

# **EVALUER L'ACTION HUMANITAIRE ETAT DES LIEUX DES REFLEXIONS ET PRATIQUES**

Document préparé par François Grünewald, Groupe URD, pour le Groupe de Travail sur les Evaluation de la  
Délégation à l'Action Humanitaire<sup>1</sup>

---

<sup>1</sup> Ce texte est un état des lieux revenant sur une quinzaine d'années de travail sur ce thème (voir annexe 1)

## SOMMAIRE

SOMMAIRE.....	2
INTRODUCTION.....	3
A. LEÇONS TIREES, LEÇONS APPRISES, LEÇONS OUBLIEES.....	4
1. L'ANALYSE DES POINTS CRITIQUES TELLE QUE PERMISE PAR L'EVALUATION .....	4
2. LEÇON 1 : DES DIAGNOSTICS INADEQUATS .....	4
3. LEÇON 2 : UN MANQUE DRAMATIQUE DE PARTICIPATION DES POPULATIONS .....	4
4. LEÇON 3 : DES PROGRAMMES PLAQUES SUR DES REALITES COMPLEXES .....	4
5. LEÇON 4 : DES PRESSIONS FORTES DES BAILLEURS .....	4
6. LEÇON 5 : UN TURN OVER DES EQUIPES ELEVE .....	4
7. LEÇON 6 : UN FAIBLE LIEN ENTRE L'URGENCE ET LE DEVELOPPEMENT.....	5
8. LEÇON 7 : UNE FAIBLE PRISE EN COMPTE DES QUESTIONS TRANSVERSALES .....	5
9. LEÇON 8 : DES COORDINATIONS SOUVENT DEFAILLANTES .....	5
10. LEÇON 9 : DES MANDATS, DROITS ET PRINCIPES ENCORE DANS LE FLOU .....	5
11. LEÇON 10 : DES MEMOIRES INSTITUTIONNELLES DEFAILLANTES .....	5
B . LA TENUE DES DEBATS METHODOLOGIQUES ACTUELS .....	6
1. CONTROLE, RESPONSABILITE ET EVALUATION.....	6
2. EVALUATION ET APPRENTISSAGE .....	6
3. CRITERES, STANDARDS OU NORMES, INDICATEURS .....	7
3.1. Les grandes familles de critères et d'indicateurs .....	7
3.2. Critères d'impact.....	7
3.3. Critères de réalisation .....	8
3.4. Critères de management.....	8
3.5. Critères de situation .....	9
4. L'EVALUATION ET SA PLACE DANS LE CYCLE DE PROJET .....	9
5. INDEPENDANCE DES EVALUATEURS OU/ET EVALUATION UTILE .....	9
C. LES SYSTEMES EXISTANTS .....	11
6. LES PROCESSUS ET MECANISMES EVALUATIFS.....	11
7. LES SYSTEMES INSTITUTIONNELS .....	12
ANNEXE 1 .....	14

## INTRODUCTION

Les termes du problème posé par le Délégué à l'Action Humanitaire au Groupe de Travail sur l' Evaluation de l' Action Humanitaire pourraient se résumer de la façon suivante :

- « Comment aider à prendre les **meilleurs décisions** pour des actions humanitaires dans un cadre de **contraintes complexes et fortes**, sur la base des **expériences connues** sur les bonnes et mauvaises pratiques , ainsi que sur les pièges déjà identifiés ? »
- « Comment créer une **dynamique d'apprentissage rapide** ? L'évaluation ex-post peut-elle faire partie des outils et méthodes pour y arriver ? »
- « Comment tenir compte des expériences déjà existantes dans le domaine de l'institutionnalisation de la **fonction évaluation** pour faciliter la mise en place des structures et processus internes de la DAH ? »

Pour tenter de répondre à ces interrogations de la DAH, le présent document est structuré en trois parties :

### **Etat des lieux de l'expérience acquise : leçons apprises, leçons utilisées:**

L'évaluation de l'action humanitaire est devenu un secteur encore récent (démarré en 1992). Il s'est beaucoup développé au cours des dernières années et a entraîné la création d'une quantité impressionnante de connaissance. Les grands familles de « leçons tirées » sont présentées dans la partie A.

### **Etat des lieux des débats sur l'évaluation :**

La teneur des débats a beaucoup évolué en fonction des expériences. Cinq grands axes de réflexions se sont développés, qui seront traités dans la partie B :

- la réflexion liant contrôle, responsabilité (dont la redevabilité n'est qu'une des composantes) et évaluation;
- la réflexion liant évaluation et apprentissage ;
- la réflexion sur les critères, les standards ou normes et indicateurs ;
- la réflexion sur le rôle et la place de l'évaluation au sein du cycle de projet
- la réflexion sur l'indépendance des évaluateurs.

### **Etat des lieux sur les dispositifs : pour des institutions apprenantes**

Enfin, la partie B fera le point sur les dispositifs optimaux existants et des idées émergentes pour faire des institutions humanitaires des structures « apprenantes »

## **A. LEÇONS TIREES, LEÇONS APPRISES, LEÇONS OUBLIEES**

### 1. L'ANALYSE DES POINTS CRITIQUES TELLE QUE PERMISE PAR L'EVALUATION

Les points qui sont présentés ci-dessous sont les leçons les plus saillantes, telles qu'identifiées sur une vingtaine d'évaluations différentes réalisées :

- sur quatre continents (Amérique Centrale, Afrique, Asie, Europe de l'Est) ;
- trois grands types de situation : catastrophe naturelle de grande ampleur; conflit armé durable; immédiat sortie de crise.

Les points critiques relevés sont ceux qui soit ont une fréquence élevée et une gravité moyenne, soit ceux qui ont une fréquence rare mais un impact négatif majeur. Il en résulte les dix leçons suivantes.

### 2. LEÇON 1 : DES DIAGNOSTICS INADEQUATS

Les diagnostics incomplets, inadéquats, biaisés par ce qu'on sait déjà avant d'arriver ce qu'on va faire, sont extrêmement fréquents. L'explication donnée comporte deux parties : pas de temps, et pas de moyens (« les bailleurs ne financent pas la mission exploratoire »). Les situations d'extrême urgence, dans lesquelles il n'y a réellement qu'un temps très restreint pour le diagnostic, sont heureusement relativement rares. Les humanitaires travaillent la plupart du temps dans des contextes de crises durables et, de plus en plus de sortie de crise, et donc théoriquement devraient avoir le temps.

### 3. LEÇON 2 : UN MANQUE DRAMATIQUE DE PARTICIPATION DES POPULATIONS

La plupart des grandes agences humanitaires, et notamment les ONG françaises, sont très faibles dans leur capacité de penser et mettre en place des processus participatifs. Les excuses avancées sont nombreuses : pas de temps, pas d'acteur locaux, pas confiance, etc. L'analyse fine de la réalité montre qu'en fait il s'agit largement d'un biais culturel des humanitaires, renforcé par une faible compétence sur les méthodes participatives.

### 4. LEÇON 3 : DES PROGRAMMES PLAQUES SUR DES REALITES COMPLEXES

Le passage du diagnostic à la conception n'est pas toujours cohérent. On s'aperçoit très fréquemment que même sur la base d'un diagnostic bien fait et différencié, la tendance sera à la réplication de programmes standards. Plusieurs raisons poussent à cela : la compétence de l'ONG, l'expérience des équipes, la pression (réelle ou pensée) des bailleurs, les besoins financiers et d'image des agences humanitaires.

### 5. LEÇON 4 : DES PRESSIONS FORTES DES BAILLEURS

La plupart des acteurs sur le terrain se plaignent de la pression des bailleurs : toujours plus de rapports à faire, des formats de compte-rendu chaque fois différents, des modes imposées, etc. Il s'avère aussi que de temps en temps, les évaluations fassent ressortir que cette « pression des bailleurs » est aussi une excellente excuse pour masquer des défaillances de l'acteur humanitaire lui-même. « C'est la faute au bailleur qui voulait ça et ça... »

### 6. LEÇON 5 : UN TURN OVER DES EQUIPES ELEVE

Dans les situations de crise et post-crise, l'analyse des flux de personnel montre un turn-over élevé. Il est difficile dans ces conditions d'avoir une perduration du savoir de l'action et de pouvoir conserver la confiance des populations et des acteurs locaux. Il y a aussi une tendance forte corrélée à cette rotation rapide des équipes, d'avoir un manque de logique interne des programmes, chaque nouveau staff voulant « imprimer sa marque » sur le programme. Les conditions matérielles et financières dans lesquelles travaillent les volontaires des ONG françaises pénalisent ces dernières et il est de plus en plus fréquent de voir d'excellents cadres intermédiaires des ONG aller travailler chez les « concurrentes » nordiques ou anglo-saxonnes.

## 7. LEÇON 6 : UN FAIBLE LIEN ENTRE L'URGENCE ET LE DEVELOPPEMENT

Les évaluations font fréquemment apparaître la faiblesse des réflexions sur la « stratégie de sortie ». Ces souvent une question qui se pose quand arrive la fin des financements abondants de l'urgence et que la structure en question doit décider de rester ou de partir. Les difficultés institutionnelles (quel mandat), de savoir-faire et de méthodes apparaissent très souvent. Là encore, les compétences ne s'inventent pas

## 8. LEÇON 7 : UNE FAIBLE PRISE EN COMPTE DES QUESTIONS TRANSVERSALES

Les nombreux travaux d'évaluation montrent l'importance de certaines questions transversales : questions de genre, d'environnement, mais plus important encore de protection et de relations entre actions humanitaires et dynamiques du conflit et des déplacements de population. Il apparaît que, dans les grandes lignes, ces problèmes qui pourraient souvent être anticipés ne sont gérés que de façon réactive.

## 9. LEÇON 8 : DES COORDINATIONS SOUVENT DEFAILLANTES

La coordination est à la fois une mécanique très consommatrice de temps et de moyens, et un enjeu fondamentale pour éviter duplication, les poches de besoins non couverts, l'incohérence des pratiques. Des progrès considérables ont été réalisés au cours des dernières années, et les grandes ONG sont plutôt des acteurs « moteurs » de ces bonnes pratiques de coordination. Les difficultés viennent plutôt du groupe des petites ONG et des collectivités locales, ainsi que des systèmes bilatéraux, la France n'étant pas un très bon élève dans ce secteur.

## 10. LEÇON 9 : DES MANDATS, DROITS ET PRINCIPES ENCORE DANS LE FLOU

Les évaluations font ressortir de façon régulière l'enjeu de la clarification et du respect des mandats des différents acteurs sur le terrain. De façon récurrente, les mélanges de genre entre « humanitaires » et « militaires », entre « humanitaires » et « défenseurs des droits de l'homme », entre « humanitaires » et « constructeurs de la paix », etc. sont repérés. Il s'avère que sur la durée ces pratiques contribuent à augmenter la dangerosité des situations.

## 11. LEÇON 10 : DES MEMOIRES INSTITUTIONNELLES DEFAILLANTES

Les évaluations réalisées démontrent que ONG, mais aussi les agences des Nations Unies et la plupart des bailleurs de fond ont de très faibles mémoires institutionnelles, même si tout le monde passe beaucoup de temps à écrire des rapports. Il en résulte une répliation fréquente des mêmes erreurs. Devenir une « institution apprenante » implique de se créer un système d'enregistrement des résultats positifs et des erreurs, avec une analyse des conditions et raisons du « pourquoi ».

## B . LA TENEUR DES DEBATS METHODOLOGIQUES ACTUELS

### 1. CONTROLE, RESPONSABILITE ET EVALUATION

Pendant longtemps, les évaluations étaient essentiellement vues et utilisées comme l'outil de contrôle des bailleurs. Le débat s'est organisé récemment autour du concept anglo-saxon d'« *accountability* ». *Accountability*, ou la responsabilité de rendre compte, a été traduit par « redevabilité ». A ce stade<sup>2</sup>, il apparaît que le terme de responsabilité est plus riche. Dans un document préparé pour un colloque organisé par MSF Hollande en 1995<sup>3</sup>, il était indiqué que les agences humanitaires avaient une responsabilité :

- d'abord envers ceux pour qui nous travaillons et au nom desquels nous mobilisons des moyens,
- ensuite une responsabilité envers ceux qui nous donnent les ressources pour travailler,

De la même façon, le compte à rendre se découpe en « redevabilité vers l'amont » ou « *upward accountability* », c'est-à-dire vers les bailleurs de fonds, et de « redevabilité vers l'aval », ou « *downward accountability* », c'est-à-dire envers les bénéficiaires. Mais cette notion reste encore à interroger : de quoi devons exactement rendre compte aux bénéficiaires ? Parlons d'évaluation « participative »<sup>4</sup> ?

L'expression de cette responsabilité ne doit pas être vue seulement en terme de comptes à rendre. Les enjeux, tant pour les populations affectées par les crises que pour les donateurs (qu'ils sont privés ou publiques), sont d'abord ceux de la qualité de l'action et de son impact, puis celui du rendu. D'autres outils gèrent la question de la redevabilité : les rapports financiers annuels, certifiés par les Commissaires aux Comptes, le audits des bailleurs, les engagements pris par rapport au Comité de la Charte, etc. Les évaluations classiques de l'action humanitaire l'approchent par l'angle de l'évaluation de conformité (*compliance*), c'est-à-dire du respect des termes du contrat. Nous verrons plus loin les répercussions de cette définition en termes d'indicateurs.

Depuis quelques années, grâce à des travaux comme ceux du projet « People In Aid », il a été aussi ajouté que nous avons aussi une responsabilité envers nos propres équipes. La responsabilité par rapport à nos équipes se débat en terme d'équité de traitement, de transparence, de stratégie collective de formation, de politique de rémunération, de chance de promotion.

### 2. EVALUATION ET APPRENTISSAGE

La question de l'utilisation des évaluations comme processus d'apprentissage et de changement institutionnel est au cœur des agendas des responsables des services d'évaluation de nombreuses institutions. En effet, trop souvent, l'évaluation est vue comme d'abord de contrôle, donc craint, avant d'être vue comme un outil pour progresser. Mais l'expérience des nombreuses missions dont le rapport, qui arrivant souvent très tard après la fin du programme, reste sur une étagère et n'apporte pas vraiment de rétroaction (feedback) laisse souvent septique. Alors que les priorités devraient s'articuler autour du doublet « évaluer pour évoluer », les préoccupations tournent autour de la question « évaluer pour contrôler » ou « évaluer comme alibi ». Mais là encore, il y a eu déconnexion entre les leçons et l'action. Il est alors apparu important de rapprocher le résultat de l'évaluation de l'action elle-même. A partir de ce point de vue, plusieurs mécanismes sont nés (voir chapitre C)

<sup>2</sup> comme l'indique le texte « Responsibility in a nutshell », 2002, par le Groupe URD pour le Groupe de Travail Qualité de la Fédération Internationale de la Croix Rouge,

<sup>3</sup> Vous avez dit « évaluation », 1995, François Grünwald, document préparé pour le colloque « Evaluation of international humanitarian assistance », Amsterdam, disponible sur le site web du Groupe URD [www.urd.org](http://www.urd.org)

<sup>4</sup> Voir sur les évaluations participatives les travaux du Groupe URD pour ALNAP dans le cadre de l'Etude Globale sur la Participation des Populations Affectées, et notamment le « Guide du Praticien » (Practitioner Handbook) publié en 2003.

### 3. CRITERES, STANDARDS OU NORMES, INDICATEURS

#### 3.1. Les grandes familles de critères et d'indicateurs

Au stade actuel, un certain nombre de grandes familles de critères et d'indicateurs existent.

les critères d'évaluation du CAD /OCDE : Effectivité, Impact, Pertinence, Efficacité, Couverture, Cohérence, Durabilité. Ces critères ont été développés dans le cadre des actions de coopération pour le développement pour répondre aux besoins des bailleurs. Diverses tentatives se sont efforcées de les adapter à l'action humanitaire (ODI, 1999 ; Minear L, 1998, ALNAP, 2003).

Les Standards universels du Projet SPHERE : Il s'agit d'un système de critères, standards et indicateurs développés et promus sous l'influence d'un groupe d'ONG anglo-saxonnes, et par rapport auquel un certain nombre d'ONG françaises se sont mobilisées a sein d'une Plate-forme Qualité.

En phase de finalisation, les critères du COMPAS Qualité : Ce COMPAS Qualité (Critères et Outils pour la Mise en place et le Pilotage d'une ASistance humanitaire de Qualité) comporte tout un système de critères et outils pour piloter l'action (assurance qualité) et l'évaluer (en auto et/ou hétéro-évaluation).

Il est indispensable de tenir compte des travaux sur l'évaluation et la qualité développés dans différents secteurs, et de s'en inspirer en adaptant. Ainsi, des correspondances ont été cherchées entre le système classique en management (impact, processus, structure) et un système plus adapté à l'action humanitaire (impact, réalisation, management, situation).

Evaluation dans divers secteurs	Critères CAD	Evaluation dans le secteur humanitaire
Critères d'impact	Impact	Critères d'impact
	Durabilité	
	Pertinence	Critères de situation
Critères de Processus	Couverture	Critères de réalisation
	Effectivité	Critères de management
	Efficacité	
Critères de structure	Cohérence	Critères sur les capacités
		Critères sur les principes, les mandats,

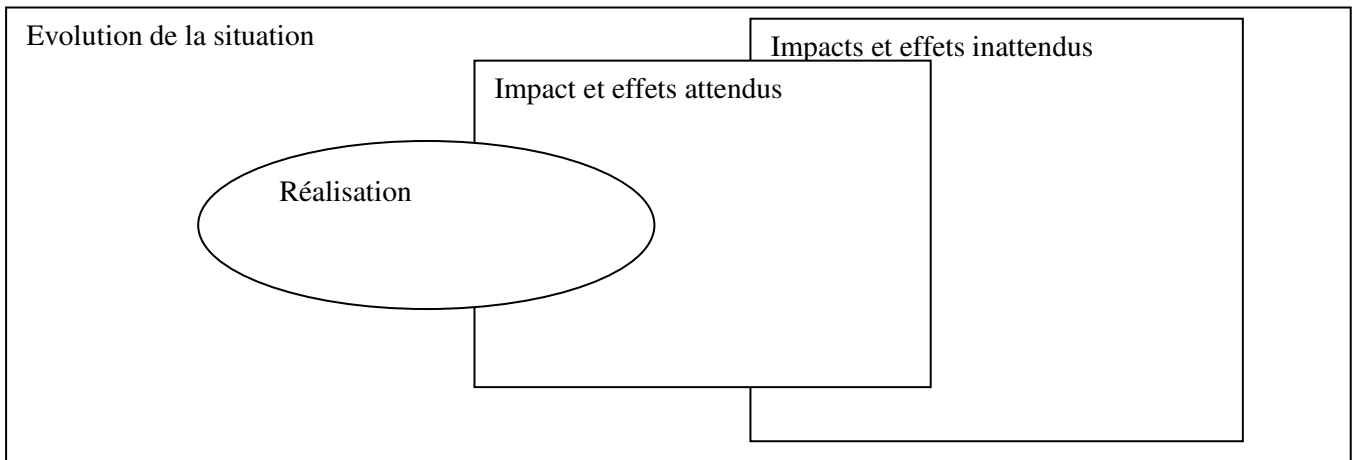
#### 3.2. Critères d'impact

L'impact global d'une intervention comprend ses effets attendus et inattendus, directs et induits. Ils peuvent positifs ou négatifs, apparaître rapidement ou au contraire avoir une lente gestation.

Les critères de qualité d'une intervention humanitaire et les indicateurs pour les mesurer peuvent être, pour des actions apparemment similaires, différents selon que l'on considère une intervention dans un contexte d'urgence aiguë, dans un contexte de crise chronique ou dans un contexte de sortie de crise. D'autre part, lorsqu'on évalue un programme sur plusieurs « pas de temps », on s'aperçoit que ce qui avait pu apparaître comme une caractéristique de qualité d'un programme à un temps donné peut devenir un indicateur de mauvaise adaptation dès que l'on se positionne sur d'autres délais d'observation. D'autre part enfin, comprendre l'impact d'une intervention demanderait théoriquement de pouvoir comparer « **situation avec et sans programme** », alors qu'au mieux, pour des raisons éthiques et pratiques, tout au plus l'évaluateur peut pratiquer une comparaison « **avant et après intervention** ».

La pratique de l'évaluation dans les contextes et sur les opérations humanitaires montre combien il est dangereux et improductif de séparer cette pratique du reste du management du cycle de projet. C'est en effet en début du cycle, lors des phases de diagnostic et de conception, que s'élaborera la stratégie de suivi et d'évaluation, et donc que se mettra en place la réflexion sur les indicateurs.

Ceci est d'autant plus important que l'action humanitaire trouvant sa place au sein de contextes fluctuants, volatiles, il n'est pas rare de voir des programmes évoluer profondément tant dans leurs objectifs que dans leur stratégie et dans la mise en place opérationnelle de celle-ci. La palette d'indicateurs prévue devra parfois être partiellement ou complètement revisitée. Ce débat a notamment entraîné toute une remise en cause de l'utilisation maintenant devenue mécanique du Cadre Logique.



### 3.3. Critères de réalisation

Ce sont les plus fréquemment utilisés dans les évaluations d'action humanitaire commanditées par les bailleurs. Il s'agit de vérifier si « on fait ce qu'on dit ». Les actions humanitaires étant souvent des actions relativement simples de logistique mises en place dans des cadres de temps relativement courts, il y a derrière l'utilisation de ces indicateurs trois hypothèses :

- une hypothèse de causalité : Elle transforme ces indicateurs de réalisation en « indicateurs de proximité » : on ne va pas mesurer l'impact d'un programme « puits » sur la santé, mais on fait l'hypothèse, sur la base de l'état des lieux des connaissances, qu'un programme X aura un effet Y (indicateurs de proximité). Il suffit alors de vérifier que les puits prévus sont bien faits, dans les délais prévus et dans l'enveloppe budgétaire allouée ;
- une hypothèse d'opacité : Elle fait des ONG des acteurs toujours à la marge de la malhonnêteté, en train de jouer avec l'argent des contribuables. Il faut donc contrôler que les contrats sont respectés;
- une hypothèse sur le contenu de l'exigence de redevabilité. A priori, les administrations sont comptables de l'argent des contribuables. Ces derniers sont perçus comme demandant de l'information sur ce qui est fait avec leurs impôts, et non de l'impact de ces réalisations.

On retrouve ici les indicateurs de conformité (compliance indicators), qui doivent mesurer l'écart entre ce qui est en engagement contractuel et la réalité des réalisations. Lorsque l'on utilise ce types d'indicateurs, on utilise rarement des pratiques d'évaluations participatives, et les formes pratiques de la mission prennent en général une forme « contrôlante » à la limite de l'audit financière.

### 3.4. Critères de management

Il s'agit là d'indicateurs devant regarder tout ou partie des procédures d'une organisation : gestion financière et des ressources humaines, stratégies de communication et modes de prises de décision. Les références méthodologiques utiles dans ce cadre sont celles présentes dans l'épais corpus d'écrits sur la théorie des organisations ainsi que sur le système ISO 9002. Un travail original a été tenté par deux ONG suisses MEDAIR et TDH. Il s'agit du test de l'utilisation des standards ISO dans le cadre de leurs institutions. Centré sur la gestion des procédures, ces expériences n'ont pas encore démontré qu'elles puissent avoir un impact mesurable sur le terrain. Mais il importe de continuer une veille sur cette tendance.



### 3.5. Critères de situation

L'action humanitaire se met en général en place dans des contextes à haute volatilité et parfois même dangerosité. Il importe de suivre ces évolutions au cours du programme, ce qui est l'objet d'une partie important du monitoring. Mais l'évaluation elle-même doit pouvoir jauger de la situation et de son évolution depuis l'élaboration du diagnostic initial. En effet, c'est seulement à l'aune de cette comparaison que l'on pourra comprendre les effets de l'intervention humanitaire. Il s'agit en tout cas d'éviter d'attribuer au programme humanitaire des bienfaits ou des difficultés qui sont le résultats d'évolutions de la situation complètement extérieures au programme lui même.

L'évaluation doit aussi tenter de comprendre les interactions réciproques et en général complexes entre l'action mise en place et l'évolution de la situation. Divers projets ont tenté de définir des méthodologies pour ce faire, et les indicateurs les accompagnant. On notera juste comme référence les divers projets de Mary Anderson (« Local Capacity for Peace », « Do no Harm », etc.), ainsi que les travaux récents du DFID britannique sur le thème de l'analyse des conflits (Peace and Conflict Assessments)

## 4. L'EVALUATION ET SA PLACE DANS LE CYCLE DE PROJET

Cette question repose :

- d'une part l'enjeu de l'évaluation précise de chacune des phases du cycle de projet ;
- d'autre part la question de la capitalisation ;
- et enfin celle des capacités d'apprentissage et de changement des institutions.

Le travail de l'évaluateur constitue en fait un exercice de « détricotage » du Cycle de Projet, pour rechercher, par une analyse des phases précédentes, des éléments explicatifs aux résultats obtenus et problèmes rencontrés. C'est la construction ex-post d'un « arbre à problème » explicatif. Cette démarche doit naturellement :

- faciliter le repérage des « points critiques » mal gérés ;
- aboutir sur des recommandations pour l'action elle-même ;
- permettre d'enregistrer les leçons tirées de l'expérience de façon ordonnée, pour faire de ces « leçons tirées » des « leçons utilisables », bref de faire d'une masse d'information parfois complexe un capital pour du « knowledge management » et du changement institutionnel.

## 5. INDEPENDANCE DES EVALUATEURS OU/ET EVALUATION UTILE

Ce sujet est assez controversé, car il est sensible et touche à la crédibilité de l'évaluation autant qu'à son impact en tant que pratique de changement. Quatre grandes écoles de pensée existent :

- celle qui prône l'indépendance des évaluateurs ;
- celle qui dit qu'au final, c'est la volonté de participer à l'amélioration des pratiques et à l'exercice de la responsabilité qui compte, pas nécessairement l'indépendance formelle ;
- celle qui voudrait développer les approches participatives de l'évaluation ;
- celle qui tente de promouvoir le mariage des savoir-faire, savoir être et positions

L'expérience montre que ce n'est pas si simple et que dans de nombreux cas, ce débat est au mieux stérile. Les évaluateurs indépendants sont en général liés à un marché : celui de l'évaluation. Hors la pratique de celle-ci est un exercice pouvant être périlleux, tant pour ceux qui la subissent que pour ceux qui la fond. De là la réticence régulièrement observée de certains évaluateurs dits « indépendants » d'aller jusqu'au bout. De plus, on opposera très souvent aux critiques de l'évaluateur extérieur son manque de compréhension du contexte et du mode de fonctionnement de l'institution. De toute façon, une fois remis son rapport, l'évaluateur extérieur est dégagé de tout « service après-vente ».

Ceci contribue à accentuer le manque d'effet réel des évaluations et le nombre de rapports inutilisés sur les étagères. Il en va tout autrement lorsque l'évaluateur interne est réellement motivé par le slogan « évaluer pour évoluer ». L'évaluation pourra alors être très pointue, menée avec une bonne connaissance de la culture institutionnelle, et voir ses résultats suivis d'effets.

## 6. EVALUATION AVEC OU PAR LES BENEFICIAIRES

Plusieurs approches méthodologiques ont été testées sur le terrain pour essayer de comprendre la vision que les bénéficiaires et populations affectées pouvaient avoir des programmes humanitaires, et notamment :

- les techniques de l'audit social ;
- l'enquête type « enquête de satisfaction »;
- l'évaluation participative ;

La première s'avère longue et coûteuse (plusieurs mois sont souvent nécessaires pour l'élaboration du protocole, la collecte de l'information et son traitement). La seconde s'avère très sensible à de nombreux biais : importance du contexte où a lieu le recueil de l'information, de l'empathie qui se crée avec l'interviewer, etc.

Reste au final les enjeux de l'évaluation participative, autour de laquelle on peut dénouer la plupart des questions autour de :

- la complexité des relations sociales et des jeux de pouvoir ;
- la complexité de ce qu'est « la satisfaction ». L'évaluation participative re-interroge en effet la relation entre l'acteur et les populations affectées de façon globale sur l'ensemble des phases du cycle de projet.
- la complexité de ce que représente « donner une voix aux sans voix ».

La pratique d'évaluation participative est très rare dans le domaine de l'action humanitaire : la culture des acteurs humanitaires a en effet plutôt tendance à considérer les populations pour lesquelles ils travaillent soit comme des victimes impotentes, soit comme des escrocs tendant d'obtenir toujours plus d'aides par des voies détournées. La pratique de l'évaluation participative dans des contextes de crise doit en tout cas prendre en compte les difficultés et parfois dangers inhérents, et notamment ceux qui créent des problèmes de sécurité tant pour les populations que pour les évaluateurs.

## C. LES SYSTEMES EXISTANTS

Différents processus, mécanismes et structurations pour procéder à des évaluations et recherches évaluatives ont été mis en place soit par les bailleurs, soit par les ONG, soit par des institutions particulières à statut variés. Chacun d'entre eux focalise sur un certain nombre d'enjeux.

### 6. LES PROCESSUS ET MECANISMES EVALUATIFS

7 processus différents sont actuellement mis en œuvre dans différents contextes et par différentes structures :

- Les « Revue post-opération » : Une initiative intéressante est celle des « *post action reviews* » ou « revues post opération ». L'outil développé, qui prend son origine dans une pratique militaire, est World Vision organise, dans les semaines ou mois après le début d'une opération, un « atelier à chaud » de mise en commun des erreurs réalisées, des leçons apprises, dont les résultats sont à la fois réinjectés dans l'opération en cours et dans la dynamique institutionnelle en amont. C'est essentiellement un outil de travail collectif en « interne ».
- Les évaluations en temps réel : Exercice finalement relativement proche du monitoring, mais en général conduit par une équipe extérieure à l'action (souvent des bureaux régionaux ou du siège, les premières tentatives d'évaluation en temps réel (real time evaluation) ont émergé, notamment dans le cadre du HCR. Il s'agit de visites évaluatives sur le terrain en cours d'action, pour pouvoir permettre des réactions nécessaires dans la conduite stratégique des opérations. Ce genre de processus résulte en général d'une demande du siège, pour pouvoir à distance « guider » une opération qui est soit fortement critiquée, soit très médiatisée.
- Les « bureaux d'apprentissage » : Cette innovation ou Learning Office a été conçue dans le cadre du réseau ALNALP (Active Learning Network for Accountability and Performance). La faisabilité de ce processus a été bien étudiée sur plusieurs cas concrets (Inde avec le cycle sur l'Orissa, Timor, la Sierra Leone, le Kosovo). Ces travaux ont abouti à une proposition qui, après un long débat et diverses hypothèses, a conduit à la mise en place d'un processus pilote au Malawi en 2002, dans le contexte de la crise alimentaire d'Afrique Australe. Un des grands enjeux de cet exercice et de voir comment il peut se mettre en place en parallèle mais sans duplication avec d'autres institutions, comme OCHA.
- Les évaluations intermédiaires : Processus en général dirigé par les sièges, ces évaluations « à mi-parcours » ont en général une problématique spécifique : soit il s'agit de corriger un problème grave identifié, mais qui demande une « validation » ; soit il s'agit de se préparer à un changement important qui demande une préparation conséquente : décision de clore le programme ou de passer à une nouvelle phase, par exemple.
- Les évaluations finales : ces dernières sont de loin les plus nombreuses, et sont en général commissionnées par les bailleurs. Elles sont la plupart du temps réalisées par des équipes externes ou mixtes sur appels d'offre. Ce sont les outils principaux du « contrôle top down ». Elles sont aussi un outil très puissant de création des mémoires institutionnelles, pour autant que le rapport est utilisé.
- Les évaluations itératives avec mini-séminaires (EIMS) : Ce processus commença à être mis au point et testé par le Groupe URD de Novembre 1998 à Mars 2002 pour tirer les leçons des actions humanitaires après l'ouragan Mitch et les tremblements de terre au Salvador. Processus complexe liant échanges, coordination des interrogations, missions d'évaluation régulièrement envoyées sur le terrain et processus complexe de retro-action envers à la fois les ONG locales et les acteurs internationaux. La phase finale a été une série d'ateliers techniques et de conférences en France, en Suisse, aux Etats Unis, des articles dans la presse spécialisée, et un certain nombre d'idées sur comment améliorer les processus d'apprentissage.

- La combinaison «évaluation-recherche action» : Cette approche a été expérimentée par différents acteurs au niveau international. L'acteur le plus avancé sur ces questions est le HPG (Humanitarian Practice Group) de l'Overseas Development Institute (ODI) de Londres. Fortement soutenu par l'administration britannique et par les grandes ONG anglaises (dont la plupart des staff d'HPN/ODI sont issus), HPG/ODI est devenu maintenant le principal « think tank » sur le sujet, avec une reconnaissance et une capacité de mobilisation de ressources internationales importantes. De facto, HPG/ODI est en situation de quasi-monopole, et la pensée anglo-saxonne modèle de façon déterminante la pensée du « système humanitaire internationale ». En France, le Groupe URD tente d'émerger comme structure alternative. Il a conduit de nombreuses recherches sur le terrain. La faiblesse du support qu'il a eu jusqu'ici des pouvoirs publics français et des ONG françaises, dont la faible volonté de travail en collectif est bien connue. Ceci est apparemment en train de changer, grâce à l'émergence de Coopération SUD dans le secteur humanitaire.

## 7. LES SYSTEMES INSTITUTIONNELS

Les grands bailleurs et institutions internationales impliqués dans l'humanitaire ont mis en place plusieurs types de mécanismes, en fonction de la place qui est donnée à l'action humanitaire dans les organigrammes :

- ***Services de l'Action Humanitaire et du Développement dans la même administration*** : dans ce type de cas, il est fréquent, mais pas la règle absolue, d'avoir un bureau d'évaluation commun.
- ***Services de l'Action humanitaire et Services de Développement sont dans des administrations différentes*** : Dans ce cas, la règle est d'avoir pour chacun son propre service d'évaluation.
- ***Les structures sont spécialisées dans l'humanitaire***. Certaines ont une Unité des Evaluations (par exemple le CICR, ou OCHA). L'une des questions qui se posent régulièrement est la place que cette unité doit avoir et son rattachement institutionnel : au plus proche de la Direction, au niveau des services techniques, « hors organigramme » ? Un nombre encore non négligeable d'ONG de tailles importantes et moyennes n'ont pas encore ni de service d'évaluation, ni même fait les premiers pas dans la direction de la création d'une « fonction évaluation » institutionnalisée.

Trois grandes questions se posent par rapport à ces dispositifs :

### Comment se déclenchent les processus évaluatifs ?

Divers types de mécanismes activent les procédures évaluatives :

- une disposition administrative : Ainsi, pour ECHO, les directives communautaires rendent obligatoire la mise en œuvre régulière d'évaluation ;
- une préoccupation de « rendu au public ». Le Disaster Emergency Committee (DEC) britannique, qui coordonne la mobilisation des fonds publics et les campagnes de « fund raising » pour les ONG britanniques, impose aux ONG récipiendaires une évaluation collective afin de rendre compte de l'utilisation des sommes collectées. Les résultats de ces évaluations sont rendus publics par le web et parfois repris dans la presse qui a participé aux campagnes de collecte ;
- une envie d'apprendre. Sur des situations compliquées ou des programmes complexes, des acteurs humanitaires peuvent déclencher des processus évaluatifs pour tirer les leçons et améliorer pratiques et fonctionnements institutionnels. De tels exercices peuvent être lancés par une institution (évaluation en temps réel du HCR sur la crise afghane en 2001-2002) ou par plusieurs (évaluation multi-bailleurs sur le Rwanda en 1995-96, évaluation lancée par le Groupe URD au sein de la Task-force Mitch en 1998-99.).

### Comment passe-t-on de l'évaluation à la capitalisation et aux processus d'apprentissage ?

Un travail collectif dans le cadre d'ALNAP a abouti à une réflexion stratégique sur les enjeux de l'évaluation pour l'apprentissage. Ils démontrent que les enjeux de l'apprentissage ont été en grande partie laissés de côté, tant grande était la pression des bailleurs en faveur de la composante « accountability ».

Comment facilite t'on le développement d'une culture d'évaluation :

Il s'avère qu'il faut un mélange subtil et puissant de carottes et de bâtons... La mise à disposition de moyens pour mettre en œuvre ces démarches est en tout cas une option qui semble efficace. La possibilité introduite dans le format actuel des dossiers de financement d'ECHO, et, de façon général, des outils financiers communautaires ouvrent des opportunités importantes. A voir comment les ONG les utiliseront.

ANNEXE 1**REFERENCES DE L'AUTEUR SUR LES QUESTIONS D'EVALUATION**QUELQUES EXEMPLES D'EVALUATIONS EX-ANTE REALISEES

Diagnostic initial pour le PAM et la FAO au Cambodge (1985-87)

Diagnostic initial pour le GRET au Cambodge (1988)

Diagnostic initial pour le HCR au Cambodge (1991)

Diagnostic initial pour le CICR au Mozambique, Rwanda, Angola, Balkans, Caucase, Afghanistan, (1992-96)

Diagnostic initial pour le HCR au Kosovo (1998)

Diagnostic initial pour la FAO au Kosovo (1999)

Diagnostic initial pour le programme de développement rural de la province de Sayaboury-Laos pour l'AFD (1997)

Diagnostic initial pour la FAO en Afghanistan (2001)

EVALUATION EN TEMPS REEL , INTERMEDIAIRES ET EIMS REALISEES

Evaluation intermédiaire du Programme de Développement Rural de Phnogsaly –Laos- pour l'AFD 1999

EIMS des programmes humanitaires et de reconstruction post-Mitch Amérique Central (1999-2001)

EIMS des programmes humanitaires et de reconstruction en Afghanistan (2002-2003)

Evaluation des programmes humanitaires et de reconstruction en Sierra Leone (2003)

QUELQUES EXEMPLES D'EVALUATIONS EX-POST REALISEES

Evaluation des programmes des Nations Unies au Cambodge , Bureau du Représentant Spécial du Secrétaire Général de l'ONU pour les Affaires Cambodgiennes ; 1987

Evaluation des programmes de coopération française en agriculture au Népal (1992)

Evaluation des programmes de sécurité alimentaire du CICR en Somalie (1993), Mozambique (1994), dans le Caucase (1996)

Evaluation des programmes de prévention du risque lié aux mines anti-personnel du CICR en Azaibaijan (1997)

Evaluation de la gestion de la crise du Bahr-El-Gazhal au Soudan (1998)

Evaluation des programmes de relance agricole de la FAO au Burundi et au Rwanda (1999)

PUBLICATIONS SUR LA QUESTION DE LA QUALITE ET DE L'EVALUATION

GRUNEWALD F.; 1998; "Situation analysis, Need assessment and programme evaluation", in *Mine awareness for Children : a discussion of Good practice*; Rädda Barnen, Sweddish Save the Children; pp 13-25

GRUNEWALD F. et de GEOFFROY V., 1999 ; “ Drawing the lessons from Kosovo ” ; in *RRN Bulletin*, ODI, London

GRUNEWALD F et al; 2000; Drawing the lessons from Mitch; HPN N°34; ODI/ London