

INTRODUCTION AUX MÉTHODES MIXTES DANS L'ÉVALUATION D'IMPACT

Michael Bamberger



Ceci est la troisième d'une série de quatre notes sur l'évaluation d'impact, élaborées par InterAction avec l'appui financier de la Fondation Rockefeller. Les autres notes de cette série sont : *Introduction à l'évaluation d'impact* ; *Faire le lien entre le suivi-évaluation et l'évaluation d'impact* et *Utilisation des résultats d'évaluations d'impact*. On peut trouver la série complète sur le site Internet d'InterAction : <http://www.interaction.org/impact-evaluation-notes>.

Photo: Alissa Everett

TABLE DES MATIÈRES

Introduction	1
1^{ère} partie : Pourquoi des méthodes mixtes ?	3
1.1. Qu'est-ce qu'une évaluation d'impact avec des méthodes mixtes ?	3
1.2. Les limites d'une utilisation exclusive d'approches QUANT ou QUAL pour l'évaluation	4
1.3. Les avantages d'une approche de méthodes mixtes	4
2^{ème} partie : L'approche des méthodes mixtes	9
2.1. Quatre décisions pour concevoir une évaluation à méthodes mixtes	9
2.2. Utiliser des approches MM à chaque phase de l'évaluation	13
3^{ème} partie : Applications des modèles à méthodes mixtes	19
3.1. Stratégies d'échantillonnage pour des évaluations MM à orientation QUANT ou QUAL	19
3.2. Utiliser des méthodes mixtes pour évaluer des interventions complexes	21
3.3. Évaluer les processus de changement de comportement	24
4^{ème} partie : Gérer les évaluations à méthodes mixtes	27
4.1. Les modèles à méthodes mixtes nécessitent une approche de gestion spéciale	27
4.2. Conseils pour permettre aux ONG ayant des ressources limitées de mobiliser l'expertise et les ressources nécessaires pour faire des évaluations à méthodes mixtes	30
5^{ème} partie : Études de cas illustrant différentes applications de modèles à méthodes mixtes	32
Références	37

TABLEAUX

Tableau 1. Les méthodes mixtes sont utilisées différemment pour des modèles d'évaluation ayant une orientation principalement QUANT ou QUAL	11
Tableau 2. Techniques de collecte de données QUANT et QUAL souvent utilisées	16
Tableau 3. Différents types de triangulation utilisés dans les évaluations à méthodes mixtes	17
Tableau 4. Exemples d'analyse des données de méthodes mixtes	18
Tableau 5. Autres domaines d'expertises dans la recherche, dont peuvent avoir besoin des équipes à orientation QUANT et QUAL pour faire des évaluations MM	29
Tableau 6. Exemples d'évaluations MM résumées en Annexe 10	36

CADRES

Cadre 1. Comment on utilise « impact » et « évaluation d'impact » dans les notes d'orientation	3
Cadre 2. Avantages opérationnels de l'utilisation de méthodes mixtes	5
Cadre 3. Analyse d'attribution, de contribution et de substitution	22
Cadre 4. Les effets des programmes sont affectés par la manière dont les bénéficiaires réagissent aux services et dont le personnel réagit aux réactions des bénéficiaires. Étude du changement de comportement dans un programme de cantines scolaires au Nicaragua.	25
Cadre 5. Planifier et budgétiser le temps et les ressources supplémentaires qui pourront être nécessaires pour une évaluation à méthodes mixtes	28
Cadre 6. Faire des évaluations d'impact à méthodes mixtes avec peu de moyens : Conseils pour les ONG qui ont peu de ressource et d'expertise en évaluation	31

FIGURES

Figure 1. Utiliser un cadre national d'échantillonnage des ménages pour garantir la représentativité et la crédibilité d'une évaluation QUAL par étude de cas	6
Figure 2. Utiliser des études ethnographiques QUAL dans les villages et les ménages pour aider à interpréter les résultats d'une enquête nationale QUANT auprès d'un échantillon représentatif	8
Figure 3. Modèle de méthodes mixtes consécutives avec une approche quantitative dominante : Étude des transferts entre les ménages comme stratégie de survie pour les ménages à faible revenu de Cartagena en Colombie	12
Figure 4. Modèle de méthodes mixtes consécutives avec une approche qualitative dominante : évaluation de l'adoption de nouvelles variétés de semences par différents types de paysans	12
Figure 5. Modèle de méthodes mixtes sur plusieurs niveaux : Évaluation des effets d'un programme de cantines scolaires sur la fréquentation et les résultats	14
Figure 6. Utiliser des modèles à méthodes mixtes pour évaluer des interventions complexes	23

ANNEXES (disponibles en anglais sur : <http://www.interaction.org/impact-evaluation-notes>)

Annexe 1. <u>Points forts et faiblesses des modèles quantitatifs d'évaluation</u>
Annexe 2. <u>Points forts et faiblesses des modèles qualitatifs d'évaluation</u>
Annexe 3. <u>Exemples de modèles d'évaluation à chaque point du continuum QUANT-QUAL</u>
Annexe 4. <u>Caractéristiques des approches QUANT et QUAL à différentes phases de l'évaluation</u>
Annexe 5. <u>Comment les approches QUANT et QUAL se complètent à différentes étapes d'une évaluation</u>
Annexe 6. <u>Comparaison des méthodes aléatoires et ciblées d'échantillonnage</u>
Annexe 7. <u>Une gamme d'approches quantitatives, qualitatives et théoriques pour définir le scénario contrefactuel</u>
Annexe 8. <u>Stratégies pour réduire les coûts de la collecte et de l'analyse des données</u>
Annexe 9. <u>Exemple de triangulation : comparer les estimations du revenu des ménages et de la pauvreté venant de sources différentes</u>
Annexe 10. <u>Études de cas de modèles d'évaluation MM avec des orientations principalement QUANT, QUAL et équilibrées</u>
Annexe 11. <u>Comment les méthodes mixtes peuvent renforcer les modèles d'évaluation QUANT</u>

Introduction

Les évaluations à méthodes mixtes (MM) cherchent à intégrer les disciplines de sciences sociales dans des approches essentiellement quantitatives (QUANT) ou essentiellement qualitatives (QUAL) de la théorie et de la collecte, de l'analyse et de l'interprétation des données. Leur objectif est de renforcer la fiabilité des données, la validité des constatations et des recommandations et d'élargir et d'approfondir notre compréhension des processus menant aux résultats et aux impacts des programmes et de la manière dont ceux-ci sont affectés par le contexte dans lequel le programme est mis en œuvre. Même si les méthodes mixtes sont maintenant largement utilisées dans l'évaluation et que les appels à proposition (RFP) d'évaluation exigent souvent leur utilisation, de nombreux évaluateurs n'utilisent pas le plein potentiel de l'approche MM.

Cette note d'orientation explique les éléments essentiels de l'approche MM et la manière dont elle peut être utilisée dans une évaluation d'impact (EI) tout en soulignant ses applications et ses avantages possibles pour les ONG. La 1^{ère} partie pose la question : « Pourquoi des méthodes mixtes ? » Nous y présentons ce qu'est un modèle MM d'évaluation d'impact, ce qui le distingue d'une évaluation d'impact QUANT ou QUAL et pourquoi cette approche aide à comprendre les évaluations du développement et les complexités de la réalité dans laquelle elles sont mise en œuvre (section 1.1). Les MM sont devenues plus populaires parce qu'on s'est rendu compte que se baser

uniquement sur des méthodes QUANT ou QUAL avait des limites (section 1.2) et qu'une combinaison appropriée des deux approches avait des avantages (section 1.3). Les MM peuvent être utilisées dans le cadre d'une grande évaluation d'impact avec un financement important, mais elles sont suffisamment souples pour pouvoir être tout aussi utiles pour de nombreuses ONG qui ont besoin d'évaluations crédibles de leurs programmes mais n'ont que des ressources et qu'une expertise limitées pour faire des évaluations d'impact.

Après ces arguments en faveur de l'approche MM, la 2^{ème} partie décrit quatre décisions clés qui

doivent être prises quand on prépare une évaluation MM (section 2.1) : à quelles étapes de l'évaluation utilisera-t-on des MM ? Est-ce que les MM sont utilisées de manière consécutive ou simultanée ? L'évaluation sera-t-elle essentiellement QUANT, essentiellement QUAL ou utilisera-t-elle une méthode équilibrée qui donne le même poids aux deux approches ? Cette conception sera-t-elle mise en œuvre sur un seul niveau ou s'agira-t-il d'une évaluation à plusieurs niveaux ? La section 2.2 explique comment les MM peuvent être utilisées à chaque phase de la conception, de la mise en œuvre et de l'analyse d'une évaluation, et non pas seulement comme une manière de diversifier les méthodes de collectes de données, comme le pensent beaucoup d'évaluateurs.

La 3^{ème} partie présente quelques applications des approches MM. La section 3.1 étudie des stratégies d'échantillonnage quand on utilise principalement des méthodes QUANT ou QUAL et montre comment les MM peuvent renforcer les deux types de choix d'échantillon. La section 3.2 présente l'utilisation des MM pour évaluer des interventions complexes de développement et la section 3.3 explique comment les méthodes MM peuvent aider à évaluer des programmes comprenant des processus de changement de comportement.

La 4^{ème} partie s'intéresse à des questions de gestion des évaluations MM. La section 4.1 explique pourquoi il faut une approche spéciale de la gestion et la section 4.2 montre comment les ONG peuvent utiliser efficacement les MM pour des évaluations qui sont faites avec peu de moyens (c'est à dire avec des contraintes budgétaires et de temps et avec une expertise limitée en recherche).

La 5^{ème} partie présente trois études de cas illustrant la manière dont les MM sont utilisées dans des évaluations principalement QUANT ou QUAL et dans une évaluation équilibrée donnant un poids égal aux approches QUANT et QUAL. L'[Annexe 10](#) présente 17 exemples d'évaluations MM illustrant une large gamme d'approches et comprenant aussi bien de grandes évaluations avec un bon financement que des évaluations faites avec des contraintes de budget, de temps et de données.

L'une des difficultés de la préparation de cette note d'orientation était le fait qu'il y a énormément de littérature sur les MM, dont certains ouvrages très techniques ou très spécialisés. Nous avons essayé de produire un texte qui soit accessible à un large public non spécialiste tout en fournissant des annexes (disponibles en anglais sur <http://www.interaction.org/impact-evaluation-notes>) qui donnent plus de détails.

1^{ère} partie : Pourquoi des méthodes mixtes ?

1.1. Qu'est-ce qu'une évaluation d'impact avec des méthodes mixtes ?

Différents évaluateurs utilisent les termes « impact » et « évaluation d'impact » différemment. Pour que nous soyons bien tous sur la même longueur d'ondes, le Cadre 1 résume la manière dont ces termes sont utilisés dans les notes d'orientation. Les méthodes mixtes ont les mêmes objectifs, posent beaucoup des mêmes questions, se servent de tous les outils et techniques d'évaluation d'impact qui sont décrits dans la première note d'orientation de cette série (N°1) et se basent sur les systèmes de suivi et évaluation décrits dans la N°2. Comme d'autres modèles d'évaluation d'impact, les MM peuvent être appliquées à n'importe quel niveau, depuis l'évaluation d'un projet travaillant dans un seul village jusqu'à une initiative nationale de développement comportant plusieurs composantes et impliquant de nombreuses agences internationales et nationales.

Il n'arrive que rarement qu'une seule méthodologie d'évaluation puisse saisir toutes les complexités de comment les programmes opèrent dans la réalité. Les évaluateurs doivent donc trouver des manières créatives de combiner différents cadres, outils et techniques d'évaluation¹ – ce qui explique l'intérêt croissant pour les approches MM. Ce qui

¹ Le choix du modèle d'évaluation approprié est un sujet important aussi. Étant donné le grand nombre de types de programmes différents qui sont évalués, les contextes variés dans lesquels ils opèrent et la diversité des questions d'évaluation qui intéressent les parties prenantes, il n'existe pas un seul « meilleur » modèle d'évaluation qui fonctionnera dans toutes les situations. Le choix de modèle d'évaluation demande une analyse approfondie de la nature du programme, de l'objectif et du contexte de l'évaluation et de l'environnement où elle est faite. Vous trouverez dans Bamberger, Rugh et Mabry (2012), aux chapitres 2 et 11 une présentation des stratégies permettant de sélectionner le modèle d'évaluation le plus approprié. Cf. aussi Stern et al (2012) pour une présentation en détail du choix de méthodes.

est particulier aux approches de méthodes mixtes, c'est qu'elles cherchent à intégrer les sciences sociales avec des approches principalement QUANT ou principalement QUAL de la théorie, de la collecte, de l'analyse et de l'interprétation des données. Même si beaucoup d'évaluateurs utilisent actuellement des méthodes variées, « Ce qui distingue l'évaluation à méthodes mixtes, c'est l'utilisation intentionnelle ou planifiée de méthodes diverses pour des objectifs particuliers aux méthodes mixtes, en utilisant des modèles particuliers aux méthodes mixtes » (Greene, 2005 : 255). La plupart du temps, des méthodes de collecte de données sont combinées pour faire une évaluation MM, mais il est aussi possible de combiner des cadres conceptuels, l'élaboration d'hypothèses, l'analyse des données ou des cadres pour l'interprétation des résultats de l'évaluation.

CADRE 1. COMMENT ON UTILISE « IMPACT » ET « ÉVALUATION D'IMPACT » DANS LES NOTES D'ORIENTATION

La Note d'orientation 1, « Introduction à l'évaluation d'impact » (page 2), définit les *impacts* comme :

les effets positifs et négatifs, prévus et imprévus, directs et indirects, primaires et secondaires, produits par une intervention. (Définition du Comité d'assistance au développement de l'OCDE)

Les impacts sont généralement vus comme apparaissant plus tard que les effets intermédiaires, et résultant de ceux-ci. La distinction entre effets et impacts peut être relative et dépend des objectifs fixés pour l'intervention. Selon notre définition, une évaluation est :

n'importe quelle évaluation qui étudie de manière systématique et empirique les impacts produits par (qui peuvent être attribués à) une intervention.

1.2. Les limites d'une utilisation exclusive d'approches QUANT ou QUAL pour l'évaluation

Quand elles sont utilisées isolément, les deux méthodes, QUANT comme QUAL, ont des points forts et des faiblesses. L'objectif des MM est de s'appuyer sur les points forts des approches QUANT comme QUAL et de les intégrer pour surmonter leurs faiblesses. Même si les méthodes QUANT de collecte et d'analyse des données ont beaucoup d'avantages intéressants, elles ont aussi un certain nombre de limites ([Annexe 1](#)). Beaucoup critiquent le fait de réduire des données narratives à des nombres et la rigidité des modèles et procédures de collecte de données qui sont difficiles à adapter à des circonstances changeantes. Les catégories standardisées des questionnaires et le codage des données ne réussissent souvent pas à saisir les nuances dans les groupes ou les communautés étudiées et il manque à l'analyse la profondeur et les détails des méthodes QUAL. Une évaluation QUANT risque d'être *décontextualisée*, d'ignorer la manière dont les programmes sont affectés par les caractéristiques économiques, politiques, institutionnelles et socioculturelles des populations étudiées. Une autre critique fréquente de beaucoup d'évaluations QUANT est qu'elles supposent que le programme opère comme prévu et que tout le monde reçoit le même ensemble de services (aussi bien en termes de qualité que de quantité). On parle souvent d'approche « de la boîte noire » car l'évaluation ne regarde pas à l'intérieur de la « boîte noire » du projet.

Les méthodes QUAL sont aussi des outils intéressants pour la collecte et l'analyse des données. Mais quand elles sont utilisées seules, les évaluations QUAL ont aussi un certain nombre de faiblesses potentielles ([Annexe 2](#)). Les évaluations QUAL s'intéressent souvent à des sujets et des situations individuels et il est plus difficile de

généraliser les résultats. De nombreux évaluateurs QUAL pensent aussi que chaque évaluation est spécifique au contexte et qu'il n'est ni possible ni souhaitable de faire des généralisations. De nombreuses évaluations QUAL, mais certainement pas toutes, utilisent une approche holistique, ce qui fait que les éléments et facteurs individuels sont plus durs à isoler et qu'il est plus difficile de comprendre la contribution particulière de différentes composantes ou approches du programme. Certains clients sont aussi mal à l'aise avec le fait qu'on semble trop dépendre de l'opinion et la perspective de l'évaluateur, sans que le lecteur ne puisse étudier facilement la masse de données écrites et enregistrées sur lesquelles s'est basé l'évaluateur. Un dernier point est que beaucoup d'évaluations QUAL ne donnent pas le type de documentation détaillée sur la méthodologie qui est généralement présenté dans les rapports d'évaluations QUAL, ce qui fait qu'il est difficile de vérifier la validité des procédures de collecte et d'analyse des données.²

1.3. Les avantages d'une approche de méthodes mixtes³

Il y a cinq raisons principales d'utiliser des modèles à méthodes mixtes (Greene 2005:255–56) :

- *Triangulation des résultats de l'évaluation* : augmenter la validité ou la crédibilité des résultats de l'évaluation en comparant les informations obtenues par différentes méthodes de collecte

² Par exemple, de nombreuses évaluations QUAL ne comportent pas de documentation détaillée sur la manière dont les membres de groupes de consultation ont été sélectionnés et rares sont celles qui fournissent les transcriptions des entretiens (pour des raisons de coût et de temps) de sorte que le lecteur n'a généralement pas la possibilité d'étudier indépendamment les données pour voir comment on en est arrivé aux résultats et aux conclusions.

³ Pour une étude récente des avantages des approches de méthodes mixtes, cf. Adato (2012).

de données (par exemple en comparant les réponses aux questions d'une enquête avec ce que l'enquêteur a observé directement). Quand des estimations provenant de différentes sources convergent et concordent, cela augmente la validité et la crédibilité des résultats ou de l'interprétation. Quand les différentes estimations ne sont pas cohérentes, l'enquêteur continue son étude pour comprendre les raisons des incohérences (cf. [Annexe 9](#)).

- *Élaboration* : utiliser les résultats d'une méthode pour aider à élaborer l'échantillon ou les instruments pour une autre.
- *Complémentarité* : augmenter l'exhaustivité des résultats de l'évaluation grâce aux résultats de différentes méthodes qui permettent une meilleure compréhension, plus approfondie.
- *Initiation* : générer de nouvelles perspectives dans les résultats de l'évaluation grâce aux

résultats de différentes méthodes qui divergent et demandent donc une nouvelle analyse, un recadrage ou un changement de perspective pour pouvoir être réconciliés.

- *Diversité des valeurs* : incorporer une plus grande diversité de valeurs grâce à l'utilisation de différentes méthodes qui, elles-mêmes, promeuvent des valeurs différentes. Ceci encourage à prendre mieux conscience des dimensions de valeurs de l'évaluation.

Un avantage supplémentaire est qu'une approche MM a plus de chance d'être acceptée à la fois par les évaluateurs et utilisateurs qui préfèrent QUANT et par ceux qui préfèrent QUAL.

Le cadre 2 résume certains des avantages opérationnels de l'utilisation d'une approche MM.

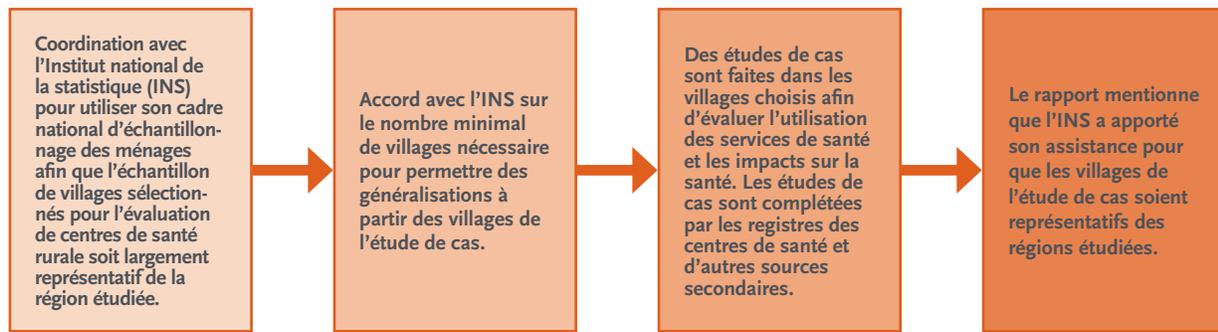
CADRE 2. AVANTAGES OPÉRATIONNELS DE L'UTILISATION DE MÉTHODES MIXTES

- Comprendre comment les facteurs contextuels aident à expliquer les variations dans la mise en œuvre et les effets du programme.
- Reconstituer les données de base pour les évaluations QUANT quand il n'a pas été possible de faire une étude de base. Beaucoup d'évaluations sont commandées vers la fin du programme et ne disposent pas d'informations très fiables sur les conditions du projet ni de groupe de comparaison au moment où le programme a commencé. Il est donc difficile de déterminer si les différences observées à la fin du projet peuvent être attribuées aux effets du programme ou si ces différences pourraient être dues, au moins en partie, à des différences préexistantes entre les deux groupes. Par exemple, des femmes qui demandent de petits prêts viennent de famille qui acceptent mieux que la plupart des familles qu'une femme soit propriétaire d'une petite entreprise, ou elles ont déjà plus d'expérience des affaires que les femmes qui ne demandent pas de prêts. Si ces différences préexistantes ne sont pas identifiées, on risque de surestimer les effets du programme de prêts. Il est souvent

possible d'utiliser des techniques QUAL comme des entretiens approfondis, des entretiens avec des informateurs clés ou des groupes de consultation pour obtenir des informations sur les caractéristiques des bénéficiaires et des non bénéficiaires du programme au moment où le programme a commencé. Ce type d'informations, qui est souvent facile et économique à collecter, peut beaucoup améliorer la validité de données d'enquêtes uniquement QUANT.

- En renforçant la représentativité d'études QUAL approfondies (par exemple en reliant la sélection des études de cas avec le cadre d'échantillonnage QUANT), il peut être plus facile de comparer les résultats avec les données d'enquêtes QUANT.
- Donner une bonne idée de la validité et de la valeur des différents types de données QUANT et QUAL.
- Promouvoir une meilleure compréhension des perspectives des parties prenantes sur la nature de l'intervention et comment elle est censée atteindre ses objectifs. Cela permet une approche plus participative et un meilleur alignement entre les parties prenantes et les évaluateurs.

Figure 1 Utiliser un cadre national d'échantillonnage des ménages pour garantir la représentativité et la crédibilité d'une évaluation QUAL par étude de cas



Pour illustrer certains de ces avantages, prenons l'exemple de l'évaluation de l'impact de centres ruraux de santé sur la santé de la population rurale, en particulier les femmes et les enfants. L'accessibilité des services de santé pour les secteurs pauvres et vulnérables de la population est une préoccupation particulière. Dans ce premier exemple, l'évaluation adopte une **approche** principalement **QUAL**. L'évaluation vise à influencer les politiques sanitaires nationales en identifiant certaines des raisons qui font que les familles pauvres ou faisant partie de minorités n'utilisent pas les centres de santé. L'équipe d'évaluation sait que le Ministère de la Santé a critiqué de précédentes évaluations QUAL en disant qu'elles s'étaient intentionnellement intéressées à des communautés connues pour des problèmes particuliers, et que les résultats n'étaient donc pas représentatifs du pays dans son ensemble et donneraient une mauvaise image du Ministère de la Santé. Les évaluateurs savaient aussi que le ministère utilisait parfois le manque de représentativité statistique comme une bonne excuse pour ignorer des critiques valables et ils voulaient être sûrs que leur étude ne serait pas rejetée pour ces raisons. Par conséquent, l'équipe d'évaluation travaille en coordination avec l'Institut national de la statistique et utilise son cadre national d'échantillonnage des ménages pour sélectionner un échantillon de

communautés qui soit largement représentatif du pays (ou de la région où l'étude est menée). L'évaluation utilise les mêmes méthodes QUAL mais il est maintenant possible d'indiquer que l'échantillon de communautés est largement représentatif de toutes les communautés de la région étudiée (cf. Figure 1).

Supposons maintenant que la même évaluation doit être menée par une équipe différente qui prévoit d'utiliser une **approche QUANT** basée sur une enquête auprès d'un échantillon national représentatif des ménages. Une enquête bien conçue peut permettre d'obtenir des estimations raisonnablement fiables de la proportion de la population qui utilise les centres de santé (bien qu'il puisse y avoir de fausses déclarations) mais les évaluateurs sont bien conscients du fait que l'enquête ne permettra pas de bien comprendre les raisons pour lesquelles les ménages utilisent ou non les centres de santé. Ils invitent donc un ethnographe à se joindre à leur équipe et à faire des études approfondies dans un petit nombre de communautés. Les études ethnographiques observeront les facteurs économiques, politiques, organisationnels, culturels et écologiques qui affectent l'opération des centres de santé dans différentes communautés. La première partie de l'analyse s'intéressera à de grandes différences culturelles qui

ont une chance d'affecter tous les centres de santé et la dernière partie (l'analyse contextuelle) aidera à expliquer les facteurs affectant la performance des différents centres (Figure 2). Le directeur de l'évaluation sait bien que les modèles utilisant les méthodes mixtes ne marchent bien que s'il y a un respect, une compréhension et un sentiment d'égalité entre les membres de l'équipe venant de professions différentes et l'ethnologue a donc été invité à se joindre à l'équipe dès la première réunion de planification. Voici quelques manières d'intégrer des approches QUANT et QUAL dans cette évaluation :

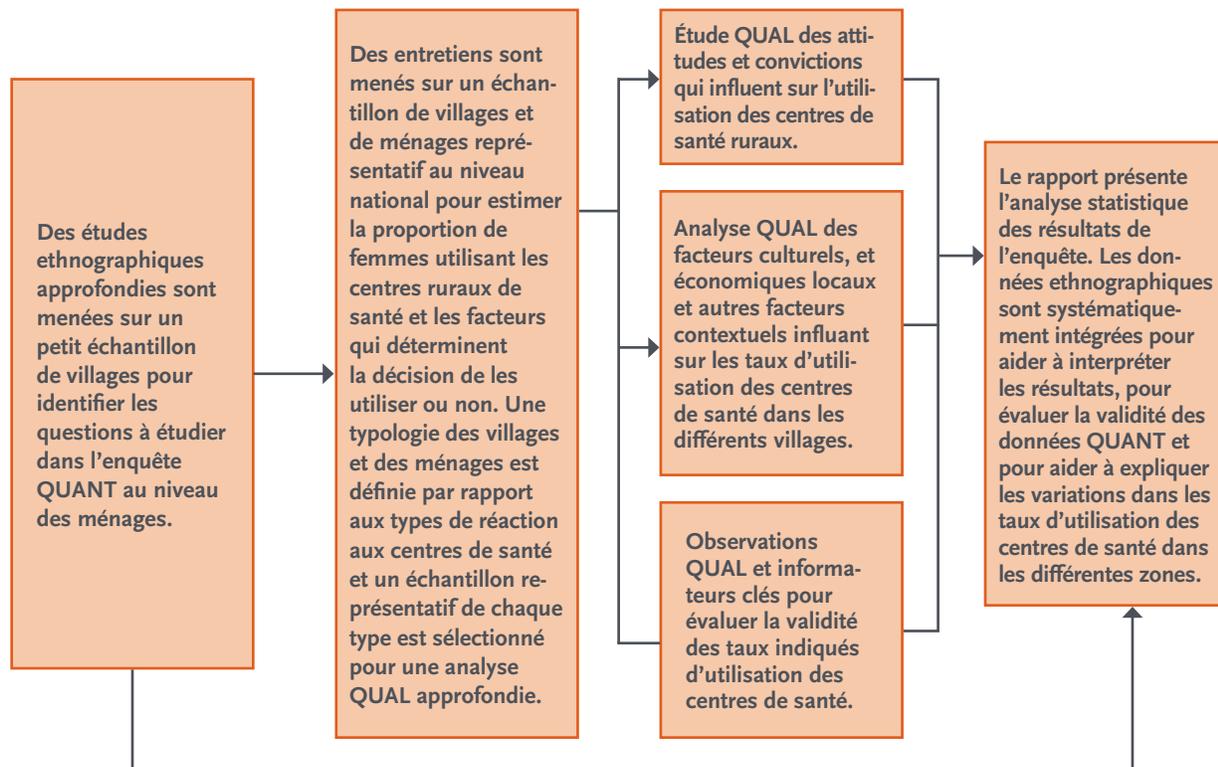
- Des études ethnographiques rapides (QUAL) sont menées dans des communautés choisies pour comprendre les questions qui doivent être abordées dans l'enquête et pour aider à formuler les questions.
- Une enquête QUANT auprès des ménages est faite, en utilisant un échantillon national représentatif.
- Une analyse de l'enquête auprès des ménages peut produire une typologie des ménages selon leur niveau d'utilisation des centres de santé ou leurs raisons de ne pas les utiliser. Un échantillon de chaque type sera sélectionné pour préparer les études de cas. Les études de cas révéleront souvent que l'utilisation ou les raisons de non utilisation qui ont été déclarées ne correspondent pas à la réalité. Par exemple, les femmes ne mentionneront généralement pas le harcèlement sexuel dans leur réponse à l'enquête et pourront donner une autre raison, comme des heures d'ouverture qui ne sont pas pratiques.
- On utilisera la triangulation pour obtenir des estimations QUANT et QUAL indépendantes pour les variables clés (comme l'utilisation des établissements de santé et les attitudes vis-à-vis de ces établissements). Une caractéristique

clé de la triangulation est que des procédures y sont intégrées pour identifier d'éventuelles incohérences dans les différentes estimations et pour rechercher la raison de ces différences. Par exemple, on comparera les réponses à l'enquête avec l'observation de la manière dont sont reçues les personnes qui entrent dans la clinique ou des discussions informelles avec des ménages pour savoir si et quand ils utilisent les cliniques.

- Des rapports QUANT et QUAL séparés seront rédigés puis les équipes se réuniront pour identifier les domaines où il semble y avoir des différences de faits ou d'interprétation. S'il y a par exemple des incohérences entre les réponses à l'enquête sur l'utilisation de centres de santé et les données venant des observations et des entretiens approfondis, les enquêteurs QUANT et QUAL se réuniront pour discuter des raisons de ces incohérences. Ils seront peut-être d'accord sur le fait que l'une ou l'autre des sources d'informations est plus fiable. Par exemple, des entretiens approfondis avec des femmes, hors de la présence d'autres membres du ménage, peuvent être considérés comme plus fiables. Cependant, si l'on ne peut pas dire clairement quelle source est plus fiable, les enquêteurs peuvent retourner sur le terrain pour collecter d'autres données ou ils peuvent chercher d'autres sources d'informations, par exemple étudier les registres des centres de santé sur les visites de patients, ou consulter des informateurs clés, comme des notables de la communauté, des infirmières, etc.

Dans ces deux cas, les MM peuvent renforcer l'évaluation. Cependant, on a un angle tout à fait différent quand on utilise les MM pour renforcer des modèles principalement QUAL (Figure 1) ou pour renforcer un modèle principalement QUANT (Figure 2).

Figure 2 Utiliser des études ethnographiques QUAL dans les villages et les ménages pour aider à interpréter les résultats d'une enquête nationale QUANT auprès d'un échantillon représentatif



Généralement, une approche MM est particulièrement utile pour :

- Examiner les interactions entre les facteurs contextuels complexes et changeants qui peuvent influencer sur la mise en œuvre et les impacts du programme.
- Définir et mesurer des indicateurs des facteurs culturels, historiques, politiques, légaux, environnementaux et psychosociaux qui influent sur la mise en œuvre. Il faut des méthodologies différentes pour les mesurer.
- Saisir des processus complexes de changement organisationnel et comportemental (Sections 3.2 et 3.3).
- Prendre en compte la manière dont les programmes changent en réponse à la manière dont ils sont perçus et utilisés par les différents secteurs de la population cible. L'expérience des premiers utilisateurs et ce

qu'ils en disent à leurs voisins peut avoir un effet spectaculaire sur la manière dont un programme évolue.⁴

- Nombreux sont les processus et les effets qui sont difficiles à observer. Dans certains cas, il est même difficile de savoir qu'ils existent. Cela est particulièrement important pour évaluer la situation de groupes vulnérables et pour des programmes qui influent sur des activités illégales ou désapprouvées par la société, comme l'usage de la drogue, la prostitution ou l'immigration illégale. Toutes ces difficultés sont multipliées pour les programmes post-conflits, humanitaires ou d'autres types de programmes d'assistance d'urgence.

⁴ *Realist evaluation* (Pawson 2006) donne un cadre utile pour l'analyse du changement de comportement et pour l'analyse de la manière dont les programmes opèrent réellement sur le terrain.

2^{ème} partie : L'approche des méthodes mixtes

2.1. Quatre décisions pour concevoir une évaluation à méthodes mixtes

Lors de la planification d'une évaluation MM, il faut prendre quatre décisions :

1. À quelle étape (ou à quelles étapes) de l'évaluation utilisera-t-on des MM?
2. Est-ce que les méthodes QUANT et QUAL seront utilisées de manière consécutive ou simultanée ?
3. Donnera-t-on aux méthodes QUANT et QUAL un poids relativement égal ou y aura-t-il une méthodologie dominante ?
4. Le modèle sera-t-il à un seul niveau ou à plusieurs ?

Décision 1 : à quelles étapes de l'évaluation utilisera-t-on des méthodes mixtes ?

La plupart des évaluations MM combinent seulement les méthodes QUANT et QUAL dans une ou peut-être deux étapes de l'évaluation, la plupart du temps dans la collecte des données. Pourtant, un modèle MM est beaucoup plus puissant si les méthodes QUANT et QUAL sont intégrées à plusieurs étapes de l'évaluation ou, dans l'idéal, à toutes les étapes. La section 2.2 explique comment les MM peuvent être utilisées pour renforcer chaque étape d'une évaluation. L'[Annexe 4](#) compare les approches QUANT et QUAL à chaque étape de l'évaluation et l'[Annexe 5](#) donne des exemples de la manière dont les approches QUANT et QUAL sont complémentaires pour renforcer chaque étape de l'évaluation.

Décision 2 : Le modèle MM est-il utilisé de manière consécutive ou simultanée ?

Modèles de méthodes mixtes consécutives

Dans les modèles consécutifs, les méthodes QUANT et QUAL sont utilisées en phases. Par exemple, l'évaluation peut commencer par une étude exploratoire pour aider à comprendre les principaux problèmes et comment ceux-ci sont perçus par les populations affectées. Cette étude permet de préparer une enquête QUANT qui est ensuite menée auprès d'un échantillon sélectionné de manière aléatoire. Les données pourraient ensuite être analysées en utilisant des méthodes d'analyse QUANT et/ou QUAL. Dans un autre exemple, une enquête QUANT rapide pourrait être utilisée pour identifier et quantifier les principaux types d'exploitations et d'activités agricoles. Ces informations seraient ensuite utilisées pour choisir un échantillon représentatif d'exploitations pour préparer des études de cas QUAL approfondies. Les études de cas seraient sans doute analysées en utilisant des méthodes QUAL et l'enquête par sondage serait analysée en utilisant des techniques QUANT. La Figure 3 est un exemple de modèle consécutif pour évaluer les transferts entre les ménages comme stratégie de survie pour les familles pauvres. Cette évaluation a commencé par une étude ethnographique (QUAL) pour comprendre les caractéristiques des communautés, suivie d'une enquête QUANT auprès des ménages et d'une analyse économétrique des résultats.

Modèles simultanés

Dans les modèles simultanés, les approches QUANT et QUAL sont utilisées en même temps. Par exemple, quand des données QUANT et QUAL sont collectées simultanément et que l'on utilise la triangulation pour comparer les informations

venant de plusieurs sources indépendantes sur les effets, les impacts et autres indicateurs clés. Autre exemple : quand des méthodes QUAL sont utilisées pour faire l'analyse contextuelle d'un site de projet (ou des zones environnantes) et que l'on fait en même temps une enquête QUANT auprès de ménages ou d'individus. Cela donne l'occasion de faire une analyse très riche mais plus compliquée, où les interactions entre la situation (contexte) et le processus de mise en œuvre du projet sont analysées.

Considérations opérationnelles pour la décision de choisir des modèles consécutifs ou simultanés

L'un des avantages des modèles consécutifs est que la logistique est souvent plus facile à organiser. Quand on utilise des questionnaires structurés pour la collecte des données, il faut souvent une importante équipe d'enquêteurs sur le terrain, qui suivent un calendrier précis pour la sélection des ménages et le nombre d'entretiens devant être conduits chaque jour. Les superviseurs sur le terrain doivent savoir où travaille chaque recenseur parce qu'il faut souvent faire des visites de suivi à un sous-échantillon des ménages pour le contrôle de qualité. Le superviseur doit aussi être disponible pour répondre aux questions des enquêteurs. Au contraire, les méthodes ethnographiques et de nombreux autres types de méthodes QUAL ont un calendrier beaucoup plus flexible du point de vue du moment et de l'endroit où les enquêteurs peuvent se trouver à n'importe quel moment. Pour ces raisons, entre autres, les modèles MM simultanés seront souvent plus difficiles à gérer, surtout pour des équipes d'évaluations qui n'ont que peu de superviseurs expérimentés. Les modèles simultanés peuvent être particulièrement problématiques dans des zones où la planification logistique (par ex. le déplacement jusqu'aux sites, les endroits où rester, la sécurité) peut être difficile à coordonner. Ils rendent aussi la gestion

du feedback plus difficile car les ajustements devront être faits plus rapidement que pour les modèles consécutifs. D'un autre côté, l'avantage des modèles simultanés est que la collecte des données et leur analyse peuvent être terminées plus rapidement.

Décision 3 : Le modèle MM sera-t-il principalement QUANT, principalement QUAL ou utilisera-t-on un modèle équilibré ?

On peut se représenter les modèles d'évaluation sur un continuum d'approches allant d'approches exclusivement QUANT en passant par des approches qui donnent un poids égal aux méthodes QUANT et QUAL jusqu'à des approches exclusivement QUAL (Bamberger et al 2012 pp. 324-34 ; Greene et Caracelli 2003). Ceci est important car des évaluateurs différents, qui ont sans doute commencé leur carrière avec une orientation soit principalement QUANT soit principalement QUAL, peuvent avoir des attentes très différentes de ce que représentera une évaluation MM. C'est aussi important parce que, du fait de l'orientation professionnelle des évaluateurs, une approche QUANT ou QUAL est dominante dans la plupart des évaluations MM (par opposition à une approche équilibrée).⁵

Le Tableau 1 illustre la manière dont les MM sont utilisées dans les évaluations où l'approche dominante est QUANT ou QUAL et l'Annexe 3 donne des exemples de modèles d'évaluation à différents points de ce continuum.

⁵ Une nouvelle génération d'évaluateurs apparaît. Ils ont été formés aux MM en tant qu'approche intégrée de l'évaluation et on commence à voir des études avec une approche plus équilibrée, sans orientation dominante. Mais celles-ci sont toujours minoritaires. On peut trouver des exemples de modèles équilibrés dans le *Journal of Mixed Method Research*.

Tableau 1. Les méthodes mixtes sont utilisées différemment pour des modèles d'évaluation ayant une orientation principalement QUANT ou QUAL

Quelle approche est dominante ?	Comment fonctionne l'approche dominante ?	Comment l'autre orientation est-elle utilisée pour renforcer le modèle ?
QUANT	L'évaluation fait remplir un questionnaire structuré à un échantillon aléatoire d'individus, de ménages, de groupes, d'institutions ou de communautés et l'analyse repose principalement sur des méthodes économétriques ou d'autres méthodes quantitatives.	On utilise des entretiens approfondis, l'observation et des entretiens de groupes pour aider à concevoir les questionnaires. De petits échantillons de cas, choisis dans l'échantillon principal, peuvent aussi permettre de comprendre plus en profondeur les relations statistiques découvertes dans l'analyse QUANT. Les cas peuvent être représentatifs de chacune des principales catégories identifiées par l'analyse ou on peut les utiliser pour étudier les exceptions ou d'autres groupes choisis de manière ciblée.*
On donne le même poids aux approches QUANT et QUAL	Des enquêtes QUANT sont combinées à différentes techniques QUAL. Quelquefois, ces dernières s'intéressent à l'analyse du processus et du contexte, dans d'autres cas elles s'intéressent à la même unité d'analyse que les enquêtes (par ex. les individus, les ménages, les communautés, les organisations) mais en utilisant des méthodes différentes pour la collectes de données.	
QUAL	Des études de cas, des entretiens approfondis et d'autres techniques QUAL sont utilisées avec de relativement petits échantillons d'individus, de ménages, de communautés ou de groupes.	On utilise une enquête QUANT rapide, soit pour identifier les questions ou les groupes qui devront être examinés dans les études QUAL approfondies, soit pour montrer que l'échantillon QUAL est raisonnablement représentatif de l'ensemble de la population.
* On trouvera dans l'Annexe 6 et Bamberger et al 2012 pp. 360-61 une description de l'échantillonnage ciblé..		

La Figure 3 décrit un modèle consécutif avec une approche QUANT dominante. C'est une étude de transferts d'argent et de biens entre les ménages comme stratégie de survie des ménages urbains pauvres de Colombie (Wansbrough, Jones et Kappaz 2000). L'objectif de l'étude était de décrire les tendances dans les transferts et d'estimer s'ils étaient suffisamment importants pour servir d'assistance sociale informelle en fournissant une aide aux secteurs les plus pauvres de la communauté en temps de besoin. Ces transferts entre les ménages sont difficiles à identifier et à mesurer

et un anthropologue a vécu dans la communauté pendant un mois pour étudier les tendances dans les transferts et aider à concevoir le questionnaire pour l'enquête QUANT qui a ensuite été faite auprès de plusieurs centaines de ménages. Les données ont été analysées en utilisant une analyse économétrique QUANT.

La Figure 4 illustre un modèle consécutif avec une approche QUAL dominante. Elle décrit l'évaluation fictive de l'adoption de nouvelles variétés de semences par différents types de familles

Figure 3 Modèle de méthodes mixtes consécutives avec une approche quantitative dominante : Étude des transferts entre les ménages comme stratégie de survie pour les ménages à faible revenu de Cartagena en Colombie

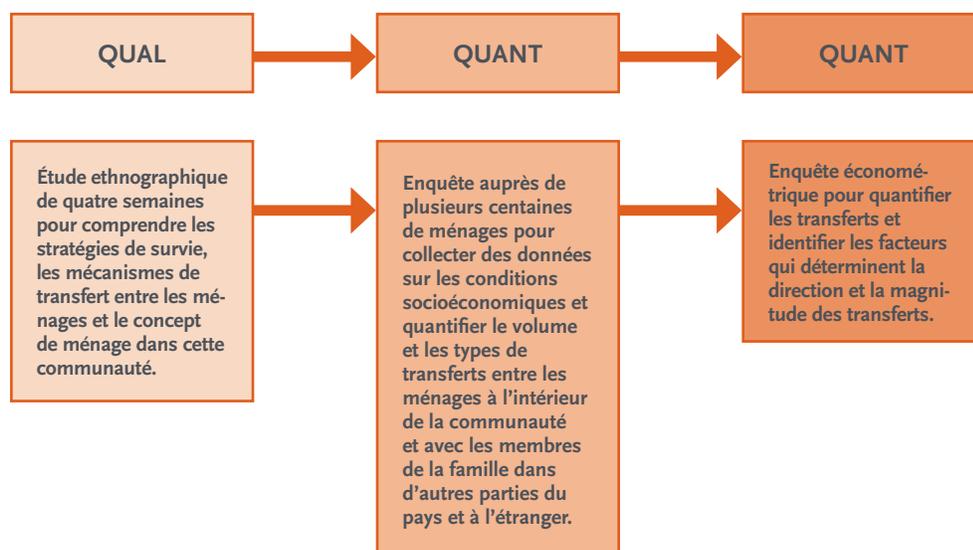
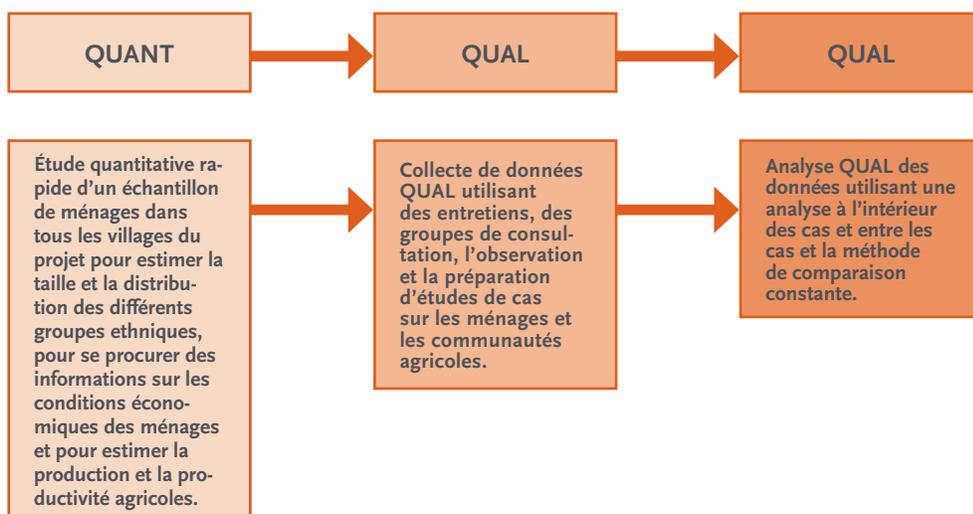


Figure 4 Modèle de méthodes mixtes consécutives avec une approche qualitative dominante : évaluation de l'adoption de nouvelles variétés de semences par différents types de paysans



rurales. Les principales méthodes de collecte de données sont qualitatives : entretiens, groupes de consultation, observation et études de cas de ménages et de communautés de petits paysans. Les principales méthodes d'analyse sont aussi qualitatives : analyse à l'intérieur des cas et entre

les cas et méthode de comparaison constante. Cependant, pour obtenir des informations sur la distribution ethnique des ménages, leurs conditions économiques et la production, l'évaluation commence par une enquête QUANT rapide auprès d'un échantillon de ménages dans tous les villages

concernés par le projet de vulgarisation agricole. Les résultats de cette étude sont utilisés pour identifier les types de ménages à étudier plus en profondeur grâce aux méthodes QUAL de collecte de données et pour que les cas sélectionnés soient largement représentatifs de la population totale de l'enquête.

L'un ou l'autre des modèles d'évaluation ci-dessus aurait pu être modifié pour donner un poids égal aux approches QUANT et QUAL. Dans le cas de l'étude des transferts entre les ménages, l'enquête auprès des ménages aurait pu être complétée par des études de cas QUAL sur les familles ou les réseaux informels de transfert. Celles-ci auraient pu être intégrées dans l'analyse pour comparer la description et l'interprétation des fonctions et de l'opération des réseaux de transferts tirés des études QUAL avec les résultats de l'analyse économétrique. Dans le second exemple, on aurait pu faire une étude QUAL ou QUANT de débouchés pour la commercialisation pour estimer les changements dans la vente des produits agricoles de la zone du projet et, peut-être les changements dans les achats de biens de consommation par les familles de la zone du projet.

L'exemple 3 de la 5^{ème} partie décrit un modèle d'évaluation MM équilibré (intégré) utilisé pour évaluer un grand programme de développement communautaire en Inde. Le modèle donne un poids égal aux approches QUANT et QUAL à toutes les phases de l'évaluation.

Décision 4 : Le modèle MM sera-t-il sur un seul niveau ou utilisera-t-on un modèle à plusieurs niveaux ?

Les modèles que nous avons présentés jusqu'à présent opèrent sur un seul niveau, par exemple l'exploitation ou le ménage. Mais les MM sont aussi un outil intéressant pour évaluer des

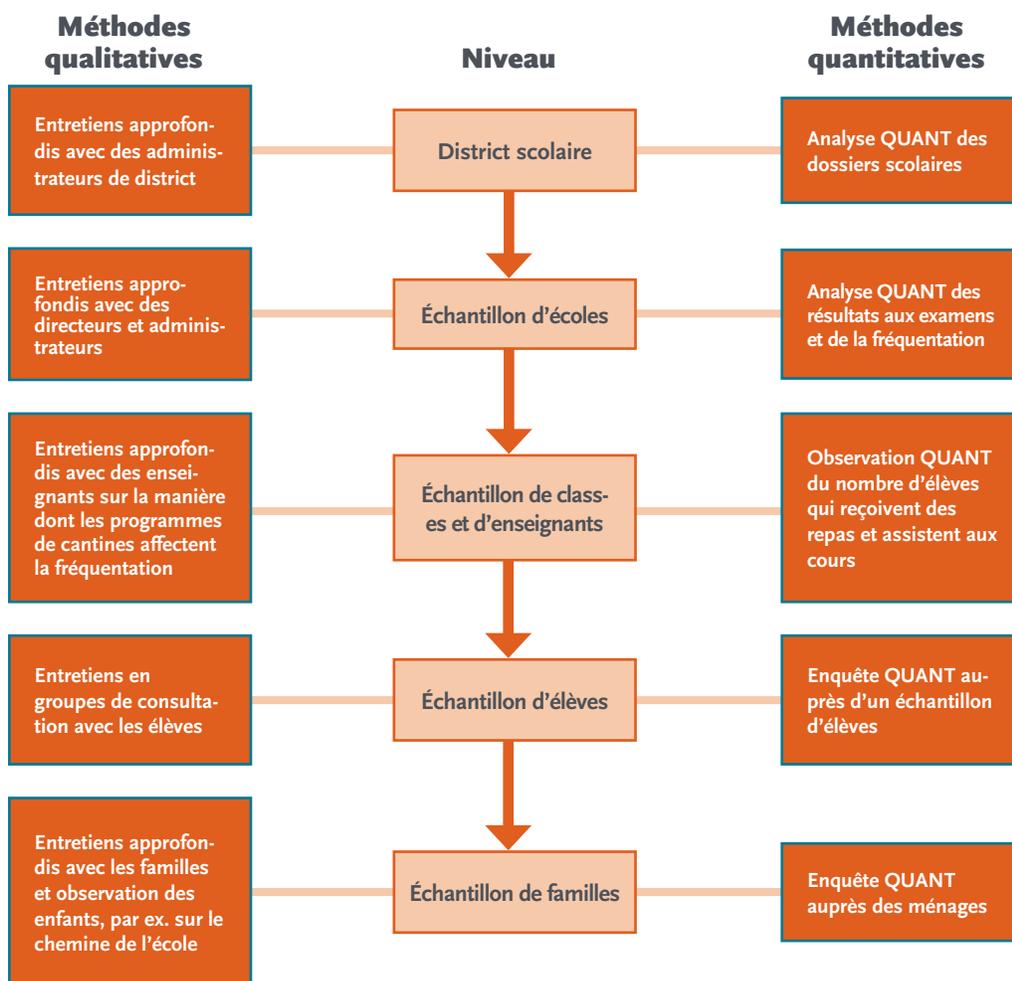
systèmes de prestation de services (par ex. les départements d'éducation dans les districts, les services nationaux de santé, un programme national de renforcement des autorités municipales) pour lesquels il faut décrire et analyser les liens entre différents niveaux. Ces évaluations peuvent être très complexes et coûteuses. Les modèles à méthodes mixtes, qui combinent des données QUANT et QUAL à chaque niveau peuvent souvent fournir des résultats valides et crédibles à partir d'échantillons plus petits et plus économiques.

La Figure 5 illustre un modèle utilisant des méthodes mixtes à plusieurs niveaux pour évaluer les effets d'un programme de cantines scolaires sur la scolarisation et la fréquentation scolaire. L'évaluation doit collecter des données au niveau du district scolaire, d'un échantillon d'écoles, d'un échantillon de classes et d'enseignants de chaque école et d'un ensemble d'élèves et de familles. À chaque niveau, des données quantitatives et qualitatives sont collectées et comparées. Les méthodes QUAL, comme l'observation, les groupes de consultation et les entretiens avec des informateurs clés, peuvent aussi permettre d'examiner les liens entre les différents niveaux (par ex. les interactions entre les fonctionnaires du district et les administrateurs et enseignants des écoles).

2.2. Utiliser des approches MM à chaque phase de l'évaluation

Cette section explique comment les approches QUANT et QUAL sont généralement utilisées à chaque phase d'une évaluation et comment les deux approches peuvent être combinées en un modèle MM. Les [Annexes 4](#) et [5](#) donnent plus de détails. [L'Annexe 11](#) donne des exemples de la manière dont les MM peuvent aider à répondre à des problèmes courants auxquels on peut être

Figure 5 Modèle de méthodes mixtes sur plusieurs niveaux : Évaluation des effets d'un programme de cantines scolaires sur la fréquentation et les résultats



confronté durant l'élaboration de l'évaluation, la collecte et l'analyse des données et aussi aider à promouvoir l'utilisation des résultats et recommandations de l'évaluation. Quand vous lisez cette section, vous devez comprendre que pour une grande évaluation avec un financement important, l'équipe comprendra peut-être un ou plusieurs membres embauchés pour leur expertise spécifique QUANT ou QUAL (par exemple l'échantillonnage statistique pour une grande étude sur la malnutrition, ou des entretiens QUAL approfondis sur des sujets sensibles comme la violence domestique). Par contre beaucoup d'évaluations (la plupart peut-être)

n'ont pas ce luxe et on demandera aux membres de l'équipe d'utiliser des approches QUANT ou QUAL, selon les besoins. Bien sûr, même pour les grandes évaluations, il est bien évidemment préférable que tous les membres de l'équipe aient une notion de toutes les méthodes d'évaluation qui seront utilisées. Les paragraphes qui suivent illustrent la manière dont des méthodes QUANT et QUAL peuvent être intégrées à différentes phases de l'évaluation.

- *Formulation d'hypothèses.* Les évaluations QUANT formulent souvent des hypothèses de

manière *déductive*, à partir de théories existantes ou de revues de la littérature, tandis que les évaluations QUAL élaborent des hypothèses de manière *inductive*, au fur et à mesure de l'évaluation de l'étude. Les MM combinent les deux approches. Par exemple, une hypothèse élaborée de manière déductive en utilisant une approche QUANT peut être examinée et affinée grâce à des approches QUAL, comme des entretiens et des observations. En revanche, les premières phases de la collecte de données QUAL peuvent décrire des processus et des problèmes qu'une approche QUANT pourra tester grâce aux données collectées dans une enquête par sondage.

- **Échantillonnage.** Les évaluations QUAL utilisent normalement un relativement petit nombre de sujets, sélectionnés de manière *ciblée* (échantillonnage théorique) pour que tous les groupes importants soient étudiés. Par contre, les évaluations QUANT utilisent normalement un échantillon aléatoire assez important, permettant une généralisation à des populations plus larges et une comparaison statistique de différents groupes (par ex. le groupe du projet et le groupe de comparaison). L'échantillonnage MM utilise le même cadre d'échantillonnage pour créer un grand échantillon pour une enquête QUANT et pour sélectionner un échantillon, petit mais représentatif, pour une analyse QUAL approfondie. *C'est l'une des contributions les plus importantes des modèles MM que de garantir que les échantillons QUAL sont raisonnablement représentatifs de l'ensemble de la population de l'échantillon.* L'exemple de l'évaluation des centres de santé montre comment les approches QUANT et QUAL de l'échantillonnage peuvent se compléter. [L'Annexe 6](#) résume les différences entre les stratégies d'échantillonnage QUANT et QUAL.
- **Modèles d'évaluation.** La plupart des évaluations QUANT utilisent un petit nombre de

modèles randomisés ou quasi-expérimentaux. Quand c'est possible, des échantillons représentatifs du groupe du projet et du groupe de comparaison sont interrogés à deux moments ou plus pendant la durée de vie du projet pour comparer les changements dans les indicateurs d'effet ou d'impact. Par contre, les évaluations QUAL visent à décrire des processus de changement en cours, qui sont souvent affectés par de nombreux facteurs différents et qui affectent les différentes personnes ou les différents groupes de manières différentes. Certaines évaluations QUAL tentent de comprendre le programme grâce à une analyse des relations entre de nombreux éléments différents de la communauté ou d'autres situations où le programme opère tandis que d'autres s'intéressent de manière plus approfondie à des individus, sans nécessairement s'intéresser au contexte plus général. Dans certaines évaluations ayant un bon financement, l'évaluateur peut vivre dans la communauté ou lui rendre de fréquentes visites pendant une certaine période, mais dans la plupart des cas, cela n'est pas possible et il faut se baser sur des groupes de consultation, des entretiens approfondis, des informateurs clés, etc. Normalement, les évaluations QUAL ne cherchent pas à établir une relation directe de cause à effet entre les interventions du projet et les effets. L'une des nombreuses manières possibles de combiner les deux approches est d'utiliser des méthodes QUAL pour étudier le processus de mise en œuvre du projet et l'influence des variables contextuelles sur la performance du projet dans certaines des communautés où l'on fait une enquête QUANT auprès des participants du projet.⁶

⁶ Il y a de nombreuses techniques permettant de transformer les descriptions QUAL de facteurs contextuels en variables QUANT (par exemple les variables nominales) qui peuvent être incorporées dans une analyse de régression (cf. Tableau 4).

Tableau 2. Techniques de collecte de données QUANT et QUAL souvent utilisées

QUANT	QUAL
<ul style="list-style-type: none"> • Enquêtes structurées dans les ménages, exploitations, auprès des utilisateurs de services publics, etc. • Guides d'observations structurées • Mesures anthropométriques de la taille et du poids • Tests d'anémie et de HIS, avec collecte d'échantillons de sang et tests • Compteurs automatiques (par ex. des personnes entrant dans un bâtiment) • Analyse sociométrique^{±**} • GIS (préparation et analyse de cartes GPS)^{**} • SIG du programme sur les données des entrées et des sorties • Étude des données de l'institution – registres des cliniques, dossiers scolaires, etc. ^{**} 	<ul style="list-style-type: none"> • Entretiens approfondis • Informateurs clés • Observation des participants • Observation des non participants^{**} • Études de cas • Interview des clients à la sortie^{**} • Études simulées de patients • Enregistrement vidéo ou audio^{**} • Photographie • Analyse de documents^{**} • Objets • Entretiens en groupes (par ex. groupes de consultation, réunions communautaires)^{**} • Techniques participatives de groupes (par ex. MRP, changement le plus significatif) • Enquêtes sur Internet
<p>[±] Techniques d'enquête pour étudier la formation de groupes, comment les informations sont diffusées, l'identification des leaders d'opinion et d'autres tendances d'organisation sociale dans une communauté ou un groupe.</p> <p>^{**} Indique que ces techniques peuvent être utilisées aussi bien de manière quantitative que qualitative. Elles sont placées dans la colonne où elles sont plus souvent utilisées.</p>	

• *Méthode de collecte et d'enregistrement des données.* Le tableau 2 présente quelques unes des techniques de collecte de données QUANT et QUAL les plus utilisées. Tandis que les évaluations QUANT collectent des données numériques standardisées, les évaluations QUAL collectent souvent les données avec des méthodes moins structurées afin d'avoir plus de flexibilité et de tenter de comprendre les complexités d'une situation. Tandis que les méthodes QUANT de collecte de données ont l'avantage de produire des données standardisées qui mesurent les changements dans le temps ou entre les groupes, elles ne sont pas bien adaptées pour saisir des informations sur des sujets sensibles ou pour interroger des groupes difficiles à atteindre. La collecte de données MM s'appuie sur les forces des données QUANT tout en approfondissant, en saisissant des données sensibles, en

étudiant des processus et des changements de comportement.

- *Triangulation.* L'utilisation systématique de la triangulation est un élément clé des MM (Annexe 9). Les évaluateurs QUAL comme les évaluateurs QUANT utilisent la triangulation pour obtenir deux estimations indépendantes, ou plus, d'importantes variables de résultats, mais les MM ont tendance à utiliser la triangulation de manière plus systématique ; c'est une partie intégrante du modèle d'évaluation.

Les évaluations QUANT utilisent la triangulation pour intégrer des vérifications de cohérence dans les instruments d'enquêtes ou pour comparer des sources de données secondaires et les informations fournies par les participants à l'enquête. Les évaluations QUAL utilisent plus largement la triangulation, mais souvent pour une meilleure compréhension, plus approfondie, grâce aux multiples perspectives obtenues auprès de différentes sources

Tableau 3. Différents types de triangulation utilisés dans les évaluations à méthodes mixtes

Méthode	Exemples
Utiliser différents cadres conceptuels	Comparer les cadres d'analyse du féminisme, des droits humains, de l'exclusion sociale ou économique (par ex. coûts bénéfice)
Différentes méthodes de collecte des données	Comparer l'enquête structurée, l'observation directe, les données secondaires, les objets
Différents enquêteurs	Comparer les effets du sexe, de l'âge, de l'origine ethnique, du niveau économique, du type d'habillement, de la langue, etc. des enquêteurs sur les réponses
Différents moments	Comparer les réponses ou les observations à différentes heures du jour, différents jours de la semaine, différents moments de l'année
Différents lieux et contextes	Comparer les réponses et les observations quand les entretiens sont faits à la maison avec d'autres personnes présentes, à des endroits où la personne interrogée peut parler plus librement, dans la rue ou d'autres endroits publics, au travail, dans la classe

d'information plutôt que pour une vérification de la cohérence. Les modèles MM font une triangulation des estimations QUANT et QUAL (cf. Tableau 3). Les MM utilisent les informations fournies par la triangulation pour : améliorer la fiabilité et la validité des estimations des indicateurs clés en comparant les informations venant de sources différentes ; mieux comprendre la signification des relations statistiques identifiées dans l'analyse quantitative ; s'assurer de bien saisir et comparer les perspectives de toutes les principales parties prenantes, en insistant particulièrement sur les groupes pauvres et vulnérables. Si les estimations obtenues auprès des différentes sources sont cohérentes, cela augmente la validité et la crédibilité des données, surtout quand les estimations sont basées sur de petits échantillons, et cela produit des estimations plus fiables que si toutes les ressources avaient été investies dans une seule technique, comme une enquête auprès des ménages.

- *Analyse des données.* Les évaluateurs QUAL utilisent un grand nombre de méthodes

différentes d'analyse des données pour déterminer les tendances et les relations générales et obtenir une idée holistique des interactions entre un projet et la situation où il est intégré. Par contre, l'objectif de l'analyse QUANT est de décrire les caractéristiques statistiques des variables clés, de déterminer la signification statistique de différences entre le groupe du projet et le groupe de comparaison et d'identifier les facteurs qui contribuent à la magnitude et à la direction du changement. L'analyse des données de méthodes mixtes utilise l'analyse QUAL pour aider à comprendre le sens que donnent différents sujets ou groupes aux associations statistiques trouvées dans l'analyse QUANT et pour fournir des études de cas et des exemples éclairant les résultats. D'un autre côté, l'analyse QUANT peut être utilisée pour évaluer dans quelle mesure les cas présentés dans l'étude QUAL représentent la population totale concernée et, le cas échéant, quels secteurs n'ont pas été étudiés.

Tableau 4. Exemples d'analyse des données de méthodes mixtes

Approche	Description	Exemple
A. Analyse des données de méthodes mixtes parallèles	Il s'agit de deux processus d'analyse séparés : les données QUANT sont analysées en utilisant des méthodes QUANT conventionnelles (comme des tableaux de fréquence, des tableaux croisés, une analyse de régression, etc.) tandis que l'on fait une analyse séparée des données QUAL en utilisant des méthodes QUAL comme l'analyse du contenu. Les résultats des deux analyses sont ensuite comparés.	Dans l'évaluation de la pauvreté faite en 2003 par la Banque mondiale au Guatemala, des équipes séparées avaient la responsabilité de collecter des données QUAL et QUANT. L'analyse QUAL a été faite sur 5 paires de villages représentant les principaux groupes ethniques. Les données QUANT concernant les mêmes groupes de villages ont été analysées séparément et les données n'ont été intégrées qu'à l'étape finale de l'analyse. La combinaison de deux analyses indépendantes a permis d'avoir plus de contexte politique et historique pour comprendre les impacts des opérations du programme (Teddlie et Tashakkori (2009) Cadre 11.5).
B. Analyse des données de méthodes mixtes par conversion	<p>a. Les données QUAL sont converties en indicateurs QUANT (« quantifier ») en utilisant des taux, des scores et des échelles* pour pouvoir utiliser des techniques d'analyse QUANT.</p> <p>b. Les données QUANT sont converties en indicateurs QUAL (« qualifier ») pour pouvoir utiliser des procédures d'analyse QUAL.</p>	<p>a. Les données sur le contexte politique, économique, social, environnemental, légal et administratif dans lequel opère un projet sont souvent présentées sous forme de narration qualitative. Les indicateurs peuvent être « quantifiés » par une conversion en variables nominales. Par exemple : « il y a une croissance de l'économie » = 1, « il n'y a pas de croissance de l'économie » = 0. Ces variables nominales peuvent ensuite être incorporées dans l'analyse de régression.</p> <p>b. Dans la figure 4, une typologie quantitative des paysans pourrait être « qualifiée » en produisant des descriptions narratives des différentes attitudes par rapport à l'adoption d'une nouvelle variété de semences.</p>
C. Analyse des données de méthodes mixtes consécutives	<p>a. Une analyse QUAL des données est suivie d'une analyse QUANT.</p> <p>b. Une analyse QUANT des données est suivie d'une analyse QUAL.</p> <p>c. Modèles MM itératifs. L'analyse comprend des étapes QUANT et QUAL consécutives</p>	a. Dans la Figure 3, l'étude des stratégies de survie commence par une analyse qualitative des rapports narratifs sur les tendances dans les transferts entre les ménages qui fournissent un appui aux ménages vulnérables. L'analyse QUAL aide à concevoir l'enquête QUANT sur les transferts entre les ménages, qui est ensuite analysée en utilisant des techniques économétriques.
D. Analyse des données de méthodes mixtes à plusieurs niveaux	Les techniques d'analyse QUANT et QUAL sont utilisées à différents niveaux d'un modèle d'évaluation à plusieurs niveaux	Le Figure 5 illustre l'utilisation d'une analyse MM à plusieurs niveaux pour évaluer les impacts d'un programme de cantines scolaires sur la fréquentation scolaire et les résultats scolaires. Des analyses QUANT et QUAL ont été faites consécutivement au niveau du district, de l'école, de la classe, de l'élève et de la famille. Cela a permis une analyse des liens réciproques entre les différents niveaux.**

* Par exemple, des rapports narratifs sur les attitudes de groupes politiques locaux par rapport à un programme de développement social pourraient être convertis en échelle numérique, avec 3 = « le groupe politique est bien disposé envers le programme », 2 : « le groupe politique n'est ni favorable, ni opposé » et 1 : « le groupe est opposé au programme. » Les classements, les scores et les échelles sont des manières légèrement différentes de faire cette conversion.

** Par exemple, l'analyse des registres au niveau du district peut déterminer quelles sont les écoles ayant une fréquentation scolaire et/ou des notes de performance au dessus ou en dessous de la moyenne. Ces informations peuvent être utilisées pour sélectionner des écoles au dessus et en dessous de la moyenne qui seront incluses dans l'échantillon. De même, des entretiens approfondis avec les enseignants pourraient être utilisés pour choisir un échantillon d'élèves qui ont des caractéristiques intéressantes particulières et devraient être inclus dans les groupes de consultation.

Source : Adapté de C. Teddlie et A. Tashakkori 2009 *Foundations of Mixed Methods Research. Chapitre 11, Sage Publications (avec autorisation)*. La plupart des exemples ont été élaborés par l'auteur de la présente note.

3^{ème} partie : Applications des modèles à méthodes mixtes

3.1. Stratégies d'échantillonnage pour des évaluations MM à orientation QUANT ou QUAL

Modèle 1 : Utiliser des méthodes mixtes pour renforcer une évaluation principalement QUANT

Dans la plupart des modèles d'évaluation QUANT, des enquêtes sur échantillon sont faites avant et après l'intervention du projet avec un groupe de contrôle ou de comparaison. On estime la taille de l'échantillon nécessaire pour les enquêtes QUANT en utilisant l'ampleur de l'effet et l'efficacité statistique.⁷ On peut utiliser des méthodes QUAL pour renforcer le modèle à l'un ou plusieurs des points suivants :

- **Étude exploratoire ou diagnostique pour comprendre le contexte et les problèmes avant l'élaboration des outils d'enquête.** Il peut s'agir d'une étude qualitative rapide ne durant que quelques jours ou d'études plus longues où un anthropologue ou un sociologue vit dans une communauté pendant plusieurs semaines ou plusieurs mois. Dans un grand projet qui opère dans des régions géographiques ou écologiques différentes, il faudra peut-être faire des études diagnostiques dans plusieurs régions ou communautés. Quelquefois, les études seront faites par un seul enquêteur (un enquêteur par communauté ou par région) tandis que, dans d'autres cas, l'enquêteur en chef pourra être assisté

⁷ L'ampleur de l'effet est l'ampleur du changement ou de l'impact qui est estimé. Plus le changement estimé est important, plus l'échantillon nécessaire pourra être petit. L'efficacité statistique est la probabilité que le test statistique détermine correctement qu'il y a un impact réel. Plus le niveau de confiance exigé est élevé, plus l'échantillon nécessaire devra être grand. Pour une présentation de l'estimation de la taille d'un échantillon, cf. Bamberger et al 2012, Chapitre 15 Section 4.

par une équipe d'assistants qui feront des enquêtes rapides ou des études des participants ou non participants, sur les systèmes communautaires de transport par exemple, la manière dont les femmes utilisent leur temps ou les systèmes de production et de marketing. Les études peuvent durer plusieurs jours ou plusieurs semaines, mais l'unité primaire d'échantillonnage sera généralement une communauté ou un groupe et on n'étudiera normalement que quelques groupes ou communautés. Cependant, on peut interroger un grand nombre de personnes en utilisant des techniques de collecte de données non structurées ou semi structurées ; dans certains cas, on peut aussi faire des enquêtes rapides sur des échantillons.

- **Groupes de consultations tenus avec différents segments de la population cible.** Ils peuvent avoir lieu soit durant l'étape préparatoire de l'évaluation, soit une fois que les enquêtes quantitatives ont été analysées et que les principaux groupes d'intérêt ont été identifiés. Dans l'idéal, il faudrait tenir trois ou quatre groupes de consultation avec chaque groupe économique ou démographique intéressant pour l'évaluation (Teddlie et Tashkkori 2009 Tableau 8.5), mais on aura souvent un plus petit nombre de groupes quand on travaille avec des contraintes de budget et de temps.
- **On peut ajouter à une enquête sur un échantillon des modules semi structurés spécialisés et les administrer à un sous échantillon de personnes interrogées.** Par exemple, l'enquête principale peut être menée auprès du chef de ménage (qui dans de nombreuses cultures sera

probablement un homme) mais, dans un sous échantillon de ménages, l'épouse pourra être interrogée. Parfois, le même enquêteur peut administrer le module spécial à l'épouse (ou à un autre membre du ménage) mais souvent, il faudra organiser un entretien séparé, souvent à un endroit et à un moment ou ni le mari, ni d'autres membres du ménage ne seront présents. Généralement, ces modules sont administrés à 10-25% de l'échantillon original.

- **Préparation d'études de cas pour compléter l'enquête.** Il est souvent utile de comparer des études de cas d'un petit échantillon des personnes interrogées au cours de l'enquête afin de comprendre plus en profondeur les questions intéressantes pour l'évaluation. Par exemple, dans l'évaluation d'un programme agricole, on pourrait faire des études de cas pour illustrer différents types de systèmes agricoles. Pour un projet d'éducation, les cas pourraient comporter des familles ayant des revenus plus ou moins élevés, d'autres qui habitent plus ou moins loin de l'école ou des familles issues de groupes religieux ou ethniques différents. Là aussi, le nombre de cas sera généralement assez réduit, même si la durée peut être assez longue. Quand on prépare des études de cas sur des organisations (comme des écoles ou des coopératives agricoles) ou concernant des communautés entières (par exemple pour illustrer les effets de l'amélioration de systèmes de transport), l'étude sera plus compliquée et on interrogera souvent un grand nombre d'individus pour chaque cas.

Modèle 2 : Utiliser des méthodes mixtes pour renforcer une évaluation QUAL

Les modèles à méthodes mixtes peuvent aussi être utilisés pour renforcer une évaluation QUAL qui utilise des groupes de consultation, l'observation des participants, l'observation des non participants

et la préparation d'études de cas. Une difficulté pour beaucoup de ces modèles est le risque de biais dû au fait que les échantillons d'individus ou de groupes ne sont pas représentatifs. Par exemple, les gens qui participent à des groupes de consultation sont souvent des personnes qui ont une opinion marquée pour ou contre un projet, qui ont le temps et les ressources leur permettant d'y participer (ils devront peut-être s'organiser pour le transport) ou, dans certains cas (souvent sans que l'évaluateur ne le sache), certains participants peuvent être envoyés par les autorités locales ou un autre groupe ayant des intérêts particuliers. Par conséquent, beaucoup d'informations précieuses et profondes seront difficiles à incorporer de manière crédible dans le rapport d'évaluation. De la même façon, de nombreux enquêteurs sont plus à l'aise quand ils parlent avec certains groupes qu'avec d'autres, de sorte qu'il peut y avoir un biais dans la sélection des études de cas. Généralement, l'échantillon est relativement petit, mais le nombre d'entretiens variera selon la taille de la population étudiée et le niveau de précision nécessaire des résultats.⁸

L'échantillonnage des méthodes mixtes garantit que les cas QUAL et les informateurs sont sélectionnés de manière à être largement représentatifs

⁸ On trouvera dans Bamberger et al. 2012 Tableau 15.6 p. 389 quelques règles générales pour estimer les tailles d'échantillon pour différentes sortes de méthodes de collecte de données QUAL. Cependant, les tailles d'échantillons varieront selon la taille et la complexité du programme évalué et le niveau de précision demandé pour les estimations. Si l'on doit faire des estimations à partir des études de cas, des groupes de consultation ou d'autres méthodes, il est important de s'assurer que les cas sont sélectionnés de manière à être raisonnablement représentatifs et qu'il y a suffisamment de cas pour que les résultats puissent être considérés comme crédibles. Un ou deux cas peuvent donc être valables pour illustrer des processus ou un comportement, mais il est rarement approprié d'utiliser un échantillon aussi petit pour faire des déclarations telles que « La plupart des paysans pensaient que... » ou « La plupart des mères trouvaient que les centres de santé... ». Il n'est pas possible d'utiliser des procédures statistiques conventionnelles pour estimer la taille appropriée d'un échantillon avec d'aussi petits nombres de cas. Il faudra donc inévitablement se fier à son jugement mais aussi se concerter avec les parties prenantes pour savoir quelle taille d'échantillon serait considérée comme raisonnable ou crédible.

de l'ensemble de la population. Cela renforce la validité de l'ensemble des résultats. Des techniques quantitatives, comme une enquête rapide sur un échantillon, peuvent être utiles pour comparer les caractéristiques socioéconomiques des personnes ou des groupes qui font l'objet des études qualitatives avec les caractéristiques de l'ensemble de la population. Généralement, la taille des échantillons QUANT sera relativement petite, car l'enquête n'est utilisée que pour vérifier que les études de cas ou autres données QUAL sont raisonnablement représentatives. Mais quand on a besoin d'estimations plus précises, on peut utiliser les procédures d'estimation de la taille des échantillons présentées dans la section précédente.

Modèle 3 : Utiliser un modèle à méthodes mixtes équilibrées (intégrées)

Si, dans la plupart des cas, les modèles à méthodes mixtes sont utilisés pour compléter un modèle principalement QUANT ou QUAL, il y a des cas où l'on peut utiliser un modèle à méthodes mixtes équilibrées, qui donne un poids égal aux deux approches. Ces modèles peuvent comporter une combinaison de techniques quantitatives et qualitatives mais aussi des techniques spécifiques aux méthodes mixtes, à différents stades de l'évaluation. L'estimation des tailles nécessaires pour les échantillons doit combiner les procédures QUANT pour les enquêtes sur échantillons et les règles générales présentées dans la section précédente pour les composantes QUAL. Pour combiner les deux, il faudra se baser sur l'expérience. L'exemple 3 de la 5^{ème} partie décrit un modèle à méthodes mixtes équilibrées (intégrées) utilisé pour évaluer un grand programme de renforcement de la capacité des communautés à gérer les subventions de développement des gouvernements des états, en Inde. Des enquêtes sur des échantillons, utilisées pour obtenir des estimations quantitatives des effets du programme, ont été combinées à une collecte de données qualitatives faite grâce à des visites régulières à un échantillon

de communautés du projet et de communautés de contrôle pour observer les processus de changement et pour interroger un petit échantillon de ménages ainsi que des autorités locales et des informateurs clés. Le même cadre d'échantillonnage a été utilisé pour les échantillons QUANT et QUAL et les échantillons QUAL ont été sélectionnés de manière à garantir leur représentativité.

3.2. Utiliser des méthodes mixtes pour évaluer des interventions complexes

On appelle généralement interventions complexes de développement⁹ des interventions qui ont certaines des caractéristiques suivantes :

- Planification et évaluation dirigées par le pays ;
- Évolution non linéaire du programme ;
- De nombreux éléments ou services et des objectifs multiples ;
- Des effets à la fois quantitatifs et qualitatifs ;
- Des populations cibles difficiles à identifier ou à suivre ;
- Un appui budgétaire général, fourni sans définition claire des services qui seront financés ;
- Plusieurs bailleurs de fonds et plusieurs agences ;
- Utilisation de processus participatifs de planification et de mise en œuvre, de sorte qu'il est difficile de définir à l'avance les effets et les impacts qui peuvent constamment changer ;
- Un contexte complexe ;
- Une situation qui change constamment et une intervention qui tente de s'adapter à ces changements ;
- Des processus complexes de changement de comportement de la population cible, des prestataires de services et d'autres agences, par exemple la police, l'armée et le système judiciaire (Section 3.3) ;

⁹ Pour une présentation des interventions complexes et de leur évaluation, cf. Funnell et Rogers (2011), Patton (2011) et Bamberger et al. (2012), Chapitre 16.

CADRE 3. ANALYSE D'ATTRIBUTION, DE CONTRIBUTION ET DE SUBSTITUTION

Si l'évaluation d'un projet peut quelquefois utiliser des modèles expérimentaux ou quasi expérimentaux qui peuvent contrôler les explications alternatives possibles des changements observés (effets ou impacts) dans le groupe du projet, il est rarement possible d'obtenir le même degré de rigueur dans l'évaluation des effets d'interventions complexes. Par conséquent, les évaluateurs des effets d'interventions complexes doivent décider lequel de ces niveaux d'analyse peut être utilisé :

Analyse d'attribution : Le groupe du projet est associé à un groupe de comparaison pour pouvoir contrôler (éliminer) les explications alternatives des changements observés. S'il y a statistiquement des différences significatives entre le groupe du projet et le groupe de comparaison, ceci est considéré comme preuve que l'intervention du projet a contribué à ces changements.

Analyse de contribution : L'analyse évalue la contribution d'une agence de développement particulière à la réalisation de l'ensemble des changements résultant des interventions techniques et financières collaboratives d'un certain nombre d'agences de développement. L'analyse inclut aussi quelquefois les programmes d'agences nationales (du gouvernement comme de la société civile).

Analyse de substitution : On fait une évaluation de l'*augmentation nette des ressources* au secteur cible ou au programme du fait de la contribution d'une agence de développement particulière. Ceci prend en compte un éventuel détournement des ressources vers d'autres activités par le gouvernement national (ce qui réduit l'augmentation nette des ressources) ou un éventuel effet multiplicateur (ce qui accroît l'augmentation des ressources).

Source : Adapté de Bamberger et al 2012 Cadre 16.3 pp 403-4.

- Si la plupart des interventions complexes sont de grande envergure, certaines sont relativement petites mais avec des processus de changement de comportement complexe ou des populations difficiles à identifier ou à étudier.

Les modèles à méthodes mixtes sont bien adaptés à l'évaluation d'interventions complexes, du fait qu'il est possible de combiner une conception, une collecte des données et une analyse standardisées à des outils qui peuvent saisir les complexités du contexte du programme, sa nature changeante et ses effets prévus ainsi que les processus de changement de comportement qui se produisent. Il est aussi possible de combiner les perspectives de différentes parties prenantes à des indicateurs quantitatifs « objectifs » et il existe toute une gamme d'outils QUAL permettant de collecter des données sensibles auprès de populations difficiles à atteindre. En même temps, les MM donnent des moyens d'estimer les impacts quantitatifs de la manière la plus crédible possible dans des circonstances difficiles et d'utiliser la triangulation pour augmenter la validité conceptuelle des différents résultats et de l'interprétation de ce qui se passe réellement sur le terrain.

La Figure 6 résume le large éventail d'outils et de techniques que peut utiliser un modèle MM pour sélectionner le modèle approprié pour évaluer une intervention complexe (toutes les techniques énumérées sur cette figure sont décrites en [Annexe 7](#)). Un élément central de l'approche des méthodes mixtes est qu'elle donne plus de manières de définir et d'estimer le scénario contrefactuel.¹⁰ Ceci se fait en deux étapes.

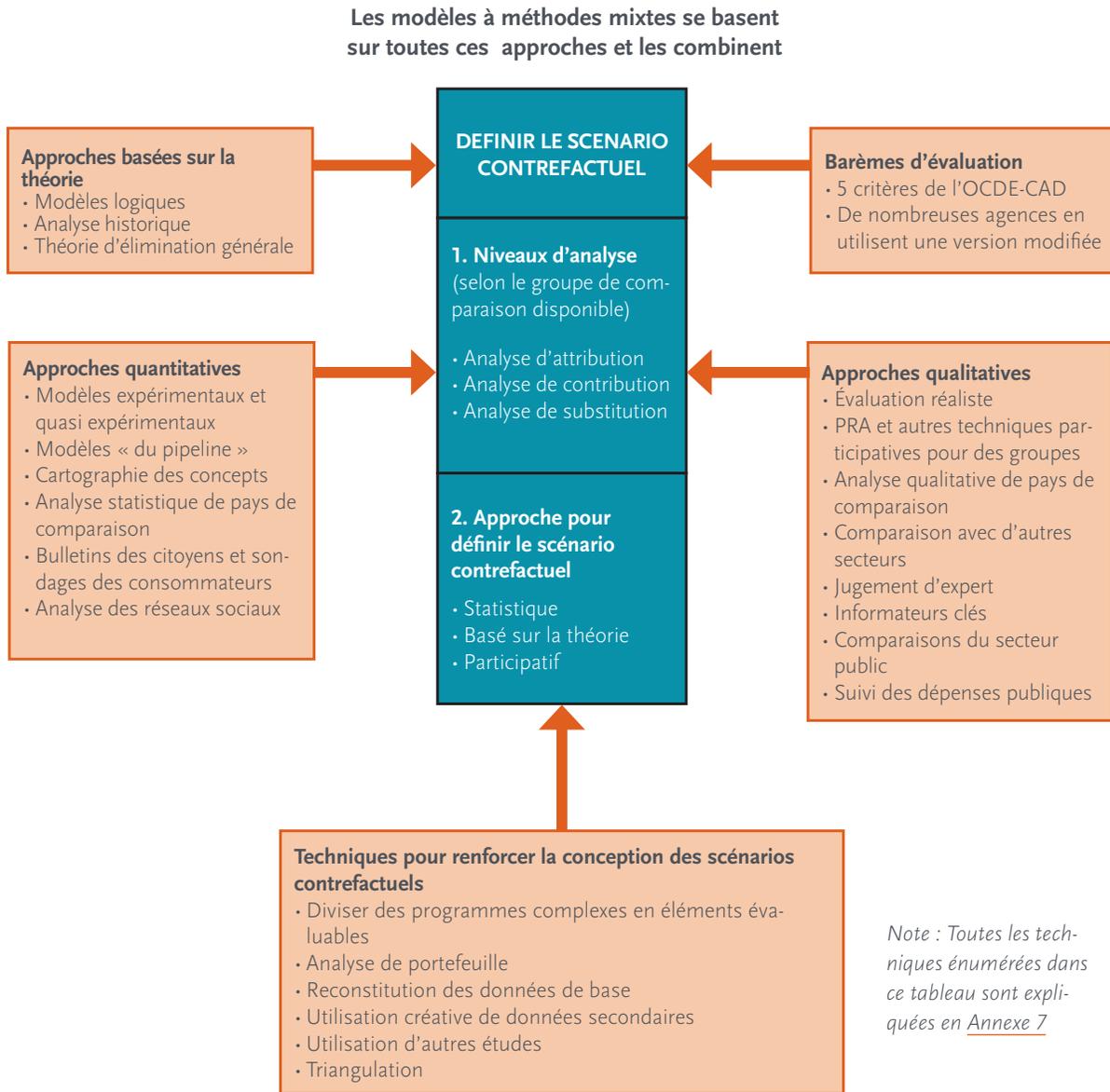
Étape 1 : Selon le type de groupe de comparaison disponible, l'évaluation fera une analyse à l'un des niveaux suivants¹¹ (Cf. Cadre 3) :

- Analyse d'attribution
- Analyse de contribution
- Analyse de substitution

¹⁰ La N^o1 (p.1) définit le scénario contrefactuel comme « une estimation de ce qui se serait passé si l'intervention n'avait pas eu lieu par exemple des comparaisons avec un groupe qui n'a pas bénéficié de l'intervention. »

¹¹ Cf. dans N^o1 Section 9 une présentation d'un large éventail d'approches pour évaluer l'attribution et la contribution. Bamberger et al (2012) pp 403-405 donne une présentation plus approfondie qui inclut l'analyse de substitution.

Figure 6. Utiliser des modèles à méthodes mixtes pour évaluer des interventions complexes



L'étape 2 sélectionne ensuite l'approche, ou la combinaison d'approches, qui sera utilisée pour définir et estimer le scénario contrefactuel :

- *Groupe de comparaison statistique* : utiliser un groupe de contrôle ou de comparaison correspondant du point de vue statistique
- *Basée sur la théorie* : le processus de mise en œuvre et les effets définis dans la théorie du changement (TdC) sont comparés à ce qui est observé sur le terrain. Au plus la réalité correspond étroitement à la TdC, au plus il est permis de supposer que l'intervention a contribué aux effets. Il est cependant également important de définir et de tester des explications alternatives aux changements observés (hypothèses rivales).¹²
- *Basée sur la participation* : les effets du programme sont estimés grâce à des évaluations faites par les bénéficiaires prévus et d'autres parties prenantes. Le changement le plus significatif est un exemple de cette approche, ainsi que l'utilisation d'évaluation rurale participative (PRA) et d'autres techniques d'analyse de groupes.
- *Barèmes d'évaluation* : on demande à des experts ou à des parties prenantes d'évaluer la performance du programme ou le degré de changement produit par l'intervention sur un ensemble de barèmes d'évaluation. Les échelles OCDE-CAD pour évaluer la pertinence, l'efficacité, l'efficience, l'impact et la durabilité ne sont qu'un exemple de barème d'évaluation couramment utilisé.

¹² Une faiblesse de nombreuses TdC est qu'elles n'identifient pas d'explications alternatives pour les effets observés. Si les changements sont cohérents avec la TdC, cela est considéré comme une preuve que ces changements peuvent être attribués aux effets du projet, ce qui n'est pas une conclusion valable. Une TdC devrait identifier des explications alternatives (des hypothèses rivales) et inclure dans le modèle de manière de tester ces explications rivales des causes des changements observés.

Les modèles à méthodes mixtes peuvent combiner, selon les cas : des approches basées sur la théorie, des approches qualitatives et statistiques et des barèmes d'évaluation. L'Annexe 7 donne une liste d'options possibles pour chacune de ces approches. Un certain nombre de techniques peuvent être utilisées pour renforcer la conception du scénario contrefactuel (« développer » les interventions complexes, analyses de portefeuille, reconstitution des données de base, utilisation créative de données secondaires et d'autres études). La validité conceptuelle des estimations tirées de toutes ces sources est ensuite renforcée en utilisant la triangulation.

Il n'y a pas une seule meilleure approche de l'évaluation par méthodes mixtes car les évaluateurs doivent sélectionner les outils et techniques les mieux adaptées au budget, au temps, aux données et aux contraintes politiques ainsi qu'aux objectifs de chaque évaluation. Les études de cas citées dans la 5^{ème} partie illustrent l'éventail d'approches de méthodes mixtes utilisées pour différentes évaluations. Pour les évaluations à méthodes mixtes, il faut de la créativité, une bonne connaissance de nombreuses approches QUANT, QUAL et théoriques différentes et la volonté de s'appuyer sur différentes disciplines.

3.3. Évaluer les processus de changement de comportement

Il y a de nombreux projets où le processus de mise en œuvre est beaucoup moins précis et linéaire qu'il ne le semble à première vue. Cela signifie que l'évaluation d'impact doit prendre en compte ces processus de changement de comportement car à cause de ceux-ci, les programmes ont souvent un certain nombre d'effets et d'impacts imprévus. Voici quelques raisons pour cela :

CADRE 4. LES EFFETS DES PROGRAMMES SONT AFFECTÉS PAR LA MANIÈRE DONT LES BÉNÉFICIAIRES RÉAGISSENT AUX SERVICES ET DONT LE PERSONNEL RÉAGIT AUX RÉACTIONS DES BÉNÉFICIAIRES. ÉTUDE DU CHANGEMENT DE COMPORTEMENT DANS UN PROGRAMME DE CANTINES SCOLAIRES AU NICARAGUA.

Un programme de petit déjeuner à l'école a été organisé pour augmenter la scolarisation dans une zone rurale du Nicaragua sujette à l'insécurité alimentaire. Dans certaines écoles, les enseignants donnaient aussi de la nourriture aux petits frères et sœurs qui accompagnaient leurs mères quand elles conduisaient les élèves à l'école. Elles sont passées du mot à d'autres mères et, comme le nombre de frères et sœurs augmentait, certains enseignants donnaient un petit déjeuner à tous les frères et sœurs et d'autres non. La décision dépendait en grande partie non seulement de l'attitude de chaque enseignant mais aussi de la bonne volonté des mères pour aider à préparer les petits déjeuners et des paysans locaux pour donner plus de nourriture. Dans certaines communautés, ces réactions comporte-

mentales ont transformé le programme de petit déjeuner à l'école en un grand programme de nutrition, avec une mobilisation de nourriture venant des paysans locaux, l'aide des hommes pour transporter de grandes quantités de nourriture, la construction de cuisines scolaires et, dans certains cas, l'installation de courant électrique dans les classes, tandis que les mères organisaient les petits déjeuners et mettaient parfois en route des potagers scolaires. Il n'a été possible de déterminer les effets finaux dans chaque école et les leçons apprises que grâce à une bonne observation du comportement des enseignants, des parents et des communautés et ces effets semblaient varier selon les attitudes individuelles, les conditions économiques locales et parfois l'origine ethnique.

- Dans la plupart des cas, les bénéficiaires prévus décident activement de participer ou non au programme, en choisissant souvent quels services ils veulent utiliser ou non.
- Les programmes sont souvent modifiés selon la perception et la réaction des premiers bénéficiaires aux services et le feedback qu'ils donnent à d'autres. La conception de nombreux programmes est modifiée suite à ces interactions à l'intérieur de la population cible et entre celle-ci et les prestataires de services (cf. Cadre 4).
- Le personnel des projets a ses propres attitudes vis à vis des différents groupes, des convictions sur qui va bénéficier ou pas et ses propres manières d'interagir avec la communauté.¹³ Même quand le personnel reçoit

une formation rigoureuse, chaque personne réagit quand même de manière légèrement différente.

- Chaque projet opère dans un contexte unique, où la mise en œuvre et les effets sont affectés par un ensemble unique de facteurs sociaux, économiques, politiques, environnementaux, légaux et autres.¹⁴

Voici quelques manières d'évaluer ces programmes avec des méthodes mixtes :

- Une étude diagnostique rapide faite au début de l'évaluation peut aider à clarifier la situation dans laquelle le programme est mis en œuvre et les facteurs contextuels qui risquent d'influer sur la mise en œuvre. L'étude peut aussi décrire la stratification sociale et identifier les groupes marginaux et vulnérables qui

¹³ Une revue des facteurs influant sur la réussite des systèmes éducatifs publics du Brésil pour améliorer l'accès des familles à bas revenu a montré que beaucoup d'enseignants pensaient que les enfants venant de ménages pauvres allaient probablement abandonner l'école et que, même s'ils restaient, ils auraient de mauvaises notes et seraient perturbateurs. Par conséquent, beaucoup d'enseignants avaient l'impression que cela ne valait pas la peine de faire l'effort d'aider ou d'encourager les enfants de milieux pauvres. Selon Bamberger and Segone (2011), beaucoup de causes d'inégalité sont déterminées socialement parce que les gouvernements et l'opinion n'ont peut-être pas envie d'encourager l'accès de différents groupes immigrés ou de minorités ethniques aux services publics.

¹⁴ L'analyse des systèmes donne des cadres utiles pour comprendre comment opèrent ces acteurs contextuels et aussi comment l'histoire affecte les attitudes et les attentes des parties prenantes et des communautés vis-à-vis de nouvelles interventions de projets. Il n'y a pas suffisamment de place ici pour une présentation de l'analyse des systèmes. On trouvera une brève introduction à l'analyse des systèmes et la manière dont on peut l'utiliser dans une évaluation d'impact dans Bamberger (2011) *How to design and manage equity focused evaluations* Section 5.2.C.

pourraient ne pas avoir accès aux bénéfices du projet. Une analyse QUAL initiale peut être combinée à une enquête rapide QUANT pour estimer, par exemple, l'importance et la distribution des groupes vulnérables.

- Le contenu du suivi normal du programme peut être étendu pour donner des informations QUANT plus détaillées sur les caractéristiques des groupes sociaux qui utilisent ou n'utilisent pas les services du projet.
- Le cadre d'analyse des goulets d'étranglement élaboré par l'UNICEF peut donner un cadre plus rigoureux pour l'analyse des facteurs qui déterminent quels secteurs ont ou non accès au projet et les facteurs du côté de l'offre et de la demande qui déterminent cet accès.¹⁵

- Il existe un certain nombre de techniques QUAL permettant d'observer ce qui se passe réellement durant la mise en œuvre du projet, comment les différents groupes réagissent au projet et comment le concept et le plan de mise en œuvre d'origine sont modifiés par l'interaction entre les différents secteurs de la population cible et le personnel du projet. Ces techniques sont, entre autres : observation des participants : études en panel avec des visites régulières à un petit échantillon de personnes, de ménages ou de communautés tout au long de la durée du projet ; groupes de consultation ; techniques d'évaluation rurale participative (PRA) et auto-rapports.¹⁶

¹⁵ Cf. Bamberger et Segone (2011) *How to design and manage equity-focused evaluations*. Pp. 45-50. UNICEF. Disponible sur : http://mymande.org/?q=equity_focused_evaluations_intro.

¹⁶ Les techniques d'auto-rapport sont par exemple de demander à un petit nombre de personnes d'écrire un journal où ils notent les expériences et les activités en lien avec le projet ou de demander à des ménages de tenir des comptes de leurs revenus et leurs dépenses. Dans certaines études, on donne un téléphone portable aux personnes interrogées à qui l'on demande d'appeler pour rendre compte, par exemple, de leurs déplacements, des endroits où ils se sentent en danger ou tentés d'acheter de la drogue, etc.

4^{ème} partie : Gérer les évaluations à méthodes mixtes

4.1. Les modèles à méthodes mixtes nécessitent une approche de gestion spéciale

Même si pour certaines applications de modèles à méthodes mixtes, il s'agit seulement d'ajouter d'autres méthodes de collecte de données à un modèle principalement QUANT ou principalement QUAL, une stratégie MM totalement intégrée demande bien plus que cela. Pour bénéficier à plein de la combinaison des approches et méthodes QUANT et QUAL, il faut planifier une stratégie d'évaluation MM dès le premier jour. Il faut pour cela un engagement des gestionnaires pour garantir le temps et les ressources supplémentaires nécessaires à une utilisation efficace de cette approche, et il faudra les défendre contre les pressions administratives demandant des coupes dans les budgets et le temps. Pour cela, le responsable de l'évaluation devra souvent faire une campagne énergique pour obtenir l'adhésion de la direction et des agences de financement. Des sessions d'informations au cours des réunions de direction, de petits rapports ou des formations sur les méthodes, des bulletins d'information, de petites vidéos, etc. peuvent être des outils de plaidoyer utiles.

L'équipe d'évaluation doit aussi pouvoir donner des preuves convaincantes que les MM produisent réellement des avantages supplémentaires. Par exemple, on peut choisir un petit nombre d'évaluations typiques auxquelles sera accordé un supplément de temps et de ressources pour tester l'approche MM et comparer les résultats avec les approches d'évaluations habituelles utilisées sur des programmes similaires. Est-il possible de démontrer que le temps et l'argent supplémentaires

nécessaires pour une approche MM augmentent réellement la valeur de l'évaluation ?

Pour gérer une évaluation MM il faut avoir suffisamment de temps et de ressources mais aussi porter une attention particulière aux domaines suivants (cf. Cadre 5) :

Composition de l'équipe de recherche. Dans l'idéal, l'équipe de recherche devrait comprendre des enquêteurs principaux issus de deux disciplines ou plus (par ex. anthropologie, médecine, droit, sociologie, économique). Cependant, pour la plupart des évaluations, ce ne sera pas possible parce que les ressources seront limitées et il faudra trouver des moyens pour que les membres de l'équipe d'évaluation puissent couvrir tous les outils et techniques QUANT et QUAL de base. Le défi pour le responsable de l'évaluation est d'aider à développer ces compétences de bases, souvent avec l'aide de consultants et d'autres agences. Quand les membres de l'équipe ont des compétences et des orientations professionnelles différentes, il est important que chaque enquêteur ait le temps et la possibilité de se familiariser avec la méthodologie des autres et qu'un respect mutuel et une confiance puisse se développer entre les membres de l'équipe. Ceci est encore plus important quand l'équipe d'évaluation comprend des spécialistes issus de disciplines différentes et qui, en plus, viennent de pays différents. Le tableau 5 montre quelques autres domaines d'expertise qui peuvent être nécessaires pour des équipes ayant une orientation QUANT ou QUAL.

Gérer des approches de méthodes mixtes durant la conception de l'évaluation. Il faudra peut-être un

CADRE 5. PLANIFIER ET BUDGÉTISER LE TEMPS ET LES RESSOURCES SUPPLÉMENTAIRES QUI POURRONT ÊTRE NÉCESSAIRES POUR UNE ÉVALUATION À MÉTHODES MIXTES

Les évaluations MM demandent souvent plus de temps (par ex. pour développer l'esprit d'équipe) et plus d'argent. Le responsable doit décider si ces ressources supplémentaires se justifient et, si oui, s'assurer qu'elles sont prévues au budget.

Composition et intégration de l'équipe de recherche. Le responsable d'évaluation a un rôle crucial à jouer pour permettre aux membres de l'équipe issus de disciplines différentes de travailler efficacement ensemble. Il faudra y consacrer plus de temps et d'effort.

- Donner aux enquêteurs le temps d'apprendre à comprendre et à respecter les disciplines et le travail les uns des autres. S'assurer que tous les membres de l'équipe connaissent la littérature de base et les débats actuels de l'autre domaine.
- Permettre le développement des mêmes liens et du même esprit d'équipe entre les enquêteurs locaux.

Approches intégrées durant la conception de l'évaluation : Le responsable d'évaluation doit activement encourager tous les membres de l'équipe à tirer pleinement parti des nouveaux cadres et outils de collecte et d'analyse des données.

- S'assurer que les enquêteurs issus des différentes disciplines sont présents dès la phase de conception de l'évaluation ou à peu près au même moment, pour que chacun ait le sentiment qu'il contribue de manière significative à l'ensemble de l'évaluation et qu'il n'est pas juste là pour « faire joli. » C'est particulièrement important quand une discipline est dominante.
- S'assurer que l'évaluation se base sur des théories et des approches issues de toutes les disciplines comprises dans l'évaluation (par ex. anthropologie, médecine, droit, sociologie, économie, démographie), chacune étant utilisée pour enrichir et approfondir les autres.
- S'assurer que les concepts et méthodes ne sont pas pris hors contexte mais qu'ils sont basés sur les débats intellectuels et les approches à l'intérieur de chaque discipline.

Collecte des données et utilisation de triangulation : De nombreuses propositions d'évaluation font référence à la triangulation mais celle-ci n'est pas souvent utilisée systématiquement. La triangulation est un élément clé d'une approche de méthodes mixtes et il est de la responsabilité du responsable de vérifier qu'elle est pleinement utilisée.

- Choisir des méthodes de collecte de données QUANT et QUAL qui se complètent et préciser comment elles seront combinées dans le travail sur le terrain et l'analyse
- Choisir au moins deux méthodes indépendantes d'estimation pour les principaux indicateurs et hypothèses.
- Faire bien documenter toutes les méthodes de sélection d'échantillons, de collecte de données et d'analyse.

Analyse des données et éventuel suivi sur le terrain : Le responsable devrait s'assurer qu'un plan d'analyse de données MM est en place dès le début de l'évaluation.

- Présenter des analyses séparées des résultats QUANT et QUAL pour mettre en lumière les différentes interprétations et les différents résultats et préparer un rapport intégré basé sur toutes les données.
- Utiliser des procédures systématiques de triangulation pour vérifier les incohérences ou les différences d'interprétation.
- Budgétiser des ressources et du temps pour des visites de suivi sur le terrain.
- Mettre l'accent sur les différentes interprétations et les différents résultats obtenus par les différentes méthodes et discuter pour voir comment cela enrichit l'étude. Des résultats différents, apparemment contradictoires, devraient être considérés comme un des principaux points forts d'une approche intégrée plutôt que comme un désagrément.
- Présenter les cas et le matériel QUAL pour illustrer ou tester les résultats QUANT.

Présentation et distribution des résultats : Le responsable devrait encourager l'équipe à élargir la gamme des méthodes de présentation et de distribution pour bien saisir toute la richesse des données MM.

- Combiner les rapports écrits conventionnels et les présentations PowerPoint à des méthodes de présentation plus participatives. Élaborer des rapports plus novateurs et faciles d'utilisation et éviter les longs rapports techniques pour les publics qui ne sont pas composés de techniciens.
- Élargir l'éventail des parties prenantes invitées à la présentation des résultats à des groupes communautaires ou de la société civile qui ne sont pas souvent consultés dans beaucoup d'évaluations QUANT.

Tableau 5. Autres domaines d'expertises dans la recherche, dont peuvent avoir besoin des équipes à orientation QUANT et QUAL pour faire des évaluations MM.

Évaluations à orientation QUANT	Évaluations à orientation QUAL
<ul style="list-style-type: none"> • Connaissance de méthodes de base de collecte de données QUAL (par ex. entretiens approfondis, techniques d'entretiens en groupes, techniques d'observation) • Connaissance de la culture locale • Expérience et évaluation de programmes de développement des communautés et des organisations • Utilisation systématique de la triangulation 	<ul style="list-style-type: none"> • Méthodes d'échantillonnage statistique • Gestion de la collecte de données QUANT, en particulier l'administration de questionnaires structurés • Analyse de données statistiques

appui en gestion pour que le cadre d'évaluation s'appuie sur toutes les disciplines concernées et que les questions et problèmes de recherche incorporent chacune de ces méthodologies. Ceci peut être important pour des organisations qui utilisaient normalement des méthodes principalement QUANT ou principalement QUAL et où l'équipe d'évaluation aura peut-être besoin d'encouragement pour donner un poids égal aux cadres et méthodes des méthodes que l'on vient d'introduire.¹⁷

S'assurer que les MM sont utilisés durant la collecte des données. La décision d'utiliser une collecte de données consécutive ou simultanées est en partie une question technique mais elle a des implications de gestion. Collecter les données QUANT et QUAL de manière consécutive peut permettre de gagner beaucoup de temps, ce qui peut aussi faire économiser de l'argent, mais il faudra des systèmes efficaces de gestion pour coordonner les équipes d'enquête QUANT et de collecte de données QUAL qui auront peut-être des méthodes d'opération différentes. Quand il y a des contraintes de sécurité et des difficultés logistiques pour la collecte des données, l'utilisation de collectes de données

¹⁷ Il arrive souvent que les professionnels des nouvelles disciplines ne soient appelés qu'une fois que la conception de l'évaluation est pratiquement terminée et qu'on leur demande seulement de collecter des données qui correspondent au modèle d'évaluation déjà défini. L'expérience montre que, souvent, les évaluations ayant une orientation QUANT embaucheront des enquêteurs QUAL seulement pour mener un groupe de consultation ou pour faire des études de cas afin de montrer que l'évaluation utilise des méthodes mixtes, mais sans vraiment inclure celles-ci dans la conception globale de l'évaluation.

consécutives nécessitera peut-être une gestion efficace du transport pour amener les enquêteurs dans les différentes zones et aller les rechercher ainsi qu'une coordination avec la police et l'armée dans les situations où il faut une autorisation pour aller dans des communautés et où on doit peut-être aussi préparer et faire approuver des itinéraires précis.

Gérer l'analyse des données de méthodes mixtes.

Avec un modèle à dominante QUANT, l'analyse des données ne commence généralement pas avant que toutes les données, ou au moins la plupart, n'aient été collectées, entrées dans une base de données et nettoyées. Cependant, l'analyse de données QUAL peut être continue, commençant peut-être après l'arrivée des enquêteurs sur le terrain. La gestion des données devient plus difficile parce qu'il s'agit d'un processus interactif plutôt que linéaire. Les résultats initiaux de l'analyse des données QUAL sont souvent utilisés pour concevoir des questions pour les enquêtes QUANT et l'analyse initiale des données QUANT est souvent utilisée pour sélectionner les échantillons pour l'analyse QUAL approfondie. Ceci pose un certain nombre de problèmes de gestion particuliers: (a) la collecte des données doit être faite plus rapidement pour ne pas retarder le début de la phase suivante de l'évaluation et de la collecte des données ; (b) les procédures de vérification de la qualité des données sont plus complexes car elles doivent évaluer la qualité de la collecte de données QUANT et QUAL ; (c) on doit utiliser systématiquement des

procédures de triangulation pour vérifier la cohérence des estimations des principaux indicateurs d'effets obtenus par différentes méthodes de collecte de données et il faut des procédures pour faire le suivi et déterminer quelles sont les raisons expliquant les incohérences entre des données venant de différentes sources. Tout ceci demande des systèmes de gestion des données plus complexes et plus efficaces.

4.2. Conseils pour permettre aux ONG ayant des ressources limitées de mobiliser l'expertise et les ressources nécessaires pour faire des évaluations à méthodes mixtes

Les évaluations à méthodes mixtes offrent un grand avantage aux ONG qui ont besoin de rapports d'évaluation crédibles et de qualité mais qui doivent faire les évaluations avec peu de moyens, peu d'expertise de l'évaluation à l'interne et/ou peu de ressources pour faire venir des experts de l'extérieur. Pour les nombreuses ONG qui font des évaluations dans ces conditions courantes, il faudra souvent faire preuve de créativité pour obtenir l'expertise nécessaire. Il peut s'agir de fournir une formation supplémentaire à l'équipe actuelle de recherche ou de trouver des moyens d'obtenir l'expertise nécessaire en collaborant avec d'autres agences. Il n'y a pas de règles absolues fixant les niveaux minimums d'expertise, de temps et de ressources nécessaires pour faire une évaluation MM rigoureuse d'un point de vue méthodologique, mais il faudrait au moins un niveau minimum d'expertise dans tous les domaines cités dans le Tableau 5. Le cadre 6 donne des conseils sur la manière d'arriver à l'expertise essentielle et de mobiliser d'autres ressources quand on travaille avec un budget limité.

Toutes ces contraintes influent sur la crédibilité des résultats de l'évaluation. Quelle est la crédibilité d'une évaluation qui a dû être faite avec un budget serré, peu de temps sur le terrain, un accès limité aux données et éventuellement avec des problèmes de sécurité ? Même s'il n'y a pas de « solution magique »

pour résoudre ces problèmes, il y a de manières d'avoir un niveau de rigueur méthodologique acceptable et de résultats crédibles grâce aux MM. Voici quelques bonnes manières d'obtenir de résultats crédibles quand on travaille avec ces contraintes.

a. Baser l'évaluation sur une théorie du changement bien formulée

(N°1 et N°2 donnent toutes deux des descriptions complètes de l'élaboration et de l'utilisation de TdC dans la conception et l'analyse de l'évaluation d'impact). Une TdC élaborée en concertation participative avec les parties prenantes peut définir les étapes et processus grâce auxquels les effets, les produits et les impacts doivent être atteints et elle peut identifier les hypothèses clés qui devront être évaluées. La TdC doit avoir des étapes importantes convenues, pour que, s'il faut faire une évaluation à un moment où il est encore trop tôt pour mesurer des effets, l'étape importante puisse aider à augmenter la crédibilité d'indices montrant que le projet est en bonne voie d'atteindre ses objectifs. Les MM peuvent renforcer la TdC en incorporant à la fois des indicateurs QUANT et des indicateurs QUAL, en étudiant ce qui se passe pendant la mise en œuvre du projet et en décrivant des processus importants de changement de comportement.

b. Concertation avec des parties prenantes pour qu'elles trouvent les méthodes d'évaluation et les indicateurs clés crédibles, et aussi pour connaître les préoccupations qu'elles peuvent avoir et y répondre.

c. Utiliser la triangulation pour maximiser la validité des estimations, sur la base de petits échantillons de sources différentes (cf. Section 2.2)

d. Utiliser un échantillonnage de méthodes mixtes pour garantir la représentativité des données QUAL tirées des études de cas, des entretiens approfondis et de l'observation (cf. Section 3.1, Modèle 2).

CADRE 6. FAIRE DES ÉVALUATIONS D'IMPACT À MÉTHODES MIXTES AVEC PEU DE MOYENS CONSEILS POUR LES ONG QUI ONT PEU DE RESSOURCE ET D'EXPERTISE EN ÉVALUATION

Même si les méthodes mixtes peuvent être utilisées dans le cadre d'évaluations d'impact rigoureuses, chères et complexes, la plupart des ONG n'ont que peu de temps et de moyens à investir dans les évaluations d'impact et n'ont qu'une expertise limitée de l'évaluation. Et pourtant, elles ont besoin d'évaluations valables et crédibles des effets et des impacts de leur programme. Heureusement, les outils et techniques MM sont très souples et peuvent aider à améliorer la qualité des évaluations d'impact pour les ONG faisant ces évaluations avec peu de moyens. Voici quelques conseils pratiques présentés dans cette note d'orientation.

Commencez progressivement et augmentez au fur et à mesure que l'expérience s'enrichit

- Il est généralement plus simple et plus économique de commencer par utiliser les MM à une seule étape de l'évaluation, par exemple en utilisant au moins deux méthodes indépendantes de collecte de données ou en faisant une étude diagnostique rapide pour aider à préparer un questionnaire. Tirez parti des données qui sont déjà produites par le système de suivi du programme. Au fur et à mesure que votre expérience augmentera, envisagez d'en faire plus.
- Il est généralement plus facile de commencer en utilisant des modèles d'évaluation d'impacts consécutifs plutôt que simultanés (cf. Section 2.2).
- Même si certaines approches de méthodes mixtes sont assez compliquées à utiliser, il existe de nombreuses autres techniques faciles et économiques et qui peuvent être facilement apprises et utilisées par des employés n'ayant que peu d'expérience en recherche.
- Intégrez le modèle MM dans le système de S&E et utilisez le plus possible les données collectées par le S&E (cf. N°2).

Vérifiez que les résultats sont à la fois rigoureux du point de vue méthodologique et crédible pour les parties prenantes

- Pour commencer, concentrez-vous sur le type de données que les parties prenantes considèrent comme crédibles. Ne dépensez pas toutes vos ressources sur une grande enquête sur échantillon si vos clients et parties prenantes s'intéressent plus aux études de cas et aux descriptions approfondies de la manière dont le programme opère réellement.
- En même temps, essayez de trouver un équilibre entre les préférences des parties prenantes d'un côté et une combinaison judicieuse de données QUANT et QUAL et l'utilisation de la triangulation et d'autres tech-

niques permettant d'améliorer la validité de l'autre. Les études MM collectent un large éventail de données quantitatives et qualitatives de différents types de sorte que, même à partir d'une petite évaluation, on peut sélectionner le type d'indices que les parties prenantes trouveront les plus convaincants.

- L'échantillonnage MM peut permettre de sélectionner un petit nombre d'études de cas qui soient largement représentatives de toute la population cible, ce qui augmente la validité et la crédibilité des résultats (cf. Section 2.3).
- En combinant de manière créative les indices obtenus de différentes sources, la triangulation peut augmenter la crédibilité d'estimations tirées de petits échantillons (cf. Section 2.2 et Annexe 9).

Restez à l'intérieur de votre budget et à votre niveau d'expertise

- Il y a de nombreuses manières de réduire les coûts de la collecte des données tout en collectant des données fiables (cf. Annexe 8).
- De nombreux départements d'universités enseignent les méthodes mixtes et il sera peut-être possible de collaborer, gratuitement ou pour un prix symbolique, avec un enseignant ou des étudiants pour l'un de leurs travaux sur le terrain. De nombreux étudiants de troisième cycle pourraient envisager de préparer une étude de cas dans le cadre de leur travail pour un cours.

Faites-vous aider

- Il existe beaucoup de séminaires gratuits en ligne et d'autres ressources sur Internet sur la conception et l'utilisation des MM. (<http://www.interaction.org/resources/training> est un bon point de départ).
- Nombreuses sont les grandes ONG internationales qui ont un département permanent d'évaluation qui pourrait accepter de donner gratuitement des conseils au téléphone ou en personne si quelqu'un passe par votre pays.

Exemples concrets de modèles à méthodes mixtes simples et économiques

- L'étude de cas No 7, évaluation du projet d'éducation de l'UNICEF au Timor Oriental et l'étude de cas No 11, évaluation du fonds de développement communautaire en Érythrée, illustrent toutes deux la manière dont les méthodes mixtes peuvent être utilisées quand les données sont difficiles à collecter et que les ressources sont limitées.

5^{ème} partie : Études de cas illustrant différentes applications de modèles à méthodes mixtes

Les trois évaluations décrites dans cette section illustrent les méthodologies utilisées dans : une évaluation essentiellement QUANT, une évaluation essentiellement QUAL et une évaluation donnant le même poids à la méthode QUANT et à la méthode QUAL. Ensemble, ces trois exemples illustrent le grand nombre d'approches d'évaluation MM qui peuvent être utilisées. Le Tableau 6 (à la fin de cette section) cite 17 exemples de modèles d'évaluation MM qui sont résumés en [Annexe 10](#). Les évaluations correspondent à des projets appuyés par des ONG, des agences des Nations unies et la Banque mondiale.

Exemple 1 : Une évaluation à orientation QUANT : Évaluation d'un programme de reconstruction après conflit au Libéria (DFDI et International Rescue Committee)

Cette évaluation, terminée en décembre 2008, évaluait l'impact d'un programme de reconstruction dirigée par la communauté (CDR), mis en œuvre en 2006-07 par International Rescue Committee dans le Nord du Libéria, qui sortait d'une période de conflit. Dans 42 communautés, le programme CDR a utilisé des subventions forfaitaires, des institutions décisionnaires élues, la planification participative et des projets de développement communautaire pour améliorer le bien-être socio-économique, la gouvernance locale et la cohésion de la communauté. L'évaluation déterminait si le programme avait atteint son objectif.

Le modèle utilisé était un modèle d'évaluation à méthodes mixtes. Ce qui motivait ce choix, c'était les difficultés d'évaluer, en se basant uniquement sur des données d'enquêtes QUANT, si les

changements d'attitudes déclarés reflétaient de réels changements. L'idée était que, du fait de leur exposition au projet, les participants risquaient d'avoir une idée de ce qu'on voulait leur entendre dire sur leur attitude vis-à-vis de la coopération et de la démocratie, sans que leur capacité ou leur inclination à s'impliquer dans une action communautaire collective ne soit vraiment affectée. L'évaluation était conçue pour remédier à ces difficultés en combinant des données d'enquête QUANT sur les changements d'attitudes et le comportement déclaré avec des données expérimentales sur le terrain qui pourraient mesurer les vrais changements de comportement. L'enquête comme les méthodes expérimentales sur le terrain étaient basées sur l'essai randomisé contrôlé : au début, les communautés étaient sélectionnées de manière aléatoire pour recevoir des traitements CFC (développement de comités communautaires et appui pour des programmes à impact rapide) tandis qu'un nombre approximativement égal de communautés de contrôle n'ont pas reçu les programmes CRC. D'autres méthodes QUANT ont été utilisées : un questionnaire sur le capital social, pour évaluer à la fois comment les effets étaient affectés par le capital social et comment le programme affectait le capital social, et une collecte de données d'enquêtes sur l'organisation communautaire et la cohésion sociale.

Les méthodes expérimentales sur le terrain et les méthodes QUAL comprenaient la collecte de données sur l'organisation communautaire et la cohésion sociale, grâce à des entretiens approfondis et l'observation du comportement individuel dans un jeu sur les biens publics. Six mois après la fin du programme CDR, toutes les communautés, ayant reçu

le traitement et les communautés de contrôle, ont eu l'occasion de lever des fonds auprès d'une ONG libérienne pour mettre en œuvre un projet de biens publics. La seule condition était que les communautés choisissent un projet à l'avance et choisissent trois personnes pour gérer l'argent. On a aussi dit aux communautés que la somme d'argent qu'elles recevraient (allant de 0\$ à environ 500\$) dépendrait des résultats d'un jeu de biens publics au niveau de tout le village, où 24 personnes sélectionnées de manière aléatoire pourraient décider individuellement de garder une somme d'argent pour leur usage personnel ou de la verser dans un fonds communautaire (avec une contribution équivalente versée par les enquêteurs) qui serait utilisé pour le bien public. Les enquêteurs ont ensuite collecté des données sur la manière dont les communautés ont choisi leurs projets et leurs représentants et ont observé les tendances dans le jeu des biens publics. L'idée était que les jeux pouvaient donner une estimation plus précise des attitudes et des comportements car agir pour faire plaisir aux personnes extérieures avait un coût pécuniaire réel.

Exemple 2 : Une évaluation à orientation QUAL : Évaluation d'un programme de prévention de la violence à l'égard des femmes au Salvador (Oxfam Amérique)

Ce programme décennal de lutte contre la violence à l'égard des femmes (VAW) a été lancé en 2008. Il opère aux niveaux de la nation, des municipalités et des communautés, en combinant des campagnes de sensibilisation du public et l'organisation des femmes et de la société civile. Il promeut une coordination plus efficace entre les autorités publiques et les femmes organisées. L'une des principales difficultés est l'attribution des changements d'attitude et de comportement aux stratégies nationales et municipales de plaidoyer et de campagnes, particulièrement entre les agences du secteur public. La question du groupe de contrôle est difficile. Une autre difficulté

est le manque de bonnes données QUANT sur les niveaux de violence à l'égard des femmes et sur les changements dans cette violence. Les données de la police, de la justice et de la santé étaient à la fois peu fiables et sous-estimées. Comme toujours, les victimes étaient peu disposées à signaler les abus. Par conséquent, l'équipe d'évaluation a décidé que la manière la plus juste de mesurer l'influence de la compagne était de faire des études de cas comparatives approfondies, en se concentrant sur les objectifs intermédiaire, avec autant que possible des données QUANT pour les justifier.

Le modèle d'évaluation MM était basé sur une théorie du changement. Celle-ci identifiait comment des effets multiples promus en coalition pourraient s'ordonner et se combiner pour produire les impacts prévus : coordination et échanges entre les agences régionales ; nouvelles politiques et normes (législation, programmes scolaires, politiques municipales de prévention) ; changements dans les relations sociales et les comportements sociaux grâce à la sensibilisation et à une meilleure application de la loi ; bien-être et action (augmentation des connaissances ; femmes plus confiantes pour agir ; femmes bénéficiant de meilleurs services ; influence des femmes, des jeunes et de leurs alliés sur les décideurs). Un ensemble de 11 indicateurs à long terme a été défini pour évaluer les effets et les impacts au cours de la période de 10 ans et des « cibles jalons » à court terme pour mesurer les progrès ont été fixées tous les trois ans.

L'évaluation était basée sur des principes de processus (équipe pluridisciplinaire et approche de méthodes mixtes, institutions de recherche régionale, conception en collaboration et concertation avec tous les partenaires et utilisation de triangulation et de concertation pour valider les résultats). Les éléments principaux de l'approche méthodologique étaient :

- Études de cas comparatives approfondies dans deux municipalités
- Grande synthèse des documents des différentes parties prenantes
- Entretiens sur l'efficacité des choix de plaidoyer au niveau national
- Étude des liens entre les différents niveaux d'intervention (national, municipal et local)
- Différentes échéances pour l'évaluation de différents effets.

Les directions de recherche et les principales sources de données étaient les suivantes :

- Donner un contexte pour les tendances générales, en collectant des indicateurs aux niveaux national et municipal (statistiques gouvernementales, statistiques sur la violence à l'égard des femmes, promulgation et application de lois sur la violence à l'égard des femmes, promesses budgétaires).
- Deux études de cas comparatives approfondies, sur les résultats dans des municipalités particulières, les impressions des femmes sur la sécurité, l'évaluation de l'efficacité des actions de prévention, la confiance des femmes pour agir sur les questions de violence à l'égard des femmes. Les sources de données étaient des entretiens approfondis, l'observation, des statistiques et tendances au niveau municipal et des groupes de consultation où les femmes parlaient des changements de conditions.
- Évaluation du plaidoyer: Faire un suivi de l'influence de la campagne, grâce à l'étude de documents et à des entretiens avec des décideurs politiques et d'autres acteurs sur les principaux événements et points de décision et demander si les progrès auraient pu avoir lieu en l'absence de campagnes de plaidoyer. Évaluer si la campagne a influé sur la version finale de la législation et évaluer l'efficacité des

mécanismes de la société civile pour influencer sur les politiques de prévention. Utiliser des données secondaires pour comparer les tendances dans des domaines où les campagnes ont opéré ou non.

L'analyse et l'interprétation des résultats combinaient :

- Analyse de la contribution au niveau du programme en utilisant les données des composantes des projets, des études de cas et le repérage de l'influence (sur la base du cadre de la théorie du changement). La causalité et l'attribution peuvent être étudiées pour chaque projet de la coalition du projet, mais l'intérêt, au niveau du programme, est de tester les hypothèses de base d'un travail extrêmement complexe.
- La théorie du changement était un élément clé pour définir un scénario contrefactuel qui donnait un cadre pour comparer les changements et processus observés avec le processus de changement prévu.

Exemple 3 : Un modèle MM équilibré: Évaluation des impacts du Programme de réforme du développement communautaire des Gram Panchayats en Inde (Banque mondiale)

L'objectif du programme était de donner du pouvoir à des gouvernements villageois élus et de promouvoir une participation plus importante des femmes et des castes répertoriées. Il devait permettre une utilisation plus efficace des fonds, grâce à une participation plus importante de la communauté. Le gouvernement fournissait des subvention qui seraient gérées par la communauté locale pour mettre en œuvre des projets d'infrastructures économiques et sociales qui devaient être sélectionnés par la communauté. Le programme évalué était un programme de formation de deux semaines pour les citoyens sur la planification et la prise de décisions participatives.

Ce programme était conçu pour améliorer l'efficacité de leur utilisation des subventions. Le modèle du programme comportait une distribution aléatoire des communautés dans le groupe recevant le traitement et le groupe de contrôle (qui recevait les subventions mais pas la formation).

L'évaluation utilisait un modèle équilibré de méthodes mixtes comprenant les étapes suivantes:

- Sélection de 200 villages (Gram Panchayats) avec une distribution aléatoire entre le groupe du projet et le groupe de contrôle (QUANT).
- Recherche exploratoire sur la propriété foncière, la propriété des biens publics, la participation et les réseaux sociaux (QUAL).
- Enquête de base avant les programmes de formation (QUANT).
- Analyse approfondie des processus dans cinq zones de projet et cinq zones de contrôle (QUAL) pour observer les changements dans la manière dont les communautés organisent, planifient et gèrent les projets et les effets sur la participation et l'implication des femmes et des castes répertoriées. Cette analyse consistait en des visites d'un à deux jours chaque semaine sur une période de deux ans. Les visites combinaient l'observation au niveau de la communauté, des entretiens avec des politiciens et des autorités communautaires et des entretiens approfondis avec 20 ménages de chaque village. L'observation concernait : la dynamique politique et sociale, la corruption, le changement économique et l'appartenance à des réseaux.
- Répétition de l'enquête de base après deux ans (QUANT).
- Tout au long de l'évaluation, la triangulation a été utilisée pour comparer les estimations

QUANT et QUAL des changements et de l'impact.

Voici quelques uns des avantages qu'a eu le modèle MM :

- On ne peut pas se fier à la mémoire pour rendre compte des processus et des discussions lors des réunions villageoises et les observateurs ont donc assisté à ces réunions et noté les procédures et la dynamique de la réunion.
- Des études antérieures avaient montré que les enquêtes donnent des estimations extrêmement variées de l'inégalité (la variance est trop importante pour que les résultats puissent être utilisés pour une analyse statistique), on a donc utilisé, pour compléter les données de l'enquête, des techniques de PRA, où des participants ont été sélectionnés pour être représentatifs de toute la population.
- Le développement de la communauté, la participation, la corruption et d'autres processus clés sont difficiles à appréhender dans des enquêtes. Les données de l'enquête ont donc été comparées à l'observation et à des entretiens avec des informateurs clés et des groupes participatifs.
- En même temps, il était essentiel d'avoir des estimations quantitatives de la rentabilité des programmes de formation, car il s'agissait d'un programme pilote sur lequel seraient basées les décisions de répliquer ou non ce programme dans d'autres états.¹⁸ Par conséquent, les stratégies d'échantillonnage garantissaient la possibilité de pouvoir relier les études de cas et autres données qualitatives à l'enquête QUANT sur échantillon afin de pouvoir généraliser les résultats de l'évaluation.

¹⁸ L'évaluation comprenait une comparaison avec quatre autres états (qui n'est pas incluse dans ce résumé).

Tableau 6. Exemples d'évaluations MM résumées en [Annexe 10](#)

Titre	Agences qui appuient le programme et font ou font faire l'évaluation.
A. Modèle d'évaluation à orientation quantitative	
1. Reconstruction après le conflit au Libéria.	DFID et International Rescue Committee
2. Évaluation à long terme du programme Tostan pour réduire l'excision dans les villages du Sénégal.	UNICEF
3. Évaluation d'un programme conditionnel de transfert d'argent au Kazakhstan.	Save the Children
4. Évaluation de l'impact d'un travail d'urgence et de réhabilitation de la FAO dans les zones rurales de RDC.	FAO
B. Modèles d'évaluation à orientation qualitative	
5. Évaluation de l'impact d'un programme de prévention de la violence liée au genre au Salvador.	Oxfam Amérique
6. Programme de sécurité alimentaire vie et moyens d'existence au Bangladesh.	USAID, Save the Children et TANGO International
7. Évaluation du programme d'éducation de l'UNICEF au Timor Oriental.	UNICEF
8. Évaluation des résultats pour l'équité du projet d'éducation pour tous au Népal.	UNICEF, NORAD et autres partenaires
9. Évaluation des résultats pour l'équité du projet communautaire d'assainissement total au Cambodge.	UNICEF
10. Évaluation inter-agences en temps réel de la réponse humanitaire à la crise de déplacement de 2009 au Pakistan.	UNICEF et divers partenaires
11. Évaluation de la composante construction de routes du fonds de développement communautaire en Érythrée.	Banque mondiale
12. Évaluation du projet d'écoles communautaires en Égypte.	UNICEF
13. Évaluation du projet de facilitation de la justice communautaire en Tanzanie.	UNICEF
14. Évaluation de la réponse de l'UNICEF dans le domaine de la protection en Indonésie, pour le tsunami de 2004 dans l'Océan Indien.	UNICEF
C. Modèles équilibrés d'évaluation, donnant un poids égal aux approches QUANT et QUAL	
15. Évaluation du projet de développement de Kecamatan, en Indonésie.	Banque mondiale et Gouvernement indonésien
16. Évaluation du programme de réforme des Panchayats en Inde.	Banque mondiale
D. Méta-analyse (analyse secondaire de plusieurs évaluations pour identifier les résultats généraux)	
17. Rapport d'évaluation d'impact de Care International, 2005-2010, Amérique Latine et Caraïbes.	CARE International

Références

- Adato, M (2012) « Combining quantitative and qualitative methods for program monitoring and evaluation: Why are mixed methods designs best? » in Lopez-Acevedo, G; Krause, P & Mackay, K (editors) Building better policies : The nuts and bolts of monitoring and evaluation systems. Banque Mondiale.
- Bamberger, M. (2010). *Reconstructing baseline data for impact evaluation and results measurement*. N°4. Série The Nuts and Bolts of M&E. Poverty Reduction and Equity Department. Banque mondiale. Disponible sur : <http://siteresources.worldbank.org/INTPOVERTY/Resources/335642-1276521901256/premnoteME4.pdf>.
- Bamberger, M., Rao, V., & Woolcock, M. (2010). *Using mixed-methods in monitoring and evaluation: Experiences from international development*. In Tashakkori A. et Teddlie C. Sage Handbook of Mixed- Methods in Social and Behavioral Research. Sage Publications. pp 613-642.
- Bamberger, M., Rugh, J. & Mabry, L. (2012). *RealWorld Evaluation* Deuxième édition. Sage Publications.
- Chen, H. (2005) *Practical Program Évaluation : Assessing and improving planning, implementation, and effectiveness*. Sage Publications.
- Donaldson, S., Christie, C. & Mark, M. (2009). *What counts as credible evidence in applied research and evaluation practice ?* Sage Publications.
- Funnell, S. & Rogers, P. (2011). *Purposive programme theory : effective use of theories of change and logic models*. Jossey-Bass Publications.
- Mayne, J. (2008) « Contribution analysis : An approach to exploring cause and effect. » Rome : Institutional learning and change initiative, ILAC Brief N° 16. Mai 2008. http://www.cgiar-ilac.org/files/publications/briefs/ILAC_Brief16_Contribution_Analysis.pdf
- Organization for Economic Cooperation and Development. 2010 (Seconde édition). *Evaluating Development Cooperation : Summary of Key Norms and Standards*. Paris : OCDE. <http://www.oecd.org/dataoecd/12/56/41612905.pdf>
- Patton, M.P. (2011). *Developmental Evaluation: applying complexity concepts to enhance innovation and use*. Guilford Press.
- Proudlock, K & Ramalingam, B (2009) *Improving humanitarian impact assessment : Bridging theory and practice*. ALNAP
- Rieper, O Leeuw, F & Ling, T (2010). *The evidence book : concepts, generation and use of evidence*. Transaction Publishers.
- Salmen, L. (1987). *Listen to the People : Evaluation of Development Projects*. New York. Oxford University Press.
- Stern, E ; Stame, N ; Mayne, J ; Forss, K ; Davis, R & Befani, B (2012) *Broadening the range of designs and methods for impact evaluation*. DFID

UNICEF (2009a). Children and the 2004 Indian Ocean Tsunami : Evaluation of UNICEF's response in Indonesia, Sri Lanka and Maldives, 2005-08. Rapport général de synthèse.

Weiss 2001. Theory based evaluation: Theories of change for poverty reduction programs pp. 103-14 in Evaluation and Poverty Reduction édité par O. Feinstein et R. Piccitto. New Brunswick. Transaction

Williams B. (2005). Systems and Systems Thinking in Mathison, S (editeur) Sage Encyclopedia of Evaluation pp. 405-412.

Williams, B. & Imam I. (Éditeurs), (2007). Systems Concepts in Evaluation : An expert anthology. American Evaluation Association.