**Réalisation d’une évaluation**

*Aide - mémoire[[1]](#footnote-1)*

**Méthodologie**

Comment faire pour recueillir les informations définies dans le contenu? Quelles techniques choisir? Comment tenir compte des premiers constats pour modifier le cours de l’évaluation tout en restant fidèle aux objectifs de départ? Comment garder une certaine souplesse pour tenir compte de l’imprévisible ou de facteurs “cachés”, nouveaux? Cherche-t-on un diagnostic quantitatif ou qualitatif? Large ou profond?

L’évaluation est grandement facilitée si un référentiel de formation a été établi au préalable avec des indicateurs **pertinents**, **mesurables** ou facilement observables, **fiables**, **compréhensibles** et, si possible, liés entre eux, corrélés. D’un bilan à l’autre, choisir un corps d’indicateurs identiques pour pouvoir mesurer les changements: **continuité de l’observation dans une approche comparative**.

Le tableau suivant donne un aperçu des techniques les plus couramment utilisées. Chacune a ses avantages et ses limites: il est important de les combiner pour une évaluation de qualité. Quelques conseils pratiques pour bien maîtriser ces outils.

| Technique | | Description | Echantillon | Utilisation |
| --- | --- | --- | --- | --- |
| **Observation** | Pré-enquête avec carnet de bord,  participante ou non | recueil de notes, propos, événements, questions | en continu  aléatoire | - permet de centrer l’étude sur les faits significatifs  - vision globale et profonde du contexte  - spontanéité  - hétérogénéité des infos  - comparaisons difficiles  - fidélité invérifiable |
| **Documentation** | Fiche de lecture | ouvrages édités et inédits | principe de saturation | - enrichit et renforce la pertinence du contenu  - appuie l’argumentation  - permet d’éviter les répétitions et d’avoir une vision globale |
|  | Fermé | Oui/Non | 1000 u maximum | - quantification rapide  - définition d’une tendance  - questions sas  - première typologie  - corrélation rapide |
| **Questionnaire** | QCM | plusieurs réponses proposées | 500 u maxi | - plus libre  - plus profond et précis |
|  | Ouvert | espace de réponse | 100 u maxi | - plus créatif, plus riche  - analyse de contenu  - traitement qualitatif des données  - permet de définir QCM et questions fermées a posteriori |
|  | Structuré | questions ordonnées préétablies | 50 u maxi | - permet les comparaisons  - intéressant pour un nombre important de personnes |
| **Entretien** | Centré | questions préétablies dans le désordre | 30 u maxi | - plus de liberté  - interprétation plus compliquée  - analyse de contenu |
|  | Libre | questions qui se surajoutent au fil de l’entretien | 10 u maxi | - degré de liberté le plus grand  - analyse de contenu  - pré-enquêtes |
| **Récits de vie,** | Croisés | un seul thème pour plusieurs vies | 4- 5 u maxi | - permet de quantifier les invariants |
| **biographies** | Uniques | oral, provoqué et intégral | 1 u maxi | - pour évaluation ex-post 4 à 5 ans après |
| **Reportages sonores, audiovisuels** | Photos, films, cassettes audio… | enregistre-ments pour documenter observations et analyses | S.O. | - facilite la compréhension des changements  - permet de mémoriser plus fidèlement |

**👍** Questionnaires

. Le **choix des mots** est une étape difficile et importante. Puiser dans les pré-enquêtes les termes les plus clairs et les plus compréhensibles par les enquêtes (…et les enquêteurs). Sinon, imprécisions et incompréhensions sont des sources importantes d’erreurs. Tester les questionnaires.

. Une **bonne question** suscite une réponse contenant l’information cherchée et n’exerce pas d’influence sur le sens de la réponse. Tous les enquêteurs doivent avoir la même compréhension : s’ils n’ont pas testé eux-mêmes les questionnaires, prévoir une réunion de clarification sur les formulaires.

. Ne pas dépasser 35 à 40 questions, soit 1.00 à 1.30 d’entretien grand maximum.

. Progresser des questions **les moins impliquantes**, les plus “faciles” aux questions **les plus personnelles**, les plus complexes.

. **L’ordre** des questions doit être **logique** et se succéder naturellement.

. Préparer l’entretien en **exposant les buts de la recherche** et en se présentant.

. Si possible, habiller les questions avec un fait connu qui s’y rattache.

. L’évaluateur doit être perçu comme quelqu’un qui comprend l’enquêté.

**👍** Entretiens

. Ce sont les questionnaires qui permettent de choisir les personnes à interroger.

. La difficulté de l’entretien réside dans la qualité de la communication, **l’empathie** enquêteur/enquêté qui permet d’éviter les blocages, le tri inconscient des réponses et les non-réponses. C’est beaucoup plus un art qu’une technique.

. L’évaluateur doit éveiller **l’intérêt de l’enquêté**, le rassurer, le mettre en confiance et **garantir l’anonymat**.

. Il doit faire **partager son enthousiasme** pour l’évaluation, qualité première.

. Dilemme entre *intervenir le moins possible* au risque de voir se faner l’entretien et le *relancer* avec le risque d’influencer les réponses.

. Les entretiens suscitent divers types de réponses:

1. les réponses spontanées: la question est large, ouverte et neutre. C’est la partie d’échauffement de l’entretien. Les réponses stéréotypées, banales, communes se rangent dans cette catégorie.
2. les réponses obtenues après relance, explications, répétitions avec un degré minimum d’intervention.
3. les réponses suscitées lors de conduites d’évitement, quand l’enquête “tourne autour du pot”, “botte en touche” ou répond à côté. Ne pas insister mais revenir plus tard, en reformulant la question. Noter la “résistance”.
4. les réponses provoquées, par un jeu de miroir comme “Et si je vous disais que certains pensent exactement le contraire de vous…”

. Il est préférable que l’enquêté soit **seul** (pas facile, n’est-ce pas?).

. Obligatoire: **Restituer les résultats** de l’évaluation, au moins à celles et ceux qui l’ont alimentée. C’est une question de respect, de déontologie et de…communication !

**👍** Echantillonnage

. Le terme *population* désigne l’ensemble des éléments à évaluer: hommes, plantes, objets, animaux… Dans la plupart des évaluations, nous étudions la population, soit la quasi totalité des acteurs, etc. La question de la représentativité ne se pose pas.

. Pour une population importante, qui nécessiterait des moyens démesurés pour une évaluation, il faut n’en étudier qu’une partie et déterminer un *échantillon*.

. Pour pouvoir généraliser, extrapoler, étendre les constats faits sur l’échantillon, il est indispensable de s’assurer de sa *représentativité*.

. La taille de l’échantillon ne dépend pas de la taille de la population; elle dépend de:

1. l’**homogénéité** ou de l’**hétérogénéité** de la population
2. la précision sur les données à recueillir
3. la finesse de l’analyse

. Pour choisir un échantillon représentatif, il faut réfléchir, émettre des hypothèses sur les facteurs qui peuvent modifier un constat; ces facteurs s’appellent les *variables*. Ces variables sont définies avant l’évaluation. Elles dépendent des questions qu’on se pose, du type d’action observée… Exemple: année de démarrage du programme, sexe, âge, lieu de résidence, fonction, activités… A chaque évaluation, il faut donc redéfinir une à trois variables susceptibles d’influencer les résultats.

. Les variables étant définies, il faut calculer le pourcentage, la proportion de la population concernée par chaque variable. Il faut *stratifier* l’échantillon, puis choisir au hasard les éléments de la population qui, additionnés, respectent les proportions observées dans la population.

. Toute population inférieure à 100 ne relève pas de l’enquête statistique.

. Généralement, un échantillon représentatif est constitué par 5 à 10% de la population.

. Plus la population observée est homogène, plus l’échantillon peut être petit. A l’extrême, nous trouvons la méthode empirique (et non probabiliste) qui consiste à réaliser un *sondage* basé sur des *quotas*. Le sondage est surtout utilisé à l’échelle d’un pays.

**Dépouillement**

. S’installer dans un endroit calme, confortable, entièrement consacré pour quelques jours au dépouillement.

. Le dépouillement nécessite un minimum de concentration: en équipe, il est préférable de se répartir le travail par groupe d’acteurs ou par thèmes, puis de **travailler seul**.

. A contrario, une saisie informatique de chiffres peut être valablement réalisée à deux.

. Vérifier le nombre de fiches (entretiens, questionnaires…) recueillies par rapport aux prévisions des termes de références:

1. **justifier les écarts** avec les termes de référence: disparition, oubli, erreur…
2. **vérifier la représentativité** de votre échantillon

. Classer logiquement et numéroter chaque fiche.

. Choisir une grille standard de dépouillement, reporter le N° des fiches, puis les photocopier.

. Si possible, toutes les informations quantitatives doivent être saisies et traitées sur logiciel (ex : Excel): les calculs sont fiables, rapides et peuvent rapidement être illustrés sur graphiques.

. Les chiffres parlent… si on les fait parler: observer en particulier la distribution des résultats et les moyennes pour donner plusieurs angles aux analyses.

. Les questions fermées et les QCM ne présentent aucune difficulté: ne pas oublier de rajouter une ligne “sans réponse”.

. Questions ouvertes

Les réponses sont plus riches, diverses donc plus compliquées à traiter: **trouver un équilibre** entre des *items trop réducteurs* (penser aux objectifs de l’évaluation) et des *réponses trop détaillées* (quel est l’intérêt, la portée opérationnelle de cette nuance? Est-ce pertinent?). Partir large, puis additionner certaines réponses voisines en fin de dépouillement.

. Dès qu’une question est traitée, additionner les réponses (sous-totaux, puis total): observez la distribution des résultats suivant les acteurs ou les établissements par exemple.

. Calculer immédiatement les pourcentages: par pitié, pas de décimales! Arrondir par excès ou défaut, puis s’assurer que la somme est au moins égale à 100%.

. Noter sur feuille libre les hypothèses, explications, idées et recommandations qui vous viennent à l’esprit. Plus tard, il est possible que vous en ayez oublié une partie! De plus, vous vous faciliterez les futures analyses et la rédaction.

. Pour les questions ouvertes, notez dans la marge les **réponses** qui vous paraissent **intéressantes, représentatives** ou **humoristiques**: ce sont de précieux matériaux pour argumenter et renforcer la lisibilité d’une étude.

## Analyses et synthèse

. L’observation attentive de chaque résultat amène l’évaluateur à le commenter. Il se pose la question **“Pourquoi un tel résultat ?”**. Il cherche quels sont les facteurs qui ont pu l’influencer et émet des *hypothèses*.

. Pour extrapoler un résultat à la population, il est parfois utile de calculer la limite de confiance ou *intervalle de confiance*.

**Pp = Po +/- 1,96 x racine carrée de (Po x Q): N**

Pp: probabilité pour la population

Po: probabilité observée à partir de l’échantillon

1, 96: coefficient de précision qui correspond à 5% de marge d’erreur (2,58 = 1% d’erreur)

Q: complément de Po, soit 1 - Po ou 100 - Po

N: taille de l’échantillon

. Quand il y a de fortes différences entre les différentes composantes d’un échantillon, le calcul de la moyenne arithmétique par catégorie est plus pertinent que la moyenne générale. On observe alors la *distribution des fréquences* qui peut être illustrée par un histogramme (ex: nombre moyen de projets financés/Université).

. Chaque indicateur observé doit être **synthétisé**, **mis en relation** avec d’autres indicateurs et, si possible, **comparé** aux fréquences observées dans une précédente évaluation.

**Rédaction**

. Le plan du rapport est directement inspiré et adapté du “contenu” défini dans les termes de référence.

. L’évaluation est souvent destinée à un public hétérogène (niveau d’instruction, culture, centres d’intérêt): bailleurs de fonds, partenaires, acteurs évalués, décideurs politiques, étudiants, chercheurs… Tant pour le plaisir du lecteur que pour la lisibilité du document final, **varier les styles**:

1. ***descriptif*** (50 à 60% du texte): n’oublions pas qu’une évaluation est une forme de capitalisation. Une bonne description permet au lecteur de se faire sa propre opinion sur le sujet. Illustrer avec des photos, cartes, schémas…
2. ***analytique*** (30 à 40%): l’auteur de l’étude est dans une position privilégiée pour commenter les résultats. Veiller à la cohérence entre les recommandations et les observations enregistrées. Classer les recommandations par étapes, selon des degrés d’urgence, d’importance… et de faisabilité.
3. ***narratif*** (10 à 20%): déclarations, proverbes, parties d’entretiens donnent de la vie, du poids au texte et ils stimulent l’intérêt du lecteur. Mieux vaut une série de témoignages qu’une docte et ennuyeuse analyse de contenu égrenant une série de fréquences.

. Prévoir une présentation et une mise en page qui permette une **lecture rapide**: résumé, table des matières, titres résumant le contenu du chapitre, intertitres bien visibles…

. Les annexes peuvent utilement **compléter le corps du rapport**: sigles, termes de références, cartes, entretiens in extenso, base de données sur les apprenants, guide méthodologique, contrats…

1. Par Rémy CASTERAN, sur la base de ses connaissances et de son expérience en évaluation [↑](#footnote-ref-1)