Pour une évaluation conjointe de grande échelle, il peut être pertinent d'adopter une approche en deux phases où la première phase ressemblerait à une phase initiale étendue. Lors de cette phase, l'équipe d'évaluation définit le périmètre de l'évaluation (par exemple, en organisant des réunions et des consultations préliminaires avec les organisations parties prenantes), une revue bibliographique et une visite préliminaire dans le pays ou la région où le travail sur le terrain va s'effectuer. Cela peut aider

à bien orienter l'évaluation. Durant cette phase, l'équipe doit également élaborer la méthodologie complète qui devra être revue et approuvée par le groupe de pilotage. Lors de la deuxième phase, la consultation complète et le travail sur le terrain sont effectués, les données analysées, et les résultats rédigés et présentés.

Les méthodes utilisées lors d'une évaluation conjointe sont souvent similaires à celles d'une évaluation d'une seule organisation. Cependant, davantage de ressources signifie davantage d'opportunités de consulter les populations affectées, en combinant par exemple méthodes qualitatives (groupes de discussion) et méthodes quantitatives (enquêtes sur le terrain). Ces méthodes doivent être soigneusement planifiées et segmentées, en particulier lorsque des méthodes mixtes sont utilisées.

L'OCDE/CAD a publié un guide sur la gestion des évaluations conjointe ([OECD/DAC, 2006](#)), qui sont également au centre d'un chapitre de la septième revue ALNAP de l'action humanitaire ([Beck and Buchanan-Smith, 2008](#)).

6.2.4 *Comment assurer le suivi*

Le nombre élevé de parties prenantes aux attentes différentes dans le cadre d'une évaluation conjointe peut rendre le suivi plus difficile qu'il ne l'est pour l'évaluation d'une seule organisation – par exemple, si les recommandations sont génériques et que les organisations choisissent celles qui les intéressent parmi elles. Il existe des moyens de contourner cette difficulté, par exemple en s'assurant que toutes les recommandations sont clairement ciblées sur des groupes d'utilisateurs, voire sur des organisations individuelles. L'adhésion au niveau national, depuis le début du processus d'évaluation, est essentielle pour l'adoption des résultats. Il est donc crucial de bien planifier en avance la diffusion et de la doter de ressources conséquentes. Ce processus doit aller bien au-delà de la distribution du rapport : pensez à organiser des ateliers pour encourager la réflexion et les discussions autour des résultats de l'évaluation.



Exemple
de bonne
pratique 19

S'impliquer au pays

Une initiative sectorielle a abouti à un certain nombre d'évaluations de santé inter-agences, impliquant des agences onusiennes, des ONG, des universitaires et des gouvernements bénéficiaires..

Apprendre de l'expérience d'évaluation conjointe rapide a conduit à l'IHE adapter son approche pour atteindre une plus grande participation au niveau des pays, la création de comités de direction et de faire une visite préliminaire par les membres du groupe de travail de base de l'IHE. Un an après la fin de l'évaluation, a organisé un atelier de planification dans le pays, dirigé par le groupe de pilotage pour discuter de suivi avec toutes les parties intéressées. Les évaluateurs retournés à participer à cet atelier (voir [Beck and Buchanan-Smith, 2008](#))

Comme pour les autres évaluations, une matrice de réponse de la direction peut aider à s'assurer que les recommandations de l'évaluation sont suivies. Chaque organisation impliquée peut préparer sa propre réponse en termes de gestion, mais il est possible d'avoir une réponse conjointe.

6.3 Les évaluations en temps réel

Dans une évaluation en temps réel (ETR), l'objectif principal est de fournir, de manière participative et durant le travail de terrain, des retours à tous ceux qui sont impliqués dans l'exécution et la gestion de l'intervention humanitaire ([Cosgrave et al., 2009b](#)). Une augmentation significative des ETR a été observée ces dernières années. Cependant, nombre d'entre elles ont eu lieu bien après le début de l'intervention et pourraient être plutôt qualifiées d'évaluations à mi parcours.

6.3.1 Les avantages

Les ETR présentent trois avantages potentiels : opportunité/rapidité (« timeliness »), interactivité et perspective.

- **Opportunité/rapidité** – *ETR sont réalisées lorsque des décisions opérationnelles et politiques clés sont prises. Elles peuvent signaler des problèmes importants négligés au moment de répondre dans la précipitation aux besoins les plus urgents. Par exemple, l'ETR du Catholic Relief Service (CRS) au Pakistan en 2010 s'est déroulée seulement 9 mois après le début de la réponse de l'organisation et comprenait un jour de réflexion avec l'équipe et les partenaires dans divers lieux, journée*

durant laquelle des plans d'action immédiats ont été élaborés. (Hagens and Ishida, 2010).

- **Interactivité** – s ETR impliquent l'instauration d'un dialogue durable avec les membres de l'équipe, à la fois sur le terrain et au siège, et peuvent servir de canal de communication entre les personnels terrain et siège pour contourner les blocages bureaucratiques. Par exemple, l'ETR de 2010 portant sur l'intervention du IASC suite aux inondations au Pakistan comprenait une seconde visite durant laquelle les résultats, les conclusions et recommandations étaient débattus de manière participative avec les parties prenante (Polastro et al., 2011b).
- **Perspective** – une équipe d'ETR peut aborder une urgence en utilisant de nombreux angles différents, discutant avec le personnel à tous les échelons, dans différents pays, avec les populations affectées, ainsi qu'avec les partenaires et les représentants du gouvernement. Cela peut apporter une vision de l'opération moins influencée par la résolution de problèmes quotidiens.

6.3.2 Caractéristiques clés

Les ETR sont différentes des autres évaluations humanitaires en raison de leur timing, des principales parties prenantes et de leur objectif (Tableau 21).

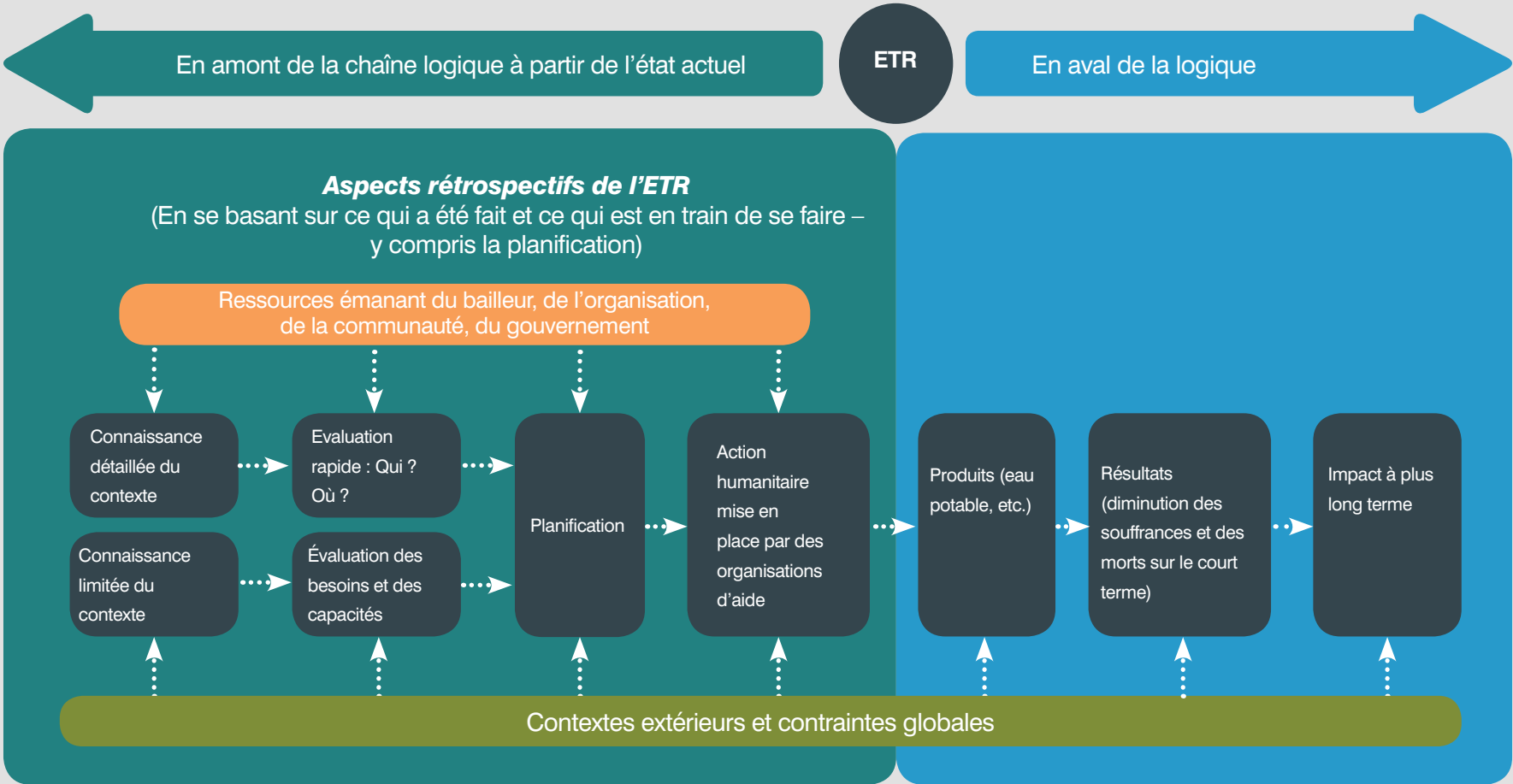
Tableau 21 : Les caractéristiques d'une évaluation en temps réel

Caractéristique	Description
Timing	<p>Les ETR se déroulent tôt dans une opération. Cela nécessite que les membres de l'équipe d'évaluation évaluent non seulement ce qui a été fait, mais se penchent également sur les conséquences potentielles de ce qui est en train d'être fait. Les ETR comportent donc des éléments qui regardent à la fois vers le futur et vers le passé (Schéma 16). Les membres de l'équipe doivent avoir suffisamment d'expérience dans le domaine opérationnel pour comprendre quels sont les probables produits, résultats et impacts des actions en cours.</p> <p>Astuce : <i>Vous pouvez gérer ces contraintes de temps en déterminant des événements déclencheurs pour les ETR. Utiliser des Termes de Référence génériques comme modèle peut accélérer le développement de Termes de Référence spécifiques, de même que présélectionner des évaluateurs potentiels (ou utiliser des évaluateurs internes, comme le fait UNHCR) peut accélérer le déploiement de l'équipe.</i></p> <p>Les évaluations standards allouent souvent un mois ou plus après le travail de terrain pour rédiger une première version du rapport alors que les rapports d'ETR doivent être achevés (ou presque) au moment où l'équipe quitte le terrain.</p> <p>Astuce : <i>Planifiez l'évaluation de manière à ce que l'équipe rédige son rapport avant de quitter le terrain. Les équipes d'ETR peuvent avoir besoin de rester sur le terrain plus longtemps que d'autres équipes d'évaluation; le minimum est probablement de 3 à 4 semaines. Utiliser un tableau de présentation des données probantes (Chapitres 5.2.4 et 5.2.6) peut aider à accélérer le travail de l'équipe.</i></p>
Parties prenantes	<p>La principale partie prenante de la plupart des évaluations humanitaires est la direction de l'organisation évaluée. Les principales parties prenantes des ETR sont l'équipe sur le terrain et ceux qui gèrent les opérations depuis le siège. L'équipe d'évaluation doit faire part des résultats à l'équipe sur le terrain, même si peu d'entre eux auront le temps de lire un rapport d'évaluation classique.</p> <p>Astuce : <i>Prenez le temps d'inclure des débriefings sur le terrain dans votre plan d'évaluation. En général, une ETR nécessite au moins trois briefings : un pour présenter l'évaluation, un à mi-parcours pour traiter des problèmes émergents avec l'équipe terrain, et un briefing final.</i></p>

Caractéristique	Description
Objectif	<p>Les ETR se concentrent avant tout sur l'apprentissage : elles sont conçues pour aider à améliorer la qualité des réponses en cours, à un moment où il existe en général trop peu d'éléments pour permettre de faire une évaluation tournée vers la redevabilité qui ait du sens.</p> <p><i>Astuce : Lors de votre arrivée, demandez à l'équipe de terrain ce qu'elle souhaiterait savoir à propos de son opération.</i></p> <p><i>Astuce : Le fait de collaborer sur le terrain avec un membre de l'équipe qui est habitué à la façon dont l'organisation fonctionne contribuera à améliorer l'utilisation des recommandations.</i></p>

Une équipe d'ETR va inévitablement mobiliser des ressources qui seraient normalement dirigées vers les populations affectées. Cela crée une tension entre ce que peut apporter l'équipe aux opérations, et le temps et les ressources qu'elle prend à la réponse. Maintenez l'équipe d'ETR en petit nombre pour simplifier la logistique et réduire l'utilisation des ressources. Ceci est particulièrement important si l'ETR a lieu tôt lors d'une opération humanitaire.

Schéma 16 : Aspects prospectifs et rétrospectifs des évaluations en temps réel



6.3.3 Les méthodes

Les méthodes utilisées dans le cadre d'autres EAH ne sont pas toujours adaptées aux ETR, en particulier à celles qui ont lieu durant les premières phases d'une intervention (**Tableau 22**). Une recherche bibliographique est moins utile que d'habitude ; le peu de documentation qui existe porte surtout sur les intentions et les moyens, et non sur les résultats. Les enquêtes sont en général impossibles à conduire en raison des contraintes de temps et de l'absence d'un bon cadre d'échantillonnage.

Les ETR se focalisent fortement sur l'utilisation directe ; une matrice de réponse de la direction est un outil particulièrement utile pour suivre la mise en œuvre des recommandations.

Tableau 22 : Les méthodes utilisées dans le cadre des évaluations en temps réel

Méthode	Utilisation
Groupes de discussions (focus groups)	Plus difficiles à organiser car les populations affectées sont en général occupées à répondre à leurs besoins pour survivre. Un groupe de discussion géré correctement prend en général deux heures ou plus.
Entretiens avec des informateurs clés	Ils sont d'une importance clé compte tenu du contexte opérationnel où les personnes interrogées peuvent ne pas avoir eu le temps de "digérer" leur expérience de manière à pouvoir l'exprimer en dehors d'un entretien.
Ateliers et revues après action	Ils peuvent être particulièrement utiles car ils fournissent au personnel sur le terrain une opportunité d'analyser et de comprendre comment l'opération s'est développée.
Observations	Elles sont encore plus importantes et utiles que d'habitude ; en général, il y aura de nombreuses opportunités d'observer la mise en œuvre.
Entretiens avec des bénéficiaires	Ils sont encore plus importants que d'habitude car ils peuvent fournir des informations vitales pouvant être traduites en actions. L'équipe d'évaluation aura plus d'occasions de parler aux bénéficiaires qu'aux responsables sur le terrain qui doivent concentrer leur énergie sur les problèmes émergents. Les entretiens demandent moins de temps aux participants que les groupes de discussion et sont bien plus simples à mener, même dans des environnements chaotiques.
Outils d'évaluation participative rapide	Ils peuvent être assez utiles mais les outils qui se basent sur des estimations collectives sont limités par le fait que les populations déplacées n'ont pas encore forcément un sens de la communauté.

6.4 Les évaluations d'impact

Le terme d'évaluation d'impact peut être considéré comme impropre car ce type d'évaluation s'intéresse plus précisément aux produits et résultats ([White, 2009a, p. 8](#)). Le terme demeure cependant largement utilisé pour faire référence à une évaluation d'impact de la chaîne des résultats. Selon le programme, elle peut avoir lieu après quelques mois (par exemple, pour un programme alimentaire d'urgence) ou après plusieurs années (par exemple, pour un programme de soutien aux moyens de subsistance).

[Roche \(2000, pp. 545–546\)](#) fournit une définition de l'impact utile et largement utilisée : « changements durables ou significatifs – positifs ou négatifs, intentionnels ou non – dans la vie des individus occasionnés par une action en particulier ou une série d'actions ». Au lieu de mettre l'accent sur le fait que l'aide attendue ait été délivrée ou non (efficacité), cette définition se concentre sur le fait que les résultats constituent ou non une amélioration pour les populations en terme de qualité et de sécurité (voir [Proudlock et al, 2009](#)). Cependant, il est important de noter que les évaluations d'impact ne se focalisent pas sur cette définition plus large mais plutôt sur la question plus précise qui

consiste à déterminer si les produits et résultats de l'intervention peuvent être attribués à l'intervention.

Les évaluations d'impact sont de plus en plus fréquentes dans le secteur du développement ([Jones et al., 2009, p. v](#)). Dans le cadre de l'action humanitaire, les preuves de l'impact d'un programme sont rares, et à l'heure actuelle, il n'y a encore que peu d'exemples d'évaluations uniquement focalisée sur l'impact. L'une des raisons semble être le manque de capacité ([Proudlock et al., 2009](#)). L'intérêt porté aux évaluations d'impact est cependant croissant compte tenu de la pression exercée par les bailleurs (entre autres) pour estimer les impacts qu'ont les actions humanitaires financées au niveau international sur les vies et les moyens de subsistance des bénéficiaires auxquelles elles viennent en aide.

Une évaluation d'impact a plus de chance d'être tournée vers la redevabilité si elle est impulsée par un bailleur et inversement, si elle est commanditée par l'organisation en charge de la mise en œuvre, elle aura plus de chance d'être tournée vers l'apprentissage. Comme avec les autres types d'évaluation, il est important de clarifier l'objectif principal ([Beck, 2011, p. 16](#)).

Le Feinstein Center a été un pionnier dans les approches participatives d'évaluation d'impact qui génèrent un partenariat d'apprentissage entre le bailleur et les partenaires chargés de la mise en œuvre (Catley et al., 2009).

Des conseils plus détaillés sont disponibles dans les ouvrages suivants (3iE, 2008a, 2008b; White, 2009a, 2009b), Poulos (2006), Roche (1999), Catley et al. (2009), et (Bamberger, 2012; Perrin, 2012; Rogers and Evaluation, 2012).

6.4.1 Les évaluations d'impact conjointes

Les évaluations d'impact conjointes peuvent aider à s'assurer que l'impact global de l'action humanitaire est exploré plutôt que d'isoler de façon artificielle le travail d'une seule organisation (Beck, 2009, p. 1). Elles permettent également de comparer différentes approches, ce qui peut aider à évaluer l'impact relatif de différentes interventions. Dans le secteur humanitaire, la plupart des organisations qui planifient une évaluation d'impact le font sous la forme d'une action conjointe.

6.4.2 Comment faire une évaluation d'impact

L'évaluation d'impact implique des jugements de valeur, par exemple sur les types de changements

qui sont significatifs et pour qui. Il est important d'impliquer les populations affectées pour récolter leurs points de vue et identifier ainsi les indicateurs d'impact appropriés, lesquels peuvent différer de ceux choisis par les autres parties prenantes (Proudlock et al., 2009). Des évaluations d'impact approfondies ont tendance à nécessiter un travail de terrain plus long et intensif que les autres types d'EAH.

Une approche comparative et quantitative de l'évaluation d'impact est rarement faisable dans le cadre d'EAH car l'environnement dynamique et fluide dans lequel la plupart des interventions humanitaires ont lieu rend difficile, voire impossible, de définir avec précision la cause des différents résultats (Proudlock et al., 2009). Une approche plus commune est celle qualitative et basée sur une théorie qui examine en profondeur un cas particulier pour expliquer comment une intervention pourrait être responsable de changements bien précis. Cette approche triangule des informations provenant de différentes sources, y compris des bénéficiaires et des informateurs clés. Le plus grand nombre possible de preuves doit être collecté pour soutenir des hypothèses plausibles de causalité.

Une méthode axée sur une théorie comprend trois étapes :

1. *Identifier* les causes possibles des résultats et les impacts intéressants.
2. *Identifier* les conditions nécessaires pour que chaque cause potentielle ait un effet.
3. *Déterminer* si ces conditions sont présentes ou non. Établir une liste réaliste de causes potentielles pour lesquelles les conditions nécessaires sont présentes ([Scriven, 2008](#)) 3. comme mentionné dans ([Proudlock et al., 2009, p. 30](#)).

Le manque de données de base dans les zones affectées par des crises humanitaires représente un défi bien particulier pour l'évaluation d'impact qui peut ainsi dépendre grandement des perceptions et jugements des populations affectées. Il est possible que ces dernières soient les plus à même de juger des changements dans leurs vies, de l'importance de ces changements, et de ce qui les a provoqués. Des données de base rétrospectives peuvent être générées en se fondant sur les souvenirs qu'ont les personnes interrogées des situations antérieures à la crise ou à l'intervention.

L'évaluation d'impact peut être grandement facilitée par la disponibilité de données de suivi systématiques et fiables en matière d'indicateurs de résultats, comme par exemple les données de malnutrition et de morbidité pour les programmes de santé et de distribution alimentaire. Dans la pratique, de telles données font souvent défaut, et les données existantes sont bien souvent quantitatives et focalisées sur les processus et la distribution plutôt que sur les résultats et l'impact.

Le plus grand défi rencontré par l'évaluation d'impact est d'imputer de façon concluante tout changement à l'action humanitaire plutôt qu'à d'autres facteurs, par exemple un changement dans l'environnement externe ou dans les décisions des populations locales. Roche en conclut que « souvent, le mieux qui puisse être fait est de démontrer, grâce à une argumentation raisonnée, qu'un moyen donné mène en toute logique à un changement donné, même si cela ne peut être prouvé de manière statistique » ([1999, p. 33](#)).

Ci-dessous sont présentées des méthodes d'évaluation d'impact supplémentaires :

- *Une étude longitudinale qui reflète les changements dans les vies des populations*

affectées sur une période donnée (un an par exemple) et leurs causes.

- *Des méthodes qualitatives, comme des groupes de discussion, des entretiens de groupe, des entretiens avec des informateurs clés, des entretiens individuels et de foyers, soit en tant qu'activités ponctuelles, soit dans le cadre d'une étude longitudinale.*
- *Des méthodes quantitatives, par exemple des enquêtes formelles sur les ménages, ou la collecte et l'analyse d'indicateurs quantitatifs comme les prix du marché (probablement accompagnées de méthodes qualitatives pour établir les causes et les effets).*

Récemment, une consultation portant sur l'évaluation conjointe d'impact humanitaire a recommandé d'adopter une approche basée sur les deux méthodes et de mettre l'accent sur les méthodes qualitatives, ce qui devrait également encourager davantage de participation. Dans l'idéal, les méthodes mixtes se complètent de manière synergétique (Adato, 2011). Proudlock et al. concluent que « la conception et la mise en œuvre d'évaluations d'impact nécessitent des compétences seulement disponibles dans le cadre de partenariats à long terme entre universitaires, bailleurs, gouvernements, praticiens et bénéficiaires

ciblés » (2009, p. 7).

Lors de la planification d'une évaluation d'impact, les questions suivantes doivent être prises en compte :

- *Comment définir l'impact humanitaire – impact sur quoi, et sur quelle échelle de temps ?*
- *Comment mesurer un impact ? Quels sont les indicateurs appropriés, et pour quels données de références ou groupes de comparaisons ? Comment prouver que les effets observés ou reportés sont en réalité causés par une intervention en particulier ? Quelles sont les méthodes appropriées au contexte et comment gérer les problèmes liés aux données, aux seuils de références et au timing ?*
- *Comment analyser et interpréter les données concernant l'impact, et quel rôle les populations affectées devraient-elles jouer dans ce processus ?*

NOTES

EVALUATIONS A DISTANCE, CONJOINTES, EN TEMPS REEL ET D'IMPACT

Partagez vos commentaires:

Aidez-nous à rendre le Guide plus pratique et focalisé sur l'utilisateur en partageant vos commentaires sur son contenu et la navigation. Les commentaires peuvent inclure des questions, des suggestions, des ressources utiles ainsi que des conseils pratiques. Vous pouvez partager vos idées en cliquant sur le bouton de commentaires en bas de chaque page. Ceci génèrera un courriel confidentiel à l'équipe EAH. Pour plus d'informations sur la façon dont vous, votre équipe ou votre organisation peut s'engager plus activement dans le pilotage du Guide contactez-nous directement à eha@alnap.org

PRODUITS, DIFFUSION ET ADOPTION

Contenu de ce chapitre

7.1	Les produits clés	216
7.2	Le rapport d'évaluation	220
7.3	Faire circuler et commenter le rapport provisoire	229
7.4	Approuver et finaliser l'évaluation	231
7.5	La diffusion	231
7.6	Faciliter l'adoption des recommandations d'une évaluation	234
7.7	Synthèses d'évaluation, revues thématiques et méta-analyses	235

Comment utiliser ce chapitre

Ce chapitre offre des éclaircissements sur les phases finales du processus d'évaluation. Les chapitres **7.1** et **7.2** traitent du rapport et des produits de l'évaluation. Les chapitres **7.3** et **7.4** couvrent les phases de révision et de finalisation, le chapitre **7.5** la diffusion, le chapitre **7.6** les façons d'encourager l'adoption des recommandations et le chapitre **7.7** les autres manières de rendre publics les résultats de l'évaluation, comme les synthèses et les méta-analyses.

Tableaux, schémas et encadrés

Tableau 23:	Les produits de l'évaluation	204
Encadré 9:	Les trois publics des rapports d'évaluation	215

7.1 Les produits clés

Les produits de l'évaluation comprennent le rapport de démarrage, les ateliers de débriefing, les conseils dispensés par l'équipe d'évaluation directement sur le terrain, les ateliers, le rapport d'évaluation et les événements de diffusion, notamment des ateliers. Les produits distincts du rapport (comme les briefings aux bureaux régionaux) sont souvent à prendre en compte dans le budget et l'emploi du temps. Le **Tableau 23** résume les principaux produits de l'évaluation ainsi que leur *timing* et processus de validation.



Les interactions personnelles sur le terrain, lors des briefings et durant les ateliers, sont plus susceptibles d'amener de l'apprentissage que les rapports écrits. Assurez-vous de les planifier, en particulier dans le cadre d'évaluations orientées vers l'apprentissage

Tableau 23 : Les produits de l'évaluation

Produit	Timing	Processus de validation
Rapport de démarrage	Pas moins de 2 semaines avant le travail de terrain pour laisser du temps aux commentaires et aux corrections	<ul style="list-style-type: none"> • Circulation au sein des parties prenantes pour commentaires • Réunion entre le groupe de référence ou de pilotage et l'équipe d'évaluation
Conseils dispensés directement sur le terrain	A mesure que les sujets apparaissent sur le terrain et en réponse à des observations	Aucun.
Ateliers de débriefing	<p>A mi-parcours de l'évaluation en temps réel pour les membres de l'équipe d'évaluation dans le but de tester les résultats émergents</p> <p>A la fin du travail de terrain pour tous les autres types d'évaluation</p>	Les parties prenantes valident ou questionnent les résultats et les conclusions des évaluateurs.
Note de débriefing	A la fin du travail de terrain (peut être aussi simple que la présentation donnée durant l'atelier de débriefing)	La version finale doit refléter les commentaires réalisés lors de la réunion de débriefing.

Produit	Timing	Processus de validation
Première version du rapport d'évaluation	Trois à quatre semaines après la fin du travail de terrain, un peu plus pour les rapports plus complexes (possibilité de plusieurs aller-retour)	<p>Vérification initiale de la qualité par le responsable d'évaluation</p> <p>Circulation au sein du groupe de référence ou du groupe de pilotage, des parties prenantes clés, et des personnes interrogées pour recueillir leurs commentaires</p> <p>Réunion entre le groupe de référence ou de pilotage et l'équipe d'évaluation pour discuter des points clés</p>
Atelier des parties prenantes	Après la première version du rapport	Les parties prenantes peuvent développer des recommandations et des conclusions.
Version finale du rapport d'évaluation	Environ 2 semaines après la réception des commentaires finaux sur l'avant-dernière version	
Séquences vidéo sur les messages clés	Au moment de la présentation du rapport final	Examen par le responsable de l'évaluation et le groupe de référence ou le groupe de pilotage
Matrice de gestion des recommandations	Après la soumission du rapport final	Préparé par l'équipe de gestion de l'organisation
Rapport de synthèse (messages clés de l'évaluation)	Environ une semaine après la soumission du rapport final	Examen par le responsable de l'évaluation

Produit	Timing	Processus de validation
Ateliers de diffusion	A n'importe quel moment après la finalisation du rapport	Aucun
Réunions informelles dites "brown-bag meetings" (courtes présentations des leçons clés à l'heure du déjeuner)	A n'importe quel moment après la finalisation du rapport	Aucun
Autres activités de diffusion (comme l'utilisation des résultats dans les formations)	A n'importe quel moment après la soumission du rapport	Aucun

7.2 Le rapport d'évaluation

Plus le rapport est court, plus il aura de chances d'être lu, mais il devra fournir suffisamment de données probantes pour étayer les résultats, la conclusion et les recommandations. Moins il y a de questions d'évaluation dans les TdR, plus l'équilibre entre ces deux exigences sera facile à trouver¹⁰.



S'il est attendu que le rapport d'évaluation corresponde à un format particulier, cela doit être notifié dans les TdR. Il peut s'agir de détails comme la police de caractère, les tailles de marge et la longueur.

Le guide *ALNAP Quality Proforma (ALNAP, 2005)* propose un modèle de rapport d'évaluation efficace et le Groupe d'Evaluation des Nations unies a mis à disposition une liste de vérification des enjeux de qualité (*UNEG, 2010*).

Préparez un plan du rapport avant de commencer le travail de terrain : cela vous aidera à structurer la collecte des données pour répondre aux exigences du rapport final, et donc à réduire la quantité de

¹⁰ Le contenu du rapport ici décrit s'inspire de *ALNAP Quality Proforma (ALNAP, 2005)*.

travail nécessaire à la production du rapport. Cela se révèle particulièrement utile pour collecter et analyser les données qualitatives recueillies lors d'entretiens.

La plupart des rapports sont structurés autour du cadre et des critères de l'évaluation. Les découpages les plus courants comprennent soit des chapitres pour chaque élément du cadre avec des sous-chapitres par critère, soit des chapitres pour chaque critère avec des sous-chapitres pour les éléments du cadre.

7.2.1 Les pages liminaires

Les pages liminaires comprennent en général les points suivants :

- **La page de titre** – comprend une date.
- **La page de données (facultative)** – peut comprendre les détails administratifs de l'évaluation, les remerciements et une citation suggérée dans le cadre de l'évaluation.
- **Le résumé exécutif** – peut comprendre un résumé des principales recommandations. Il ne doit pas être inclus dans les premières versions du rapport afin d'éviter des commentaires sur le résumé plutôt que sur le rapport complet.

- **La table des matières** – le cas échéant, ajoutez également une liste des tableaux et des schémas. La plupart des logiciels de traitement de texte peuvent générer automatiquement ces éléments si les titres et les légendes sont mis aux bons formats (par exemple, en utilisant des styles Microsoft Word).
- **Liste des acronymes** – appropriée si le document contient un nombre important d'acronymes. S'il contient aussi des termes rares (comme du vocabulaire local pour un élément géographique particulier ou une source d'eau), ajoutez une liste séparée pour les définitions de ces termes.
- **Carte(s)** – produisez vos propres cartes en utilisant un logiciel de cartographie ou cherchez des cartes libres de droits.



Utilisez des acronymes seulement quand un terme se retrouve de façon si fréquente que le désigner sous une forme courte est utile pour les lecteurs ou quand il est plus facile de le reconnaître sous sa forme d'acronyme. Explicitez toujours les acronymes et définissez les termes rares lors de leur première occurrence dans le texte, même si vous pensez qu'ils seront connus de nombreux lecteurs. Ces pratiques ne facilitent pas seulement la lecture d'un document mais aident aussi à souligner que le rapport a été écrit pour toutes les parties prenantes, et pas seulement pour les initiés de l'organisation ou du secteur.

7.2.2 Le corps du texte

Le corps du texte d'un rapport contient en général les éléments suivants :

- **Une introduction** qui fixe le périmètre et l'objectif de l'évaluation, la composition de l'équipe et la structure du rapport.
- **Un chapitre sur la méthodologie** qui décrit les principales méthodes utilisées, les contraintes rencontrées et tous les biais inhérents à l'équipe d'évaluation. Cela peut faire partie de l'introduction et peut être un simple résumé qui renvoie les lecteurs à une analyse plus détaillée en annexe.
- **Un chapitre de contexte.** Il peut faire partie de l'introduction et peut être complété par une chronologie en annexe mais il doit être détaillé. Le contexte est un sujet clé de l'action humanitaire : le fait que les actions soient adaptées, efficaces, efficientes ou cohérentes dépend beaucoup du contexte.
- **Les principaux chapitres**, qui présentent les éléments de preuve et les résultats, organisés par critères d'évaluation ou tout autre cadre de référence. Les conclusions et recommandations sont présentées à la fin de chaque chapitre ou bien dans un ou deux chapitres distincts à la fin du rapport.



Le fait de présenter les conclusions et recommandations à la fin de chaque chapitre souligne (et aide à assurer) leur bon ancrage dans les éléments de preuve, les résultats et les conclusions



Exemple
de bonne
pratique 20

Utiliser un tableau pour résumer des actions multiples de partenaires multiples

Une évaluation réalisée pour le CBHA du travail de relèvement qui a suivi les inondations au Pakistan résume dans un tableau d'une page les diverses expériences de six organisations ayant apporté un appui en espèces, en bons d'achat et en nature à travers trois secteurs liés aux moyens de subsistance (Leturque et al., 2012, p. 45). Ce tableau résume également les avantages et inconvénients théoriques des trois modalités.

	Avantages et inconvénients théoriques	Résultat 1 – Intrants agricoles	Résultat 2 – Bétail	Résultat 3 – Actifs d'entreprises
Subventions conditionnelles en espèces	<p>Avantages :</p> <ul style="list-style-type: none"> • L'organisation peut influencer les dépenses des bénéficiaires. • Les bénéficiaires peuvent avoir d'autres priorités. <p>Inconvénients :</p> <ul style="list-style-type: none"> • Forte charge administrative (suivi des dépenses avant paiement) 	<p>CARE International Réussites : Les bénéficiaires sont capables d'adapter leurs achats d'intrants en fonction du rythme des transferts d'argent.</p>	<p>CARE International Bien que les bénéficiaires aient le choix, la subvention en espèces est presque toujours utilisée pour des intrants agricoles plutôt que pour le bétail.</p>	<p>CARE International, Concern, IRC, Oxfam Réussites et difficultés : Indications qu'environ la moitié ou plus du financement a été dépensée à d'autres fins.</p>

	Avantages et inconvénients théoriques	Résultat 1 – Intrants agricoles	Résultat 2 – Bétail	Résultat 3 – Actifs d'entreprises
Bons d'achats	<p>Avantages :</p> <ul style="list-style-type: none"> • Meilleur contrôle sur la réalisation des objectifs du marché. <p>Inconvénients :</p> <ul style="list-style-type: none"> • Nécessité d'une planification importante • Stimule les entreprises locales • Exclut certains commerçants • Un commerçant sélectionné peut augmenter ses prix • Limite les choix des bénéficiaires 	<p>ACF, Oxfam</p> <p>Réussites :</p> <ul style="list-style-type: none"> • Intrants ciblés reçus • Exécution de l'activité au moment opportun. <p>Difficultés :</p> <ul style="list-style-type: none"> • Des bénéficiaires ont rapporté des prix faussés par les fournisseurs retenus. 	<p>Non utilisé par les partenaires de CBHA</p>	<p>SC</p> <p>Difficultés :</p> <ul style="list-style-type: none"> • Des bénéficiaires ont rapporté qu'ils auraient préféré des subventions en espèces. • Choix limités et prix gonflés par des fournisseurs retenus.

	Avantages et inconvénients théoriques	Résultat 1 – Intrants agricoles	Résultat 2 – Bétail	Résultat 3 – Actifs d'entreprises
En nature	<p>Avantages :</p> <ul style="list-style-type: none"> • Objectifs plus ciblés • Contrôle sur la qualité des intrants <p>Inconvénients :</p> <ul style="list-style-type: none"> • Risque associé au transport • Coûts logistiques élevés • Limite les choix des bénéficiaires. 	<p>Concern, IRC, Oxfam, SC</p> <p>Réussites :</p> <ul style="list-style-type: none"> • Importantes économies sur les achats en gros • Aucun retard majeur de livraison notifié. 	<p>Concern, IRC, Oxfam, SC</p> <p>Réussites :</p> <ul style="list-style-type: none"> • Economies sur l'achat en gros <p>Difficultés :</p> <ul style="list-style-type: none"> • Retard de livraison (Oxfam) • Forte mortalité (livraison en retard, transport) • Les lots d'aliments pour les animaux sont utiles mais pas durables • Ratio mâle/femelle difficile à contrôler • Nécessité d'une formation et chronologie mal adaptée au renforcement des compétences. 	<p>Formation professionnelle : non analysée lors de cette évaluation</p>

Les tableaux sont seulement une façon de communiquer rapidement des informations. Des photographies choisies avec soin, des diagrammes bien construits et d'autres schémas peuvent aussi communiquer des résultats clés. Tufts (1990; 2001; 1997) dispense de bons conseils sur la présentation des informations visuelles.

7.2.3 Recommandations

Les recommandations sont utiles seulement si elles sont:

- **Spécifiques** – *ce qui est recommandé doit être extrêmement clair.*
- **Liées à des actions vérifiables** – *il doit être possible de dire si la recommandation a été mise en œuvre ou non.*
- **Dirigées** – *la personne ou l'entité responsable de la mise en œuvre de la recommandation doit être identifiée ; la responsabilité peut être davantage clarifiée dans la réponse de la direction au rapport.*
- **Faisables** – *les recommandations peuvent impliquer des manières de travailler innovantes ou inhabituelles mais elles doivent garder à l'esprit les ressources et autres contraintes.*
- **Limitées dans le temps** – *un calendrier de mise en œuvre des recommandations*

doit être donné quand cela est possible.

- **Cohérentes** – *les recommandations ne doivent pas se contredire ou sembler se contredire.*
- **Priorisées** – *clarifier quelles sont les recommandations de premier et de second ordre.*
- **Economiques** – *les actions recommandées doivent clairement apporter des bénéfices proportionnels à leurs coûts.*

Les recommandations doivent également être limitées en nombre. Moins un rapport contient de recommandations, plus un client peut l'utiliser facilement. Cependant, un évaluateur expérimenté peut remarquer des douzaines de problèmes de performance et a le devoir éthique de les signaler. Diverses stratégies peuvent aider à résoudre ce dilemme :

- *Proposer une recommandation générale et indiquer ensuite en détail comment elle peut être mise en œuvre par différents acteurs.*
- *Emettre les recommandations mineures à l'oral et les placer dans l'annexe du rapport.*
- *Classer les recommandations par importance.*
- *Regrouper les recommandations par cible.*

7.2.4 Les annexes

Les annexes sont utilisées pour apporter des éléments complémentaires et des détails qui surchargerait le rapport principal. Elles aident les lecteurs à se faire une idée de la qualité de l'évaluation en démontrant l'étendue du travail de terrain et la variété des méthodes utilisées. Les annexes peuvent inclure les éléments suivants :

- **TdR** – par convention, c'est la première annexe d'un rapport d'évaluation.
- **Chronologie du sujet ou de l'action en cours d'évaluation** – essentielle pour toutes les évaluations dans lesquelles l'opportunité/rapidité (« timeliness ») des actions est un critère.
- **Liste des personnes rencontrées ou consultées** – souligne le nombre important des acteurs consultés par l'équipe et aide les lecteurs à se forger une opinion quant à la fiabilité des résultats.
- **Itinéraire de l'équipe** – fournit également aux lecteurs une idée de l'ampleur de la recherche effectuée par l'équipe.
- **Biographies de l'équipe** – donne aux lecteurs une idée de l'expérience des membres de l'équipe, ainsi que des biais que ces profils peuvent avoir sur l'évaluation.
- **Approche méthodologique** – définit en détail les méthodes utilisées et mentionne toutes les contraintes et limites potentielles. Cette partie est en général complétée par des exemples des différents outils d'entretiens et des formulaires d'enquête utilisés.
- **Matrice de réponse de la direction** – peut être jointe au rapport en annexe même si elle est plus souvent imprimée comme document distinct.
- **Autres annexes** – peuvent répondre à des sujets pour lesquels une étude détaillée serait hors de propos au sein du rapport principal ou présenter des résultats issus de méthodes spécifiques (comme un résumé des réponses obtenues lors d'un sondage en ligne).
- **Bibliographie** – peut inclure certaines sources clés consultées en plus des sources citées dans le rapport d'évaluation.

Une tendance actuelle consiste à rendre disponible la plupart ou la totalité des annexes sur un site Internet plutôt que de les publier avec le rapport. Malheureusement, il est rarement possible de trouver ces annexes après plusieurs années. Par

exemple, les annexes de l'évaluation du IASC en Somalie ([Polastro et al., 2011a](#)) ne peuvent plus être trouvées sur le site Internet des Nations unies pour la Somalie.



Même si les annexes ne feront pas partie de la plupart des exemplaires papier du rapport, produisez une copie électronique complète qui puisse être postée sur le site Internet du client

Encadré 9 : Les trois publics des rapports d'évaluation

Tous les rapports d'évaluation ont trois grandes cibles :

1. Les responsables n'ont pas le temps de lire de longs rapports. Leurs besoins sont souvent mieux satisfaits par un bon résumé exécutif de quelques pages. (De nombreux résumés exécutifs sont bien trop longs pour atteindre cet objectif.)
2. Les parties prenantes peuvent lire le rapport complet.
3. Les experts techniques comptent sur les informations détaillées fournies en annexes.



Chaque fois que vous écrivez un point du rapport, pensez au public qu'il intéressera le plus.

7.3 Faire circuler et commenter le rapport provisoire

On fait en général circuler les rapports d'évaluation provisoires via le responsable de l'évaluation aux parties prenantes clés et éventuellement à toutes les personnes interrogées. Donner aux personnes interrogées la possibilité de répondre aux erreurs factuelles ou de compréhension est une mesure efficace de contrôle de la qualité. Cela nécessite que l'équipe d'évaluation enregistre les adresses de courrier électronique de toutes les personnes interrogées.

Faire circuler le rapport permet également que les parties prenantes puissent attirer l'attention sur des erreurs de fait, de compréhension ou d'analyse du rapport. Cela peut aider à promouvoir l'appropriation de l'évaluation par les parties prenantes et à encourager son utilisation.

Accordez le temps suffisant – en général 1 à 3 semaines – à cette phase. Les parties prenantes qui reçoivent cette version provisoire doivent la faire circuler au sein de leur organisation et ensuite collecter les commentaires qu'elles reçoivent.

Faire circuler des rapports dans des formats de traitement de texte (en général Microsoft Word) génère des problèmes. La tentation est d'intégrer ses commentaires en suivi de modification, ce qui devient difficile à suivre si plus de deux ou trois personnes laissent des commentaires sur le même document. Les différents paramètres des utilisateurs peuvent également causer des changements dans la pagination et la mise en page, ce qui peut créer de la confusion. Il est préférable de faire circuler le rapport en format .pdf, ce qui fige la mise en page et offre des options de commentaires plus lisibles.



Numérotez les paragraphes du rapport provisoire afin de faciliter l'écriture des commentaires qui se réfèrent à des passages spécifiques du texte



Compilez les commentaires et travaillez de façon systématique de manière à ce que tous les commentaires d'un chapitre donné puissent être pris en compte en même temps.



Exemple
de bonne
pratique 21

Collectez les commentaires de façon transparente

Une évaluation du travail de relèvement du CBHA qui a suivi les inondations au Pakistan résume dans un tableau d'une page les diverses expériences de six organisations en matière d'appui en espèces, en bons d'achats et en nature à travers trois secteurs liés aux moyens de subsistance (Leturque et al., 2012, p. 45). Ce tableau résume également les avantages et inconvénients théoriques des trois modalités.

Page 42 para: 151 – *La désignation d'un envoyé spécial des Nations unies de très haut niveau, l'ancien président Bill Clinton, s'est avérée une innovation particulièrement réussie. Alors qu'à l'origine, ce choix avait pu être motivé par le souhait qu'un tel Envoyé puisse débloquer les cordons de la bourse des bailleurs (Clinton, 2006), il s'est avéré très fructueux lorsque l'objectif est devenu la promotion de l'efficacité de la réponse.*

Agence des Nations unies 1 – Ce point doit également mentionner la création et le rôle du Global Consortium on Tsunami Recovery présidé par Bill Clinton, ainsi que son action essentielle en matière de communication et de coordination innovante auprès du IASC-UNDG.

Agence des Nations unies 2 – Le rapport de coordination loue également la nomination d'un Représentant Spécial pour le Tsunami – le DERC d'OCHA (Margareta Wahlstrom). Cela a été perçu comme un autre rôle de coordination positif. Je ne sais pas pourquoi vous n'avez pas choisi celui-là... mais cela relève bien entendu des prérogatives de l'équipe.

Relecteur A – La base à partir de laquelle il a été conclu que l'Envoyé Spécial était utile n'est pas claire.

ONG 1 – Il peut être utile d'explicitier ce qui a fait de la désignation de Clinton un "succès".



Faire savoir aux relecteurs que vous allez collecter et faire circuler leurs commentaires encourage les remarques plus réfléchies.

Une bonne pratique consiste pour l'équipe d'évaluation à indiquer comment elle a répondu aux commentaires des relecteurs en ajoutant des petites notes comme :

- *Corrigé*
- *Nuancé*
- *Détail ajouté*

- *Supprimé*
- *Refusé [avec une explication détaillée]*

7.4 Approuver et finaliser l'évaluation

Le rapport d'évaluation est le principal produit d'une évaluation. Il doit :

- *Répondre aux questions de l'évaluation.*
- *Fournir des éléments de preuve pour appuyer ses conclusions et ses recommandations.*
- *Présenter des recommandations claires et identifier la personne qui doit les mettre en œuvre.*
- *Etre cohérent et libre des contradictions internes.*
- *Utiliser les méthodes spécifiées dans les TdR et le rapport de démarrage (sinon, expliquer clairement pourquoi).*
- *Satisfaire toutes les exigences de mise en page et de format.*
- *Contenir suffisamment d'informations sur le processus et les méthodes pour permettre aux lecteurs de juger de sa fiabilité.*

Les lecteurs du rapport d'évaluation se feront leur

propre jugement sur la qualité de l'évaluation et le crédit à apporter aux recommandations. Un bon rapport donne une bonne image du responsable de l'évaluation et de l'équipe d'évaluation et inversement pour un mauvais rapport.

7.5 La diffusion

Bien trop souvent, la planification de la diffusion des résultats et des recommandations de l'évaluation est reportée après la finalisation du rapport d'évaluation. Pourtant, si vous visez une utilisation sérieuse, l'ensemble de l'évaluation – y compris la diffusion – doit être planifiée et budgétée en gardant à l'esprit ce besoin d'anticiper (Chapitre 2). Le rendu du rapport d'évaluation final n'est pas la fin de l'évaluation, c'est simplement le début de la phase de diffusion.

La diffusion doit être guidée par une stratégie de communication qui explicite les points suivants :

- *Les groupes clés à qui les résultats et les recommandations de l'évaluation doivent être communiqués*
- *Les meilleurs moyens de communiquer avec ces groupes*

- *Qui est responsable de la communication*
- *La plupart des personnes apprennent et se souviennent plus en écoutant et en discutant qu'en lisant. Se contenter de faire circuler le rapport ne suffit donc pas. Il existe de nombreuses autres solutions complémentaires pour communiquer des résultats d'évaluation :*
 - *Des réunions ciblées peuvent être organisées en présence de l'équipe d'évaluation et des utilisateurs attendus. Si celles-ci sont conçues sur le mode d'un atelier et modifiées pour s'adapter aux besoins de chaque groupe d'utilisateurs, elles peuvent contribuer à un plus fort engagement de la part des utilisateurs.*
 - *Des briefings individuels peuvent être proposés à des utilisateurs clés par l'équipe d'évaluation ou le responsable de l'évaluation.*
 - *Des résumés brefs et accessibles ou des documents de briefing sont plus susceptibles d'être lus par des décideurs débordés qu'un long rapport d'évaluation.*
 - *Des courriers électroniques courts peuvent communiquer les résultats, conclusions et recommandations à des utilisateurs clés.*
 - *Le rapport d'évaluation ou son résumé peuvent être traduits dans des langues locales et diffusés aux parties prenantes locales.*
- *Des membres de l'équipe d'évaluation peuvent retourner sur le site pour rapporter les résultats et recommandations aux parties prenantes clés, comme le personnel du programme et les membres de la population affectée.*
- *Les résultats et les recommandations peuvent être communiqués en utilisant des séquences vidéo de l'évaluation ou des entretiens enregistrés avec l'équipe d'évaluation.*
- *Des podcasts (enregistrements audio pouvant être écoutés sur des lecteurs MP3 et des ordinateurs) abordant les points clés de l'évaluation peuvent être mis à disposition).*



Exemple
de bonne
pratique 22

Utiliser la vidéo pour diffuser des messages clés

Un certain nombre d'organisations expérimentent l'usage de vidéos pour accompagner leurs évaluations – par exemple, l'UNHCR, la Croix-Rouge et le Groupe URD. Celles-ci peuvent fournir de puissants témoignages de la part des bénéficiaires sur des sujets clés et sont plus susceptibles de marquer les esprits que de longs rapports d'évaluation. La vidéo du Groupe URD sur l'évaluation en temps réel en Haïti peut être vue en ligne à l'adresse suivante : <http://vimeo.com/15198053>.

Les médias sociaux s'annoncent comme des moyens prometteurs pour communiquer les résultats d'évaluation (**Kaplan and Haenlien, 2010**), notamment des projets collaboratifs comme les wikis, les sites de réseautage social comme LinkedIn, les blogs et les communautés de contenu comme YouTube. Par exemple :

- *Des détails concernant la progression de l'évaluation sur l'intégration de la diversité Age et Genre de l'UNHCR en Colombie (Mendoza et Thomas, 2009) ont été mis à jour sur un blog intitulé **It Begins with Me. It Begins with You. It Begins with Us: Thoughts and Actions Around Age, Gender and Diversity Mainstreaming.***

- *La Société européenne de l'évaluation mène des discussions sur LinkedIn.*

Des mises à jour d'évaluation peuvent être communiquées régulièrement via microblogging. Le plus connu probablement de ces services est Twitter : il limite les messages à 140 caractères, un format qui convient mieux à des petites mises à jour en lien avec un processus dynamique qu'à des informations détaillées.

Les médias sociaux continuent d'évoluer rapidement, de nouvelles formes et de nouvelles utilisations ne cessent d'émerger.

7.6 Faciliter l'adoption des recommandations d'une évaluation

Impliquer les principales parties prenantes tout au long du processus d'évaluation est un élément clé de l'utilisation. Une phase de diffusion bien planifiée et bien financée est également importante. Voici d'autres manières de faciliter l'adoption :

- *Impliquer certains des acteurs clés dans la rédaction des recommandations.*
- *Définir clairement les responsabilités en matière de suivi, par exemple par la mise en place d'une matrice formelle de réponse de la direction.*
- *Transformer les résultats de l'évaluation en contenu de formation, par exemple sous la forme d'études de cas.*



Exemple
de bonne
pratique 23

Impliquer les parties prenantes clés tout au long de l'évaluation

L'évaluation ECB de la réponse au tremblement de terre de Yogyakarta (Wilson et al., 2007) s'est fortement concentrée sur l'utilisation. Parmi les aspects clés, on retrouvait les suivants :

- discussions au début de l'évaluation avec les membres du comité de pilotage au niveau du pays par rapport à leurs attentes
- dialogue entre le chef d'équipe et le comité de pilotage avant que le chef d'équipe n'arrive dans le pays
- discussions entre l'équipe d'évaluation et le comité de pilotage à propos des méthodes d'évaluation (par exemple, l'interprétation la plus adaptée des critères du DAC).

L'équipe d'évaluation a présenté un résumé des résultats de l'évaluation au comité de pilotage et au personnel de terrain de manière à ce que les principales parties prenantes puissent travailler avec l'équipe pour rédiger les conclusions et les recommandations. Une réunion finale a été organisée en présence de l'équipe d'évaluation, des membres du comité de pilotage, des représentants du gouvernement et d'habitants de la région pour réviser et amender les premières conclusions et établir de nouvelles recommandations.

7.7 Synthèses d'évaluation, revues thématiques et méta-analyses

La longueur et le détail d'un rapport d'évaluation peut décourager les décideurs à le lire et à prendre en compte ses leçons dans de futures interventions. Une manière appréciée et efficace de s'assurer que les résultats clés atteignent leur public cible est de produire une synthèse des évaluations passées, par exemple un document court et facile d'accès sur les leçons apprises. Un certain nombre d'organisations ont réalisé de telles méta-analyses, parfois de façon thématique, et elles sont en général bien accueillies par les utilisateurs potentiels. Voir l'exemple de bonne pratique 24.



Exemple
de bonne
pratique 24

Synthétiser les leçons apprises

En 2008, en réponse au tremblement de terre en Chine, ALNAP a rédigé un document sur les activités humanitaires qui a distillé les leçons d'un certain nombre d'évaluations de réponses post-séisme (**Cosgrave, 2008**). Dans les jours qui ont suivi le tremblement de terre de 2010 en Haïti, ce document a été téléchargé plus de 3.500 fois.

CARE a réalisé une méta-analyse des évaluations et des revues après action à partir de 24 réponses d'urgence, explorant dans quelle proportion CARE avait internalisé les recommandations et les leçons des évaluations. L'objectif était de réfléchir à la manière dont CARE pourrait faire un usage plus efficace des résultats de ses évaluations (**Oliver, 2007**).

World Vision a réalisé un DVD contenant plus de 50 résumés de leçons apprises à travers une large gamme de secteurs (**Montbiot, 2006**).

NOTES

Partagez vos commentaires:

Aidez-nous à rendre le Guide plus pratique et focalisé sur l'utilisateur en partageant vos commentaires sur son contenu et la navigation. Les commentaires peuvent inclure des questions, des suggestions, des ressources utiles ainsi que des conseils pratiques. Vous pouvez partager vos idées en cliquant sur le bouton de commentaires en bas de chaque page. Ceci générera un courriel confidentiel à l'équipe EAH. Pour plus d'informations sur la façon dont vous, votre équipe ou votre organisation peut s'engager plus activement dans le pilotage du Guide contactez-nous directement à eha@alnap.org

ALNAP

c/o Overseas Development Institute
111 Westminster Bridge Road
London SE1 7JD
United Kingdom

Tel +44 (0)20 7922 0300
Fax +44 (0)20 7922 0399
Email alnap@alnap.org